HAL open science

Assimilation de données par filtrage pour les systèmes hyperboliques du second ordre - Applications à la mécanique cardiaque

Philippe Moireau

▶ To cite this version:

Philippe Moireau. Assimilation de données par filtrage pour les systèmes hyperboliques du second ordre - Applications à la mécanique cardiaque. Modélisation et simulation. Ecole Polytechnique X, 2008. Français. NNT : . tel-00358582

HAL Id: tel-00358582

https://pastel.hal.science/tel-00358582

Submitted on 3 Feb 2009

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.

ÉCOLE DOCTORALE DE L'ÉCOLE POLYTECHNIQUE

Thèse présentée en vue de l'obtention du grade
de Docteur de l'École Polytechnique
Spécialité Mathématiques Appliquées

Assimilation de données par filtrage pour les systèmes hyperboliques du second ordre. Applications à la mécanique cardiaque.

Philippe MOIREAU

soutenue publiquement le 12 décembre 2008 devant un jury composé de :

Jacques BLUM	Président
Alain BENSOUSSAN	Rapporteur
Frédéric BOURQUIN	Rapporteur
Grégoire ALLAIRE	Examinateur
Cristian LORENZ	Examinateur
Annie RAOULT	Examinatrice
Michel SORINE	Invité
Dominique CHAPELLE	Directeur de thèse
Patrick LE TALLEC	Directeur de thèse

à M. C. & F.

Remerciements

Un seul être vous manque, et tout est dépeuplé

— Alphonse de Lamartine *L'isolement*

En premier lieu, j'aimerais exprimer mes plus profonds remerciements à Dominique Chapelle qui m'a tellement apporté au fil de ces trois ans tant au point de vue scientifique qu'humain. Sa facilité à pouvoir analyser chacun des problèmes qui s'offrait à nous restera une source d'inspiration et d'admiration pour toute ma carrière. Un immense merci aussi à Patrick Le Tallec qui, en plus d'avoir été à l'initiative de ma rencontre avec l'équipe de Dominique, a toujours su renforcer la dynamique de notre travail par son expérience mélangeant tous les aspects de la mécanique et des mathématiques appliquées. Cette thèse ne serait certainement pas ce qu'elle est aujourd'hui sans la complémentarité et la complicité qui se sont nouées entre nous tous dont je garde au moins autant de satisfaction que de ce qui a trait aux aspects scientifiques. D'ailleurs, Patrick doit se souvenir avec fierté que c'est finalement son fils, Yann, que l'on devrait remercier ici pour nous avoir réunis autour d'un dîner un soir de fin décembre 2004.

Je tiens aussi à remercier chaleureusement mes deux rapporteurs, Alain Bensoussan et Frédéric Bourquin qui m'ont fait l'honneur de relire avec beaucoup d'attention ce manuscrit. La façon d'enseigner le filtrage d'Alain Bensoussan a largement inspiré le premier chapitre de cette thèse et j'ai beaucoup appris des remarques de Frédéric Bourquin sur la contrôlabilité exacte pour les structures. Un grand merci aussi aux autres membres du jury, Annie Raoult, Cristian Lorenz, Jacques Blum, et Grégoire Allaire qui m'a vu évoluer depuis les bancs de l'Ecole Polytechnique et a largement contribué, à l'époque avec Patrick, à mon orientation. Enfin, j'ai une pensée particulière pour Michel Sorine qui a suivi mes trois années de thèse avec la disponibilité d'un quasi directeur de thèse. C'est un plaisir de voir à l'oeuvre, autour d'un café, toute sa créativité sur des sujets aussi variés que l'automatique, les EDP, la physiologie, etc... sans parler de la pomme.

L'INRIA a organisé son fonctionnement autour d'équipes-projets de petite taille dans lesquelles se développe une véritable dynamique de travail. Ce travail a ainsi largement bénéficié de l'aide des membres de Macs, Jacques Sainte-Marie et Marina Vidrascu. Mais au delà des aspects professionnels, j'ai eu le sentiment de m'intégrer dans une véritable petite famille à laquelle on peut sans doute associer tous les « cousins » de Reo avec notamment Miguel Fernandez, Jean-Frédéric Gerbeau mais aussi Céline Grandmont et tous les autres. A Rocquen-

court, l'isolement n'est que de façade quand on voit les liens qui se tissent entre les différentes équipes, en particulier au *bâtiment 16* avec les équipes Bang et Gamma. Et puis, pour des raisons qui n'échapperont à personne, la logique des bâtiments se fissure aussi avec toutes les relations fructueuses et amicales qui se sont nouées avec l'équipe-projet Poems de Patrick Joly et Anne-Sophie Bonnet Ben-Dhia. En particulier, merci à Jérôme Le Rousseau pour m'avoir fait découvrir la richesse de l'analyse micro-locale dont j'espère profiter dans mes recherches futures sur les aspects de contrôles sous-jacents au filtrage. Je voudrais aussi étendre mes remerciements à tous les membres des équipes INRIA de l'action transversale Cardiosense3D avec qui j'ai eu la chance de travailler, dans un cadre fertile spécialement dédié à la recherche en modélisation cardiaque.

Mais au delà de l'INRIA, j'ai une pensée particulière pour Elsa Angelini de l'institut Telecom et Columbia University pour la collaboration dynamique et chaleureuse que nous avons pu mettre en place, avec l'espoir qu'elle continue durablement. J'en profite également pour exprimer tout mes remerciements à Guillaume Bal pour tout le temps qu'il m'a accordé avant le début de cette thèse dans ce qui était ma première expérience de recherche. Elle m'a beaucoup apporté pour la suite et c'est certainement la meilleure occasion de le remercier a posteriori.

Reste sans doute le plus important, ma maman qui, en plus de tout ce qu'elle représente, me donne la joie de comprendre ce que je fais et d'y croire presque autant que moi, avec le sentiment qu'elle continue de me transmettre les valeurs et l'attachement de mon père. Merci aussi à mes grands frères qui m'accompagnent avec cette affection bienveillante qui vous donne l'impression malgré la distance que vous faites toujours les bons choix. Je pense également très fort à Bénédicte qui me fait profiter de son expérience et Clarinette qui me laisse en faire de même avec elle. Merci aussi à tous mes amis qui ont développé une patience sans limite à voir le film, relire (ou revoir...) le manuscrit et écouter ma thèse. Je ne leur dis certainement pas assez combien j'en suis touché, alors encore une fois merci pour tout à Thierry, JC, Paul, Guillaume, Céline, Flo et Christelle mais aussi Marc & Sarah, Ophélie, Delphine, Manu, Sophie, Céline, Ôrel, Aurélien, Thomas, Hugues, Benjamin, et bien sûr Samuel, Florian et Quentin. Mention spéciale à Manon qui, à 2 ans, m'a ouvert de nouvelles perspectives dans mes recherches bibliographiques ! J'ai une pensée particulière aussi pour Nicolas qui d'une certaine façon m'a longtemps montré la voie, jusqu'au choix de format de présentation de ce manuscrit. Enfin, comment remercier suffisamment Sonia qui m'apporte à elle seule tout ce qu'on peut imaginer : son soutien, sa générosité, sa patience pendant ces trois ans alors même qu'elle devait faire face au même long fleuve agité du cheminement scientifique et hélas bien plus. Elle a réussi à toujours trouver le temps de m'écouter raconter inlassablement les mêmes histoires et de comprendre, et c'est unique, mes théories scientifiques (ou pas !) les plus farfelues. Je pourrais vivre à Paris et partir travailler de l'autre côté de la Terre tous les matins si Sonia pouvait le faire avec moi.

Résumé

L'objectif de cette thèse est de formuler des méthodes d'assimilation de données adaptées à la simulation du comportement mécanique du coeur tout au long d'un battement cardiaque. La motivation est de bénéficier dans la compréhension du comportement de l'organe à la fois du développement des techniques d'imagerie cardiaque et de l'interêt croissant des cliniciens pour la modélisation et la simulation numérique. Pour ce faire, il est nécessaire d'adapter (c'est à dire d'estimer des quantités incertaines dans le modèle telles que les conditions initiales, certains paramètres et les conditions aux limites en plus de l'anatomie) des simulations génériques aux spécificités d'un patient (potentiellement malade, donc) donné.

Nous présentons une stratégie nouvelle adaptée à l'estimation de systèmes mécaniques, et plus généralement de systèmes hyperboliques du second ordre, avec des conditions initiales et des paramètres inconnus. Notre idée est d'adopter une approche par filtrage sur la formulation en système dynamique du problème afin de construire un estimateur conjoint état-paramètres utilisant les mesures disponibles. Le fondement de cette approche est d'obtenir un système dynamique modifié qui converge vers la trajectoire cible en incorporant des termes de correction liés aux données. En premier lieu, en supposant les paramètres connus, l'estimation d'état est réalisée en utilisant *des observateurs de Luenberger* exploitant la stabilisation par *feedback* à des fin d'estimation. Ces types d'observateurs sont reconnus pour leur potentielle efficacité et robustesse. Notamment, à la différence d'approches « Kalmaniennes » classiques, ces filtres peuvent être numériquement adaptés à des systèmes issus de la discrétisation d'EDP et (bien que non optimaux au sens des moindres carrés) la stabilité exponentielle du système de l'erreur d'observation sous-jacent assure la convergence de l'estimateur. Ainsi, nous analysons en particulier la stratégie collocalisée du Direct Velocity Feedback (DVF) utilisée en contrôle des structures. Nous formulons aussi une méthode originale dans le cas de mesures de déplacements voire de positions (et par extension de contours dans une image) bénéficiant de la « dualité » entre la contrôlabilité de systèmes réels et l'observation à partir de systèmes numériques.

Lorsque les paramètres sont incertains, nous étendons alors l'estimateur en ajoutant une dynamique paramétrique. L'effet des observateurs d'état précédents est alors de restreindre l'incertitude à l'espace paramétrique (d'ordinaire de plus petite dimension que pour l'état) afin d'y appliquer *des filtres de rang réduit* H^2 ou H^∞. La convergence de l'estimateur résultant peut alors être mathématiquement démontrée et nous l'illustrons en estimant (à partir de données synthétiques précisément modélisées) des paramètres distribués de type raideurs et contractilités avec la perspective d'aide au diagnostic de régions infarcies du muscle cardiaque.

Abstract

The objective of this thesis is to formulate data assimilation methodologies adapted to the simulation of the heart mechanical behavior over a complete beat. The motivation is to benefit from both the development of cardiac medical imaging and the increased interest of clinicians for *in-silico* modeling and simulation. In this respect, it is now necessary to be able to adapt – that is, not only to adapt the anatomy of the patient but also to estimate unknown quantities in the model such as initial conditions, parameters, boundary conditions – generic simulations to the behavior of a specific – potentially ill – patient.

We present a novel strategy to perform estimation for mechanical systems, and more generally for second-order hyperbolic equations with uncertain initial conditions and parameters. We decide to adopt a filtering approach on the dynamical system formulation to construct a joint state-parameter estimator that uses some measurements available in standard operating conditions. Namely, the aim is to obtain a modified dynamical system converging to the reference by incorporating correction terms using the data. First, in the case of known parameters, state estimation is performed with a filter based on Luenberger observers which use feedback control strategies for estimation purposes. These types of state estimators can be selected according to their particular effectiveness and robustness. In particular, unlike the classical Kalman approach, these filters can be computationally tractable for numerical systems arising from the discretization of PDEs and – although optimality is lost – the exponential stability of the corresponding controlled system gives exponential convergence of the estimator. In this respect, we more specifically analyze collocated strategies as the Direct Velocity Feedback (DVF) arising in the control of mechanical structures. We then formulate an original observer approach when the available measurements are displacements or positions – with extension to contours tracking in images – taking full advantage of the "duality" between an actual controlled system and the *in-silico* observer system.

With uncertain parameters we extend the estimator by incorporating the parameters in an augmented dynamical system. The effect of the first stage state filter then consists in essence in circumscribing the uncertainty to the parameter space – which is usually much "smaller" than the state space – and allows for an H^2 or H^∞ type filter in the resulting low rank space. This second step is related to *reduced rank filtering* procedures in data assimilation and observer theory. The convergence of the resulting joint state-parameter estimator can be mathematically established, and we demonstrate its effectiveness by identifying localized parameters. The test problems considered, motivated by cardiology, are based on synthetic measurements available in medical imaging. They focus on the identification of stiffness and contractility parameters in the mechanical model, with the perspective of diagnosis assistance for infarcted regions in the cardiac tissue.

Table des matières

Introduction **1**

CHAPITRE 1 Principe d'estimation moindres carrés **13**

1.1 Introduction à l'estimation bayesienne 14
1.2 Inverses généralisés au sens des moindres carrés 20
1.2.1 Inverse généralisé d'un opérateur d'observation 21
1.2.2 Terme d'ébauche 23
1.2.3 Estimation récursive 25
1.3 Estimation pour les systèmes dynamiques linéaires 29
1.3.1 Cas discret en temps : Cadre déterministe 29
1.3.2 Cas discret en temps : Cadre probabiliste 34
1.3.3 Cas continu en temps 35
1.3.4 Bruits de modèle 42
1.3.5 Notions d'observabilité, de contrôlabilité et d'identifiabilité 45
1.4 Extensions autour de la théorie LQ 48
1.4.1 Filtre non-linéaire déterministe 48
1.4.2 Contrôle robuste 53
1.4.3 Filtre EKF 55
1.4.4 Filtre UKF 59
1.5 Conclusion 65

CHAPITRE 2 Modélisation cardiaque **67**

2.1 Le cœur 68
2.2 Eléments de physiologie cardiaque 69
2.2.1 Anatomie Cardiaque 69
2.2.2 Cycle cardiaque 71
2.2.3 Formation et conduction de l'excitation électrique 73
2.2.4 La fonction cardiaque 75

2.2.5 Le réseau vasculaire 78
2.2.6 Physiologie détaillé des tissus musculaires cardiaques 80
2.3 Modélisation électromécanique cardiaque 85
2.3.1 Formulation mécanique en Lagrangien total 85
2.3.2 Loi de comportement mécanique 88
2.3.3 Modèles de circulation 100
2.3.4 Modèle cardiaque complet 105
2.3.5 Conditions aux limites 105
2.3.6 Eléments de modélisation électrique 107
2.4 Analyse numérique et simulations 112
2.4.1 Modèles géométriques 112
2.4.2 Discrétisation 124
2.4.3 Calibration du modèle 130
2.4.4 Simulations numériques 131
2.5 Observations cardiaques 136
2.5.1 IRM 136
2.5.2 Echographie 143
2.5.3 Mesures de pression 152
2.5.4 Mesures électriques 152
2.6 Modèles et observateurs en 1D 153
2.7 Conclusion 156

CHAPITRE 3 Estimation d'état pour une classe de systèmes hyperboliques 157

3.1 Introduction 158
3.2 Estimation d'état moindres carrés 161
3.2.1 Formulation variationnelle spécifique 161
3.2.2 Approche séquentielle et complexité 165
3.2.3 Problématique 167
3.3 Filtrage asymptotique de Luenberger 168
3.3.1 Système d'erreur 168
3.3.2 Fonctions de Liapounov 171
3.3.3 Stabilisation et contrôle 175
3.4 Estimation d'état en élastodynamique à partir de vitesses 177
3.4.1 Rappels sur la position du problème 177
3.4.2 Définition de l'estimateur DVF 182
3.4.3 Estimations en dimension infinie issues du contrôle 183
3.4.4 Problème aux valeurs propres quadratiques 187
3.4.5 Estimation d'erreur en fonction des bruits 197
3.4.6 Observateur frontière 206
3.4.7 Extension en dynamique non-linéaire 212

3.5 Estimation d'état en élastodynamique à partir de déplacements 215
3.5.1 Définition de l'estimateur 216
3.5.2 Analyse de l'estimateur 219
3.5.3 Schéma en temps 224
3.5.4 Résultats d'estimation 227
3.5.5 Cas d'observations eulériennes 227
3.5.6 Résultats 233
3.5.7 Extension en dynamique non-linéaire 233
3.5.8 Résultats d'estimation d'état pour le cœur 237
3.6 Conclusion et perspectives 243

CHAPITRE 4 Estimation conjointe état-paramètres pour les systèmes distribués 245

4.1 Introduction 246
4.2 Cas linéaire 248
4.2.1 Approche variationnelle 249
4.2.2 Estimation état-paramètre par filtrage réduit 250
4.2.3 Application à l'estimation paramétrique 252
4.2.4 Analyse de l'estimateur 255
4.2.5 Schéma numérique associé et analyse de stabilité 263
4.2.6 Résultats 265
4.2.7 Opérateur frontière et Formulation H infini 267
4.3 Cas bilinéaire et extensions du filtre SEEK 276
4.3.1 Expression à partir de la sensibilité 276
4.3.2 Analyse d'erreur 277
4.3.3 Premiers résultats 277
4.3.4 Amélioration des performances de l'estimateur conjoint en non-linéaire . 280
4.3.5 Equivalence avec EKF et le filtrage non-linéaire 282
4.4 Cas non-linéaire et filtre SEIK 283
4.4.1 Filtre SEIK 284
4.4.2 Application à l'estimation paramétrique 293
4.5 Perspectives autour du système cardiaque 299
4.5.1 Estimation pour *des paramètres dynamiques* 299
4.5.2 Implications pour l'estimation cardiaque 301

Conclusion et perspectives 305

BIBLIOGRAPHIE 309

Introduction

Même obtenue à l'aide d'images possibles, quoi de plus réel, bien que contingent, qu'une guérison.

— Michel Serres, *La Recherche, janvier 2006*

En essayant d'imiter dans sa démarche la rigueur inflexible d'on ne sait quel automate, tout en portant son plateau avec une sorte de témérité de funambule, en le mettant dans un équilibre perpétuellement instable et perpétuellement rompu, qu'il rétablit perpétuellement d'un mouvement léger du bras et de la main [...] Il s'applique à enchaîner ses mouvements comme s'ils étaient des mécanismes se commandant les uns les autres.

— Jean Paul Sartre, *L'Être et le Néant, « Le garçon de café »*

Le titre de cette thèse évoque la conjonction de différents champs scientifiques que sont la médecine, la mécanique et l'assimilation de données. La médecine tout d'abord, qui est un champ d'applications de plus en plus présent dans les sciences de l'ingénieur. Les récents efforts de repositionnement des politiques scientifiques de l'INRIA, ou du CEA, autour de la simulation pour le vivant en sont des exemples français parmi d'autres. Plus particulièrement, il existe aujourd'hui de plus en plus de simulations numériques sur des systèmes biologiques, notamment humains : propagation d'ondes dans le cerveau, écoulements sanguins ou pulmonaires, ostéoporose, croissance cellulaire, et dans notre cas la simulation de la contraction cardiaque. Cette émergence d'applications nouvelles est portée par le développement sans précédent des outils d'acquisition de données, notamment d'imagerie médicale (scanner, IRM, échographie 3D...) qu'on souhaite désormais coupler à une modélisation fine des phénomènes physiques comme ce fut le cas pour d'autres systèmes industriels. En particulier, la mécanique des structures trouve autour de la modélisation du myocarde un exemple original de matériau complexe à la fois multi-échelles, composite, anisotrope, actif et même vivant qui représente un nouvel enjeu scientifique avec une double exigence méthodologique de modélisation et d'identification quantitative de ses composantes. Quant à l'assimilation de données, ce terme a pour origine la météorologie et plus généralement les sciences de l'environnement. Il désigne toute la gamme de problèmes inverses consistant à identifier les grandeurs inconnues d'un problème d'évolution à partir d'observations réelles (toujours plus nombreuses et riches). Ainsi la dynamique cardiaque, suivie au travers des nouvelles modalités d'imagerie dans l'ob-

jectif de l'identification des caractéristiques spécifiques de l'organe, notamment pathologiques, semble se prêter parfaitement à l'extension de cette discipline à un nouveau champ d'application, la médecine, et disciplinaire, l'élastodynamique au sens large. C'est donc une thèse pluridisciplinaire qui, entre biomécanique et « météo du cœur » souhaite d'abord proposer des méthodes mathématiques originales pour l'assimilation de données sur des systèmes caractéristiques en élastodynamique, linéaire et non-linéaire, et plus généralement les systèmes hyperboliques du second ordre.

Contexte

□ ENJEUX DE SOCIÉTÉ AUTOUR DES MALADIES CARDIO-VASCULAIRES – Les maladies cardio-vasculaires sont désormais en Europe la deuxième cause de mortalité après les maladies cancéreuses avec 4,35 millions de décès par an, et encore la première aux Etats-Unis. D'ici 2010, elle représentera la première cause de décès dans les pays émergents. Les maladies cardio-vasculaires les plus fréquentes sont les accident vasculaires cérébraux (AVC, qui ne nous occuperont pas dans cette thèse), les maladies coronariennes conduisant à l'infarctus du myocarde, les troubles du rythme ou arythmies cardiaques, mais il existe aussi un très grand nombre de pathologies moins fréquentes, notamment celles liées à des malformations congénitales au niveau des ventricules, oreillettes ou valves cardiaques. Ces maladies ont un impact sur l'économie européenne d'environ 169 milliards d'euros par an (voir Petersen et al (2005)).

Or le diagnostic précoce de ces maladies constitue un enjeu majeur pour assurer leur recul. Il a d'ailleurs d'ailleurs permis, depuis deux décennies, de limiter statistiquement le développement des maladies cardio-vasculaires face aux maladies cancéreuses. Cependant, ce sont principalement les accidents vasculaires cérébraux qui ont chuté (d'un tiers entre 1982 et 1994) avec une meilleure prise en charge de l'hypertension artérielle ; *a contrario* le taux de mortalité des maladies cardiaques n'a diminué que de 8% sur la même période et reste un défi majeur malgré la généralisation des appareils d'imagerie (notamment IRM) dans les pays développés.

□ MODÉLISATION MATHÉMATIQUE EN PHYSIOLOGIE CARDIAQUE – Le cœur, organe vital par excellence, est le siège de phénomènes physiques extrêmement divers qui éveillent naturellement la curiosité de tout modélisateur. A l'origine de la contraction de l'organe, on trouve une activité électrique régulière correspondant à la propagation d'une onde de dépolarisation à la surface du coeur. Cette activité, on la « ressent » quelquefois (notamment dans des situations pathologiques telles que la fibrillation ventriculaire qui précède l'infarctus) mais surtout on la mesure à travers l'examen électrocardiographique (ECG). Répondant à cette activation, le coeur en tant qu'entité mécanique se contracte afin d'assurer sa fonction : celle de pomper le sang côté poumons pour qu'il soit réoxygéné, et vers le reste des organes pour les alimenter. Ce moteur, au sens thermodynamique du terme a, de plus, la particularité qu'il s'auto-alimente en oxygène via ses coronaires ; on parle alors de perfusion. Evidemment cette activité mécanique et fluide, on la ressent et on la mesure lors de l'auscultation, la prise de tension et désormais l'imagerie médicale notamment échographique et IRM. Autant de phénomènes donc qui, si on souhaite simuler le comportement cardiaque, mobilisent de nombreuses compétences en modélisation physique, en calcul scientifique et en traitement du signal.

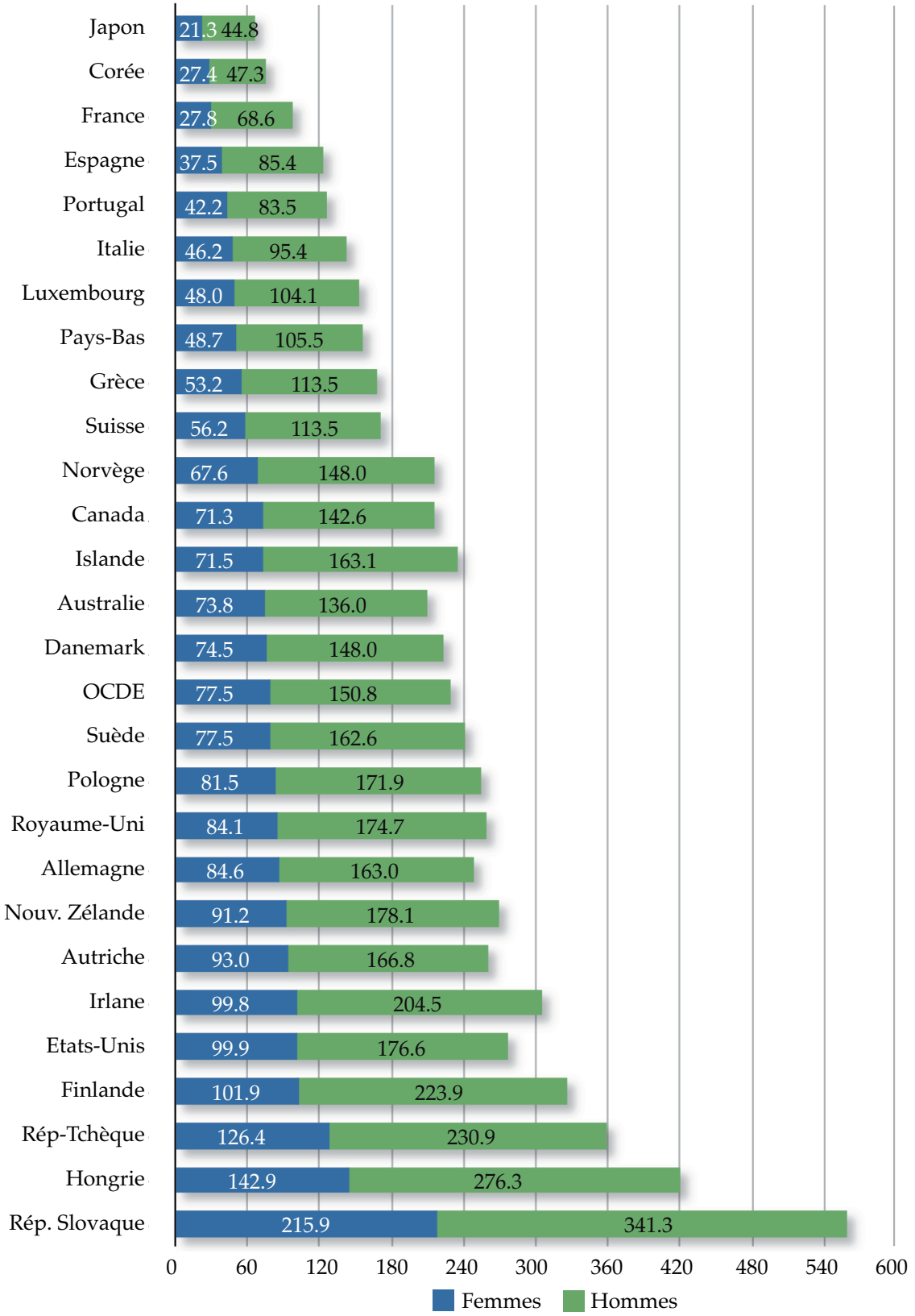


FIGURE 1: Maladies Cardiovasculaires, taux de mortalité sur 100 000 habitants en 2002, ***source ECO-SANTE OCDE 2005***

L'analyse mathématique des phénomènes biologiques et notamment physiologiques (c'est-à-dire traitant des fonctions des organes), qu'on appelle parfois physiologie mathématique, date déjà de plus d'un siècle avec les travaux de Helmoltz (dès 1851 en ophtalmologie) et Frank (plus près de notre sujet sur le muscle cardiaque en 1899). La notion de modèle au sens physique et mathématique du terme est donc ancienne et connaît sans doute en cardiologie sa plus forte avancée avec les travaux de A.F. Huxley et Hodgkins (Prix Nobel de Medecine en 1963) en électrophysiologie nerveuse et cardiaque, puis A.F. Huxley seul pour son modèle de contraction cardiaque à l'échelle de la fibre en 1957. Très tôt pour l'activité électrique cardiaque, les aspects à la fois de modélisation à l'échelle cellulaire mais aussi de compréhension de la propagation de l'onde électrique notamment pour ses conséquences pathologiques (la fibrillation notamment) à la surface de l'organe ont été étudiés intensivement. Ainsi d'un côté on progressait sur des modèles d'équations différentielles ordinaires (EDO) de comportement cellulaire, mais aussi sur les aspects distribués, donc équations aux dérivées partielles (EDP), du phénomène à travers des modèles de réaction-diffusion que Kolmogorov en 1937 appliquait déjà en électrophysiologie. En revanche, pour ce qui est des EDP en mécanique, la notion de mécanique des milieux continus intégrant la loi locale de contraction cardiaque a suscité nettement moins d'enthousiasme, comme si elle semblait acquise.

Une fois ces questions de modélisation révélées, l'intérêt de la communauté du calcul scientifique pour ces phénomènes physiologiques au même titre que d'autres phénomènes d'EDP datent des années 70. Sans rentrer dans les détails côté électrique on peut notamment citer les travaux précurseurs de Colli Franzone et al (1980) et de Panfilov et Winfree (1985) un peu plus tard. Côté mécanique c'est surtout Hunter (1975) en mécanique du solide et Peskin (1982) à la fois côté solide mais aussi fluide sanguin à qui on doit les premières contributions majeures.

□ UNE RELATIVEMENT FAIBLE PERCÉE APPLICATIVE – Etonnamment, alors que dans de nombreuses autres applications du calcul scientifique, des résultats de 1970 pourraient aujourd'hui sembler obsolètes au vu de l'amélioration conjointe des machines, des algorithmes et parfois des méthodes, ces résultats en mécanique cardiaque font encore figure de références. Les raisons sont à mon sens de plusieurs natures :

- D'abord les aspects numériques du comportement physiologique de l'organe coeur sont complexes car ils mettent en jeu des EDP de natures très diverses (visco-élastodynamique non-linéaire, Navier Stokes, réaction diffusion, électrostatique pour l'ECG). Autrement dit, même après une modélisation convaincante, la physiologie intégrative qui consiste à mettre en musique l'ensemble est un défi en soi et nécessite sans doute des capacités numériques *hautes performances*.

- Ensuite parce qu'il est très difficile d'être très confiant sur les qualités intrinsèques des modèles en l'absence d'une modélisation, fine notamment en mécanique des milieux continus, et d'expérimentations tissulaires. Or il est clair qu'en biologie le paramétrage et la quantification des constantes physico-chimiques constitutives du problème étudié sont à la fois essentiels et complexes. Se pose en effet la question de l'obtention de données sur lesquelles recaler les modèles de phénomènes in-vivo. Côté électrique, la communauté de la physiologie mathématique s'est toujours appuyée sur des données expérimentales permettant l'analyse de la dépolarisation chimique au niveau de la

membrane cellulaire. Côté mécanique, les expérimentations de type éprouvettes (au sens mécanique) à l'image de Huxley sont plus rares même si on peut citer les nombreuses expériences de Hunter (1975) côté loi de comportement active et passive pour les tissus cardiaques, de Fung (1993) sur les tissus en général, ou par la suite des expériences au niveau cellulaire à partir de pinces optiques (voir Cazorla et al (2003)), sorte d'éprouvettes microscopiques.

- Le dernier élément, sans doute lié aux deux premiers mais pas uniquement, est le relativement faible intérêt des médecins pour ce type de simulations. Autrement dit, à la différence d'autres disciplines où l'industrie cherche des solutions auprès de la communauté du calcul scientifique, ici c'est plutôt l'inverse où cette dernière cherche à convaincre les praticiens d'un nouveau besoin, avec un succès mitigé. Pour l'instant la seule communauté de mathématiques appliquées intégrée dans la chaîne médicale reste la communauté du traitement de signal et d'images, donc de l'analyse des données, avec en toile de fond trois grands industriels, Philips, General Electrics et Siemens, fournisseurs d'instruments de mesure tant au niveau machines que logiciels. Ces industriels répondent d'ailleurs aux attentes de la communauté médicale en l'aidant à formaliser l'analyse des données obtenues toujours plus riches et complexes mais sans chercher à apporter une réponse « physiologique donc médicale ».

Ainsi à l'heure actuelle, la simulation numérique en physiologie cardiaque est certes très riche et donne lieu à de nombreux développements mais constitue avant tout un outil explicatif, pédagogique au sens noble du terme. En revanche elle n'est pas encore perçue comme un outil prédictif potentiel pour la communauté médicale. C'est à mon sens la raison qui fait que ces résultats pionniers restent des références dans l'attente que la modélisation franchisse le nouveau cap de l'aide au diagnostic.

□ COUPLAGE MODÈLES-DONNÉES – Face à la difficulté de mettre en oeuvre des validations aux niveaux cellulaire et tissulaire du point de vue expérimental mais surtout parce que intrinsèquement on parle d'un organe qui doit être considéré dans son ensemble et dans son milieu in-vivo ; face aussi à la difficulté de définir des conditions aux limites raisonnables pour l'organe qui est bien évidemment entouré de tous les autres ou celle de définir dans quelle configuration de référence il est ; face à l'ensemble de ces incertitudes de modélisation donc qui réduisent définitivement la modélisation à être descriptive plus que prédictive, nous souhaitons utiliser les données, les observations issues de modalité de mesure considérée : imagerie pour la cinématique, ECG ou sondes pour l'électrique, cathéter ou dispositifs de mesures de pressions artérielles pour nourrir et adapter la modélisation. Cette notion d'assimilation de données existe aussi dans les communautés du calcul scientifique, notamment celles pour qui la confrontation modèle-données est absolument nécessaire : météorologie, climatologie, géophysique ou exploration pétrolière, et désormais le transport de polluants. Dans ces différentes situations on utilise les données pour pallier les incertitudes de modélisation comme en traitement d'images à base de modèle on utilise les modèles pour palier les manques et les bruits de mesures.

On constate alors que dans ces disciplines, modélisation, données, expérimentation et validation s'interconnectent au lieu de se succéder. Il y a un siècle c'était l'expérimentation qui validait une théorie physique. De nos jours, si comme Michel Serres le dit la simulation nu-

mérique est *le troisième pilier de la recherche*, elle reste un moyen soit d'expérimenter un modèle physique qui ne pourrait être soumis à l'expérimentation (très petites et très grandes échelles) soit elle est la représentation du modèle et exige d'être, elle aussi, validée par l'expérimentation (soufflerie en aéronautique). Pour autant, le référent est toujours défini. En assimilation de données on pourrait penser qu'il n'y a plus de référent puisque modèle, mesure et simulations sont si fortement interconnectés qu'on ne sait plus qui détermine ce qui est « vrai » de ce qui ne l'est pas. Le référent est alors transporté au niveau de l'application. Schématiquement, si le modèle et sa simulation recalée sur les données prouvent qu'ils peuvent fournir au praticien une information qui statistiquement est différenciante pour telle ou telle pathologie, alors ils sont pertinents vis-à-vis des données fournies et pour le type de pathologie étudiée. C'est le sens de la citation de Michel Serres en tête de cette introduction : la contingence c'est ce qui est nécessaire et donc une théorie n'est nécessaire que lorsqu'elle est confrontée à la solution qu'on veut qu'elle apporte. Cette idée est d'ailleurs très ancienne et très générale en épistémologie. Elle n'implique pas un relativisme logique avec en ligne de mire la simple efficacité statistique à l'image des premiers simulateurs médicaux basés sur le data mining dans les années 60. L'objectif était en effet de créer une sorte de médecin virtuel qui puisse détecter certaines pathologies simplement par analogie dans une base de cas. Ici, on doit aller le plus loin possible dans la propre validation physique de chacun des aspects théoriques : la mesure sur ses moyens d'acquisition et de traitement à partir de fantômes, la modélisation sur ses paramètres et ses simplifications à partir d'expériences cellulaires et tissulaires, l'assimilation de données via des problèmes inverses synthétiques. Mais en tout état de cause l'application, c'est-à-dire le diagnostic (dans quel état est le patient aujourd'hui) ou le pronostic (comment va-t-il évoluer), ne sera possible qu'avec ces trois éléments couplés ensemble et validés par rapport aux résultats de diagnostic et de pronostic. Avant cela, il ne peut y avoir d'application, ce qui explique sans doute la relative retenue de la communauté médicale à propos de la simulation numérique pour le vivant.

Enjeux théoriques

Le sujet de cette thèse s'inscrit donc dans ce contexte scientifique et dans la volonté de coupler modèles et données avec la perspective de pouvoir transformer l'outil pédagogique qu'est un modèle numérique cardiaque en un outil potentiellement de diagnostic et de pronostic, car adapté au patient considéré. L'originalité thématique est donc d'introduire la notion d'assimilation de données dans la communauté de la modélisation pour les sciences du vivant, comme elle l'est déjà dans d'autres communautés d'ingénierie.
Cependant, nous n'avons pas voulu faire de ce travail la simple mise en musique de méthodes déjà disponibles, mais véritablement revisiter ces méthodes dans le cadre de la mécanique cardiaque. *A fortiori*, nous ne sommes pas à un stade d'intégration de méthodes pour l'ensemble des phénomènes physiques concernés : ce n'est sans doute pas encore possible du point de vue méthodologique à cause des nombreux couplages electro-mécaniques et fluides-structures et, si ça l'était, cela deviendrait alors un défi d'ingénierie. Ici, nous particularisons le problème global à celui des aspects propres à la mécanique des milieux continus et nous envisageons d'y définir de nouveaux outils d'assimilation de données dont la portée espère dépasser le cadre cardiaque pour la mécanique des solides en général, voire d'autres problèmes d'évolu-

tions pour des EDP hyperboliques du second ordre (équation des ondes par exemple).

□ CONTEXTE THÉORIQUE – Avant de présenter notre démarche, j'aimerais donner quelques éléments de contexte théorique autour de l'assimilation de données. Ce terme est profondément associé aux disciplines dont j'ai parlé plus haut, notamment la météorologie, mais correspond à des théories plus anciennes d'automatique et de contrôle optimal, et plus particulièrement de théorie de l'observation. Celle-ci est en quelque sorte le dual de la notion de contrôle en automatique. En effet, en contrôle on considère le problème de commander un système dynamique pour l'amener d'un point A à un point B, et éventuellement de manière optimale c'est-à-dire en minimisant par exemple le temps pour y parvenir, l'énergie déployée, ou tout autre critère. En observation, la problématique est de déterminer l'état d'un système à partir de mesures partielles disponibles. D'ailleurs de la notion d'observation émerge la notion d'identification, c'est-à-dire celle de retrouver un ensemble de paramètres à partir des observations. On parle souvent d'état du système dynamique pour les variables qui évoluent au cours de la dynamique et qui caractérisent le système mais cette notion est toute relative puisque les paramètres peuvent aussi être une partie « statique » de l'état du système. La frontière entre identification et observation (d'état) est ainsi très ténue et on parlera aussi d'estimation au sens large pour regrouper toutes ces notions. En observation comme en contrôle, on peut aussi définir une notion d'optimalité si on souhaite minimiser une certaine norme entre ce qu'on a observé initialement du système et ce qu'il donnerait comme observations dans l'état. On rejoint alors les principes de l'estimation de type moindres carrés au sens de Gauss pour des normes bien choisies.
Pour résumer schématiquement cette dualité contrôle-observation on peut donc dire que dans le premier on connaît l'état et on cherche une trajectoire et dans l'autre on connaît une trajectoire et on cherche un état. Si le contrôle est « exogène » au sens où il est indépendant de l'état du système on parle de commande en boucle ouverte. Mais le plus souvent on s'intéresse à des contrôles qui s'ajustent en fonction de l'état du système appelés commandes en boucle fermée. Dans ce cas, les notions de contrôle et d'observation se rejoignent puisque le système effectue des observations pour déterminer son état et assurer le contrôle. C'est typiquement le cas d'une poursuite de missile où le missile adapte sa trajectoire en fonction de l'écart entre sa position et celle de l'avion. En plus de ces deux notions une troisième apparaît rapidement qui est la notion de stabilisation. La stabilisation correspond à maintenir un système dynamique dans une position d'équilibre. Autrement dit si les jambes d'un enfant jouant à la balançoire lui permettent de contrôler l'amplitude, il peut les utiliser pour arrêter la balançoire plus rapidement, c'est-à-dire la stabiliser sur sa position de repos. Stabilisation, contrôle et observation sont finalement intimement liés puisque contrôler un système c'est stabiliser l'erreur entre le système et la cible et de même pour l'observation.
Nous avons donc quatre notions fondamentales en automatique : contrôle (ou commande), observation et identification (estimation), et enfin stabilisation.

Initialement concentrée sur les systèmes dynamiques de dimension finie (donc d'EDO), l'automatique a aussi permis le développement de théories autour des équations aux dérivées partielles d'évolution grâce notamment aux travaux de Lions (1968) en contrôle et Bensoussan (1971) en observation. D'ailleurs la citation du garçon de café de Sartre est un clin d'oeil aux travaux de Coron (2007) en stabilisation non-linéaire sur les équations de Saint-Venant. L'assimilation de données dans les disciplines évoquées plus haut prend alors racine dans ces

concepts d'automatique pour les équations aux dérivées partielles par l'approche du contrôle optimal puisque l'idée est bien de chercher le « meilleur » état x à un instant t satisfaisant un certain nombre d'observations z au cours du temps.
Supposons que nous ayons des observations disponibles sur $[0, T]$, en suivant la terminologie rappelée par Bensoussan (1971), on parle de *lissage* si on souhaite estimer les meilleurs $x(t)$ sur $[0, T]$, et de *filtrage* si on ne souhaite que $x(T)$. La détermination de $x(t)$ pour $t > T$ est ce qu'on conçoit classiquement comme de la *prédiction*. Pour nous en médecine, en imaginant qu'on dispose des mesures sur un intervalle $[-T, 0]$ par rapport au présent, le lissage correspond à retrouver l'historique, le filtrage correspond donc au diagnostic et la prédiction au pronostic.
Que ce soit pour l'identification en géophysique, ou l'observation en météorologie, la méthode la plus naturelle en assimilation de données au vu des systèmes d'EDP sous-jacents est de minimiser directement la fonctionnelle par des méthodes de gradient donnant directement la solution de lissage. On a appelé ces méthodes *méthodes variationnelles* car elles minimisent explicitement le critère optimal et on en trouve une description exhaustive dans Banks et Kunisch (1989) ; Chavent (2008) en identification et Blum et al (2008) en observation. Pour le filtrage, le problème avait été résolu par Wiener (1949) sur des systèmes de dimension finie, puis en 1962 Kalman établit une propriété très forte de récursivité dans le filtre de Wiener. Il démontre que le filtre de Wiener peut être obtenu en considérant le problème de minimisation sur des intervalles successifs. Ainsi quand T augmente il n'est pas nécessaire de recalculer toute la minimisation mais il suffit de minimiser sur la nouvelle fenêtre à partir de l'état filtré obtenu précédemment. Cette propriété de récursivité est extrêmement intéressante en terme opérationnel. Ces filtres peuvent s'appliquer dans le cadre des EDP (voir notamment Bensoussan (1971)) mais en contrepartie leur coût de mise en oeuvre numérique pour les systèmes distribués en limite l'utilisation. Le filtrage est donc moins développé pour les EDP, ou alors avec des versions dites *réduites* qui ne vont stabiliser l'erreur d'estimation que partiellement. En océanographie par exemple, les travaux de Pham et al (1997) ; Hoteit (2001) suivent cette idée.

□ EDP, ANALYSE NUMÉRIQUE ET FILTRAGE – Ces différents éléments théoriques nous permettent de positionner clairement notre recherche. L'objectif est ainsi de proposer des méthodes pertinentes pour mettre en oeuvre les principes du filtrage pour des systèmes d'EDP en mécanique cardiaque. Une fois les filtres imaginés, la propriété de récursivité nous conduira à résoudre un problème numérique dont la structure reste similaire à l'approximation numérique classique (sans utiliser les observations) du phénomène physique modélisé. Cette structure nous amène alors à nous poser les questions d'analyse numérique usuelles de convergence de notre observateur numérique vers le système réel. Nous verrons alors qu'à l'aide des outils standard d'analyse numérique, le filtrage nous permet de justifier l'utilisation de mesures dès que possible pour améliorer la qualité de l'approximation numérique. Enfin, comme dans les autres méthodes d'assimilation de données, nous aurons alors la possibilité de prendre en compte l'ensemble des incertitudes de modélisation, ce que, en mécanique cardiaque, nous mettrons à profit dans l'identification des paramètres mécaniques du tissu avec un objectif de diagnostic de pathologies de la contraction du myocarde.

Organisation du document

□ CHAPITRE 1 – Le premier chapitre, introductif, constitue une synthèse des approches classiques en estimation. J'ai voulu à la fois présenter les résultats sous l'angle déterministe (contrôle optimal) mais aussi probabiliste (estimation bayesienne). L'accent est sans doute nettement plus marqué sur le formalisme déterministe mais les considérations probabilistes nous ont aussi offert un cadre d'interprétation complémentaire et riche. A ce titre, mon introduction à l'estimation bayesienne en tout début de chapitre doit beaucoup au cours donné par A. Bensoussan à l'école d'été CEA-EDF-INRIA 2006 sur l'assimilation de données.

La plupart des éléments présentés préparent les démonstrations des résultats originaux des chapitres 3 et 4. J'ai de plus souhaité compléter cette présentation en formulant la solution du problème de filtrage dans un contexte non-linéaire par les méthodes de programmation dynamique. Ces résultats sont classiques, mais dans mon expérience ne se trouvent pas facilement sous forme synthétique. Bien que leur application soit peu envisageable pour des systèmes non-linéaires de grande taille, on peut au moins s'en inspirer pour formuler ensuite des extensions robustes ou sous-optimales.

□ CHAPITRE 2 – A l'autre extrémité du spectre, le chapitre 2 débute par une synthèse des travaux de modélisation et d'analyse numérique réalisés depuis plusieurs années à l'INRIA, notamment au sein de l'action Cardiosense3D. Il m'apparaissait nécessaire de réaliser cette synthèse, plus détaillée que celle réalisée en 2006 avec le film « *le cœur numérique* » (voir Chapelle et al (2006)), pour deux raisons. La première est qu'il est incontournable de présenter le registre (catégorie d'EDP, type de paramétrage, comportement qualitatif des solutions...) dans lequel la modélisation directe opère avant de s'attaquer aux problèmes inverses. La seconde est que cette synthèse a conduit à revisiter un certain nombre des composants de la modélisation directe (schémas en temps, conditions aux limites, modélisation anatomique, voir article en préparation Chapelle et al (2008)), à comparer d'ailleurs, pour mesurer le chemin parcouru, avec ce qui était présenté en 2006, mais aussi à réaliser un code numérique[1], aujourd'hui utilisé à la fois en simulation mécanique cardiaque directe et inverse, code dont je suis l'auteur principal.

Dans la seconde partie du chapitre, nous nous concentrons sur une présentation des différents types de mesure employés en cardiologie. Au delà de la description de ces techniques, nécessaire à la formalisation des opérateurs d'observation des chapitres suivants, nous nous sommes intéressés à certains aspects de modélisation des mesures, avec la motivation d'en générer des synthétiques réalistes, notamment en imagerie cardiaque. Celles-ci nous donneront les moyens d'évaluer nos méthodes inverses indépendamment de la qualité prédictive du modèle. Mais par ailleurs, l'enjeu peut s'étendre à la communauté de l'imagerie pour l'évaluation de ses méthodes de traitement, comme le montrent notre collaboration avec E. Angelini de Telecom ParisTech (voir Duan et al (2007)) et le coencadrement du stage de Master de O. Talcoth (voir Talcoth (2007)).

[1]Code `heartLab` (30 000 lignes) architecturé sur la librairie élément finis `OpenFEM` (www.openfem.net)

□ CHAPITRE 3 – Nous attaquons dans ce chapitre le cœur du sujet en formulant un certain nombre de filtres efficaces pour l'observation de l'état d'un système mécanique distribué, donc hyperbolique du second ordre, en linéaire puis non-linéaire. Leur principe réside dans l'utilisation de lois de feedback robustes, performantes, de faible coût numérique et à portée de main dans les codes de simulation, y compris commerciaux, pour construire des *observateurs de Luenberger*. Le feedback collocalisé de type *Direct Velocity Feedback* (DVF) nous en a fourni un premier exemple naturel dans les configurations où l'on peut mesurer des vitesses. Mais l'originalité des résultats réside surtout dans la possibilité d'imaginer des lois de contrôles qui ne sont envisageables que parce que l'observateur n'est pas un système réel mais un objet numérique. Ainsi, il a été possible de concevoir un filtre, de performance équivalente au DVF, utilisant directement des mesures de déplacement, bien plus répandues en pratique. Ces filtres ont été analysés en profondeur dans leur version linéaire, avant de donner lieu à des extensions non-linéaires, que ce soit pour la dynamique ou les mesures.

Au delà d'avoir ainsi obtenu un observateur compatible avec la mécanique cardiaque, nous avons pu formuler une nouvelle méthode pour exploiter des images (donc des mesures profondément non-linéaires) dans l'estimation de trajectoires de systèmes élastodynamiques, avec des applications naturelles en asservissement visuel. L'impact potentiel de telles méthodes dépasse donc largement le cadre strict de l'assimilation de données en cardiologie, pour s'étendre d'une part vers des systèmes (mécaniques) plus généraux, mais aussi de l'autre vers le traitement de séquences d'images à base de modèles (version faible, en quelque sorte, de l'assimilation de données, où le modèle n'a pas essentiellement vocation a être prédictif). Ces différents résultats ont donné lieu à l'article (Moireau et al (2008a)) soumis à *Inverse Problems*.

□ CHAPITRE 4 – Ce dernier chapitre ajoute l'identification paramétrique à l'estimation d'état, ou plutôt s'appuie sur les résultats du chapitre 3 pour formuler des estimateurs paramétriques robustes. La principale difficulté provient de notre abandon (nécessaire), dans la formulation des observateurs d'état, des approches de type contrôle optimal Kalmanien qui sont connues pour *s'étendre* au cas de l'estimation conjointe état-paramètres. Cependant, la théorie du filtrage adaptatif offre les moyens de contourner cette première difficulté. Notre principal résultat, est à nos yeux, d'avoir pu établir l'équivalence entre certains types de filtres adaptatifs et la famille des filtres « Kalmaniens » réduits. Ceux-ci, souvent utilisés sur des estimateurs dont on peut assurer que seule une faible partie de la dynamique doit être contrôlée, apparaissent naturellement dans nos formulations « une fois » que nos observateurs de Luenberger initiaux ont restreint l'incertitude sur l'espace paramétrique. Cette équivalence nous a permis de démontrer la performance (à la fois en termes de convergence d'erreur et d'optimalité au sens variationnel H^2 ou H^∞) des filtres proposés, de formuler des schémas en temps adéquats, et d'imaginer des extensions non-linéaires efficaces. Au passage, nous avons démontré que le filtre non-linéaire SEIK de Pham et al (1997) est une version particulière de filtre UKF (*Unscented Kalman Filter* proposé par Julier et al (1995) et d'utilisation pratique courante, car notoirement plus robuste que le filtre EKF en présence de non-linéarités) sous une forme réduite dont nous proposons l'extension dans le cas général (résultat semble-t-il original). Là encore ces travaux ont conduit aux deux articles (Moireau et al (2008b) ; Chapelle et al (2009)) déjà publiés.

Eléments de bibliographie personnelle

Pour finir cette introduction, je me permets de résumer quelques éléments bibliographiques me concernant qui sont particulièrement cités dans le reste du document et qui par l'hétérogéneité des sujets impliquent des auteurs d'origines très diverses. Ce résumé a ainsi vocation a éviter toute confusion dans le reste du manuscrit avec d'autres citations.

Journaux

- Sermesant et al (2006) – Article introductif entre les équipes-projets Asclepios et Macs de l'INRIA et le King's College London sur les défis autour de l'adaptation de modélisation à un patient :
  M. SERMESANT, P. MOIREAU, O. CAMARA, J. SAINTE MARIE, R. ANDRIANISIMIAVONA, R. CIMRMAN, D.L.G. HILL, D. CHAPELLE, R.RAZAVI – *Cardiac function estimation from MRI using a heart model and data assimilation : Advances and difficulties* – Medical Image Analysis, 10(4) :642–656, 2006.

- Moireau et al (2008b) – Article présentant le couplage entre un filtre collocalisé en vitesses *Direct Velocity Feedback* et un estimateur paramétrique :
  P. MOIREAU, D. CHAPELLE, P. LE TALLEC – *Joint state and parameter estimation for distributed mechanical systems* – Computer Methods in Applied Mechanics and Engineering, 197 : 659–677, 2008.

- Chapelle et al (2009) – Extension de l'article précédent au cas de mesures frontières au travers de l'estimation H^∞ :
  D. CHAPELLE, P. MOIREAU, P. LE TALLEC – *Robust filtering for joint state-parameter estimation in distibuted mechanical systems* – Discrete and Continuous Dynamical Systems – Series A, 23(1&2), 2009 (sous presse).

- Moireau et al (2008a) – Article spécifiquement consacré à l'estimation d'état et présentant des filtres de spécifiques pour des observations de types déplacements Lagrangien ou positions Eulériennes à partir d'images :
  P. MOIREAU, D. CHAPELLE, P. LE TALLEC – *Filtering for distributed mechanical systems using position measurements : Applications in medical imaging*, accepté pour publication dans Inverse Problems.

- Chapelle et al (2008) – Article présentant les avancées en matière de modélisation électromécanique en partenariat avec l'équipe-projet Reo de l'INRIA :
  D. CHAPELLE, M.A. FERNÀNDEZ, J.F. GERBEAU, P. MOIREAU, J. SAINTE-MARIE, N. ZENZEMI – *A 3D model for the electromechanical activity of the heart*, en préparation.

Conférences

- Duan et al (2007) – Article de conférence en partenariat avec Telecom ParisTech et Columbia University sur l'utilisation de modèle mécanique comme fantômes pour un simulateur d'échographie 3D :
  Q. DUAN, P. MOIREAU, E.D. ANGELINI, D. CHAPELLE, A.F. LAINE – *Simulation of 3D Ultrasound with a Realistic Electro-Mechanical Model of the Heart* – (FIMH'07), 4466(LCNS) : 463–470, 2007. Springer.

Multimedia

- Chapelle et al (2006) – Film présentant l'effort de modélisation, de traitements des mesures et d'assimlation de données au sein de l'action Cardiosense3D :
  D. CHAPELLE, P. MOIREAU, M. SERMESANT, M. FERNANDEZ, H. DELINGETTE – Film : *Le Cœur Numérique (The Digital Heart)* – Réalisation : A. PAOURI – Production : INRIA MULTIMEDIA

CHAPITRE

1 Principe d'estimation moindres carrés

Que c'est à lui que les sciences de l'observation sont redevables d'une méthode de calcul qu'il a nommé méthodes des moindres carrés des erreurs*, et dont Laplace a montré tout l'avantage probable sous le rapport de la précision des résultats ;*

— Siméon Denis Poisson, *Discours prononcé aux funérailles de Adrien-Marie Legendre*

Une intelligence qui à un instant déterminé devrait connaître toutes les forces qui mettent en mouvement la nature, et toutes les positions de tous les objets dont la nature est composée, [...] pour une telle intelligence nul serait incertain et le propre futur comme le passé serait évident à ses yeux.

— Pierre-Simon Laplace, *Essai philosophique sur les probabilités*

Sommaire du chapitre

1.1 Introduction à l'estimation bayesienne **14**
1.2 Inverses généralisés au sens des moindres carrés **20**
1.2.1 Inverse généralisé d'un opérateur d'observation 21
1.2.2 Terme d'ébauche 23
1.2.3 Estimation récursive 25
1.3 Estimation pour les systèmes dynamiques linéaires **29**
1.3.1 Cas discret en temps : Cadre déterministe 29
1.3.2 Cas discret en temps : Cadre probabiliste 34
1.3.3 Cas continu en temps 35
1.3.4 Bruits de modèle 42
1.3.5 Notions d'observabilité, de contrôlabilité et d'identifiabilité 45
1.4 Extensions autour de la théorie LQ **48**
1.4.1 Filtre non-linéaire déterministe 48
1.4.2 Contrôle robuste 53
1.4.3 Filtre EKF 55
1.4.4 Filtre UKF 59
1.5 Conclusion **65**

1.1 Introduction à l'estimation bayesienne

On s'intéresse à un système physique accessible uniquement au travers de *mesures* aussi appelées *observations*. Ce système est caractérisé par un ensemble de grandeurs (supposé fini pour l'instant) de dimension N le caractérisant. Nous appellerons ces grandeurs *variables d'état* du système puisqu'elles rendent compte effectivement de l'état dans lequel le système se trouve. Certains ouvrages, notamment en statistique, parlent de *paramètres* du système mais nous garderons ce terme pour caractériser le sous ensemble des variables d'état qui n'ont pas de dynamique propre... Ce *vecteur d'état* rassemblant l'ensemble des variables, est déterministe mais inconnu. Cependant les mesures sont souvent bruitées au sens probabiliste usuel et donc on pourra juger pertinent ou non, suivant les situations, de considérer l'état lui-même comme probabiliste à travers sa *probabilité a priori* $\mathbb{P}_X$. Si on agrège les observations sous forme d'un vecteur de dimension finie m (car les mesures sont toujours en nombre fini), l'approche probabiliste nous permet de considérer la *probabilité a posteriori* $\mathbb{P}_{X|Z}$ qui n'est autre que la probabilité conditionnelle que $X = x$ sachant $Z = z$ où x et z sont des réalisations particulières[1] pour les vecteurs aléatoires considérés. A partir de l'identité

$$\mathbb{P}(X, Z) = \mathbb{P}(X|Z)\mathbb{P}(Z) = \mathbb{P}(Z|X)\mathbb{P}(X),$$

la règle de *Bayes* donne cette probabilité a posteriori sous la forme

$$\mathbb{P}(X|Z) = \frac{\mathbb{P}(Z|X)\mathbb{P}(X)}{\mathbb{P}(Z)}. \tag{1.1}$$

Dans le cas où les variables aléatoires sont définies par leur densité de probabilité $p(X)$, $p(X|Z)$ (avec les mêmes conventions de notation), la règle de Bayes s'applique directement sur les densités et par exemple

$$p(Z) = \int p(Z|X)p(X)\, dX.$$

Dans tous les cas, le principe de l'estimation consiste à trouver une fonction de Z noté $\hat{X}(Z)$ qui résume l'information sur X contenue dans la mesure Z. Cette notion est d'ailleurs à rapprocher de celle de statistique suffisante (en anglais) ou exhaustive de Fisher (cf. Casella et Berger (2002)) qui exprime le fait que connaître $\hat{X}(Z)$ à partir de Z est suffisant vis-à-vis de X

$$p(Z|\hat{X}, X) = p(Z|\hat{X}).$$

Cet estimateur $\hat{X}(Z)$ admet donc une densité de probabilité en tant que fonction de la variable aléatoire Z. Insistons sur le fait qu'il y a ici deux types d'évaluation probabiliste dans la manipulation de l'estimateur. Tout d'abord, celui-ci est défini au travers d'évaluation de X sachant $Z = z$ fixé ; puis c'est en tant que fonction de Z de nouveau variable aléatoire que l'on peut analyser sa qualité statistique. En effet, l'estimateur, comme fonction de Z, est qualifié de bon estimateur s'il est

- **sans biais**, ce qui signifie que l'erreur $\tilde{X} = X - \hat{X}$ entre l'état et l'estimateur est en moyenne nulle

$$\mathbb{E}(\tilde{X}) = 0;$$

[1] on omettra les réalisations quand il n y a pas d'ambiguité pour écrire simplement $\mathbb{P}(X)$ au lieu de $\mathbb{P}(X = x)$ etc...

- **consistant** c'est-à-dire que, plus le nombre d'observations est grand, plus l'estimateur converge vers l'état réel du système. Cette notion est conditionnelle à la notion d'observabilité du système, c'est-à-dire qu'on assure que dans Z on ajoute des observations nouvelles vis-à-vis de l'estimation ;
- **efficace**, par exemple au travers d'une covariance d'erreur minimale. Ainsi $\hat{X}$ est efficace si
$$\forall \phi, \quad \mathrm{Cov}(\tilde{X}) \leq \mathrm{Cov}(X - \phi(Z)),$$
où $\mathrm{Cov}(\cdot, \cdot)$ est la matrice de covariance de deux vecteurs aléatoires entre eux[2]. Autrement dit, ce critère signifie que la covariance de l'erreur d'estimation est minimale parmi toutes les statistiques $\phi(Z)$. En fait, cette dernière notion contient la première car on démontre qu'un estimateur biaisé ne peut-être efficace.

Nous allons maintenant donner trois exemples d'estimateurs parmi les plus classiques.

□ MMSE (MINIMIUM MEAN-SQUARE ESTIMATOR) – Dans le cas où on associe à X une densité de probabilité a priori, le MMSE est l'estimateur qui minimise
$$\hat{X} = \operatorname{argmin} \mathbb{E}((X - \phi(Z))^T (X - \phi(Z)))$$

Rappelons la propriété classique de type projection orthogonale de l'espérance conditionnelle,

LEMME 1.1.1 (PROJECTION ORTHOGONALE)

$$\forall \phi, \quad \mathbb{E}((X - E(X|Z))^T \phi(Z)) = 0. \tag{1.2}$$

■ ***Démonstration :***

$$\begin{aligned}\mathbb{E}((X - \mathbb{E}(X|Z))^T \phi) &= \int \mathbb{E}((X - \mathbb{E}(X|Z))^T \phi | Z) p(Z) dZ \\ &= \int \mathbb{E}(X^T \phi | Z) p(Z) dZ - \int \mathbb{E}(\mathbb{E}(X|Z)^T \phi | Z) p(Z) dZ.\end{aligned}$$

Mais ϕ et $\mathbb{E}(X|Z)$ ne dépendant que de Z, prendre leur espérance sachant Z signifie qu'ils sont constants vis-à-vis de cette espérance donc

$$\mathbb{E}(X^T \phi | Z) = \mathbb{E}(X|Z)^T \phi = \mathbb{E}(\mathbb{E}(X|Z)^T \phi | Z)$$

ce qui, réinjecté dans la formule plus haut, conduit au résultat. ■

Par conséquent on peut préciser le MMSE :

THÉORÈME 1.1.2
Le MMSE est l'espérance conditionnelle de X sachant Z

$$\hat{X} = \mathbb{E}(X|Z), \tag{1.3}$$

et en particulier

$$\forall \phi, \quad \mathbb{E}((X - \hat{X})^T \phi) = 0.$$

[2]Lorsque la covariance n'a qu'un seul argument, elle correspond à la covariance du vecteur avec lui-même (la variance en scalaire).

■ ***Démonstration*** *: Le lemme de projection orthogonale permet de formuler une démonstration autour d'une formule de Pythagore du type*

$$\begin{aligned}\forall\phi,\quad \mathbb{E}((X-\phi)^T(X-\phi)) &= \mathbb{E}((X-\hat{X}+\hat{X}-\phi)^T(X-\hat{X}+\hat{X}-\phi))\\ &= \mathbb{E}((X-\hat{X})^T(X-\hat{X}))+\mathbb{E}((\hat{X}-\phi)^T(\hat{X}-\phi))\\ &\geq \mathbb{E}((X-\hat{X})^T(X-\hat{X})).\end{aligned}$$

■

REMARQUE 1.1.1
Il est possible de démontrer que la minimisation de l'erreur en espérance $\mathbb{E}((X-\phi(Z))^T(X-\phi(Z)))$ est compatible avec la définition de l'efficacité à partir de la variance d'erreur $\mathrm{Var}(X-\hat{X})$. Ces notions sont plus généralement toutes reliées aux bornes de Cramer-Rao (voir Legland (2002)) structurant les notions d'efficacité.

□ LSE (LEAST SQUARE ESTIMATOR) – ou en français *estimateur moindres carrés*. Dans le cas où X est aléatoire, le LSE est la statistique ϕ qui minimise la fonctionnelle

$$\hat{X} = \mathrm{argmin}\left\{J(\phi) = \begin{pmatrix}\phi\\ Z\end{pmatrix}^T \mathcal{P}^{-1}\begin{pmatrix}\phi\\ Z\end{pmatrix}\right\}, \tag{1.4}$$

où $\mathcal{P}$ est la matrice de covariance associée au couple aléatoire (X,Z)

$$\mathcal{P} = \begin{pmatrix}P^X & P^{ZX}\\ P^{XZ} & P^Z\end{pmatrix} = \begin{pmatrix}\mathrm{Cov}(X) & \mathrm{Cov}(X,Z)\\ \mathrm{Cov}(X,Z) & \mathrm{Cov}(Z)\end{pmatrix}.$$

THÉORÈME 1.1.3
Si P^Z est inversible le LSE est donné par

$$\hat{X} = \mathbb{E}(X) + P^{XZ}(P^Z)^{-1}(Z-\mathbb{E}(Z)). \tag{1.5}$$

■ ***Démonstration*** *: Quitte à opérer un shift dans la fonctionnelle, on peut supposer les variables centrées. Dans ce cas, en développant le critère quadratique sous la forme $J = \phi^T A\phi + 2\phi^T BZ + Z^T CZ$ donc le minimum est atteint en*

$$\hat{X} = -A^{-1}BZ = P^{ZX}(P^Z)^{-1}Z.$$

■

Si P^Z n'est pas inversible, il est possible de définir un inverse généralisé (au sens classique de Moore-Penrose Penrose (1955), cf. Section 1.2.1) de P^Z tel que cette formule reste vraie. On pourra notamment se reporter à Lipster (2008) pour une preuve exhaustive. Nous exploiterons pour notre part au chapitre 4 des développements autour de cette généralisation laissée de côté pour l'instant.

□ BLUE (BEST LINEAR *unbiased* ESTIMATOR) – Dans la formulation de ce troisième estimateur, on choisit de restreindre le type de fonction ϕ en cherchant la meilleure prédiction linéaire en Z pour la covariance d'erreur. On définit ainsi l'estimateur par

$$\hat{X} = \mathbb{E}(X) + K(Z-\mathbb{E}(Z)) \text{ où } K = \mathrm{argmin}\,\mathbb{E}(\tilde{X}^T\tilde{X}). \tag{1.6}$$

On peut aussi démontrer que, assez logiquement, cet opérateur linéaire K est, en un certain sens, un opérateur de projection orthogonale. Pour ce faire, commençons par un rappel

sur les **variables** aléatoires. L'ensemble $\mathcal{L}(\Omega, \mathcal{F}, P)$ des **variables** aléatoires de carré intégrable définies sur $(\Omega, \mathcal{F}, P)$ est un espace de Hilbert (en fait simplement son espace quotienté par le noyau de P) pour le produit scalaire

$$\forall (x, y) \in \mathcal{L}(\Omega, \mathcal{F}, P)^2, \ < x, y >= \mathbb{E}(xy).$$

Il est donc possible de définir sur n'importe quel ensemble convexe fermé $\mathcal{V}$ une projection orthogonale minimisant la distance à $\mathcal{V}$

$$\forall x, \quad \|x - \Pi_{\mathcal{V}}(x)\|^2 = \inf_{y \in \mathcal{V}} \|x - y\|^2 = \inf_{y \in \mathcal{V}} \mathbb{E}((x - y)^2).$$

En particulier sur l'espace Vect$\{1, z\}$, dont une base orthogonale est clairement $(v_1, v_2) = (1, z - \mathbb{E}(z))$, on obtient

$$\hat{x} = \sum_{k=1,2} \frac{\mathbb{E}(xv_k)}{\mathbb{E}(v_k v_k)} v_k = \mathbb{E}(x) + P^{xz}(P^z)^{-1}(z - \mathbb{E}(z)).$$

Sur l'espace des **vecteurs** aléatoires cette fois, on peut soit écrire ce même résultat sur un espace tensorisé (voir Lipster (2008)), ou se rappeler qu'à défaut d'avoir un espace de Hilbert, on a un module de Hilbert sur $S_N(\mathbb{R})$ muni de la relation

$$\forall (X, Y) \in \mathcal{L}(\Omega, \mathcal{F}, P)^2, \ < X, Y >= \mathbb{E}(XY^T) \in S_N(\mathbb{R}).$$

Or, sur les modules de Hilbert la notion de projecteur orthogonal continue d'exister, et on obtient une relation similaire (mais matricielle)

$$\hat{X} = \sum_{k=1,2} \mathbb{E}(XV_k^T)\mathbb{E}(V_k V_k^T)^{-1} V_k = \mathbb{E}(X) + P^{XZ}(P^Z)^{-1}(Z - \mathbb{E}(Z)).$$

avec toujours $(V_1, V_2) = (1, Z - E(Z))$. On constate que, en tout rigueur, l'inverse de $\mathbb{E}(V_1 V_1^T)$ par exemple n'est pas défini. Cependant, dans ce cas, elle est définie au sens généralisé (toujours de Moore Penrose) et par exemple ici vaut $\frac{1}{N}$ (voir un exemple similaire remarque 1.2.1 de la Section 1.2.1 sur les inverses généralisés, et plus précisément Lipster (2008)). Cette expression caractérise alors le BLUE sur n'importe quel type d'espace de vecteurs aléatoires avec $K = P^{ZX}(P^Z)^{-1}$.

Enfin, concernant son efficacité, on peut démontrer que ce minimum de variance d'erreur est aussi un minimum au sens de la covariance d'erreur $K = \operatorname{argmin} \mathbb{E}(\tilde{X}\tilde{X}^T)$. En effet, pour tout opérateur linéaire Γ, l'erreur pour ce choix d'estimateur

$$\mathbb{E}(\tilde{X}_\Gamma \tilde{X}_\Gamma^T) = P^X + \Gamma P^Z \Gamma^T - \Gamma P^{ZX} - P^{XZ}\Gamma^T.$$

Puis avec la structure de $K = P^{ZX}(P^Z)^{-1}$, estimateur optimal, et quelques manipulations algébriques

$$\mathbb{E}(\tilde{X}_\Gamma \tilde{X}_\Gamma^T) = \mathbb{E}(\tilde{X}_K \tilde{X}_K^T) + (K - \Gamma)P^Z(K - \Gamma)^T > \mathbb{E}(\tilde{X}_K \tilde{X}_K^T).$$

Cet estimateur est donc bien le plus efficace des estimateurs linéaires au sens des covariances et la covariance d'erreur est donnée par une expression que nous allons retrouver tout au long de ce mémoire :

$$\mathbb{E}(\tilde{X}^T \tilde{X}) = P^X + P^{XZ}(P^Z)^{-1}P^{ZX}. \tag{1.7}$$

Nous constatons avec les différentes caractérisations de ces trois estimateurs naturels qu'ils sont étroitement liés. En particulier les expressions des deux derniers conduisent immédiatement au théorème suivant :

THÉORÈME 1.1.4
Supposons P^Z est inversible, alors le meilleur estimateur linéaire est l'estimateur moindres carrés.

Ainsi, on pourra par la suite utiliser soit le critère LSE, soit le critère BLUE, soit la notion de projection orthogonale pour rechercher ce même estimateur. De plus, nous avons vu que BLUE-LSE vérifiait une propriété d'orthogonalité analogue à celle du MMSE (i.e l'espérance conditionnelle). Ainsi on notera le meilleur estimateur linéaire sous la forme (introduite aussi dans Lipster (2008))

$$\hat{X} = \hat{E}(X|Z) \tag{1.8}$$

pour respecter cette analogie.

A fortiori, la comparaison ne s'arrête pas là entre ces deux estimateurs (BLUE, LSE d'un coté et MMSE de l'autre) dans le cas des variables gaussiennes. En effet, on démontre le résultat suivant :

THÉORÈME 1.1.5
Si X et Z sont des variables gaussiennes, alors

$$\hat{X} = \mathbb{E}(X|Z) = \hat{E}(X|Z) = \mathbb{E}(X) + P^{XZ}(P^Z)^{-1}(Z - \mathbb{E}(Z)).$$

■ ***Démonstration :*** *Il existe plusieurs démonstrations de ce résultat. Nous choisissons ici la démonstration sans doute la plus directe et technique, et pas nécessairement la plus élégante. Ce choix est motivé par le fait qu'elle permet en revanche de présenter quelques manipulations algébriques que nous verrons réapparaître tout au long de la présentation. Là encore, on peut se ramener à l'étude de variable centrées par une simple translation donc on considère $\mathbb{E}(X) = 0$ et $\mathbb{E}(Z) = 0$.*

Les distributions gaussiennes s'expriment simplement à partir des deux premiers moments sous la forme, par exemple pour Z,

$$p(Z) = (2\pi)^{-\frac{m}{2}} \det(P^Z)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} Z^T (P^Z)^{-1} Z\right),$$

et de même pour la loi de X. De plus les deux vecteurs aléatoires étant gaussiens leur loi conjointe est elle-même gaussienne de densité

$$p(X, Z) = (2\pi)^{-\frac{N+m}{2}} \det(\mathcal{P})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \begin{pmatrix} X \\ Z \end{pmatrix}^T \mathcal{P}^{-1}, \begin{pmatrix} X \\ Z \end{pmatrix}\right),$$

où comme nous l'avons déjà introduit

$$\mathcal{P} = \begin{pmatrix} P^X & P^{XZ} \\ P^{ZX} & P^Z \end{pmatrix}.$$

L'inverse de cette matrice bloc 2×2 symétrique est donnée par les relations usuelles (« à l'image des cofacteurs dans $\mathbb{M}_2(\mathbb{R})$ »)

$$\mathcal{P}^{-1} = \begin{pmatrix} (P^a)^{-1} & -(P^a)^{-1} P^{XZ} (P^Z)^{-1} \\ (-(P^a)^{-1} P^{XZ} (P^Z)^{-1})^T & (P^Z - P^{ZX} P^X P^{XZ})^{-1} \end{pmatrix} = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix},$$

où $P^a = P^X - P^{XZ}(P^Z)^{-1}P^{ZX}$. Ainsi la loi conditionnelle de X sachant Z est donnée par

$$p(X|Z) = \frac{p(X,Z)}{p(Z)} = (2\pi)^{-\frac{N}{2}} \left(\frac{\det(\mathcal{P})}{\det(P^Z)}\right)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\left(\begin{smallmatrix} X \\ Z \end{smallmatrix}\right)^T \left(\begin{smallmatrix} A & B \\ B^T & C-(P^Z)^{-1} \end{smallmatrix}\right)\left(\begin{smallmatrix} X \\ Z \end{smallmatrix}\right)\right).$$

Procédons maintenant à quelques manipulations algébriques. Tout d'abord, le terme dans l'exponentielle peut être simplifié par l'utilisation du lemme d'inversion matricielle suivant

LEMME 1.1.6 (INVERSION MATRICIELLE)
Soit M_1, M_{12}, M_{21}, M_2 des matrices avec M_1 et M_2 et $M_2 - M_{21}M_1^{-1}M_{12}$ inversibles, alors $M_1 - M_{12}M_2^{-1}M_{21}$ est inversible et

$$(M_1 - M_{12}M_2^{-1}M_{21})^{-1} = M_1^{-1} + M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}M_{21}M_1^{-1}.$$

■ ***Démonstration :*** *La preuve la plus classique de ce lemme repose sur la résolution du système $Y = MX$ décomposé par bloc*

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} M_1 & M_{12} \\ M_{21} & M_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

En effet, on peut réduire ce système à un seul bloc sur X_1 (resp. sur X_2) par des compléments de Schur $M_1 - M_{12}M_2^{-1}M_{21}$ (resp. $M_2 - M_{21}M_1^{-1}M_{12}$). La solution du système étant unique, on en déduit le lemme en identifiant le bloc $(1,1)$ de M^{-1}.

Il existe aussi une démonstration directe de ce lemme par la manipulation suivante

$$\begin{aligned}&\big(M_1^{-1} + M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}M_{21}M_1^{-1}\big(M_1 - M_{12}M_2^{-1}M_{21}) = \\ &\quad \mathbb{1} + M_1^{-1}M_{12}M_2^{-1}M_{21} + M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}M_{21}M_1^{-1}(M_1 - M_{12}M_2^{-1}M_{21}),\end{aligned}$$

or le dernier terme se simplifie avec l'avant dernier si on remarque que

$$\begin{aligned}&M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}M_{21}M_1^{-1}(M_1 - M_{12}M_2^{-1}M_{21}) \\ &\qquad = M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}(M_{21} - M_{21}M_1^{-1}M_{12}M_2^{-1}M_{21}) \\ &\qquad = M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}(M_2 - M_{21}M_1^{-1}M_{12})M_2^{-1}M_{21} \\ &\qquad = M_1^{-1}M_{12}M_2^{-1}M_{21},\end{aligned}$$

ce qui assure bien que

$$\big(M_1^{-1} + M_1^{-1}M_{12}(M_2 - M_{21}M_1^{-1}M_{12})^{-1}M_{21}M_1^{-1}\big(M_1 - M_{12}M_2^{-1}M_{21}) = \mathbb{1}. \qquad ■$$

En effet, ce lemme permet alors de prouver que

$$(P^Z - P^{ZX}P^XP^{XZ})^{-1} - (P^Z)^{-1} = (P^Z)^{-1}P^{ZX}(P^a)^{-1}P^{XZ}(P^Z)^{-1},$$

donc

$$\begin{pmatrix} X \\ Z \end{pmatrix}^T \begin{pmatrix} A & B \\ B^T & C-(P^Z)^{-1} \end{pmatrix} \begin{pmatrix} X \\ Z \end{pmatrix} = (X - P^{XZ}(P^Z)^{-1}Z)(P^a)^{-1}(X - P^{XZ}(P^Z)^{-1}Z).$$

Deuxièmement, la décomposition LU de la matrice de covariance conjointe donne

$$\begin{pmatrix} P^X & P^{ZX} \\ P^{XZ} & P^Z \end{pmatrix} = \begin{pmatrix} P^a & P^{ZX} \\ 0 & P^Z \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ (P^Z)^{-1}P^{XZ} & \mathbb{1} \end{pmatrix},$$

donc

$$\det(\mathcal{P}) = \det(P^a)\det(P^Z).$$

Ainsi avec $\hat{X} = P^{XZ}(P^Z)^{-1}Z$, la densité conjointe s'exprime sous la forme

$$p(X|Z) = (2\pi)^{-\frac{N}{2}} \det(P^a)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(X-\hat{X})^T(P^a)^{-1}(X-\hat{X})\right),$$

assurant que la probabilité conditionnelle est bien une gaussienne de moyenne $\hat{X}$ et de covariance P^a, ce qu'il fallait démontrer. ■

Il est remarquable que dans le cas des variables gaussiennes, la covariance associée à la loi de propabilité de X sachant Z soit exactement la matrice de covariance de l'erreur d'estimation du BLUE issue de la moyenne sur Z. Cette propriété est en fait la conséquence du fait que P^a est indépendant de Z ce qui implique pour les variables gaussiennes que la covariance d'erreur pour tout Z et celle sachant un Z donné sont égales.

REMARQUE 1.1.2
On parle beaucoup aussi d'estimateur dit de maximum a posteriori qui maximise $p(X|Z)$. Ce que nous venons de voir montre que dans le cas gaussien, ce maximum est le LSE. Inversement, on pourrait aussi définir un estimateur en maximisant $p(Z|X)$ qui on le rappelle est donné par la règle de Bayes

$$p(Z|X) = \frac{p(X|Z)p(Z)}{p(X)}.$$

Cet estimateur est bien connu puisque c'est le maximum de vraisemblance. Nous utiliserons peu cette notion pour lui préférer le plus souvent des estimateurs LSE plus simples à définir, y compris dans le cadre déterministe.

1.2 Inverses généralisés au sens des moindres carrés

Cette introduction pose les bases très générales de la notion d'estimation statistique et permet de se familiariser avec un certain nombre de concepts, notations, et dénominations. Nous allons désormais préciser la notion d'estimateur dans le cas où l'on dispose d'une structure liant la mesure Z et le vecteur d'état X[3]. En effet, les observations sont, à l'origine, le résultat d'un processus de mesure sur l'état. Cet état nous est peut-être inaccessible (la mécanique quantique est pleine de ce genre de situations) mais il est à l'origine de cette mesure et algébriquement on peut considérer qu'il existe un opérateur, appelé *opérateur d'observation*, générant Z à partir de X. Quand bien même on ne connaît pas l'état précisément, on sait décrire raisonnablement l'opérateur de mesure. Si cet opérateur est noté H, alors mathématiquement on considère la situation

$$Z = H(X) + \chi, \tag{1.9}$$

[3]Dans cette situation, il est courant en statistique d'appeler paramètre ce vecteur X mais nous conservons ce vocable pour un usage plus spécifique

où χ représente l'ensemble des bruits associés à la mesure. Si l'opérateur d'observation est bien *modélisé* (car là aussi il y a en fait un travail de modélisation physique) alors on peut raisonnablement supposer le bruit de mesure de moyenne nulle $\mathbb{E}(\chi) = 0$. En toute généralité, il faudrait considérer que le bruit n'est pas nécessairement additif $Z = H(X, \chi)$ mais nous n'étudierons pas cette situation en détails. En revanche, nous verrons plus loin dans l'exposé que, dans certaines situations, on ne connaît pas de H qui fournit explicitement la mesure mais qu'on sait simplement dire si une mesure est en adéquation avec le système à travers une fonction de vérification implicite

$$D(Z, X) = \chi.$$

Toujours est-il que, dans un premier temps, la situation classique (1.9) sera privilégiée dans les calculs et les trois sections suivantes vont nous permettre de préciser nos estimateurs dans ce cadre. De plus, nous allons commencer à voir que l'estimation se formule aussi dans un cadre déterministe. Les notions de moyenne et de bruit se transforment alors en notions d'a priori et d'incertitude. Subtilement, les résultat algébriques finaux seront similaires mais le cadre fonctionnel et le formalisme diffèrent assez nettement entre les deux. L'ambition de ce chapitre, et même du reste de cet exposé, n'est pas de présenter dans chaque cadre l'ensemble des résultats. Nous ne le ferons que pour les quelques paragraphes qui suivent avant de rebondir, dans les situations plus délicates, d'un formalisme à l'autre pour obtenir les résultats critiques vis-à-vis de nos méthodologies. On notera sans doute un certain tropisme déterministe, comme le suggère le titre de cette section et des suivantes.

1.2.1 Inverse généralisé d'un opérateur d'observation

Commençons tout de même par l'approche probabiliste et supposons, de plus, que H linéaire et que le bruit χ est indépendant de X et de covariance $\mathbb{E}(\chi\chi^T) = W$. On considère en outre que X n'est pas aléatoire mais simplement inconnu et nous cherchons le meilleur estimateur linéaire $\hat{X} = KZ$. L'estimateur étant non biaisé on a

$$\mathbb{E}(\hat{X}) = X \Rightarrow KH = \mathbb{1}.$$

Le lemme suivant fournit alors l'estimateur

LEMME 1.2.1
Soit $K = (H^T W^{-1} H)^{-1} H^T W^{-1}$ alors

$$\forall \Gamma \,|\, \Gamma H = \mathbb{1}, \quad \mathit{Cov}(X - KZ) \leq \mathit{Cov}(X - \Gamma Z).$$

■ ***Démonstration :*** *Tout d'abord on remarque que*

$$\forall \Gamma \,|\, \Gamma H = \mathbb{1}, \quad \Gamma W K^T = K W \Gamma^T = (H^T W^{-1} H)^{-1} = K W K^T,$$

(ce qui ressemble encore à un relation de type produit scalaire sur un module de Hilbert). Donc

$$\forall \Gamma \,|\, \Gamma H = \mathbb{1}, \quad \Gamma W \Gamma^T = K W K^T + (\Gamma - K) W (\Gamma - K)^T,$$

(ce qui ressemble cette fois à une forme de théorème de Pythagore sur un module de Hilbert). De plus pour tout Γ tel que $\Gamma H = \mathbb{1}$, on a $\mathbb{E}((X - \Gamma Z)(X - \Gamma Z)^T) = \Gamma W \Gamma^T$ donc K est bien l'opérateur qui minimise cette covariance. ■

REMARQUE 1.2.1
Un exemple instructif d'application de cet estimateur est la situation élémentaire où l'on souhaite estimer une grandeur scalaire x à partir de m échantillons x_i de celle-ci. Le résultat est évidemment de choisir la moyenne empirique comme estimateur

$$\hat{x} = E_m(x) = \tfrac{1}{d} \sum_{1 \leq i \leq m} x_i.$$

Pour retrouver ce résultat, on définit ici

$$Z = (x_1, \cdots, x_m)^T, \text{ et } H = (1, \cdots, 1)^T,$$

tel qu'on ait bien $Z = Hx$ et on choisit $W = \mathbb{1}$. Dans ce cas on retrouve bien

$$\hat{x} = (H^T H)^{-1} H^T Z = \frac{1}{d}(1, \cdots, 1) \cdot (x_1, \cdots, x_m)^T = E_m(x).$$

Du point de vue déterministe, on reconnaît dans les expressions ci-dessus les expressions classiques de l'inverse généralisé de H pour la norme $\|\cdot\|^2_{W^{-1}} =< .,. >_{W^{-1}}$. D'ailleurs en notant $H' = H^T W^{-1}$ on a bien

$$\hat{X} = (H'H)^{-1} H' Z.$$

Or l'inverse généralisé correspond en fait à la minimisation d'un critère quadratique

$$\min_X \left\{ J(X) = \frac{1}{2} \|Z - HX\|^2_{W^{-1}} \right\}, \tag{1.10}$$

où W^{-1} est une norme[4] de type L^2, à définir, caractéristique de l'incertitude potentielle entre Z et X. La solution de ce problème de minimisation est obtenue simplement en différentiant le critère,

$$\mathrm{d}J(X) = (Z - HX)^T W^{-1} H,$$

tel que le minimum est donné par $\mathrm{d}J(\hat{X}) = 0$ soit

$$\hat{X} = (H^T W^{-1} H)^{-1} H^T W^{-1} Z. \tag{1.11}$$

Du point de vue des notations, on notera dans toute la suite comme pour l'erreur d'estimation $\tilde{X} = X - \hat{X}$, l'erreur d'observation (ou résidu) $\tilde{Z} = Z - H\hat{X}$. Elle vérifie d'ailleurs l'identité

$$\tilde{Z} = (\mathbb{1} - H(H'H)^{-1} H')Z.$$

Si nous revenons au point de vue « probabiliste », la minimisation d'un tel critère s'interprète également et devient

$$\min \left\{ J(\chi) = \frac{1}{2} \|\chi\|^2_{W^{-1}} \right\}. \tag{1.12}$$

Dans ce cas, $W = \mathbb{E}(\chi\chi^T)$ est *la matrice de covariance du bruit* χ. Par exemple si les différentes composantes de la mesure ne sont pas corrélées alors

$$\mathbb{E}(\chi) = 0, \quad \mathbb{E}(\chi\chi^T) = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{pmatrix}.$$

[4] L'utilisation de l'inverse pour définir une norme est purement conventionnelle et permettra d'interpréter W à partir de covariances. Pour autant en dimension infinie c'est bien directement la norme qu'on saura véritablement définir et non son inverse.

et J s'écrit sous une forme bien connue en estimation moindres carrés depuis Gauss

$$J(\chi) = \frac{1}{2} \sum_{1 \leq i \leq m} \left(\frac{\chi_i}{\sigma_i} \right)^2 .$$

Inversement l'approche déterministe nous permettra rapidement d'obtenir une extension de ce résultat dans le cas où H est non-linéaire. On remarque aussi que finalement le LSE ici est bien un estimateur efficace (au sens statistique défini en introduction) puisqu'il vérifie bien une propriété de minimisation de covariance. En supposant Z gaussien, on peut montrer que c'est aussi le maximum de vraisemblance.

1.2.2 Terme d'ébauche

□ APPROCHE PROBABILISTE – Dans cette deuxième situation, on imagine disposer, en plus des observations, d'informations sur la loi de probabilité de X traduisant les connaissances a priori. Les hypothèses sur le bruit de mesure restent inchangées. En général, cette loi est connue uniquement à travers sa moyenne et covariance, ce qui naturellement la caractérise complètement dans le cas gaussien. On reprend stricto sensu, le cas présenté en introduction avec simplement la structure d'observation (1.9). Il est alors possible de préciser les covariances de Z et X

$$P^Z = \mathbb{E}(\chi\chi^T) + H\mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^T)H^T = W + HP^XH^T,$$

et

$$P^{XZ} = \mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^T)H^T = P^XH^T.$$

Dans ce cas on obtient directement

$$\hat{X} = \mathbb{E}(X) + P^XH^T(W + HP^XH^T)^{-1}(Z - H\mathbb{E}(X)). \tag{1.13}$$

Le lemme d'inversion matricielle nous fournit une autre structure sous la forme

$$\hat{X} = \mathbb{E}(X) + ((P^X)^{-1} + H^TW^{-1}H)^{-1}H^TW^{-1}(Z - H\mathbb{E}(X)). \tag{1.14}$$

De plus la covariance de l'erreur d'estimation est

$$P^a = P^X - P^{XZ}(P^Z)^{-1}P^{ZX} = P^X - P^XH^T(W + HP^XH^T)^{-1}P^XH^T \tag{1.15}$$

$$= ((P^X)^{-1} + P^{XZ}(P^{ZZ})^{-1}P^{ZX})^{-1}. \tag{1.16}$$

Rappelons que, dans le cas gaussien, c'est aussi la covariance de la loi de probabilité $p(X|Z)$ pour tout Z.

□ APPROCHE DÉTERMINISTE – Cette connaissance sur X se traduit du point de vue déterministe par l'existence d'un a priori X_0 sur X. Ce dernier est alors souvent appelé *terme d'ébauche* dans la littérature de l'assimilation de données (voir entres autres Talagrand (2003)) et traduit une connaissance simplement imparfaite sur le système. Le problème de l'estimation consiste alors à trouver l'indétermination ζ^X sur X telle que

$$\begin{cases} X = X_0 + \zeta^X \\ Z = HX + \chi \end{cases} \tag{1.17}$$

à partir des données d'ébauche X_0 et d'observation Z. Cette situation peut se ramener à la définition d'un inverse généralisé dans un espace d'observation augmenté en définissant une nouvelle observation notée

$$\mathcal{Z} = \tilde{H}X + \tilde{\chi},$$

avec

$$\mathcal{Z} = \begin{pmatrix} X_0 \\ Z \end{pmatrix}, \quad \tilde{H} = \begin{pmatrix} I \\ H \end{pmatrix}, \quad \tilde{\chi} = \begin{pmatrix} -\zeta^X \\ \chi \end{pmatrix},$$

où la norme sur l'incertitude est simplement tensorisée par rapport aux normes relatives sur χ et ζ^X

$$\tilde{W} = \begin{pmatrix} P_0 & 0 \\ 0 & W \end{pmatrix}.$$

On doit donc minimiser $J(X) = \frac{1}{2}\|\mathcal{Z} - \tilde{H}X\|^2_{\tilde{W}^{-1}}$, qui réécrit sur les variables initiales devient

$$\min_X \left\{ J(X) = \frac{1}{2}\|Z - HX\|^2_{W^{-1}} + \frac{1}{2}\|X_0 - X\|^2_{P_0^{-1}} \right\}. \tag{1.18}$$

On retrouve alors le même critère que pour l'estimation probabiliste, la tensorisation des normes traduisant formellement la notion d'indépendance des bruits-incertitudes. Si on associe la norme déterministe P_0 à la covariance P_X, les équations pour l'estimateur sont les mêmes ici que dans le cadre probabiliste mais apparaissent naturellement sous la forme (1.14)

$$\hat{X} = X_0 + (P_0^{-1} + H^T W^{-1} H)^{-1} H^T W^{-1}(Z - HX_0). \tag{1.19}$$

Pour obtenir la première forme (1.13), il faut évidement utiliser le lemme d'inversion matricielle « en sens inverse ». Pour cela on définit P_0^+ par

$$P_0^+ = (P_0^{-1} + H^T W^{-1} H)^{-1}, \tag{1.20}$$

d'où

$$\hat{X} = X_0 + P_0^+ H^T W^{-1}(Z - HX_0), \tag{1.21}$$

et on remarque immédiatement que P_0^+ est le P^a obtenu dans le cadre probabiliste c'est-à-dire la *covariance d'analyse* ou *covariance a posteriori*. On peut alors définir

$$K = P_0^+ H^T W^{-1} = (P_0^{-1} + H^T W^{-1} H)^{-1} H^T W^{-1} = P_0 H^T (W + HP_0H^T)^{-1},$$

tel que $\hat{X} = X_0 + K(Z - HX_0)$ est bien un estimateur linéaire en X_0.

Il est intéressant de constater que, autant dans l'approche probabiliste il est plus naturel de directement considérer cette situation d'estimation où X et Z sont aléatoires, autant dans celle déterministe le résultat découle directement de l'inverse généralisé de la section précédente. En effet, on prend conscience, avec cette dernière démonstration, que toute l'information peut initialement être considérée comme une observation au lieu d'une connaissance a priori sur le système. Ces notions d'observation et de connaissance sont donc relatives à un contexte donné, et l'objectif de la fonctionnelle J est de permettre la comparaison des différents éléments de connaissance sur le système dans les espaces qui les caractérisent. Notamment, l'importance donnée à chaque partie du critère est entièrement contenue dans le choix des matrices W et P_0 définissant les normes. Nous avons vu que W s'interprétait comme la covariance du bruit

d'observation, et P_0 est appelée aussi *covariance a priori*. Par exemple, si W traduit sur χ un bruit de faible amplitude alors la partie observation du critère sera prépondérante et la connaissance jouera simplement le rôle d'une régularisation de *Tikhonov*. En effet, réécrivons l'estimateur sous la forme

$$\hat{X} = (P_0^{-1} + H^T W^{-1} H)^{-1}(P_0^{-1} X_0 + H^T W^{-1} Z), \tag{1.22}$$

alors si P_0^{-1} est petit devant $H^T W^{-1} H$, celui ne sert qu'à rattraper un éventuel défaut d'injectivité de l'opérateur H qui ne permettrait pas de définir l'inverse généralisé de la section précédente ($H'H$ n'est pas inversible dans ce cas). Ce défaut d'injectivité définit d'ailleurs le caractère *mal posé* du problème inverse qu'on cherche à résoudre ici.

Afin de préciser le cas limite $P_0^{-1} \to 0$, il est possible de reformuler le problème de minimisation sous la forme d'un problème mixte. En effet, on a

$$dJ(X) = 0 \Leftrightarrow P_0^{-1}(X - X_0) - H^T W^{-1}(Z - HX) = 0.$$

En introduisant alors la variable

$$\Lambda = W^{-1}(Z - HX),$$

on est donc ramené à résoudre le système étendu de taille $m + N$

$$\begin{pmatrix} W & H \\ H^T & -P_0^{-1} \end{pmatrix} \begin{pmatrix} \Lambda \\ X \end{pmatrix} = \begin{pmatrix} Z \\ -P_0^{-1} X_0 \end{pmatrix}. \tag{1.23}$$

Or « P_0^{-1} petit devant W^{-1} fixé » est asymptotiquement équivalent à « W petit devant P_0 fixé » ce qui donne à la limite dans la formulation mixte

$$\begin{pmatrix} 0 & H \\ H^T & P_0^{-1} \end{pmatrix} \begin{pmatrix} \Lambda \\ X_0 \end{pmatrix} = \begin{pmatrix} \Pi_{\mathrm{Im(H)}} Z \\ P_0^{-1} X_0 \end{pmatrix}. \tag{1.24}$$

où $\Pi_{\mathrm{Im(H)}}$ est la projection orthogonale selon la norme W^{-1} sur l'espace image de H. Ce problème est alors équivalent à la minimisation sous contrainte

$$\min_{X \,|\, \Pi_{\mathrm{Im(H)}} Z = HX} \left\{ J_0(X) = \frac{1}{2}(X - X_0)^T P_0^{-1}(X - X_0) \right\}. \tag{1.25}$$

Notons que, dans le cas où H est injectif, la contrainte $\Pi_{\mathrm{Im(H)}} Z = HX$ caractérise directement la solution et on retrouve l'inverse généralisé (y compris lorsque H n'est pas surjectif du fait de la projection). Cette minimisation a donc vocation à fournir une solution limite générale y compris pour H non injectif et/ou non surjectif.

Cependant, dans tout ce qui suit, rien ne permet de préjuger de l'importance de la contribution d'observation par rapport à l'a priori. Ainsi, on peut très bien considérer que notre confiance est prépondérante dans X_0 et que les observations ne vont être qu'une forme de régularisation.

1.2.3 Estimation récursive

Dans la dernière situation de la présente introduction, sans doute plus artificielle que les précédentes mais néanmoins instructive sur la forme, on suppose disposer d'une série de mesures

obtenues séquentiellement par rapport au système. Autrement dit, soit $(Z_n)_n$ une suite de mesures obtenues à différents instants t_n par l'intermédiaire d'un opérateur d'observation H_n que l'on suppose différent à chaque instant.

$$Z_n = H_n X + \chi_n.$$

On souhaite construire un estimateur et, si possible tel que celui-ci puisse lui aussi être calculé récursivement afin de bénéficier des temps $t < t_n$ directement dans l'estimation à t_n.

□ APPROCHE DÉTERMINISTE – Plaçons nous à l'instant t_n, et écrivons l'estimateur complet en imaginant qu'il n'y ait pas d'ébauche. On pose

$$\mathcal{Z} = \tilde{H} X + \tilde{\chi},$$

avec

$$\mathcal{Z} = \begin{pmatrix} Z_0 \\ \vdots \\ Z_n \end{pmatrix}, \quad \tilde{H} = \begin{pmatrix} H_0 \\ \vdots \\ H_n \end{pmatrix}, \quad \tilde{\chi} = \begin{pmatrix} \chi_0 \\ \vdots \\ \chi_n \end{pmatrix}.$$

Si on suppose de plus les incertitudes de mesures indépendantes entre elles, la norme W est tensorisable et donc diagonale bloc

$$W = \begin{pmatrix} W_0 & & 0 \\ & \ddots & \\ 0 & & W_n \end{pmatrix}.$$

Comme dans la remarque de la Section 1.2.1, on obtient immédiatement

$$\hat{X}_n = \Big(\sum_{k=0}^{n} H_k^T W_k^{-1} H_k\Big)^{-1} \sum_{k=0}^{n} H_k^T W_k^{-1} Z_k. \tag{1.26}$$

Posons alors la suite d'opérateurs symétriques $(P_n)_n$ définie par

$$P_n^{-1} = \sum_{k=0}^{n} H_k^T W_k^{-1} H_k. \tag{1.27}$$

Clairement, $(P_n)_n$ vérifie la relation de récurrence

$$P_n^{-1} = H_n^T W_n^{-1} H_n + P_{n-1}^{-1}. \tag{1.28}$$

Cette relation de récurrence peut être réécrite directement sur P_n en utilisant le lemme d'inversion matricielle. On obtient ainsi dans notre cas

$$P_n = P_{n-1} - P_{n-1} H_n^T (W_n + H_n P_{n-1} H_n^T)^{-1} H_n P_{n-1}. \tag{1.29}$$

Ainsi $\hat{X}_n$ se déduit récursivement de $\hat{X}_{n-1}$ par

$$\begin{aligned} \hat{X}_n &= P_n \Big(\sum_{k=0}^{n} H_k^T W_k^{-1} Z_k\Big) = P_n P_{n-1}^{-1} \hat{X}_{n-1} + P_n H_n^T W_n^{-1} Z_n \\ &= \hat{X}_{n-1} + P_n H_n^T W_n^{-1} Z_n - K_n H_n \hat{X}_{n-1}, \end{aligned}$$

où on définit

$$K_n = P_{n-1}H_n^T(W_n + H_nP_{n-1}H_n^T)^{-1}. \tag{1.30}$$

Quelques manipulations matricielles inspirées du lemme d'inversion matriciel permettent de simplifier l'expression ci-dessus :

$$\begin{aligned}
K_n &= P_nP_n^{-1}P_{n-1}H_n^T(W_n + H_nP_{n-1}H_n^T)^{-1} \\
&= P_n(H_n^TW_n^{-1}H_n + P_{n-1}^{-1})P_{n-1}H_n^T(W_n + H_nP_{n-1}H_n^T)^{-1} \\
&= P_n(H_n^TW_n^{-1}H_nP_{n-1}H_n^T + H_n^T)(W_n + H_nP_{n-1}H_n^T)^{-1} \\
&= P_nH_n^TW_n^{-1}(H_nP_{n-1}H_n^T + W_n)(W_n + H_nP_{n-1}H_n^T)^{-1} \\
&= P_nH_n^TW_n^{-1}. \qquad (1.31)
\end{aligned}$$

L'estimateur s'écrit finalement

$$\hat{X}_n = \hat{X}_{n-1} + K_n(Z_n - H_n\hat{X}_{n-1}). \tag{1.32}$$

Evidemment, toutes ces manipulations ne sont possibles que sous condition que P_n est inversible. C'est la condition nécessaire d'existence de l'inverse généralisé. On remarque de plus que l'expression (1.32) est algébriquement la même que l'équation (1.21). Ainsi, dans cette configuration, l'estimateur moindres carrés peut s'exprimer séquentiellement en considérant à chaque nouveau pas de temps $\hat{X}_{n-1}$ comme le nouveau terme d'ébauche. D'ailleurs on peut pousser ce raisonnement plus loin et oublier l'expression (1.27) pour ne garder que les expressions récursives (1.28) et (1.29) initialisées en fonction des connaissances a priori à t_0 (cf. Remarque 1.2.3). Dans ce cas, ce terme d'ébauche supplémentaire assure l'inversibilité de P_n à tout instant comme expliqué à la section précédente 1.2.2.

REMARQUE 1.2.2
Dans le cas où P_n n'est pas inversible, il est cependant possible d'exprimer un filtre à partir d'un inverse généralisé de P_n à l'image de ce que nous avions abordé en introduction. Ce cas sera largement étudié au quatrième chapitre et nous le laissons donc de côté pour l'instant.

Une dernière manipulation algébrique permet finalement d'exprimer P_n en fonction de K_n :

$$\begin{aligned}
P_n &= P_{n-1} - P_{n-1}H_n^T(W_n + H_nP_{n-1}H_n^T)^{-1}H_nP_{n-1} \\
&= (\mathbb{1} - K_nH_n)P_{n-1}. \qquad (1.33)
\end{aligned}$$

Cette formulation peut-être symétrisée par utilisation de nouveau du lemme d'inversion matricielle et donne, tout calcul fait, une expression que nous n'utiliserons pas mais qu'on retrouve régulièrement dans les ouvrages

$$P_n = (\mathbb{1} - K_nH_n)P_{n-1}(\mathbb{1} - K_nH_n)^T + K_nW_nK_n.$$

REMARQUE 1.2.3 (INITIALISATION ET TERME D'ÉBAUCHE)
Dans le cas où il n'y a pas d'a priori et où on dispose d'observation dès l'instant 0, la récurrence est intialisée de telle sorte que

$$P_{n=0}^{-1} = H_0^TW_0^{-1}H_0.$$

Dans le cas où on dispose d'un terme d'ébauche, lui-même pouvant être considéré comme une observation supplémentaire à l'instant initial, on initialise suivant

$$X_{n=0} = X_0 + P_{n=0}H_0^TW_0^{-1}(Z_0 - H_0X_0),$$

avec

$$P_{n=0}^{-1} = H_0^T W_0^{-1} H_0 + P_0^{-1}.$$

Ainsi, la forme « non récursive » (1.26) est modifiée par la prise en compte de l'a priori suivant

$$\hat{X}_n = \Big(P_0^{-1} + \sum_{k=0}^{n} H_k^T W_k^{-1} H_k\Big)^{-1} \Big(X_0 + \sum_{k=0}^{n} H_k^T W_k^{-1} Z_k\Big). \tag{1.34}$$

□ APPROCHE PROBABILISTE – Encore une fois, insistons sur le fait que cette écriture séquentielle n'est possible que parce que W est diagonale bloc. Du point de vue probabiliste, cela signifie que la covariance entre les bruits d'observation est diagonale donc que les bruits χ_n sont indépendants les uns des autres. Dans l'hypothèse supplémentaire où chaque χ_n est gaussien, on appelle ce type de bruit un *bruit blanc* en temps. Nous utilisons ici la notion de projection orthogonale vue en introduction[5] pour obtenir l'estimateur directement de manière récursive. En effet, introduisons l'espace défini à t_n par

$$\mathcal{V}_n = \text{Vect}(1, Z_0, \ldots, Z_n) = \mathcal{V}_{n-1} \oplus \text{Vect}(Z_n - \Pi_{\mathcal{V}_{n-1}}(Z_n))$$

le vecteur aléatoire $c_n = (Z_n - \Pi_{\mathcal{V}_{n-1}}(Z_n))$ de $\mathcal{V}_n$ orthogonal à $\mathcal{V}_{n-1}$ est appelé *composante d'innovation*. Ainsi notre meilleur estimateur linéaire est directement obtenu récursivement

$$\hat{X}_n = \hat{X}_{n-1} + K_n(Z_n - H_n \hat{X}_{n-1}),$$

où

$$K_n = \text{Cov}(X, c_n)\text{Cov}(c_n)^{-1}.$$

Or l'indépendance des bruits donne

$$c_n = Z_n - \Pi_{\mathcal{V}_{n-1}}(Z_n) = Z_n - H_n \hat{X}_{n-1} = \tilde{Z}_n$$

donc

$$\text{Cov}(X, c_n) = \mathbb{E}((X - \hat{X}_{n-1})\tilde{Z}_n^T) = \mathbb{E}((X - \hat{X}_{n-1})(X - \hat{X}_{n-1})^T) H_n^T = P_{n-1} H_n^T,$$

avec $P_{n-1} = \text{Cov}(X - \hat{X}_{n-1})$. De même

$$\text{Cov}(c_n) = \text{Cov}(\tilde{Z}_n) = H_n P_{n-1} H_n^T + W_n,$$

donc finalement

$$\hat{X}_n = \hat{X}_{n-1} + P_{n-1} H_n^T (H P_{n-1} H_n^T + W_n)^{-1} (Z_n - H_n \hat{X}_{n-1}). \tag{1.35}$$

On vérifie de plus directement que P_n satisfait les mêmes relations de récurrence que précédemment

$$P_n = P_{n-1} - P_{n-1} H_n^T (W_n + H_n P_{n-1} H_n^T)^{-1} H_n P_{n-1},$$

et pour le filtre

$$\begin{aligned} K_n &= P_{n-1} H_n^T (H_n P_{n-1} H_n^T + W_n)^{-1} \\ &= P_n H_n^T W_n^{-1}. \end{aligned}$$

Pour le MMSE, si initialement les probabilités sont gaussiennes, la linéarité de l'opérateur d'observation permet le même type de projection à partir, cette fois, de l'espérance conditionnelle (qui est opérateur linéaire). BLUE et MMSE ont donc la même formulation récursive sous l'hypothèse de variables gaussiennes.

[5] voir la définition sur un module de Hilbert quand nous avons introduit le BLUE

1.3 Estimation pour les systèmes dynamiques linéaires

Nous allons désormais considérer un problème d'estimation pour un système dynamique alternativement discret en temps ou continu. A ce stade, nous ne mènerons plus à terme l'exercice, trop ambitieux, de démontrer les résultats d'estimation à la fois dans le cadre probabiliste et déterministe. De notre point de vue, nous privilégierons l'approche déterministe, notamment car le point de vue probabiliste en temps continu nous obligerait à introduire le formalisme des ODS (puis EDS d'ailleurs). Or nous n'utiliserons au final (chapitre 3) qu'une partie des notions de cette théorie en ne considérant que quelques aspects des processus de Wiener et des variables cylindriques introduites par Bensoussan (1971). Ainsi l'aspect probabiliste nous fournira préférentiellement un cadre d'interprétation formel par lequel on obtient rapidement la formulation algébrique du filtre de Kalman. Sa version déterministe, elle, nous conduira directement à établir des schémas en temps pertinents pour le problème d'estimation en temps continu. A noter que dans ce premier chapitre introductif les mesures sont supposées être disponibles à tout instant de l'évolution du système dynamique. Ainsi en toute généralité dans le cadre continu en temps, le problème que nous considérons est l'estimation du système

$$\begin{cases} \dot{X} = A(X, \omega(t), t) \\ X(0) = X_0 + \zeta^X, \end{cases} \tag{1.36}$$

à partir de mesures toujours sous la forme $Z = H(X) + \chi$. Dans cette modélisation, ω représente potentiellement une incertitude sur la dynamique que nous ne considérerons pas la plupart du temps pour la raison suivante. Cette erreur de dynamique s'interprète comme une erreur de modélisation, c'est-à-dire qu'on ne donne à l'estimateur du système qu'une connaissance imparfaite de la vraie dynamique du système qu'il estime. On sait que ces erreurs de modèle sont souvent critiques en assimilation de données comme le montrent les travaux réalisés dans les disciplines historiques de l'assimilation (voir, parmi de nombreux exemples, Mallet et Sportisse (2006) en qualité de l'air). Cependant elle peuvent aussi permettre de masquer le véritable comportement des estimateurs car « sous elles » on peut cacher ou, au contraire, surestimer beaucoup de choses. En d'autres termes, il apparaît parfois que toute estimation bonne ou mauvaise d'un phénomène dynamique aurait pu être une bonne estimation pour un autre modèle et un certain bruit. Donc, si on veut juger un estimateur et sans doute par la suite notre capacité de modélisation de la dynamique du phénomène qu'on estime, il vaut mieux, dans un premier temps, ne pas accepter d'erreurs de modèles dans l'estimation et dire qu'on fait l'effort d'avoir une connaissance fine de la dynamique du système.

Enfin, précisons que dans cette section nous nous limitons pour l'instant à une dynamique linéaire.

1.3.1 Cas discret en temps : Cadre déterministe

On considère donc le système dynamique linéaire $(X_n)_n$ défini à partir d'un opérateur de transition $A_{n+1|n}$,

$$X_{n+1} = A_{n+1|n} X_n + R_n, \tag{1.37}$$

pour lequel on suppose incertaine la condition initiale $X_{n=0}$. Ainsi, il existe ζ^X tel que cette condition initiale s'exprime à partir de l'a priori X_0 sous la forme

$$X_{n=0} = X_0 + \zeta^X.$$

Le système dynamique est donc entièrement défini par

$$\begin{cases} X_{n+1} = A_{n+1|n} X_n + R_n \\ X_{n=0} = X_0 + \zeta^X \end{cases} \tag{1.38}$$

et les observations sont elles-mêmes données par un opérateur d'observation linéaire

$$Z_n = H_n X_n + \chi_n.$$

La démarche que nous allons suivre est similaire à celle exposée Section 1.2.3 où on a introduit un estimateur à partir d'un inverse généralisé avant de le reformuler sous forme récursive. Pour cela, nous commençons par exprimer l'ensemble des états et des observations à partir de l'incertitude « originelle » qu'est la condition initiale. Soit, d'une part, $\mathcal{R}_{n|0}$ la résolvante discrète du système définissant l'état X_n à partir de $X_{n=0}$, on a

$$X_n = \mathcal{R}_{n|0} X_{n=0} = A_{n|0} X_{n=0} + \sum_{k=0}^{n-1} A_{n|k+1} R_k. \tag{1.39}$$

où $A_{n|k+1} = A_{n|n-1} \circ \cdots \circ A_{k+2|k+1}$ et avec la convention $A_{n|n} = \mathbb{1}$. D'autre part, les observations peuvent être représentées par un vecteur d'observations augmenté du type

$$\mathcal{Z} = \tilde{H} X_0 + \tilde{\chi} \text{ avec } \mathcal{Z} = \begin{pmatrix} Z_0 \\ \vdots \\ Z_n - H_n \sum_{k=0}^{n-1} A_{n|k+1} R_k \end{pmatrix}, \ \tilde{H} = \begin{pmatrix} H_0 \\ \vdots \\ H_n \cdot A_{n|0} \end{pmatrix}.$$

où apparaît implicitement la dynamique de (X_n). Si on suppose les bruits de mesure indépendants au sens déterministe où la norme associée à $\tilde{X}$ est la tensorisation des W_n (cf. Section 1.2.3), on peut donc définir $\check{X}_n$ l'estimation de $X_{n=0}$ à partir, seulement, de l'observation $\mathcal{Z}$ soit (cf. expression (1.34)),

$$\check{X}_n = \Big(P_0^{-1} + \sum_{k=0}^{n} \tilde{H}_k^T W_k^{-1} \tilde{H}_k \Big)^{-1} \Big(X_0 + \sum_{k=0}^{n} \tilde{H}_k^T W_k^{-1} \mathcal{Z}_k \Big).$$

Cette même Section 1.2.3 nous a permis alors d'exprimer $\check{X}_n$ sous une forme récursive

$$\check{X}_n = \check{X}_{n-1} + \check{K}_n (\mathcal{Z}_n - \tilde{H}_n \check{X}_{n-1})$$

et

$$\check{K}_n = \check{P}_n H_n^T W_n^{-1}.$$

Notre objectif est maintenant de formuler une version récursive de l'estimation X_n lui-même. Pour ce faire, on introduit

$$\hat{X}_n^+ = \mathcal{R}_{n|0} \check{X}_n = A_{n|0} \check{X}_n + \sum_{k=0}^{n-1} A_{n|k+1} R_k,$$

et on vérifie alors l'identité

$$\hat{X}_n^+ = \hat{X}_n^- + A_{n|0}\check{K}_n(Z_n - H_n\hat{X}_n^-),$$

en définissant

$$\hat{X}_n^- = A_{n|n-1}\hat{X}_{n-1}^+ + R_{n-1}. \tag{1.40}$$

Reformulons finalement le filtre à partir des relations de récurrence sur $\check{P}_n$ que nous rappelons

$$\check{P}_n^{-1} = \tilde{H}_n^T W_n^{-1}\tilde{H}_n + \check{P}_{n-1}^{-1} = A_{n|0}^T H_n^T W_n^{-1} H_n A_{n|0} + \check{P}_{n-1}^{-1}.$$

En définissant

$$P_n^+ = A_{n|0}\check{P}_n A_{n|0}^T,$$

et

$$P_n^- = A_{n|n-1}(P_{n-1}^+)A_{n|n-1}^T, \tag{1.41}$$

on en déduit directement que

$$(P_n^+)^{-1} = H_n^T W_n^{-1} H_n + A_{n|n-1}^{-T}(P_{n-1}^+)^{-1}A_{n|n-1}^{-1} = H_n^T W_n^{-1} H_n + (P_n^-)^{-1} \tag{1.42}$$

Ainsi, il existe un filtre K_n tel que

$$\hat{X}_n^+ = \hat{X}_n^- + K_n(Z_n - H_n\hat{X}_n^-), \tag{1.43}$$

s'exprimant sous une forme similaire à celles présentées Section 1.2.3

$$\begin{aligned} K_n = A_{n|0}\check{K}_n = A_{n|0}\check{P}_n A_{n|0}^T H_n^T W_n^{-1} &= P_n^+ H_n^T W_n^{-1} \\ &= P_n^- H_n^T (W_n + H_n P_n^- H_n^T)^{-1}, \end{aligned}$$

où la dernière transformation utilise la même technique algébrique que dans l'équation (1.31).

L'ensemble des équations (1.40),(1.41),(1.42) et (1.43) définissant respectivement $\hat{X}_n^-$, $\hat{X}_n^+$, P_n^- et P_n^+ construisent un procédé d'estimation dynamique plus connu sous le nom de *filtrage de Kalman-Bucy* d'après l'article historique de 1960 Kalman et Bucy (1961). La démonstration est d'ailleurs assez éloignée de celle-ci puisqu'elle repose sur le filtre optimal de Wiener dont Kalman propose un version récursive dans un cadre probabiliste (cf. section suivante pour un résumé des grandeurs probabiliste sous-jacentes). C'est un estimateur appelé aussi prédicteur-correcteur (voir schéma (1.1)) où on note la prédiction $\hat{X}_{n+1}^-$, ou encore $\hat{X}_{n+1|n}$, et la correction $\hat{X}_{n+1}^+$ ou encore $\hat{X}_{n+1|n+1}$. De même, $P_{n+1|n}$ et $P_{n+1|n+1}$ et sont appelées respectivement covariance de prédiction et d'analyse d'après l'interprétation que nous en ferons à la section suivante. A noter enfin que le filtre de Kalman est *stricto-sensu* l'opérateur K_n mais on utilise parfois cette dénomination pour décrire l'estimateur global.

REMARQUE 1.3.1 (INTIALISATION)
Avec un terme d'ébauche et des observation à l'instant 0, la procédure commence par initialisation de

$$(P_0^+)^{-1} = P_0^{-1} + H_0^T W_0^{-1} H_0,$$

et

$$X_{n=0}^+ = X_0 + K_0(Z_0 - H_0 X_0).$$

Mais on pourrait tout aussi dire que $P_0^- = P_0$ et $X_{n=0}^- = X_0$. Les expressions peuvent être modifiées en conséquence en l absence d'a priori où d'observation initiale. En particulier, dans de nombreux cas, l'a priori prend déjà en compte l'observation à t_0 et c'est pourquoi l'algorithme est présenté le plus souvent en commençant par l'étape de prédiction.

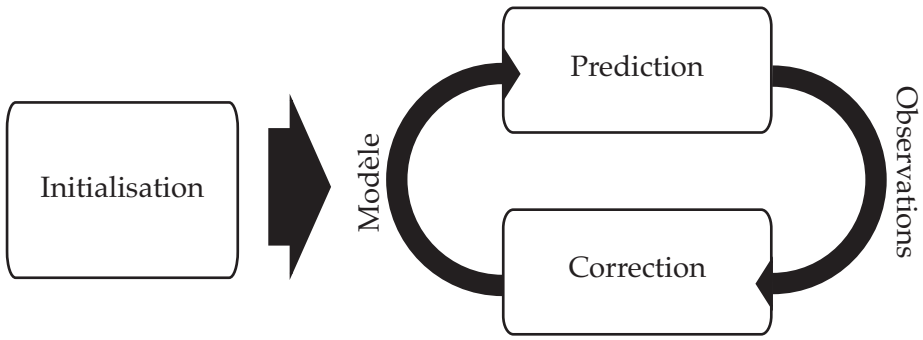


FIGURE 1.1: Filtre de Kalman discret en temps vu comme un algorithme prédiction correction.

Pour revenir à une structure de système dynamique plus conventionnelle, il est évidemment possible de réécrire sous forme séquentielle classique cette procédure d'estimation-correction en définissant

$$\hat{X}_n = \hat{X}_n^-, \quad P_{n+1} = P_{n+1}^-.$$

La dynamique devient alors

$$\begin{cases} K_{n+1} = P_{n+1} H_{n+1}^T (H P_{n+1} H_{n+1}^T + W_{n+1})^{-1} \\ \hat{X}_{n+1} = A_{n+1,n} \hat{X}_n + R_{n+1|n} + K_{n+1}\big(Z_{n+1} - H_{n+1}(A_{n+1|n}\hat{X}_n + R_{n+1|n})\big) \\ P_{n+1} = A_{n+1|n} P_n A_{n+1|n}^T - A_{n+1|n} P_n H_n^T (H_n P_n H_n^T + W_n)^{-1} H_n P_n A_{n+1|n}^T \end{cases} \tag{1.44}$$

Cette dynamique sur la prédiction (il existe de même une dynamique uniquement à partir de la correction) sera appelée dynamique « *one-step* », à la différence des expressions prédiction-correction vues plus haut.

REMARQUE 1.3.2
Nous venons de formuler les équations du filtrage de Kalman uniquement à partir d'inverse généralisé. Ce type d'arguments se retrouve dans le très complet Hassibi et al (1999) qui en profite d'ailleurs pour proposer l'extension au cas W non positive où l'inverse généralisé est posé sur un espace de Krein et non de Hilbert. Ceci lui permet d'introduire directement l'estimation H^∞ que, pour notre part, nous présenterons classiquement Section 1.4.2 via la théorie du contrôle optimal.

En anticipant un peu sur la suite, l'approche que nous venons de présenter justifie aussi les théorèmes relatifs à l'observabilité des systèmes que nous présenterons Section 1.3.5.

Comme précédemment le filtre de Kalman est aussi un estimateur moindres carrés. Pour présenter la fonctionnelle associée on considère pour la même dynamique (1.37) une suite $(X_n(\xi))_n$ tel que $X_{n=0} = X_0 + \xi$. Etant données les observations Z_n et l'a priori sur la condition initiale X_0, on cherche alors le paramètre ξ optimal pour le critère quadratique

$$\min_\xi \left\{ J_n(\xi) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{k=0}^{n} \|Z_k - H_k X_k(\xi)\|^2_{W_k^{-1}} \right\}. \tag{1.45}$$

PROPOSITION 1.3.1
L'estimateur formulé par le filtre de Kalman, dont les équations « one step » sont données par (1.44)*, vérifie le critère optimal discret*

$$(\hat{X}_n)_n = (X_n(\operatorname*{argmin}_{\xi} J_n))_n.$$

■ ***Démonstration :*** *Dans ce cadre récursif, $X_n(\xi)$ est une fonction affine de ξ et on démontre sans difficulté*[6] *qu'ainsi J est convexe et dérivable par apport à ξ. Dans ce cas, J admet un unique minimum obtenu une fois encore par $\mathrm{d}J = 0$. Or*

$$\mathrm{d}J_n(\xi) \cdot \delta\xi = \xi^T P_0^{-1} \delta\xi - \sum_{k=0}^{n} (Z_k - H_k X_k)^T W_k^{-1} H_k \, \mathrm{d}_\xi X_k \cdot \delta\xi,$$

où

$$\mathrm{d}_\xi X_{k+1} = A_{k+1|k} \, \mathrm{d}_\xi X_k, \quad \mathrm{d}_\xi X_0 = \mathbb{1}.$$

On introduit alors une suite ***finie*** $(p_k)_{1 \le k \le n+1}$ *vérifiant la dynamique adjointe de celle de X_k entre $0 \le k \le n$*

$$\begin{cases} p_k - A_{k+1|k}^T p_{k+1} = H_k^T W_k^{-1} (Z_k - H X_k) \\ p_{n+1} = 0 \end{cases} \tag{1.46}$$

L'expression de la différentielle se simplifie alors

$$\begin{aligned} \mathrm{d}J_n(\xi) \cdot \delta\xi &= \xi^T P_0^{-1} \delta\xi - \sum_{k=0}^{n} (Z_k - H_k X_k)^T W_k^{-1} H_k \, \mathrm{d}_\xi X_k \cdot \delta\xi \\ &= \xi^T P_0^{-1} \delta\xi + \sum_{k=0}^{n} (-p_k + A_{k+1|k}^T p_{k+1})^T \, \mathrm{d}_\xi X_k \cdot \delta\xi \\ &= \xi^T P_0^{-1} \delta\xi - \sum_{k=0}^{n} p_k^T \, \mathrm{d}_\xi X_k \cdot \delta\xi + \sum_{k=0}^{n} p_{k+1}^T \, \mathrm{d}_\xi X_{k+1} \cdot \delta\xi \\ &= \xi^T P_0^{-1} \delta\xi - p_0^T \delta\xi. \end{aligned}$$

On obtient donc $\bar{\xi} = P_0 p_0$.

A partir des équations de Kalman « one-step » définies précédemment, on vérifie alors par récurrence que

$$X_k(\bar{\xi}) = \hat{X}_k^- + P_k^- p_k, \quad 1 \le k \le n. \tag{1.47}$$

En effet, en utilisant la covariance a posteriori $P_{k+1}^- = A_{k+1|k} P_k^+ A_{k+1|k}^T$, on a

$$\begin{aligned} \hat{X}_{k+1}^- &= X_{k+1} - A_{k+1|k} P_k^+ A_{k+1|k}^T p_{k+1} \\ &= X_{k+1} - A_{k+1|k} P_k^+ (p_k - H_k^T W^{-1} (Z_k - H_k X_k)) \\ &= A_{k+1|k} (\hat{X}_k^- + P_k^- p_k) + R_k - A_{k+1|k} P_k^+ p_k + A_{k+1|k} P_k^+ H_k^T W^{-1} (Z_k - H_k (\hat{X}_k^- + P_k^- p_k)) \\ &= A_{k+1|k} \hat{X}_k^- + R_k + A_{k+1|k} P_k^+ H_k^T W^{-1} (Z_k - H_k \hat{X}_k^-) \\ &\qquad + (A_{k+1|k} P_k^- + A_{k+1|k} P_k^+ - A_{k+1|k} P_k^+ H_k^T W^{-1} H_k P_k^-) p_k. \end{aligned}$$

[6] voir aussi la démonstration pour le problème en temps continu qui demande un tout petit peu plus de soin

Or la matrice multipliant p_n est nulle car, en la multipliant par $A^T_{k+1|k}$, on a alors

$$\begin{aligned}
&A_{k+1|k}P_k^- A^T_{k+1|k} + A_{k+1|k}P_k^+ A^T_{k+1|k} - A_{k+1|k}P_k^+ H_k^T W^{-1} P_k^- A^T_{k+1|k} \\
&\quad = A_{k+1|k}P_k^- A^T_{k+1|k} + P^-_{k+1} - A_{k+1|k}K_k P_k^- A^T_{k+1|k} \\
&\quad = A_{k+1|k}P_k^- A^T_{k+1|k} + P^-_{k+1} - A_{k+1|k}P_k^- H_k^T (H_k P_k^- H_k^T + W_k)^{-1} H_k P_k^- A^T_{k+1|k} \\
&\quad = 0,
\end{aligned}$$

ce qui démontre la relation (1.47)*. En particulier on a*

$$\forall n, \quad \hat{X}^-_{n+1} = X_{n+1}(\mathit{argmin} J_n),$$

où n indice à la fois la suite et le critère. ■

Cette présentation peut sembler un peu « laborieuse ». En fait, la Section 1.3.3, en continu, apportera un éclairage plus naturel sur cette preuve qui, en discret, a le gros avantage de ne présenter aucune difficulté théorique mais d'être simplement technique du point de vue algébrique.

1.3.2 Cas discret en temps : Cadre probabiliste

Une approche probabiliste pemettant de retrouver rapidement les équation du filtre de Kalman s'appuie sur les mêmes arguments qu'à la Section 1.2.3 où nous avions vu comment le MMSE ou le meilleur estimateur linéaire pouvaient s'obtenir récursivement à partir de la composante d'innovation lorsque l'ensemble des observations augmentait. De même, il est élémentaire de déterminer comment l'estimateur est propagé par l'opérateur de transition. Afin d'écrire le MMSE, on utilise les notations suivantes définies pour tout n :

estimateur a priori : $\hat{X}^-_{n+1} = \mathbb{E}(X_{n+1}|Z_0, \ldots, Z_n)$,

covariance a priori : $P^-_{n+1} = \mathbb{E}((X_{n+1} - \hat{X}^-_{n+1})(X_{n+1} - \hat{X}^-_{n+1})^T)$,

estimateur a posteriori : $\hat{X}^+_{n+1} = \mathbb{E}(X_{n+1}|Z_0, \ldots, Z_{n+1})$,

covariance a posteriori : $P^+_{n+1} = \mathbb{E}((X_{n+1} - \hat{X}^+_{n+1})(X_{n+1} - \hat{X}^+_{n+1})^T)$,

dont l'interprétation nous est déjà apparue au cours des sections précédentes. Pour le BLUE, on peut définir les mêmes quantités mais dans ce cas l'espérance conditionnelle est simplement remplacée par $\hat{E}(X|Z)$ avec les mêmes propriétés de projecteur orthogonal. On a, au travers de la dynamique,

$$\begin{aligned}
\hat{X}^-_{n+1} &= \mathbb{E}(X_{n+1}|Z_0, \ldots, Z_n) \\
&= A_{n+1|n}\mathbb{E}(X_n|Z_0, \ldots, Z_n) + R_n \\
&= A_{n+1|n}\hat{X}^+_n + R_n,
\end{aligned} \tag{1.48}$$

et de même

$$\begin{aligned}
P^-_{n+1} &= \mathbb{E}((X_{n+1} - \hat{X}^-_{n+1})(X_{n+1} - \hat{X}^-_{n+1})^T) \\
&= A_{n+1|n}\mathbb{E}((X_n - \hat{X}^+_n)(X_n - \hat{X}^+_n)^T)A^T_{n+1|n} \\
&= A_{n+1|n}P^+_n A^T_{n+1|n}.
\end{aligned} \tag{1.49}$$

Puis l'intégration de la nouvelle observation donne pour le MMSE dans le cas de variables gaussiennes ou pour le BLUE en considérant $\hat{E}$,

$$\begin{aligned}\hat{X}^+_{n+1} &= \mathbb{E}(X_{n+1}|Z_0,\dots,Z_{n+1})\\ &= \hat{X}^-_{n+1} + K_{n+1}(Z_{n+1} - H_{n+1}\hat{X}^-_{n+1}),\end{aligned} \tag{1.50}$$

avec dans l'ordre (voir la démonstration de la section 1.2.3), plusieurs relations définissant le filtre

$$\begin{aligned}K_{n+1} &= P^-_{n+1}H^T_{n+1}(H_{n+1}P^-_{n+1}H^T_{n+1} + W_{n+1})^{-1}\\ &= P^+_{n+1}H^T_{n+1}W^{-1}_{n+1},\end{aligned} \tag{1.51}$$

et

$$\begin{aligned}P^+_{n+1} &= (\mathbb{1} - K_{n+1}H_{n+1})P^-_{n+1}\\ &= ((P^-_{n+1})^{-1} + H^T_{n+1}W^{-1}_{n+1}H_{n+1})^{-1}\\ &= (\mathbb{1} - K_{n+1}H_{n+1})P^-_{n+1}(\mathbb{1} - K_{n+1}H_{n+1})^T + K_{n+1}W_{n+1}K^T_{n+1}\end{aligned} \quad . \tag{1.52}$$

Ainsi, dans le cadre probabiliste, on formalise le filtrage de Kalman sous la forme du meilleur estimateur linéaire dans le cas d'opérateurs de transition et d'observation linéaires, ou le MMSE pour les variables gaussiennes[7].

1.3.3 Cas continu en temps

On considère le système dynamique suivant défini sur $[0, T]$

$$\begin{cases}\dot{X} = A(t)X + R(t)\\ X(0) = X_0 + \zeta^X\end{cases} \tag{1.53}$$

La condition initiale exacte est inconnue au sens où X_0 est une donnée mais pas ζ^X. En revanche, on dispose d'observations au cours du temps qui s'expriment en fonction de l'état par l'application d'un opérateur toujours supposé linéaire

$$Z = HX + \chi.$$

où $H = H(t)$ peut dépendre du temps bien que nous ne le précisons pas dans les notations.

Pour estimer ce système on souhaite trouver la trajectoire qui minimise la fonctionnelle de coût

$$\min_{\xi}\left\{J_T(\xi) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\int_0^T \|Z - HX_\xi\|^2_{W^{-1}}\,dt\right\}, \tag{1.54}$$

pour toute trajectoire de même dynamique que (1.53) mais de condition initiale

$$X_\xi(0) = X_0 + \xi.$$

Ce problème de minimisation est ainsi appelé **problème LQ** (*linéaire-Quadratique*) car la dynamique du système est linéaire pour un critère moindres carrés à minimiser.

[7] D'ailleurs les variables sont alors gaussiennes tout au long de l'estimation ce qui est nécessaire pour la récurrence

On suppose que $A(t)$ et $R(t)$ sont continus sur $[0,T]$ si bien que le système (1.53) admet une solution globale et que son flot s'exprime, dans un cadre linéaire, à partir de la résolvante du système homogène notée

$$\mathcal{R} : \left\{ \begin{array}{rcl} [0,T]^2 \times \mathbb{R}^N & \to & \mathbb{R}^N \\ (t_2, t_1, X_1) & \mapsto & X(t_2) \text{ solution de } \dot{X} = A(t)X,\ X(t_1) = X_1 \end{array} \right.$$

On rappelle que si le système est autonome (i.e. $A(t) = A$ est indépendant du temps) alors $\mathcal{R}(t_2, t_1) = e^{A(t_2 - t_1)}$. De plus, en toute généralité, $\mathcal{R}(t,t) = \mathbb{1}$ et la composition suit

$$\mathcal{R}(t_2, t_1) \circ \mathcal{R}(t_1, t_0) = \mathcal{R}(t_2, t_0).$$

Ainsi la résolvante est inversible d'inverse

$$\mathcal{R}^{-1}(t,s) = \mathcal{R}(s,t).$$

Enfin, on obtient la solution du problème non homogène par la méthode de variation de la constante

$$X(t) = \mathcal{R}(t,0)X(0) + \int_0^t \mathcal{R}(t,s)R(s)\,ds. \tag{1.55}$$

Afin de retrouver, dans le cadre continu, la dynamique du filtre de Kalman à partir de la minimisation de la fonctionnelle (1.54), nous suivons une démonstration très classique en contrôle optimal composée des étapes suivantes.

□ EXISTENCE DE TRAJECTOIRES MINIMISANTES – Commençons tout d'abord par nous assurer que le problème est bien posé, la preuve du lemme suivant offrant un éclairage supplémentaire sur la nécessité d'intégrer dans le critère un a priori sur ξ.

LEMME 1.3.2
Supposons que $P_0 > 0$, alors il existe une trajectoire minimisante pour le problème LQ

■ ***Démonstration :*** *Soit $(\xi_n)_n$ une suite minimisante pour la fonction coût J_T. Alors*

$$\forall \xi, \quad J_T(\xi) \geq \alpha \|\xi\|^2$$

implique que $(\xi_n)_n$ est borné et on peut extraire une sous-suite convergente. D'après la formule de variation de la constante, la suite de fonctions $(X_{\xi_n}(t))$ définie à partir de ξ_n converge simplement et on peut passer à la limite dans l'équation. On a alors

$$\bar{X}(t) = \mathcal{R}(t,0)(X_0 + \bar{\xi}) + \int_0^t \mathcal{R}(t,s)R(s)ds,$$

qui permet de passer à la limite dans le critère. Ainsi

$$J_T(\bar{\xi}) \leq \lim_n \left\{ \|\xi_n\|^2_{P_0^{-1}} + \int_0^T \|Z - HX_{\xi_n}\|^2_{W^{-1}}\,dt \right\} = \min_\xi J_T,$$

donc $\bar{\xi}$ minimise bien la fonctionnelle de coût. ■

□ CONVEXITÉ DE J_T – En plus de l'existence, on a unicité de la trajectoire minimisante car

LEMME 1.3.3
La fonction $\xi \mapsto J_T$ est strictement convexe

■ ***Démonstration :*** *Tout d'abord $\xi \mapsto \|\xi\|^2_{P_0^{-1}}$ est strictement convexe. De plus, par linéarité de la résolvante $\xi \mapsto X_\xi$ est clairement convexe. Enfin l'application $x \mapsto \|Z - HX\|^2_{W^{-1}}$ est strictement convexe et l'intégrale préserve la convexité. Donc par composition J_T est strictement convexe par rapport à ξ.* ■

□ CARACTÉRISATION DU MINIMISANT – Il est alors possible de caractériser le minimisant en introduisant la variable adjointe :

THÉORÈME 1.3.4
Pour la trajectoire optimale $X_{\bar{\xi}}$ on définit l'état adjoint $p : [0, T] \to R^N$ tel que

$$\begin{cases} \dot{p} + A(t)^T p = -H^T W^{-1}(Z - HX_{\bar{\xi}}) \\ p(T) = 0 \end{cases} \tag{1.56}$$

et alors

$$\bar{\xi} = \underset{\xi}{\operatorname{argmin}}\, J_T = P(0)p(0). \tag{1.57}$$

■ ***Démonstration :*** *Dans le cadre LQ, on peut directement dériver J_T au sens de Fréchet et le minimum vérifie*

$$\mathrm{d}J_T(\bar{\xi}) = 0$$

Par unicité du minimum global cette condition est nécessaire et suffisante.
Or on a pour tout accroissement $\delta\xi$

$$\mathrm{d}J_T \cdot \delta\xi = \xi^T P_0^{-1}\delta\xi - \int_0^T (Z - HX_\xi)^T W^{-1} H\, \mathrm{d}X_\xi \cdot \delta\xi\, dt,$$

où $\mathrm{d}X_\xi$ (dérivée par rapport à ξ de X_ξ) vérifie

$$\begin{cases} \dot{\widehat{\mathrm{d}X_\xi}} = A(t)\, \mathrm{d}X_\xi \\ \mathrm{d}X_\xi(0) = \mathbb{1} \end{cases}$$

Le système (1.56) *est bien défini sur $[0, T]$ pour tout X_ξ au second membre solution du système de Cauchy* (1.53). *En remplaçant alors $-(Z - HX_\xi)^T W^{-1} H$ par $\dot{p}^T + p^T A(t)$ et en intégrant par parties on obtient*

$$\begin{aligned} \mathrm{d}J_T \cdot \delta\xi &= \xi^T P_0 \delta\xi + \int_0^T (\dot{p} + A(t)^T p)^T\, \mathrm{d}X_\xi \cdot \delta\xi\, dt \\ &= \xi^T P_0^{-1}\delta\xi - \int_0^T p(t)^T\, \dot{\widehat{\mathrm{d}X_\xi}} \cdot \delta\xi\, dt + \left[p(t)^T\, \mathrm{d}X_\xi.\delta\xi\right]_0^T + \int_0^T p(t)^T A(t)\, \mathrm{d}X_\xi \cdot \delta\xi\, dt \\ &= \xi^T P_0^{-1}\delta\xi - p(0)^T \delta\xi, \end{aligned}$$

donc au minimum

$$\bar{\xi} = P_0 p(0).$$ ■

On remarque que cette équation adjointe est cohérente avec celle utilisée en temps discret. En effet, pour pouvoir comparer temps discret et continu, il suffit de poser $A_{n+1|n} = \mathbb{1} + \Delta t A$ et de considérer la limite $\Delta t \to 0$. Dans ce cas on retrouve pour l'équation adjointe discrète

$$p_k - (1 + \Delta t A^T)p_{k+1} = p_k - p_{k+1} - \Delta t A^T p_{k+1} = H^T W_k^{-1}(Z - HX),$$

ce qui donne à la limite avec $W_k = \frac{1}{\Delta t} W$ (nous verrons plus loin la justification de ce choix),

$$-\dot{p} - A^T p = H^T W^{-1}(Z - HX).$$

□ EQUATION DE RICCATI ET CONTRÔLE EN BOUCLE FERMÉE – L'objectif est maintenant d'exprimer le contrôle optimal $\bar{\xi}$ en boucle fermée. En effet, le problème de minimisation s'exprime finalement sous la forme d'un problème aux deux bouts

$$\begin{cases} \dot{X}_{\bar{\xi}} = A(t)X_{\bar{\xi}} + R(t) \\ X(0) = X_0 + P(0)p(0) \\ \dot{p} + A(t)^T p = -H^T W^{-1}(Z - HX_{\bar{\xi}}) \\ p(T) = 0 \end{cases} \tag{1.58}$$

et nous aimerions le reformuler comme un problème de Cauchy pour un calcul séquentiel. En fait, ceci est possible au temps final d'après le théorème suivant.

THÉORÈME 1.3.5
Il existe $r(t, X_0) \in R^N$, et $P(t) \in \mathbb{M}_N(\mathbb{R})$ tel que

$$X_{\bar{\xi}}(t) = r(t) + P(t)p(t), \tag{1.59}$$

où r est indépendant de p et vérifie

$$\begin{cases} \dot{r} = A(t)r + R(t) + PH^T W^{-1}(Z - Hr), \\ r(0) = X_0 \end{cases} \tag{1.60}$$

et où P est solution sur $[0, T]$ de l'équation matricielle de Riccati

$$\begin{cases} \dot{P} - PA^T - AP + PH^T W^{-1} HP = 0 \\ P(0) = P_0 \end{cases} \tag{1.61}$$

de plus, au temps final d'estimation

$$X_{\bar{\xi}}(T) = r(T). \tag{1.62}$$

■ ***Démonstration :*** *Soit $\mathcal{R}_*$ la résolvante du système de Cauchy linéaire $(X(0), p(0)) \mapsto (X, p)$:*

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_*^{\mathrm{xx}} & \mathcal{R}_*^{\mathrm{xp}} \\ \mathcal{R}_*^{\mathrm{px}} & \mathcal{R}_*^{pp} \end{pmatrix}.$$

Par définition

$$\begin{cases} X(t) = \mathcal{R}_*^{\mathrm{xx}}(t,0)X(0) + \mathcal{R}_*^{\mathrm{xp}}(t,0)p(0) + C_{\mathrm{x}}(t) \\ p(t) = \mathcal{R}_*^{\mathrm{px}}(t,0)X(0) + \mathcal{R}_*^{pp}(t,0)p(0) + C_p(t) \end{cases}$$

où $C_{\mathrm{x}}(t)$, $C_p(t)$ correspondent à la variation de la constante sur les seconds membres. Or $p(0) = P_0^{-1}(X(0) - X_0)$, donc

$$\begin{cases} X(t) = (\mathcal{R}_*^{\mathrm{xx}}(t,0) + \mathcal{R}_*^{xp}(t,0)P_0^{-1})X(0) - \mathcal{R}_*^{xp}(t,0)P_0^{-1}X_0 + C_{\mathrm{x}}(t) \\ p(t) = (\mathcal{R}_*^{\mathrm{px}}(t,0) + \mathcal{R}_*^{pp}(t,0)P_0^{-1})X(0) - \mathcal{R}_*^{pp}(t,0)P_0^{-1}X_0 + C_p(t) \end{cases}$$

On omettra les variables temporelles de $\mathcal{R}$ tant qu'il n y a pas d'ambiguïté. Le caractère bien posé du problème LQ assure l'inversibilité de $(\mathcal{R}_*^{px} + \mathcal{R}_*^{pp}P_0^{-1})$ d'où on déduit l'existence d'une relation affine entre X et p de la forme

$$\begin{aligned} X(t) = &\Big(-\mathcal{R}_*^{\mathrm{xp}}P_0^{-1} + (\mathcal{R}_*^{\mathrm{px}} + \mathcal{R}_*^{pp}P_0^{-1})^{-1}\mathcal{R}_*^{pp}P_0^{-1}\Big)X_0 \\ &+ \Big(C_{\mathrm{x}}(t) - (\mathcal{R}_*^{\mathrm{xx}} + \mathcal{R}_*^{\mathrm{xp}}P_0^{-1})(\mathcal{R}_*^{\mathrm{px}} + \mathcal{R}_*^{pp}P_0^{-1})^{-1}C_p(t)\Big) \\ &+ \Big((\mathcal{R}_*^{\mathrm{xx}} + \mathcal{R}_*^{\mathrm{xp}}P_0^{-1})(\mathcal{R}_*^{\mathrm{px}} + \mathcal{R}_*^{pp}P_0^{-1})^{-1}\Big)p(t) \end{aligned}$$

Notons cette relation sous la forme condensée (1.59)

$$X(t) = r(t) + P(t)p(t),$$

où r est indépendant de p. On dérive cette relation par rapport au temps en utilisant les dynamiques respectives de X et p. Le couple (r, P) vérifie alors la relation différentielle

$$\dot{r} + PH^TW^{-1}Hr - Ar + (\dot{P} - PA^T - AP + PH^TW^{-1}HP)p = R + PH^TW^{-1}Z.$$

L'indépendance de r par rapport à p conduit à

$$\dot{P} - PA^T - AP + PH^TW^{-1}HP = 0,$$

et

$$\dot{r} = Ar + R + PH^TW^{-1}(Z - Hr),$$

et donc aux dynamiques (1.60) *et* (1.61) *respectivement de P et r. De plus, la relation affine* (1.59) *appliquée au temps final assure bien que*

$$X_{\bar{\xi}}(T) = r(T). \tag{1.63}$$

A noter enfin que P s'exprime donc en fonction de la résolvante du système et assure ainsi l'existence d'une solution pour l'équation de Riccati. ■

Notons finalement $\hat{X} : T \mapsto r(T)$, alors $\hat{X}$ vérifie la dynamique suivante

$$\begin{cases} \dot{\hat{X}} = A(t)\hat{X} + R(t) + K(Z - H\hat{X}) \\ \hat{X}(0) = X_0 \end{cases} \tag{1.64}$$

où

$$K = PH^TW^{-1}, \tag{1.65}$$

et P vérifie l'équation de Riccati. Cet estimateur utilise l'information disponible sous forme d'un feedback, c'est pourquoi nous parlions bien d'un contrôle en boucle fermée à partir de

la formulation variationnelle. Enfin à tout temps T, $\hat{X}(T)$ correspond au minimum de la fonctionnelle

$$J_T(\xi) = \|\xi\|^2_{P_0^{-1}} + \int_0^T \|Z - HX_\xi\|^2_{W^{-1}}\, dt$$

Le fait que c'est simplement en fin de fenêtre en temps (cf. expression (1.63)) que le filtre rejoint la trajectoire minimisante est absolument fondamental dans la compréhension de l'équivalence entre les deux stratégies et nous l'illustrons Figure 1.2. D'ailleurs en non-linéaire nous réutiliserons ce principe pour passer du contrôle optimal au filtre (voir fin de Section 1.4.1).

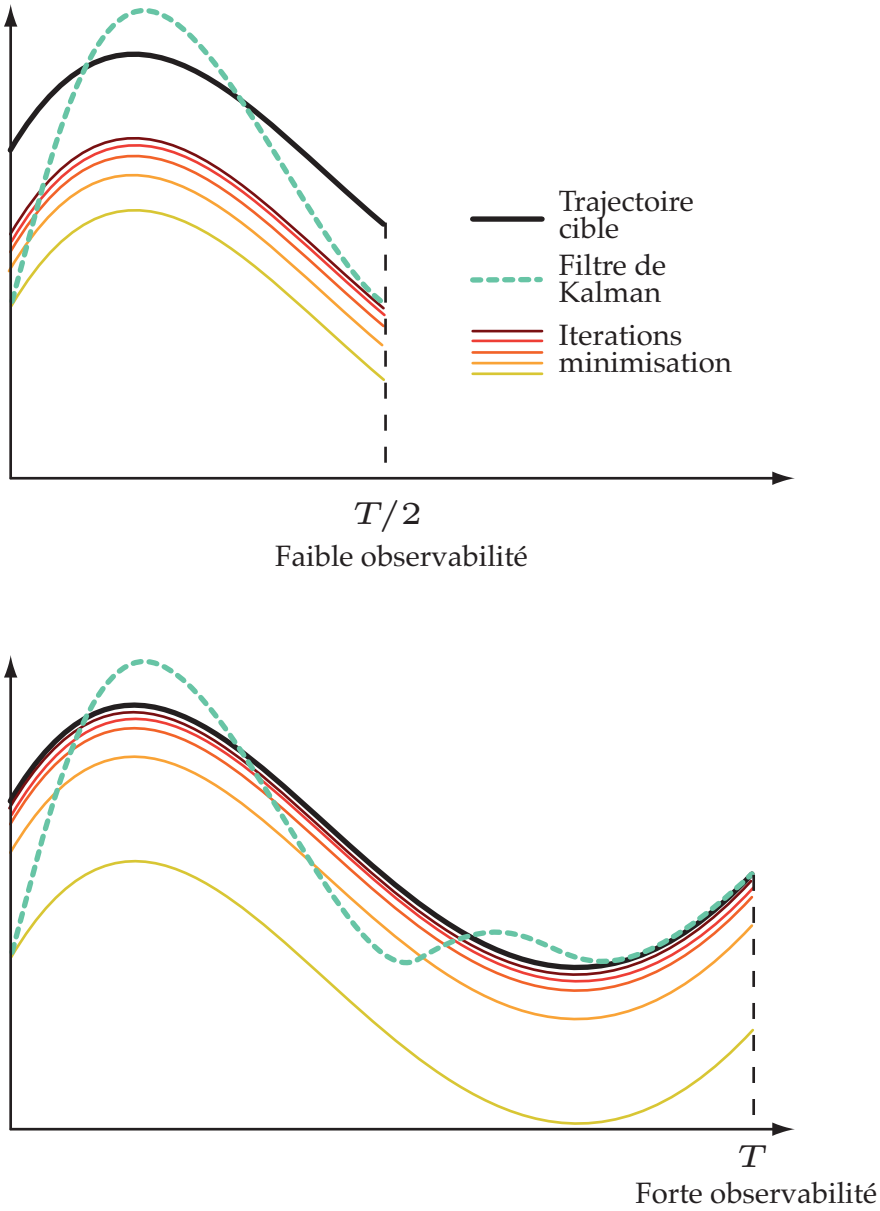


FIGURE 1.2: Exemple de trajectoire entre l'approche variationnelle et l'approche séquentielle. Avec des observations comprises entre $[0, T/2]$ le variationnel et le séquentiel n'ont pas encore convergé par manque d'observabilité (voir Section 1.3.5).

En fait, le changement de variable $(X, p) \to (r, p)$ est une façon de trigonaliser inférieure-

ment par bloc la dynamique du problème couplé état–état adjoint puisque la dynamique de l'estimateur séquentielle est indépendante de p. Matriciellement, la dynamique couplée s'écrit

$$\frac{d}{dt}\begin{pmatrix} X \\ p \end{pmatrix} = \begin{pmatrix} A & 0 \\ H^T W^{-1} H & -A^T \end{pmatrix}\begin{pmatrix} X \\ p \end{pmatrix} + \text{terme cst}, \quad \begin{pmatrix} X(0) \\ p(0) \end{pmatrix} = \begin{pmatrix} \mathbb{1} & P_0 \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} X_0 \\ p_0 \end{pmatrix}.$$

qui n'est pas trigonalisée inférieurement du fait de la condition initiale (voir aussi plus loin le cas plus flagrant avec bruit de modèle). Après changement de variable on recherche donc une dynamique de la forme

$$\frac{\mathrm{d}}{\mathrm{dt}}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} X \\ p \end{pmatrix} = \begin{pmatrix} * & 0 \\ * & * \end{pmatrix}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} X \\ p \end{pmatrix} + \text{terme cst},$$

soit

$$\begin{pmatrix} 0 & \dot{P} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}^{-1} + \begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} A & 0 \\ H^T W^{-1} H & -A^T \end{pmatrix}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}^{-1} = \begin{pmatrix} * & 0 \\ * & * \end{pmatrix}.$$

Cette relation est satisfaite sur le bloc $(1,2)$ si et seulement si

$$(2,1) : -\dot{P} + PA^T + AP - PH^T W^{-1} HP = 0.$$

Nous venons ainsi de voir en temps continu ou en discret l'équivalence entre deux approches classiques : d'un côté l'approche dite *variationnelle* qui consiste à minimiser une fonction coût pour trouver le contrôle optimal (ici la condition initiale) caractérisant la trajectoire du système cible ayant fourni les observations. De l'autre, une approche séquentielle dite de *filtrage* qui consiste à poursuivre la trajectoire cible à partir de la condition initiale moyenne en utilisant les observations disponibles pour corriger la dynamique. Ces deux approches sont radicalement différentes dans leur esprit de mise en oeuvre mais pourtant elle correspondent au même point en fin (relativement au critère) de trajectoire T. Nous avons vu que, du point de vue probabiliste, cette équivalence est analogue à l'équivalence entre le meilleur estimateur linéaire et l'estimateur moindres carrés ce qui d'ailleurs permet aussi de procéder aux mêmes démonstrations que ce que nous venons de voir dans le cadre des équation différentielles stochastiques. Enfin, malgré la présentation suivie dans cette section que nous avons souhaitée crescendo dans les aspects théoriques, la démonstration la plus naturelle du point de vue formel est certainement celle en temps continu. Dans ce cas on peut alors légitimement se poser la question de sa mise en oeuvre informatique qui nécessite forcément la définition d'un schéma en temps. Or, la version discrète en temps nous fournit directement ce schéma (notamment pour l'équation de Riccati) car elle minimise aussi un critère variationnel qui serait la version discrétisée par une simple méthode des trapèzes du critère continu. En effet, supposons que X_n approche $X(n\Delta t)$ alors

$$J_T(\xi) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\int_0^T \|Z - HX_\xi\|^2_{W^{-1}}\, ds \simeq \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{k=0}^{n}\|Z_k - H_k X_k(\xi)]\|^2_{W_k^{-1}},$$

si on choisit une norme sur le bruit de mesure discret de la forme

$$W_k = \frac{1}{\Delta t} W \tag{1.66}$$

□ CADRE PROBABILISTE – Pour comprendre la singularité qui apparaît dans (1.66) vis-à-vis du pas de temps, et les liens entre problèmes continu et discret, il est utile d'introduire quelques concepts de calcul stochastique. On définit tout d'abord un processus de Wiener comme une variable aléatoire fonction du temps $w(t)$ gaussienne et centrée pour tout $t \geq 0$ et telle que

$$\mathrm{Cov}(w(t), w(t')) = \min(t, t')W, \quad \forall t \geq 0, t' \geq 0,$$

où W est une matrice symétrique positive. On introduit ensuite le bruit blanc $\chi(t)$ associé au Wiener $w(t)$ par dérivation en temps :

$$dw = \chi dt,$$

et on montre que

$$\mathrm{Cov}(\chi(t), \chi(t')) = \delta(t - t')W. \tag{1.67}$$

Pour une observation continue en temps, il est alors classique de modéliser l'incertitude à partir d'un tel bruit blanc, ce qui donne un sens probabiliste à l'identité

$$Z(t) = HX(t) + \chi(t).$$

On retrouve le lien avec le bruit blanc discret en notant qu'une mesure n'est en réalité jamais instantanée, et donc on aurait plutôt, à intervalles réguliers,

$$Z_k = Z(t_k) = \frac{1}{\Delta t}\int_{t_k-\Delta t}^{t_k} \big(HX(s) + \chi(s)\big)\, ds = \frac{1}{\Delta t}\int_{t_k-\Delta t}^{t_k} HX(s)\, ds + \chi_k,$$

auquel cas on retrouve bien (1.66) car

$$\mathrm{Cov}(\chi_k) = \frac{1}{\Delta t^2}\int_{t_k-\Delta t}^{t_k}\int_{t_k-\Delta t}^{t_k} \mathbb{E}(\chi(t)\chi(t'))\, dtdt' = \frac{1}{\Delta t^2}\int_{t_k-\Delta t}^{t_k} W\, dt = \frac{1}{\Delta t}W, \tag{1.68}$$

et évidemment $\mathrm{Cov}(\chi_k, \chi_{k'}) = 0$ si $k \neq k'$.

1.3.4 Bruits de modèle

Même si en pratique nous l'utiliserons peu dans les chapitres suivants, nous rappelons succinctement comment les théorèmes présentés ci-dessus s'étendent au cas d'une erreur de modèle si souvent utilisée en pratique en assimilation de données. Comme précisé dans (1.36) l'erreur de modèle, c'est-à-dire l'indétermination sur la dynamique, peut être comprise de manière duale comme la présence d'un contrôle sur l'ensemble de la trajectoire, et pas seulement sur la condition initiale. Dans le cas linéaire étudié ici, le problème (1.36) s'écrit

$$\begin{cases} \dot{X} = A(t)X + B\omega(t) + R(t) \\ X(0) = X_0 + \zeta^X \end{cases} \tag{1.69}$$

et le critère minimisé devient

$$J_T(\xi, \omega) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\int_0^T (\|Z - HX_\xi\|^2_{W^{-1}} + \|\omega\|^2_{S^{-1}})\, dt.$$

Toute la théorie que nous venons de présenter reste valide pour les questions d'existence de trajectoire minimisante, de convexité de la fonctionnelle, de formalisme adjoint et d'expression d'un estimateur en boucle fermé à partir des observations. Elle est cependant plus technique :

□ EXISTENCE – La preuve nécessite plus d'attention car $\omega \in L^2([0,T],R^N)$. On considère toujours des suites minimisantes (ξ_n,ω_n). Or pour ω_n, on généralise en fonction de l'espace fonctionnel ce que nous avions dit pour ξ_n. En effet,

$$J_T(\xi,\omega) \geq \alpha\|\omega\|_{L^2}^2$$

permet de déduire que (ω_n) est bornée dans $L^2([0,T],R^N)$ donc converge faiblement vers $\bar{\omega}$. Puis, la variation de la constante s'écrit désormais

$$X_n(t) = \mathcal{R}(t,0)(X_0+\xi_n) + \int_0^t \mathcal{R}(t,s)(B(s)u_n(s)+R(s))\,ds.$$

si bien que, à une sous suite près, on a toujours la convergence simple de la suite $(X_n)_n$. La fin de la preuve est alors identique :

$$J_T(\bar{\xi},\bar{\omega}) \leq \liminf_n J_T(\xi_n,\omega_n)$$

car

$$\int_0^T \|\bar{\omega}\|_{S^{-1}}^2\,dt \leq \liminf_n \int_0^T \|\omega_n\|_{S^{-1}}^2\,dt.$$

Donc J_T est minimum pour $(\bar{\xi},\bar{\omega})$. La stricte convexité étant toujours assuré, ce minimum est aussi unique.

□ ADJOINT – L'équation adjointe n'est pas modifiée mais le critère ayant deux arguments, il doit donc être dérivé par rapport à ses deux variables. On a toujours

$$\mathrm{d}_\xi J_T = 0 \Leftrightarrow \bar{\xi} = P(0)p(0),$$

mais aussi

$$\mathrm{d}_\omega J_T = \int_0^T (p^TB + \omega^TS^{-1})\,dt = 0 \Leftrightarrow \bar{\omega} = SB^Tp.$$

Donc le système optimal couplé l'est désormais tout au long de la trajectoire, et pas simplement aux deux bouts

$$\begin{cases} \dot{X}_{\bar{\xi}} = A(t)X_{\bar{\xi}} + BSB^Tp + R(t) \\ X(0) = X_0 + P(0)p(0) \\ \dot{p} + A(t)^Tp = -H^TW^{-1}(Z - HX_{\bar{\xi}}) \\ p(T) = 0. \end{cases} \tag{1.70}$$

□ RICCATI – Réécrire ce problème sous forme de Cauchy est toujours possible, mais l'équation de Riccati est modifiée et, afin d'éviter des développements techniques redondants par rapport au cas « modèle parfait », nous utilisons le principe de diagonalisation par blocs pour obtenir directement les nouvelles relations. En effet, trigonaliser sous la forme

$$\begin{pmatrix} 0 & \dot{P} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}^{-1} + \begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} A & BSB^T \\ H^TW^{-1}H & -A^T \end{pmatrix}\begin{pmatrix} \mathbb{1} & P \\ 0 & \mathbb{1} \end{pmatrix}^{-1} = \begin{pmatrix} * & 0 \\ * & * \end{pmatrix}$$

n'est possible que si cette fois

$$-\dot{P} + PA^T + AP - PH^TW^{-1}HP - BSB^T = 0,$$

et le feedback est toujours

$$K = PH^TW^{-1}.$$

□ CADRE DISCRET EN TEMPS – Du point de vue probabiliste les démonstrations du filtrage de Kalman sont exactement les mêmes pour peu qu'on suppose que ω est lui aussi un bruit blanc. Ainsi, pour le cas discret en temps, il n'y a pas de difficulté et l'algorithme prédiction-correction est simplement modifié **au niveau de la covariance de prédiction** avec

$$P_n^- = A_{n+1|n}P_n^+A_{n+1|n}^T + B_nS_nB_n^T. \tag{1.71}$$

REMARQUE 1.3.3 (LAGRANGIEN)
Un formalisme classique en estimation est de considérer le Lagrangien afin d'obtenir directement l'équation adjointe. On pourra se reporter à Chavent (2008) pour une présentation détaillée sous cet angle. En effet, notre problème de minimisation initiale est en fait sous contrainte de l'équation d'état, soit

$$\begin{cases} \min\limits_{X,\omega} J_T(X,\omega) \\ \text{avec } \dot{X} = A(X) + B\omega, \quad X(0) = X_0 + \xi. \end{cases}$$

La minimisation opère sur l'ensemble de l'état (puisqu'il est contraint), donc en particulier sur ξ, ce qui lui donne un statut particulier entre variable de contrôle et condition initiale de l'état. Ce problème admettant une solution peut être remplacé par la formulation min-max

$$\min_{X,\omega}\max_{p}\left\{L(X,\omega,p) = J_T(X,\omega) + \int_0^T p^T(\dot{X} - A(X) - B\omega)\,dt\right\},$$

sous des hypothèses de régularité suffisantes sur X (par exemple « C^1 en temps »). Dans ce cas, une condition nécessaire est la stationnarité du Lagrangien. Or

$$\mathcal{L}(X,\omega,p) = J_T(X,\omega) + [p^TX]_0^T + \int_0^T -\dot{p}^TX - p^T(A(X) + B\omega)\,dt$$

donc

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial X} = 0 \quad &\Rightarrow \quad \dot{p} + \frac{\partial A}{\partial X}^T p + \frac{\partial H}{\partial X}^T W^{-1}(Z - H(X)) = 0, \\
\frac{\partial\mathcal{L}}{\partial X(0)} = 0 \quad &\Rightarrow \quad P_0^{-1}(X(0) - X_0) - p(0) = 0, \\
\frac{\partial\mathcal{L}}{\partial X(T)} = 0 \quad &\Rightarrow \quad p(T) = 0, \\
\frac{\partial\mathcal{L}}{\partial p} = 0 \quad &\Rightarrow \quad \dot{X} - A(X) - B\omega = 0.
\end{aligned}$$

Les dérivations par rapport à $X(0)$ et $X(T)$ sont légitimes. Elles sont en fait partie intégrante de la dérivation variationnelle par rapport à X pour des fonctions test en temps « concentrées » en 0 ou T.

A noter que nous avons anticipé sur la Section 1.4 en traitant le cas général d'opérateurs non-linéaires.

1.3.5 Notions d'observabilité, de contrôlabilité et d'identifiabilité

Nous avons pour l'instant passé sous silence la question de l'efficacité de ces estimateurs dynamiques. Or, ce n'est pas parce qu'il sont définis à partir d'un critère moindres carrés, et qu'ils sont donc optimaux relativement à ce critère, qu'ils permettent d'obtenir une solution pertinente pour le problème d'estimation. Autrement dit, à un instant T donné, le problème admet bien une solution mais cette solution est peut-être loin de la cible, ce que nous avions illustré Figure 1.2. La question de l'efficacité se traduit pour les problèmes d'évolution via les conditions d'observabilité que nous avons déjà mentionnées. Elles se définissent rigoureusement comme suit.

DÉFINITION 1.3.6
Un système continu en temps est dit observable si pour toute condition initiale et pour tout temps T, la condition initiale peut être déterminée de manière unique à partir des seules observations disponibles $Z(t)$, $t \in [0, T]$.

Cette définition admet un pendant dans le cadre discret en temps (d'ailleurs moins restrictif).

DÉFINITION 1.3.7
Un système discret en temps est dit observable pour toute condition initiale il existe un entier n tel que la condition initiale peut être déterminée de manière unique à partir des seules observations disponibles $Z_n, 0 \leq k \leq n$.

Il existe un certain nombre de tests d'observabilité permettant de déterminer a priori si le problème est finalement bien posé ou non. Le plus classique dans le cas où la dynamique est linéaire autonome est le critère de Kalman :

THÉORÈME 1.3.8 (CRITÈRE DE KALMAN)
Soit Q la matrice d'observabilité définie par

$$Q = \begin{pmatrix} H \\ HA \\ \vdots \\ HA^{N-1} \end{pmatrix},$$

alors le système est observable si et seulement si le rang de Q est N.

Ce théorème est valide en temps discret et continu. Pour le cas discret, la formalisation du filtre à partir de Hassibi et al (1999) que nous avons vue en remarque 1.3.2 fournit directement le critère, puisque finalement elle permet de lier la dynamique en temps à l'inversion statique d'un opérateur de dimension nN où on a concaténé tous les pas de temps. En temps continu, en fait le résultat est valable même si nous ne pouvons pas a priori « augmenter l'état continûment » pour nous ramener au cas statique. Cependant, le théorème de Cayley-Hamilton permettant d'exprimer toute puissance de A fournit à la fois le résultat en discret et en continu.

Cependant, le calcul du rang de cette matrice d'observabilité Q est bien trop coûteux en pratique pour les dimensions de système que nous allons considérer et on lui préfère un critère équivalent qui peut de plus se généraliser pour les systèmes non autonomes.

DÉFINITION 1.3.9 (GRAMMIENNE D'OBSERVABILITÉ)
On définit la grammienne d'observabilité du système observé

$$G(t_0, t_1) = \int_{t_0}^{t_1} \mathcal{R}(t_0, s)^T H(s)^T H(s) \mathcal{R}(t_0, s)\, ds$$

où $\mathcal{R}$ est la résolvante du système linéaire.

Dans le cas autonome, $\mathcal{R}(t_2, t_1) = e^{A(t_2 - t_1)}$ donc la grammienne est donnée par la formule classique

$$G(t_0, t_1) = \int_{t_0}^{t^1} e^{A^T(s-t_0)} H(s)^T H(s) e^{A(s-t_0)}\, ds.$$

THÉORÈME 1.3.10 (CONDITION D'OBSERVABILITÉ)
Le système est observable sur $[0, T]$ si et seulement si il existe $t \in [0, T]$ tel que la grammienne d'observabilité $G(0, t)$ est définie positive.

Cette propriété se définit aussi en temps discret. Soit Q_n définie par

$$Q_n = \begin{pmatrix} H \\ HA_{1|0} \\ \vdots \\ HA_{n|n-1} \dots A_{1|0} \end{pmatrix},$$

alors la grammienne d'observabilité est donnée par

$$G_n = Q_n^T Q_n.$$

Dans le cas où $\forall k, \quad A_{k+1|k} = A$ alors Q_N est la matrice d'observabilité et la grammienne devient simplement pour tout n

$$G_n = \sum_{k=0}^{n} (A^k)^T H^T H A^k.$$

A noter que, classiquement, on définit les grammiennes sans préciser la norme d'observation W. Cependant c'est bien par rapport à cette norme qu'il faut déterminer comment évolue l'observabilité, car plus les valeurs propres de G augmentent plus on peut espérer converger rapidement sur la trajectoire cible. Donc, soit on introduit W dans la matrice d'observabilité, soit on calcule ces valeurs propres par rapport à W.

Enfin cette condition d'observabilité est liée aux matrices « de covariance » par l'équation de Riccati sous la forme

THÉORÈME 1.3.11
Le système est observable si et seulement si l'équation de Riccati

$$\begin{cases} \dot{P} - PA^T - AP + H^T W^{-1} H = 0 \\ P(0) = 0 \end{cases}$$

admet une solution définie positive pour $t \in [0, T]$.

On comprend mieux la différence entre les exigences de minimisation du critère variationnel où on voyait la covariance P_0 sur la condition initiale qui permettait de minimiser le critère en toutes circonstances et qui apparaissait donc comme condition initiale de l'équation de Riccati et le théorème ci-dessus qui ne juge l'estimateur que par rapport aux observations. C'est l'équivalent dynamique de ce que nous avions vu en statique entre inverse généralisé classique et inverse généralisé avec terme d'ébauche. A noter encore une fois que ce théorème a lui aussi un pendant discret en temps.

Ces notions d'observabilité sont duales des notions de commandabilité pour un système du type $\dot{X} = AX + BU$ avec U un contrôle à déterminer pour transférer toute condition initiale fixée $X(0)$ vers une valeur quelconque $X(t)$. On parle alors de contrôlabilité exacte du système et la matrice de commandabilité est alors

$$P = \begin{pmatrix} B \\ AB \\ \vdots \\ A^{n-1}B \end{pmatrix}.$$

et la grammienne

$$G_B = \int_{t_0}^{t_1} \mathcal{R}(t_1, s) B B^T \mathcal{R}(t_1, s)^T \, ds.$$

Ainsi, commandabilité de système et observabilité sont deux notions qui formellement sont encore une fois très étroitement liées et font appel au mêmes outils d'analyse. Les liens entres ces grammiennes de contrôlabilité et d'observabilité ainsi que l'existence d'un contrôle sur le système apparaîtront plus clairement 3.3.3. Du point de vue strict du vocabulaire, elle explique déjà pourquoi nous parlons généralement de contrôle optimal ou de commande optimale pour ξ et ω y compris en observation.

Reste la notion d'identifiabilité dont nous avons parlé en préambule et qui est réservée à l'estimation paramétrique. Pour l'instant, nous avons simplement mentionné le problème de l'estimation de condition initiale à modèle exact ou incertain. Cependant, cette incertitude est souvent la conséquence d'une connaissance imparfaite de certains paramètres physiques du modèle. Dans ce cas, mais nous le verrons plus précisément au chapitre 4, on peut formaliser l'incertitude via la dépendance paramétrique en considérant des systèmes de la forme

$$\dot{X} = A(X, \theta),$$

où $\theta = \theta_0 + \zeta^\theta$ rassemble l'ensemble des paramètres incertains. L'identifiabilité correspond alors à la condition d'observabilité pour les paramètres. D'ailleurs, une technique classique consiste à augmenter l'état X avec les paramètres sous la forme

$$\frac{\mathrm{d}}{\mathrm{dt}} \begin{pmatrix} X \\ \theta \end{pmatrix} = \begin{pmatrix} A(X, \theta) \\ 0 \end{pmatrix}.$$

L'identifiabilité paramétrique est ainsi une forme d'observabilité et suit donc les mêmes critères.

1.4 Extensions autour de la théorie LQ

Ce que nous venons de démontrer jusqu'à présent pour les systèmes dynamiques continus en temps s'intègre dans le cadre plus général du contrôle optimal (ou commande optimale), comme déjà introduit au paragraphe précédent sur l'observabilité. Cette théorie ne s'applique pas exclusivement au cas dynamique linéaire, critère quadratique (LQ). Cependant, pour une dynamique bruitée non-linéaire

$$\begin{cases} \dot{X} = A(X,t) + B\omega \\ X(0) = X_0 + \zeta^X \end{cases} \tag{1.72}$$

et un observateur non-linéaire sous forme générale

$$D(Z,X) = \chi \in \mathbb{R}^m, \tag{1.73}$$

les preuves sont bien plus difficiles que pour le cas LQ et sont fondées sur le principe du minimum de Pontryagin[8]. Nous proposons ici une présentation assez générale à partir de Mortensen (1968) ; Fleming (1997) qui permet de trouver l'existence de minimisant pour le coût

$$J_T(\xi,\omega) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\int_0^T \|D(X,Z)\|^2_{W^{-1}} + \|\omega\|^2_{S^{-1}}\,dt, \tag{1.74}$$

mais surtout de reformuler cette minimisation sous forme de boucle fermée via la programmation dynamique et donc de trouver l'extension non-linéaire du filtre optimal. En pratique, ce filtre est, nous le verrons, « impossible » à calculer explicitement et c'est pourquoi la plupart des filtres en non-linéaires divergent de la notion de contrôle optimal pour être simplement « des filtres approchés ». Nous pourrons cependant mieux motiver les deux derniers filtres que nous examinerons aux paragraphes suivants : le *filtre de Kalman étendu* EKF et le *filtre de Kalman dit unscented* UKF.

1.4.1 Filtre non-linéaire déterministe

□ PRINCIPE DE PONTRYAGIN – En toute généralité, la minimisation du critère (1.74) est donnée par le principe du Mimimum de Pontryagin (voir Th. 5.1 de Fleming et Rischel (1975)). En fait dans Fleming (1997), le théorème de Pontryagin est défini et démontré dans le cadre de ce qu'on appelle un problème de Mayer. c'est-à-dire un problème où la fonction coût est de la forme

$$g = (g_1 \ldots g_d). \quad J_T = g_1(\xi,T), \quad \forall i > 1,\ g_i(\xi) = 0,$$

sans intégrale en temps mais où $g : \mathbb{R}^N \to \mathbb{R}^d$. *A contrario*, un problème de Lagrange ne comporte que l'intégrale en temps

$$J_T = \int_0^T L(X,\omega,t)\,dt.$$

Notre cas est hybride, et appelé Mayer-Lagrange

$$J_T = g_1(\xi,T) + \int_0^T L(X,\omega,t)\,dt, \quad \forall i > 1,\ g_i(\xi) = 0,$$

[8]Principe du maximum suivant le sens où on le regarde

avec $g = g_1 = \frac{1}{2}\|\xi\|^2_{P_0^{-1}}$ (on pourrait d'ailleurs faire toute la présentation en gardant g) et

$$L(X,\omega) = \frac{1}{2}(\|D(X,Z)\|^2_{W^{-1}} + \|\omega\|^2_{S^{-1}}). \tag{1.75}$$

En fait ces trois problèmes sont équivalents. Pour passer de Lagrange à Mayer, on considère un problème étendu $X^e = \left(\begin{smallmatrix} X \\ X^d \end{smallmatrix}\right)$ où $\dot{X}^d = L(X,\omega,s)$. Réciproquement, un problème de Mayer devient un problème de Lagrange en considérant toujours un état étendu tel que $\dot{X}^d = 0$ sur la nouvelle composante, puis $g_{d+1}(\xi,T) = X^d(T) - \frac{1}{T}g_1(\xi)$. Dans ce cas la fonction $L = X^d$.

Toutes ces notions justifient la présentation de Fleming et Rischel (1975) du principe de Pontryagin pour les problèmes de Mayer que, pour rester consistant, nous réadaptons au problème de Mayer-Lagrange de type (1.74) (voir aussi Faurre et Robin (1984)).

THÉORÈME 1.4.1 (PRINCIPE DE PONTRYAGIN)
Pour X vérifiant la dynamique de (1.72)*, on définit le Hamiltonien $\mathcal{H}$ associé au critère* (1.74) *(avec L défini en* (1.75)*)*

$$\forall t \in [0,T], \quad \mathcal{H}(X,\omega,p,t) = L(X,\omega) - p^T(A(X,t) + B\omega). \tag{1.76}$$

Le critère J_T est minimum si et seulement si

$$\forall t, \quad \frac{\partial \mathcal{H}}{\partial \omega} = 0,$$

avec p vérifiant l'équation adjointe

$$\dot{p} + \frac{\partial A}{\partial X}^T p = \frac{\partial D}{\partial X}^T W^{-1} D(Z,X), \tag{1.77}$$

de condition finale $p(T) = 0$ et initiale $p(0) = P_0^{-1}\xi$.

Nous ne démontrons pas ce résultat largement hors de portée de cette introduction succincte, mais analysons simplement ses implications. Pour la dynamique et le critère considérés, on peut préciser le Hamitlonien

$$\forall t, \ \frac{\partial \mathcal{H}}{\partial \omega} = 0 \quad \Rightarrow \quad \forall t, \ \bar{\omega} = SB^T p(t)$$

On en déduit le Hamiltonien optimal $\bar{\mathcal{H}}$ par substitution dans $\mathcal{H}$. En effet

$$\mathcal{H}(X,p,\bar{\omega},t) = \frac{1}{2}p^T BSB^T p + \frac{1}{2}D(Z,X)^T W^{-1} D(Z,X) - p^T(A(X,t) + BSB^T p)$$

donne

$$\bar{\mathcal{H}}(X,p,t) = -\frac{1}{2}p^T BSB^T p + \frac{1}{2}D(Z,X)W^{-1}D(Z,X) - p^T A(X,t),$$

On retrouve alors les équations de la dynamique et de l'adjoint à partir du Hamiltonien optimale sous la forme d'*équation de de Hamilton*

$$\left\{ \begin{array}{rcccl} \dot{X} & = & -\dfrac{\partial \bar{\mathcal{H}}}{\partial p}^T & = & A(X,t) + BSB^T p \\ \dot{p} & = & \dfrac{\partial \bar{\mathcal{H}}}{\partial X}^T & = & -\dfrac{\partial A}{\partial X}^T p + \dfrac{\partial D}{\partial X}^T W^{-1} D(Z,X) \end{array} \right. \tag{1.78}$$

avec les conditions aux deux bouts

$$X(0) = X_0 + \xi, \quad p(0) = P_0^{-1}\xi.$$

□ PROGRAMMATION DYNAMIQUE – En fait ce principe de Pontryagin est directement lié à la programmation dynamique ou à la théorie de Hamilton-Jacobi. Elles consistent à analyser l'évolution du critère en temps final T pour une trajectoire dont on contraint uniquement le point d'arrivée $X(T) = X$. On pose alors

$$V(X,T) = \min_{\omega | X(T)=X} J_T(\xi, \omega),$$

souvent appelé « cost-to-come »[9]. On note que dans cette minimisation la condition initiale $X(0) = X_0 + \xi$ est obtenue par intégration rétrograde en temps à partir du point d'arrivée une fois ω choisi. On ne considère donc plus ξ comme un contrôle. Le principe de la programmation dynamique est alors de démontrer que $V : \mathbb{R}^N \times [0,T] \to \mathbb{R}$ satisfait une équation de Hamilton-Jacobi-Bellman (HJB)

$$\forall t, \quad \frac{\partial V}{\partial t}(X,t) - \bar{\mathcal{H}}(X, \nabla_X V(X,t), t) = 0. \tag{1.79}$$

On utilise la notation ∇_X pour exprimer ici un gradient,[10] c'est-à-dire le dual de la différentielle de V qui est bien de même dimension que X. Nous ne démontrerons pas cette relation dans le cas général mais nous nous contenterons d'un aperçu dans le cas discret (voir Remarque 1.4.1 ci-dessous). Cependant, une fois cette équation admise (du moins son origine), le théorème suivant lie la solution de cette équation à la commande optimale issue du principe du maximum.

THÉORÈME 1.4.2
Si l'équation de Hamilton-Jacobi-Bellman

$$\forall t, \quad \frac{\partial V}{\partial t}(X,t) - \bar{\mathcal{H}}(X, \nabla_X V(X,t), t) = 0$$

admet une solution C^1 avec pour condition initiale

$$V(X,0) = \frac{1}{2}\|X - X_0\|^2_{P_0^{-1}},$$

alors la commande $\check{\omega}(X,t) = \operatorname{argmin}(\mathcal{H}(X, \nabla_X V(X,t), \omega, t))$ est optimale pour le critère J_T.

■ ***Démonstration :*** *On commence par démontrer que*

$$\forall t, \quad \nabla_X V(\check{X}(t), t) = p(t),$$

pour $\check{X}$ l'état associé à la commande optimale $\check{\omega}$ dont on ne sait pas encore que c'est $\bar{\omega}$ donné par le principe de Pontryagin. Afin de prouver cette relation, démontrons que $\nabla_X V(X,T)$ vérifie la même dynamique que p. En dérivant par rapport à X la condition initiale sur V, on obtient déjà

$$\nabla_X V(X,0) = p(0).$$

[9] Coût pour atteindre à X alors qu'en contrôle, on introduit plutôt un coût « cost-to-go » pour partir d'un état X

[10] A ne pas confondre avec la notation $\underline{\underline{\nabla}}$ des mécaniciens utilisée au chapitre 2 simplement pour dire qu'on dérive non pas par rapport à un champ mais par rapport à un système de coordonnées.

Puis en dérivant l'équation HJB (sous condition de différentiabilité donc)

$$\nabla_X \frac{\partial V}{\partial t} - \frac{\partial \bar{\mathcal{H}}}{\partial X}^T (\check{X}, \nabla_X V, t) - \nabla_X^2 V \cdot \frac{\partial \bar{\mathcal{H}}}{\partial p}^T (\check{X}, \nabla_X V, t) = 0, \tag{1.80}$$

$$\nabla_X \frac{\partial V}{\partial t} + (\frac{\partial A}{\partial X}^T \nabla_X V - \frac{\partial D}{\partial X}^T W^{-1} D(Z, \check{X})) + \nabla_X^2 V \cdot (A(\check{X}, t) + BSB^T \nabla_X V) = 0. \tag{1.81}$$

Or

$$\frac{d}{dt} \nabla_X V = \nabla_X \frac{\partial V}{\partial t} + \nabla_X^2 V \cdot \dot{\check{X}},$$

où $\dot{\check{X}}$ correspond bien à la dynamique de $\check{X}$ pour le contrôle $\check{\omega}$. Or, comme au paragraphe précédent, on détermine dès le départ $\check{\omega} = SB^T \nabla_X V$ d'après l'expression de $\mathcal{H}$. Ainsi, $\nabla_X V$ vérifie la même dynamique que p, c'est-à-dire

$$\frac{d}{dt}(\nabla_X V) + \frac{\partial A}{\partial X}^T \nabla_X V = \frac{\partial D}{\partial X}^T W^{-1} D(Z, \bar{X}).$$

Ayant, de plus, même condition initiale, ils sont donc égaux. En particulier on obtient donc que $\check{\omega} = \bar{\omega}$ défini dans le principe de Pontryagin. ■

On peut désormais utiliser la condition au temps final sur p pour obtenir une relation intéressante sur V (alors que pour Pontryagin, c'était plutôt la condition finale sur p qui était naturelle et la condition initiale qui fournissait une information nouvelle...)

$$\forall T, \quad \nabla_X V(\bar{X}(T), T) = 0. \tag{1.82}$$

Cette relation nous permet de déduire **l'équation de l'estimateur séquentiel $\hat{X}$ comme la solution de cette équation pour tout temps**. En effet, considérant maintenant la définition implicite de $\hat{X}$ pour tout t comme solution de $\nabla_X V(\hat{X}, t) = 0$, on a alors

$$\frac{\mathrm{d}}{\mathrm{dt}} \nabla_X V(\hat{X}, t) = 0 \tag{1.83}$$

Or, les mêmes règles de composition des dérivées qu'en (1.80),(1.81) ajouté à la relation (1.83) fournissent

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{dt}} \left(\nabla_X V(\hat{X}(t), t) \right) &= \left(\frac{\partial}{\partial t} \nabla_X V \right) (\hat{X}(t), t) + \nabla_X^2 V \cdot \dot{\hat{X}} \\ &= \nabla_X \left(\bar{\mathcal{H}}(., \nabla_X V(., t), t) \right) (\hat{X}(t), t) + \nabla_X^2 V \cdot \dot{\hat{X}} \\ &= \frac{\partial \bar{\mathcal{H}}}{\partial X}^T (\hat{X}, \nabla_X V, t) + \nabla_X^2 V \cdot \frac{\partial \bar{\mathcal{H}}}{\partial p}^T (\hat{X}, \nabla_X V, t) + \nabla_X^2 V \cdot \dot{\hat{X}}, \end{aligned}$$

d'où, de nouveau en remplaçant les composantes de la différentielle du Hamiltonien, on obtient l'équation de l'estimateur

$$\begin{cases} \dot{\hat{X}} = A(\hat{X}, t) - (\nabla_X^2 V)^{-1} \frac{\partial D}{\partial X}^T W^{-1} D(Z, \hat{X}) \\ \hat{X}(0) = X_0 \end{cases} \tag{1.84}$$

Mortensen (1968) utilise classiquement une observation de type $Z = H(X) + \chi$ (H opérateur d'observation et pas Hamiltonien $\mathcal{H}$!) d'où pour lui l'expression du filtre

$$\dot{\hat{X}} = A(\hat{X}, t) + (\nabla_X^2 V)^{-1} \frac{\partial H}{\partial X}^T W^{-1} (Z - H(\hat{X})).$$

Ici, comme en théorie LQ, on a d'abord obtenu un minimiseur sur une fenêtre en temps $[0, T]$, avant de faire varier cette fenêtre afin d'obtenir le feedback en boucle fermée. On est donc dans le même principe d'équivalence que Figure 1.2. La seule difficulté de cette expression est qu'elle nécessite le calcul de la Hessienne de V qui elle-même suit une équation de Hamilton-Jacobi-Bellman en fonction de la variable X. La complexité algorithmique d'une éventuelle mise en oeuvre numérique rend inutilisable cette expression pour des systèmes de grande taille comme le nôtre. La théorie LQ est finalement un cas particulier de cette situation où justement on peut résoudre explicitement l'équation de Hamilton-Jacobi

$$\frac{\partial V}{\partial t}(X,t) + \nabla_X V^T AX + \frac{1}{2}\nabla_X V^T B^T SB\nabla_X V^T - \frac{1}{2}(Z - HX)^T H^T W^{-1} H(Z - HX) = 0.$$

de condition initiale

$$V(X,0) = \frac{1}{2}(X - X_0)^T P_0^{-1}(X - X_0).$$

Dans ce cas (seulement) l'équation se résout par séparation des variables sous la forme

$$V(X,t) = \frac{1}{2}X^T P(t)^{-1} X \tag{1.85}$$

avec P solution de l'équation de Riccati déjà présentée Section 1.3.4. On retrouve alors exactement les équation du filtre de Kalman puisque la différentielle d'ordre 2 est indépendante de X :

$$(\nabla_X^2 V)^{-1} = P.$$

REMARQUE 1.4.1 (PROGRAMMATION DYNAMIQUE EN TEMPS DISCRET)
Nous avons admis le principe de programmation dynamique entre J et V justifiant l'introduction de l'équation de Hamilton Jacobi. Celle-ci peut donc sembler tomber du ciel. Le cas continu est en fait assez complexe du point de vue théorique pour manipuler les expressions sans connaissance a priori de leur régularité (introduction des solutions de viscosité etc...). Nous proposons cependant une justification dans le cadre discret en temps tirée de Faurre et Robin (1984) pour fixer les idées.

Soit le système récursif

$$\forall n, X_{n+1} = A_{n+1|n}(X_n, \omega_n),$$

et le critère

$$J_n(\xi, (\omega_k)_{1\le k\le n}) = g(\xi) + \sum_{k=0}^{n} L_k(X_k, \omega_k).$$

On définit à n fixé

$$V_n(X) = \min_{(\omega_k)_{0\le k\le n}} J_n,$$

où X est la valeur finale $X_{n+1} = X$ de la suite récursive $(X_k)_{0\le k\le n}$ partant de $X_0 + \xi(X,\omega)$ déterminé de façon rétrograde. Afin de se ramener à un problème classique de contrôle on définit donc l'opérateur rétrograde

$$X_n = A_{n+1|n}^{-1}(X_{n+1}, w_n), \quad X_{n+1} = X.$$

Ainsi on a

$$\begin{aligned}
V_n(X) &= \min_{(\omega_k)_{0\le k\le n}} \{J_n(\zeta^X)\} = 0 \\
&= \min_{(\omega_k)_{0\le k\le n}} \left\{ g(\xi) + \sum_{k=0}^{n-1} L_k(X_k, \omega_k, k) + L_n(X_n, w_n) \right\} = 0 \\
&= \min_{\omega_n} \left\{ \min_{(\omega_k)_{0\le k\le n-1}} \left\{ g(\xi) + \sum_{k=0}^{n-1} L_k(X_k, \omega_k, k) \right\} + L_n(A_{n+1|n}^{-1}(X, \omega_n), w_n) \right\} = 0
\end{aligned}$$

soit finalement l'équation de Hamilton-Jacobi-Bellman discrète

$$V_n(X) = \min_{\omega} \left\{ V_{n-1}(A^{-1}_{n+1|n}(X,\omega)) + L_n(A^{-1}_{n+1|n}(X,\omega_n),\omega) \right\}. \tag{1.86}$$

Pour retrouver l'équation continue on peut formellement travailler par différence finie. On a vu que pour faire une comparaison temps discret - temps continu il fallait prendre

$$L_k(\cdot,\cdot) = \Delta t L(\cdot,\cdot,t_k),$$

et par exemple pour les opérateurs de transition

$$A^{-1}_{k+1|k}(\cdot,\cdot) = \mathbb{1} - \Delta t A(\cdot,\cdot,t_k).$$

Dans ce cas on obtient par différence finie

$$V(X,t_n) - V(X,t_{n-1}) = \min_{\omega} \left\{ \Delta t L(X,\omega,t_n) + V(\mathbb{1} - \Delta t A(X,\omega,t_n),t_n) - V(X,t_n) \right\},$$

d'où nous retrouvons l'équation de Hamilton-Jacobi continue en passant à la limite pour des données C^1 (ici formellement donc !).

$$\frac{\partial V}{\partial t}(X,t) + \min_{\omega} \left\{ -L(X,\omega,t) + (\nabla_x V)^T A(X,\omega,t) \right\} = \frac{\partial V}{\partial t}(X,t) - \bar{\mathcal{H}}(X,\nabla_x V(X,t),t) = 0.$$

1.4.2 Contrôle robuste

a. Formulation du problème d'optimisation

Nous savons maintenant que le filtre de Kalman faisait intervenir des hypothèses fortes sur le bruit, notamment que celui-ci devait être blanc, lorsque l'on souhaitait l'interpréter de manière probabiliste. En continu, cependant, ce que nous venons de voir est généralisable à des situations où le bruit est simplement évalué au travers de la norme W^{-1} (y compris non diagonale) et nous n'avons pas les difficultés de formalisme autour du type de bruit en temps nécessaire. En fait, toute une discipline s'est créée dans un optique de pouvoir prendre en considération tous les types de bruit dans les approches de contrôle optimale, y compris dans un cadre probabiliste. Ce type d'estimation, appelé *contrôle robuste* ou *contrôle H^∞*, suppose le bruit le plus désavantageux possible avant de déterminer un estimateur. On trouvera une description exhaustive des enjeux et fondements théoriques de cette théorie dans Başar et Bernhard (1995) et notre présentation de ses implications côté filtrage est inspirée de Başar (2002).

L'estimateur reste une fonction uniquement des observations mais l'objectif du filtrage H^∞ est donc de choisir le meilleur estimateur pour tous les bruits possibles. Il vérifie donc le critère

$$\inf_{\hat{X}} \sup_{\chi,\xi,\omega} \frac{\int_0^T \|X - \hat{X}\|_N^2 \, dt}{\|\xi\|^2_{P_0^{-1}} + \int_0^T (\|\chi\|^2_{W^{-1}} + \|\omega\|^2_{S^{-1}}) \, dt}, \tag{1.87}$$

où l'erreur d'estimation sur X (i.e. la sortie de l'estimateur) est évaluée dans une norme N de notre choix. Rapidement, on s'affranchit de ce problème *inf-sup* pour se fixer l'objectif d'un estimateur qui assure un contrôle des bruits dans le pire des cas. On se concentre donc sur trouver un $\hat{X}(Z)$ tel que

$$\sup_{\chi,\xi,\omega} \frac{\int_0^T \|X - \hat{X}\|_N^2 \, dt}{\|\xi\|^2_{P_0^{-1}} + \int_0^T (\|\chi\|^2_{W^{-1}} + \|\omega\|^2_{S^{-1}}) \, dt} \leq \frac{1}{\beta}. \tag{1.88}$$

Pour vérifier ce critère d'efficacité, fixons nous d'abord, à l'image de ce que nous avons vu en introduction de ce chapitre, un choix $\phi(Z)$ de fonction de Z fixé. Alors, on peut déjà regarder si

$$\sup_{\xi,\omega}\left\{J_T^\beta(\xi,\omega;\phi)=\frac{1}{2}\int_0^T\|X-\phi\|_N^2-\frac{1}{2\beta}\|\xi\|_{P_0^{-1}}^2-\frac{1}{2\beta}\int_0^T(\|D(Z,X)\|_{W^{-1}}^2+\|\omega\|_{S^{-1}}^2)\,ds\right\}\leq 0. \tag{1.89}$$

Ce critère est très semblable au critères moindres carrés précédents **au signe près** (ce qui est fondamental !). D'ailleurs, le critère moindres carrés usuel est appelé *critère* H^2 par comparaison. Pour calculer cette borne-sup on introduit donc, comme précédemment, la fonction

$$V(X,T;\phi)=\sup_{\omega|X(T)=X}J_T^\beta(\xi,\omega;\phi).$$

où la condition finale est fixée à X et la condition initiale est donc obtenu de manière rétrograde. Elle satisfait l'équation HJB

$$\frac{\partial V}{\partial t}(X,t)-\bar{\mathcal{H}}(X,\nabla_X V(X,t),t)=0,$$

pour

$$\bar{\mathcal{H}}(X,p,t)=-p^TA(X,t)+\frac{1}{2}\|X-\phi\|_N^2-\frac{1}{2}\|D(Z,X)\|_{W^{-1}}^2-\frac{\beta}{2}p^TS^TBSp. \tag{1.90}$$

et condition initiale

$$V(X,0;\phi)=-\frac{1}{2\beta}\|X-\bar{X}_0\|_{W^{-1}}^2.$$

Le théorème 1.4.2 permet alors de définir un estimateur noté $\hat{X}$ tel que

$$\nabla_X V(\hat{X}(T),T;\phi)=0.$$

Mais dans ce cas, l'équation de HJB nous assure qu'en choisissant $\phi=\hat{X}(T)$ on a

$$\forall T,\frac{\mathrm{d}V}{\mathrm{dt}}=\frac{\partial V}{\partial t}+\nabla_X V\dot{\hat{X}}=\frac{\partial V}{\partial t}(\hat{X}(T),T;\hat{X}(T))\leq 0.$$

Comme la condition initiale de V est négative, on a bien un estimateur vérifiant (1.89).

b. Cas particulier LQ H^∞

Reconsidérons maintenant le cas où A est linéaire pour déterminer l'extension H^∞ du problème LQ. Pour ce qui est de la résolution de HJB, la même séparation des variables (1.85) s'applique ici. On décompose donc V sous la forme

$$V(X,T)=\frac{1}{2}X^TP(t)^{-1}X,$$

où P vérifie alors une équation de Riccati particulière de la forme

$$\begin{cases}\dot{P}=AP+PA^T-(H^TW^{-1}H-\frac{1}{\beta^2}I)+PSP\\ P(0)=P_0\end{cases} \tag{1.91}$$

Cette expression introduit une contrainte sur l'existence de solution. En effet, les solutions de l'équation de Riccati ne peuvent être définies sans ambiguïté dans les matrices définies positives seulement lorsque le second membre est positif. Donc le coefficient β est contrôlé par le temps T pour lequel on souhaite une estimation.

1.4.3 Filtre EKF

Nous reprenons le fil conducteur de notre première section mais cette fois les opérateurs sont non-linéaires. Comme précédent nous repartons du cas statique car il offre déjà l'aperçu de la formulation des estimateurs, *a fortiori* quand on cherche une expression récursive. Désormais, la difficulté majeure réside dans l'absence d'équivalence variationnelle et il est donc nécessaire d'imaginer des formulations approchées.

a. Cas statique

Commençons par le cas statique, c'est-à-dire qu'on souhaite « inverser » la relation $Z = H(X) + \chi$. En suivant la même approche déterministe que Section 1.2.1, $\mathrm{d}J(\hat{X}) = 0$ impose cette fois

$$\mathrm{d}H(\hat{X})^T W^{-1}(Z - H(\hat{X})) = 0$$

cette équation implicite peut être résolue par approximation par une méthode de descente

$$\hat{X}_k = \hat{X}_{k-1} + k\,\mathrm{d}H(\hat{X}_{k-1})^T W^{-1}(Z - H(\hat{X}_{k-1})).$$

Cependant, dans ce type de méthode de gradient, on sait que leur efficacité est relative au préconditionnement devant le gradient. Or nous connaissons un préconditionneur de $\mathrm{d}J$ via l'inverse généralisé du linéarisé. Donc on choisit plutôt une approximation

$$\hat{X}_k = \hat{X}_{k-1} + k(\,\mathrm{d}H(\hat{X}_{k-1})^T W^{-1}\,\mathrm{d}H(\hat{X}_{k-1}))^{-1}\,\mathrm{d}H(\hat{X}_{k-1})^T(Z - H(\hat{X}_{k-1})),$$

cette méthode pouvant toujours être arrêtée après un petit nombre d'itérations. En particulier après une itération on a simplement considéré le problème d'observation linéaire sur le développement de Taylor à l'ordre 1 de l'observateur. Dans ce cas, nous avons en fait inversé une relation d'observation sur l'erreur linéarisée

$$Z = \mathrm{d}H(\hat{X}_n)(X - \hat{X}_n) + H(\hat{X}_n) + \chi_n$$

qui, avec $H_n^l = \mathrm{d}H(\hat{X}_n)$, peut se réécrire sous forme d'une observation linéaire

$$e_n^Z = Z - H(\hat{X}_n) = H_n^l e_n^X + \chi_n. \tag{1.92}$$

Une autre possibilité est d'utiliser une méthode du second ordre en résolvant un algorithme de Newton sur la relation $\mathrm{d}J(\hat{X}) = 0$. Dans ce cas, on obtient

$$\mathrm{d}J(\hat{X}_{k-1}) + \mathrm{d}^2J(\hat{X}_{k-1}) \cdot (\hat{X}_k - \hat{X}_{k-1}) = 0.$$

Or la différentielle d'ordre 2 de J est matriciellement

$$\mathrm{d}^2J(X) = \mathrm{d}H(X)^T W^{-1}\,\mathrm{d}H(X) - \sum_{1\le i\le m}(Z^{(i)} - H^{(i)}(X))^T W_i^{-1}\,\mathrm{d}^2H^{(i)}(X),$$

où les $H^{(i)}(.)$ sont les m composantes de $H(.)$. Donc

$$\hat{X}_k = \hat{X}_{k-1} + \Big(H_{k-1}^l W^{-1} H_{k-1}^l - \sum_{1\le i\le m}(Z^{(i)} - H^{(i)}(\hat{X}_{k-1}))W_i^{-1}\,\mathrm{d}^2H^{(i)}(\hat{X}_{k-1})\Big)^{-1} (H_{k-1}^l)^T W^{-1}(Z - H(\hat{X}_{k-1})).$$

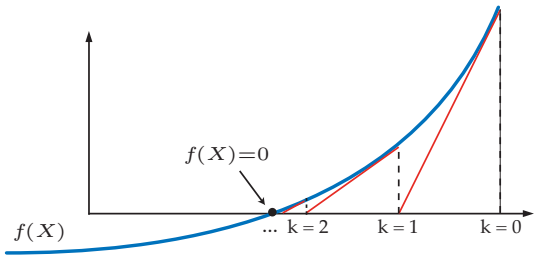


FIGURE 1.3: Principe de l'algorithme de Newton (utilisé aussi au chapitre 2).

b. Cas dynamique

Dans le cas dynamique et dans l'impossibilité de calculer explicitement $\nabla_X^2 V$, on procède au même type d'approximation qu'en statique en remplaçant la dynamique et l'équation d'observation par leur linéarisé. Dans ce cas on définit un filtre K remplaçant $\nabla_X^2 V$. C'est le cadre classique du *filtre de Kalman étendu*.

□ CAS DISCRET EN TEMPS – on suppose donc disposer d'une estimation à l'instant n, notée $\hat{X}_n^+$, associée à une covariance d'erreur d'estimation P_n^+. Commençons alors par la prédiction

$$\begin{aligned}\hat{X}_{n+1}^- &= \mathbb{E}(A(\hat{X}_n^+ + \tilde{X}_n^+)|Z_0 \dots Z_n)\\ &= \mathbb{E}(A(\hat{X}_n^+) + \frac{\partial A}{\partial X}\tilde{X}_n^+ + o(\|\tilde{X}_n^+\|)|Z_0 \dots Z_n)\\ &= A(\hat{X}_n^+) + \frac{\partial A}{\partial X}\mathbb{E}(\tilde{X}_n^+|Z_0 \dots Z_n) + o(\|\tilde{X}_n^+\|).\end{aligned}$$

Rien ne permet cependant de déterminer, dans le cas général, $\mathbb{E}(\tilde{X}_n^+|Z_0 \dots Z_n)$. En effet, la seule chose qu'on peut a priori supposer récursivement est que l'estimateur est non biaisé ce qui signifierait $\mathbb{E}(\tilde{X}_n^+) = 0$. Sinon, il faut soit supposer que jusqu'à présent on avait réussi à obtenir le MMSE soit que si on avait par exemple choisit un estimateur BLUE, les variables aléatoires étaient restées gaussiennes. On fait cependant le choix de « supposer » $\mathbb{E}(\tilde{X}_n^+|Z_0 \dots Z_n) = 0$ en plus de limiter le développement de Taylor de A à l'ordre 1. On se limite donc à

$$\hat{X}_{n+1}^- = A(X_n^+). \tag{1.93}$$

La covariance d'estimation vérifie elle

$$\begin{aligned}P_{n+1}^- &= \mathrm{Cov}(A(X_n) - A(X_n^+)|Z_0 \dots Z_n)\\ &= \frac{\partial A}{\partial X}\mathbb{E}(\tilde{X}_n^+(\tilde{X}_n^+)^T|Z_0 \dots Z_n)\frac{\partial A}{\partial X} + o(\|\tilde{X}_n^+\|),\end{aligned}$$

or là encore $\mathbb{E}(\tilde{X}_n^+(\tilde{X}_n^+)^T|Z_0 \dots Z_n)$ n'est pas P_n^+ sauf à imaginer des variables gaussiennes mais on la surestime en la remplaçant par

$$P_{n+1}^- = \frac{\partial A}{\partial X}(X_n^+)P_n^+\frac{\partial A}{\partial X}(X_n^+)^T. \tag{1.94}$$

Pour les observations, on considère la variable d'erreur d'observation

$$\begin{aligned}\tilde{Z}_{n+1} &= Z_{n+1} - H(\hat{X}_{n+1}^-)\\ &= H(X_{n+1}) - H(\hat{X}_{n+1}^-) + \chi_{n+1}\\ &= \frac{\partial H}{\partial X}(\hat{X}_{n+1}^-)\tilde{X}_{n+1}^- + \chi_n + o(\|\tilde{X}_{n+1}^-\|).\end{aligned}$$

Donc en se limitant à l'ordre 1, on peut alors appliquer un estimateur BLUE à cette variable aléatoire $\tilde{X}^-_{n+1}$

$$\hat{E}(\tilde{X}_{n+1}|\tilde{Z}_{n+1}) = \mathbb{E}(\tilde{X}^-_{n+1}) + \mathrm{Cov}(\tilde{X}^-_{n+1}, \tilde{Z}_{n+1})\mathrm{Cov}(\tilde{Z}_{n+1})^{-1}(\tilde{Z}_{n+1} - \mathbb{E}(\tilde{Z}_{n+1})).$$

Or en se limitant à l'ordre 1

$$\mathbb{E}(\tilde{Z}_{n+1}) = 0,$$

puis

$$\mathrm{Cov}(\tilde{X}^-_{n+1}, \tilde{Z}_{n+1}) = P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T,$$

et

$$\mathrm{Cov}(\tilde{Z}_{n+1})^{-1} = \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P^-_{n-1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T + W_n.$$

On obtient donc pour la phase de correction

$$\begin{aligned}\hat{X}^+_{n+1} &= X^-_{n+1} + \hat{E}(\tilde{X}_{n+1}|\tilde{Z}_{n+1}) \\ &= X^-_{n+1} + P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T\Big(W_n + \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T\Big)^{-1}\big(Z_{n+1} - H(\hat{X}^-_{n+1})\big).\end{aligned}$$

Cependant il est immédiat, en reconsidérant la démonstration du filtre de Kalman via les projections successives, que cette estimation n'est en rien $\mathbb{E}(X_{n+1}|Z_0 \dots Z_{n+1})$! Donc la récurrence ne tient pas et le filtre EKF est un filtre **heuristique**. Finalement, on peut résumer l'algorithme EKF en posant $A^l_{n+1|n} = \dfrac{\partial A_{n+1|n}}{\partial X}$ et $H^l_n = \dfrac{\partial H_n}{\partial X}$.

- **Prédiction :**

$$\begin{cases}\hat{X}^-_{n+1} = A_{n+1|n}(\hat{X}^+_n), & \text{(Estimation a priori)} \\ P^-_{n+1} = A^l_{n+1|n}(\hat{X}^+_n)P^+_n A^l_{n+1|n}(\hat{X}^+_n)^T, & \text{(Covariance a priori)}\end{cases} \tag{1.95a}$$

- **Correction :**

$$\begin{cases}C_{n+1} = H^l_{n+1}P^-_{n+1}(H^l_{n+1})^T + W_{n+1}, & \text{(Covariance de l'innovation)} \\ K_{n+1} = P^-_{n+1}(H^l_{n+1})^T C^{-1}_{n+1}, & \text{(Gain de Kalman)} \\ \hat{X}^+_{n+1} = \hat{X}^-_{n+1} + K_{n+1}\big(Z_{n+1} - H_{n+1}(\hat{X}^-_{n+1})\big), & \text{(Estimation a posteriori)} \\ P^+_{n+1} = (I - K_{n+1}H^l_{n+1})P^-_{n+1}, & \text{(Covariance a posteriori)}\end{cases} \tag{1.95b}$$

Pour obtenir le filtre de Kalman il suffit donc simplement de prendre $A_{n+1|n}(\cdot) = A^l_{n+1|n}(\cdot) = A_{n+1|n}$.

□ Cas continu en temps – Enfin, ce système récursif s'étend au cadre continu sous la forme

$$\begin{cases}\dot{\hat{X}} = A(\hat{X}, t) + K(Z - H(\hat{X})), & \text{(Estimateur)} \\ \hat{X}(0) = X_0, & \text{(A priori sur la condition initiale)} \\ K = P\dfrac{\partial H}{\partial X}^T W^{-1}, & \text{(Gain de Kalman)} \\ -\dot{P} + P\dfrac{\partial A}{\partial X}^T + \dfrac{\partial A}{\partial X}P - P\dfrac{\partial H}{\partial X}^T W^{-1}\dfrac{\partial H}{\partial X}P = 0, & \text{(Riccati)}\end{cases} \tag{1.96}$$

REMARQUE 1.4.2 (BRUIT NON ADDITIF)
Si on souhaite s'affranchir de l'hypothèse de bruit additif, c'est-à-dire

$$Z = H(X, \chi),$$

alors le filtre EKF linéarise la dépendance et se ramène au cas additif avec

$$W \to \frac{\partial H}{\partial \chi}^T W \frac{\partial H}{\partial \chi}.$$

De même, toutes ces expressions sont compatibles avec une observation de la forme $\chi = D(Z, X)$. Enfin pour un bruit de modèle non additif il suffit d'ajouter dans le Riccati le terme

$$\frac{\partial A}{\partial \omega}^T S \frac{\partial H}{\partial \omega}.$$

c. Kalman d'ordre 2

A l'image de ce que nous avons fait en statique, il est aussi possible de considérer un filtre de Kalman d'ordre 2 comme le prolongement du filtre EKF un cran plus loin dans les développements de Taylor. On peut trouver de nombreux compléments dans Jazwinsky (1970) ainsi qu'un résumé pratique et bibliographique dans Simon (2006).

□ CAS DISCRET EN TEMPS – On reprend rapidement le filtre EKF mais on développe l'opérateur de transition à l'ordre 2,

$$\hat{X}_{n+1}^- = \mathbb{E}(A(X_n^+) + \frac{\partial A}{\partial X}\tilde{X}_n^+ + \frac{1}{2}\frac{\partial^2 A}{\partial X^2} : \tilde{X}_n^+ \otimes \tilde{X}_n^+ + o(\|\tilde{X}_n^+\|^2)|Z_0 \dots Z_n),$$

donc

$$\hat{X}_{n+1}^- = A(X_n^+) + \frac{1}{2}\frac{\partial^2 A}{\partial X^2}(X_n^+) : P_{n+1}^+. \tag{1.97}$$

La matrice de covariance à l'ordre 2 est la même que pour le filtre EKF car les termes supplémentaires sont tous au moins d'ordre 2 (voir Section 1.4.4 pour un développement détaillé)

$$P_{n+1}^- = \frac{\partial A}{\partial X}(X_n^+)P_n^+\frac{\partial A}{\partial X}(X_n^+)^T. \tag{1.98}$$

Côté correction, les choses se compliquent avec le développement à l'ordre 2 de l'erreur d'observation

$$\tilde{Z}_{n+1} = \frac{\partial H}{\partial X}(\hat{X}_{n+1}^-)\tilde{X}_{n+1}^- + \frac{1}{2}\frac{\partial^2 H}{\partial X^2}(\hat{X}_{n+1}^-)\tilde{X}_{n+1}^- \otimes \tilde{X}_{n+1}^- + \chi_{n+1} + o(\|\tilde{X}_{n+1}^-\|^3).$$

On a alors

$$\hat{E}(\tilde{X}_{n+1}|\tilde{Z}_{n+1}) = \hat{X}_{n+1}^- + \mathrm{Cov}(\tilde{X}_{n+1}^-, \tilde{Z}_{n+1})\mathrm{Cov}(\tilde{Z}_{n+1})^{-1}(\tilde{Z}_{n+1} - \frac{1}{2}\frac{\partial^2 H}{\partial X^2}(\hat{X}_{n+1}^-) : P_{n+1}^-).$$

La correction de biais d'observation induit des modifications sur les matrices de covariance $\mathrm{Cov}(\tilde{X}_{n+1}^-, \tilde{Z}_{n+1})$ et $\mathrm{Cov}(\tilde{Z}_{n+1})$. Certes, on a toujours

$$\begin{aligned}\mathrm{Cov}(\tilde{X}_{n+1}^-, \tilde{Z}_{n+1}) &= \mathbb{E}(\tilde{X}_{n+1}^-(\chi_{n+1} + \frac{\partial H}{\partial X}(\hat{X}_{n+1}^-)\tilde{X}_{n+1}^- - \frac{1}{2}\frac{\partial^2 H}{\partial X^2}(\hat{X}_{n+1}^-) : P_{n+1}^-)^T) \\ &= P_{n+1}^-\frac{\partial H}{\partial X}(\hat{X}_{n+1}^-)^T,\end{aligned}$$

mais

$$\begin{aligned}\mathrm{Cov}(\tilde{Z}^-_{n+1}, \tilde{Z}_{n+1}) &= \mathbb{E}\Big((\chi_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})\tilde{X}^-_{n+1} - \frac{1}{2}\frac{\partial^2 H}{\partial X^2}(\hat{X}^-_{n+1}) : P^-_{n+1})(\ldots)^T)\Big)\\ &= W_n + \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T + \frac{1}{4}\Big(\frac{\partial^2 H}{\partial X^2} : P^-_{n+1}\Big)\Big(\frac{\partial^2 H}{\partial X^2} : P^-_{n+1}\Big)^T\\ &= W_n + \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T + \Lambda_{n+1}.\end{aligned}$$

La notation tensorielle est mélangée à la notation matricielle afin de simplifier l'écriture mais il faut se reporter au cas statique pour un notation matricielle plus rigoureuse.

Finalement, tous calculs faits, la correction d'ordre 2 est donc

$$\begin{aligned}\hat{X}^+_{n+1} = \hat{X}^-_{n+1} + P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T\Big(W_n + \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1}) + \Lambda_{n+1}\Big)^{-1}\\ \Big(Z_{n+1} - H(\hat{X}^-_{n+1}) - \frac{1}{2}\frac{\partial^2 H}{\partial X^2}(\hat{X}^-_{n+1}) : P^-_{n+1}\Big). \quad (1.99)\end{aligned}$$

et pour la covariance

$$\begin{aligned}P^+_{n+1} &= P^-_{n+1}\\ &+ P^-_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})^T\Big(W_n + \frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P_{n+1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1}) + \Lambda_{n+1}\Big)^{-1}\frac{\partial H}{\partial X}(\hat{X}^-_{n+1})P^-_{n+1}. \quad (1.100)\end{aligned}$$

□ CAS CONTINU EN TEMPS – L'estimateur continu est encore une fois obtenu en passant à la limite à partir du système récursif réecrit sous forme « one-step ». Nous n'effectuons pas ce calcul laborieux mais indiquons juste pour la suite que lorsque A est non-linéaire mais H linéaire, le filtre revient simplement à corriger la dynamique à l'ordre 2 sous la forme

$$\hat{X} = A(\hat{X}) + K(Z - HX) + \frac{1}{2}\frac{\partial^2 A}{\partial X^2}(\hat{X}) : P. \quad (1.101)$$

1.4.4 Filtre UKF

Les deux filtres que nous venons d'étudier sont donc fondés sur la qualité de l'approximation de variables aléatoires résultant de l'application d'opérateurs linéaires. Le filtre EKF est très largement utilisé mais conduit à des estimations assez grossières. Quand au filtre de Kalman d'ordre 2, il montre en pratique de fortes instabilités comme le résume Simon (2006). C'est ainsi que Julier et al (1995) proposèrent, à partir de la propagation de variables aléatoires par des opérateurs non-linéaires, non pas d'utiliser des développements de Taylor tronqués mais leur version interpolée. Ils en déduisent le filtre UKF pour *Unscented Kalman Filter*, écrit spécifiquement dans le cadre discret en temps, dont l'efficacité est avérée dans de nombreuses situations pratiques. Par exemple sur des systèmes mécaniques de faible dimension on peut déjà citer les travaux de Wu et Smyth (2006) ou Mariani et Ghisi (2007). Afin de le présenter, nous commençons par un bilan des formules des développements de Taylor que nous venons d'écrire pour les filtres EKF et d'ordre 2 avant de montrer comment l'interpolation et l'utilisation de covariances empiriques permettent de nouvelles approximations. Ensuite le filtre UKF est exactement à l'image des deux précédents pour la définition des étapes de prédiction et de correction. C'est donc lui aussi un filtre avant tout *heuristique*. C'est un reproche qui a

d'ailleurs rendu délicate l'introduction de ce filtre d'un point de vue théorique (voir notamment Lefebvre et al (2002)) avant que des mises en oeuvre pratique assurent son efficacité par rapport au filtre EKF. C'est pourquoi ce filtre datant de 1995 donne parfois l'impression d'avoir véritablement émergé cinq ans plus tard (cf. Julier et al (2000)).

a. Transformation *unscented*

Supposons donc que nous voulions estimer la moyenne $\bar{X}^f$ et la covariance P^{ff} du processus aléatoire X^f défini par l'application d'une fonction non-linéaire sur X

$$X^f = f(X). \tag{1.102}$$

En utilisant la formule de Taylor-Lagrange autour de $\bar{X}$ avec $\tilde{X} = X - \bar{X}$, on résume les expressions à l'ordre 2 déjà utilisées :

- **Moyenne**

$$\begin{aligned}\bar{X}^f &= \mathbb{E}(f(\bar{X} + \tilde{X})) \\ &= \mathbb{E}(f(\bar{X}) + \mathrm{d}f(\bar{X})\tilde{X} + \tfrac{1}{2}\,\mathrm{d}^2 f(\bar{X}) : \tilde{X} \otimes \tilde{X} + o(\|\tilde{X}\|^2)) \\ &= f(\bar{X}) + \tfrac{1}{2}\,\mathrm{d}^2 f(\bar{X}) : P + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right).\end{aligned}$$

- **Covariance**

$$\begin{aligned}P^{ff} &= \mathbb{E}((X^f - \bar{X}^f)(X^f - \bar{X}^f)^T) \\ &= \mathbb{E}((\,\mathrm{d}f(\bar{X})\tilde{X} + \tfrac{1}{2}\,\mathrm{d}^2 f(\bar{X}) : (\tilde{X} \otimes \tilde{X} - P))(\,\mathrm{d}f(\bar{X})\tilde{X} + \tfrac{1}{2}\,\mathrm{d}^2 f(\bar{X}) : (\tilde{X} \otimes \tilde{X} - P))^T) \\ &\qquad + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right) \\ &= \,\mathrm{d}f(\bar{X}) \cdot P \cdot \,\mathrm{d}f(\bar{X})^T + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right).\end{aligned}$$

Then,

Supposons maintenant qu'on souhaite approcher chacune de ces expressions à partir d'évaluations sur un nombre restreint de valeurs (particules). On définit donc r particules appelées classiquement *sigma-points* par

$$X_i = \bar{X} + \tilde{X}_i, \tag{1.103}$$

avec $r > N$ mais du même ordre de grandeur. À ces particules, on associe r coefficients α_i tels que les relations suivantes soient vérifiées

$$\begin{cases} \displaystyle\sum_{1 \le i \le r} \alpha_i = 1 \\ X_\alpha \overset{\text{def}}{=} \displaystyle\sum_{1 \le i \le r} \alpha_i X_i = \mathbb{E}(X) \\ P_\alpha^{XX} \overset{\text{def}}{=} \displaystyle\sum_{1 \le i \le r} \alpha_i (X_i - X_\alpha)(X_i - X_\alpha)^T = P^{XX} \end{cases} \tag{1.104}$$

Ces conditions impliquent alors

$$\begin{cases} \displaystyle\sum_{1 \le i \le r} \alpha_i \tilde{X}_i = 0 \\ \displaystyle\sum_{1 \le i \le r} \alpha_i \tilde{X}_i \tilde{X}_i^T = P^{XX} \end{cases} \tag{1.105}$$

et permettent d'exprimer la moyenne et variance de X^f à partir des *Sigma-points propagés* :

$$X_i^f = f(X_i). \tag{1.106}$$

En effet on vérifie alors

$$\begin{aligned}
X_\alpha^f &\overset{def}{=} \sum_{1\le i\le r} \alpha_i X_i^f \\
&= f(\bar{X}) + \sum_{1\le i\le r} \alpha_i \,\mathrm{d}f(\bar{X})\tilde{X}_i + \sum_{1\le i\le r} \alpha_i \,\mathrm{d}^2 f(\bar{X}) : \tilde{X}_i \otimes \tilde{X}_i + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right) \\
&= f(\bar{X}) + \,\mathrm{d}^2 f(\bar{X}) : P^{XX} + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right) = \bar{X}^f + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right),
\end{aligned}$$

et

$$\begin{aligned}
P_\alpha^{XX} &\overset{def}{=} \sum_{1\le i\le r} \alpha_i (f(X_i) - X_\alpha^f)(f(X_i) - X_\alpha^f)^T \\
&= \sum_{1\le i\le r} \alpha_i \left(\mathrm{d}f(\bar{X})\tilde{X}_i + \,\mathrm{d}^2 f(\bar{X}) : (\tilde{X}_i \otimes \tilde{X}_i - P^{XX}) + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right)\right) (...)^T \\
&= \,\mathrm{d}f(\bar{X}).\left(\sum_{1\le i\le r} \alpha_i \tilde{X}_i \tilde{X}_i^T\right).\,\mathrm{d}f(\bar{X}) + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right) \\
&= \,\mathrm{d}f(\bar{X}) \cdot P^{XX} \cdot \,\mathrm{d}f(\bar{X}) + o\left(\mathbb{E}\|\tilde{X}\|^2\right) = P^{ff} + o\left(\mathbb{E}(\|\tilde{X}\|^2)\right).
\end{aligned}$$

Cette approximation de la moyenne et de la covariance d'un vecteur aléatoire transformé par un opérateur non-linéaire est le point de départ de l'*Unscented Kalman Filtering* UKF.

□ CHOIX DES SIGMA-POINTS – La seule difficulté est de bien définir les sigma-points « autour » de X tels que l'interpolation de l'opérateur non-linéaire soit « bonne ». L'avantage est, qu'en fait, il suffit de déterminer les sigma-points d'un vecteur aléatoire I de même dimension, de moyenne nulle $\mathbb{E}(I) = 0$ et de variance l'identité $P^I = \mathbb{1}$. En effet, si on suppose ces sigma-points I_i définis, on peut prendre

$$\tilde{X}_i = \sqrt{P^{XX}} I_i \tag{1.107}$$

pour toute racine carrée de P^{XX}, et vérifier les conditions ci-dessus

$$\begin{cases} \displaystyle\sum_{1\le i\le r} \alpha_i \tilde{X}_i = 0 \\ \displaystyle\sum_{1\le i\le r} \alpha_i \tilde{X}_i \tilde{X}_i^T = \sqrt{P^{XX}} \cdot I_d \cdot \sqrt{P^{XX}} = P^{XX}. \end{cases} \tag{1.108}$$

Nous nommerons personnellement de tels points *sigma-points unitaires*. Il est important de noter que les sigma-points permettant le calcul de la covariance empirique peuvent être placées à une rotation unitaire près. En effet pour $U \in O_N(\mathbb{R})$, en définissant

$$\tilde{X}_i = U\sqrt{P^{XX}} I_i,$$

on a toujours

$$\sum_{1\le i\le r} \alpha_i \tilde{X}_i \tilde{X}_i^T = \sqrt{P^{XX}} \cdot U^T U \cdot \sqrt{P^{XX}} = P^{XX}.$$

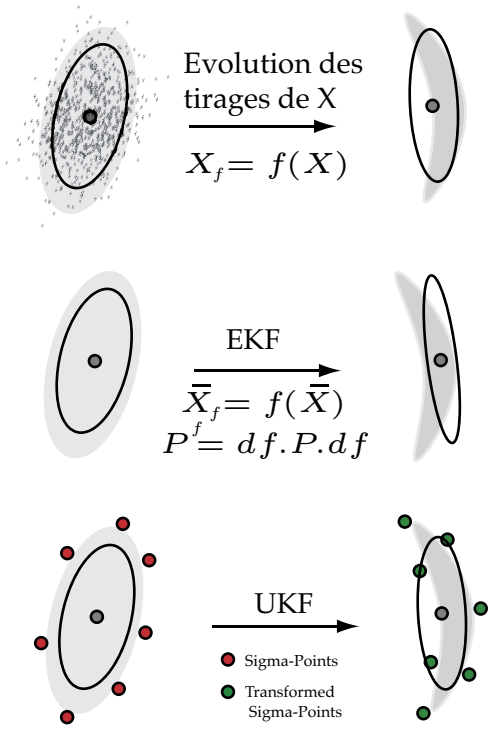


FIGURE 1.4: Principe de la transformation UKF afin de mieux représenter l'évolution non-linéaire d'une variable aléatoire. Sont représentés en gris différents tirages de la variable aléatoire (ici N=2). Les covariances des variables aléatoires sont représentées par des ellipses d'axes les composantes principales de la matrice P.

Le dernier point concernant la répartition des sigma-points est leur distance par rapport à la moyenne. Il est clair que les trois conditions sur les points et les coefficients α_i ne sont pas une contrainte très forte sur la distance de chaque point par rapport au point moyen. En effet, il n'est pas difficile d'imaginer éloigner les points des un des autres tout en compensant au niveau des coefficients α_i. Cette critique, parmi d'autres, est en particulier présente dans Lefebvre et al (2002). Il est donc crucial de choisir la distance globale des points par rapports à première propriété recherchée qui est d'interpoler à l'ordre 2 l'opérateur non-linéaire f agissant sur le processus X. Pour cela, comme présenté Figure 1.4, on essaiera autant que faire ce peut de placer les sigma-points à $\sqrt{\mathbb{E}(\|X - \mathbb{E}(X)\|)}$ de la moyenne. Or

$$\begin{aligned}\mathbb{E}(\|X - \mathbb{E}(X)\|^2) &= E((X - \mathbb{E}(X))^T (X - \mathbb{E}(X))) \\ &= \mathbb{E}(\mathrm{tr}((X - \mathbb{E}(X))^T (X - \mathbb{E}(X)))) = \mathbb{E}(\mathrm{tr}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^T)) \\ &= \mathbb{E}(\mathrm{tr}\, P).\end{aligned}$$

Donc si P est la matrice unité les tirages sont en moyenne à $\sqrt{N}$ et les sigma-points doivent l'être aussi. Afin d'illustrer ces remarques nous finissons par quelques exemples de sigma-points caractéristiques

- **sigma-points canoniques**[11]($r = 2N$) : Ils sont placés sur la base canonique de part et

[11] cette dénomination n'est pas officielle. Ces sigma-points sont en fait l'échantillon historique proposé par Julier et al (2000)

d'autre de l'origine. Leurs coefficients associés sont tout simplement $\alpha_i = \frac{1}{2N}$.

$$I_i = \begin{cases} \sqrt{N}, & \text{si } 1 \leq i \leq r \\ -\sqrt{N}, & \text{si } N+1 \leq i \leq 2r \end{cases} \tag{1.109}$$

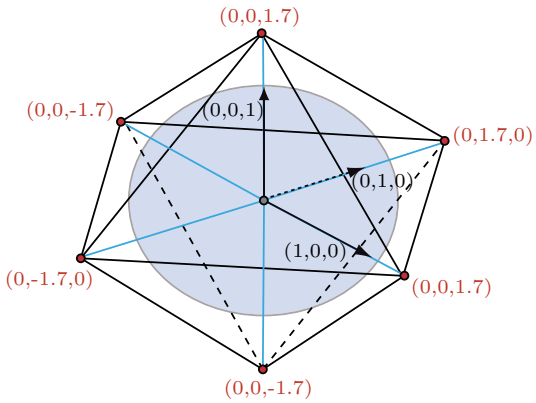


FIGURE 1.5: Sigma-points canoniques en dimension 3. Pour obtenir les sigma-points étoilés il suffit d'ajouter la moyenne. En bleu, la matrice de covariance unitaire.

- **sigma-points étoilés**[12] ($r = 2N + 1$) : Au points définies plus haut on peut ajouter un point centrale correspondant à la moyenne avec un poids particulier.

$$I_i = \begin{cases} \sqrt{N}, & \text{si } 1 \leq i \leq r \\ -\sqrt{N}, & \text{si } N+1 \leq i \leq 2r \\ 0 & \text{si } i = 2r+1 \end{cases} \tag{1.110}$$

dans ce cas, Julier, and J. K. Uhlmann (2002b) démontre que certains types de densité peuvent être représentés à un ordre supérieur à deux pour un choix de poids adéquat de α_i.

- **sigma-points simplectiques** ($r = N + 1$) : Ces $N + 1$ points correspondent au plus petit nombre nécessaire pour représenter la covariance et sont situés sur le polyèdre régulier de rayon $\sqrt{r}$. A l'image de ce qui est proposé dans Julier, and J. K. Uhlmann (2002a), on construit ces points récursivement en définissant

$$I_i = \sqrt{N+1}\tilde{I}_{i,N+1}$$

où les $\tilde{I}_{i,r}$ sont issus de la matrice de leur coordonnées $[\tilde{I}_{i,N+1}]$ définie ci-dessous. De manière générale, la matrice de $\mathbb{M}_{N,r}$ où chaque colonne correspond à un des sigma-points sera noté $[X_i]$ avec l'abus de notation sur l'indice. Pour les $\tilde{I}_{i,N+1}$, nous la construisons

[12] idem (cf. note précédente)

comme suit

$$\begin{cases} [\tilde{I}_{i,1}] = \left(-\frac{1}{\sqrt{2\alpha}} \quad \frac{1}{\sqrt{2\alpha}}\right), \quad \alpha = \frac{r}{r+1} \\ [\tilde{I}_{i,d}] = \begin{pmatrix} & & & 0 \\ & [\tilde{I}_{i,d-1}] & & \vdots \\ & & & 0 \\ \frac{1}{\alpha d(d+1)} & \cdots & \frac{1}{\alpha d(d+1)} & \frac{-d}{\alpha d(d+1)} \end{pmatrix}, \quad 2 \leq d \leq N+1 \end{cases}$$

Concernant les poids, puisque chaque point est placé sur un polyèdre régulier entourant la moyenne, ils sont donc tous égaux à $\alpha_i = \frac{1}{N+1}$.

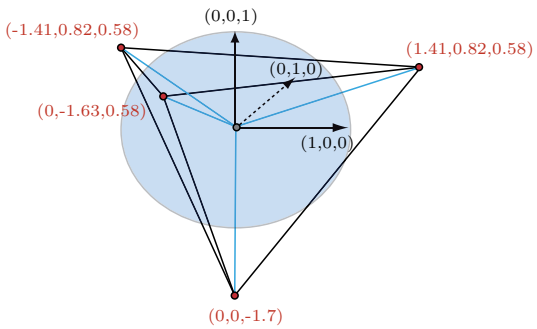


FIGURE 1.6: Sigma-points simplectiques en dimension 3 placés donc sur un tétraèdre régulier

b. Unscented Kalman Filtering

Le principe du filtre *unscented* est d'utiliser les moyennes et covariances empiriques et leur propagation en lieu et place des moyennes et covariances approchées des filtres EKF. On obtient donc directement

- **Prédiction :**

$$\begin{cases} X_{n,i}^{+} = \hat{X}_n^{+} + (\sqrt{P_n^{+}})I_i \\ \hat{X}_{n+1}^{-} = X_{n+1,\alpha}^{-} = \sum_{1 \leq i \leq r} \alpha_i A(X_{n,i}^{+}) \\ P_{n+1}^{-} = P_{\alpha}^{X_n X_n} = \sum_{1 \leq i \leq r} \alpha_i (A(X_{n,i}^{+}) - \hat{X}_{n+1}^{-})(A(X_{n,i}^{+}) - \hat{X}_{n+1}^{-})^T \end{cases} \tag{1.111a}$$

- **Correction :**

$$\begin{cases} X^-_{n+1,i} = \hat{X}^-_{n+1} + (\sqrt{P^-_{n+1}})I_i \\ Z_{n+1,i} = H(X^-_{n+1,i}) \\ Z_{n+1,\alpha} = \sum_{1\leq i\leq r} \alpha_i Z_{n+1,i} \\ P^{X\tilde{Z}}_\alpha \stackrel{\text{def}}{=} \sum_{1\leq i\leq r} \alpha_i (X^-_{n+1,i} - X^-_{n+1,\alpha})(Z^-_{n+1,i} - Z^-_{n+1,\alpha})^T \\ P^{\tilde{Z}}_\alpha \stackrel{\text{def}}{=} W + \sum_{1\leq i\leq r} \alpha_i (Z^-_{n+1,i} - Z^-_{n+1,\alpha})(Z^-_{n+1,i} - Z^-_{n+1,\alpha})^T \\ K_{n+1} = P^{\tilde{X}\tilde{Z}}_\alpha (P^{\tilde{Z}}_\alpha)^{-1} \\ \hat{X}^+_{n+1} = \hat{X}^-_n + K_{n+1}(Z_{n+1} - Z_{n+1,\alpha}) \\ P^+_{n+1} = P^-_n - P^{\tilde{X}\tilde{Z}}_\alpha (P^{\tilde{Z}})^{-1} (P^{\tilde{X}\tilde{Z}}_\alpha)^T. \end{cases} \tag{1.111b}$$

Evidemment, pour les mêmes raisons que EKF, il n'y a pas de justification rigoureuse au filtre *unscented* sauf dans le cas où les variables sont supposées gaussiennes où alors on peut considérer $\mathbb{E}(X|Z)$ au lieu de $\hat{E}(X|Z)$. Mais même dans ce cas les opérateurs non-linéaires ne permettent pas d'affirmer que les variables restent gaussiennes au cours du temps. On paye de nouveau l'absence de récursivité. Cependant ce n'est pas pour cela que le filtre n'est pas bon. Il est déjà sans biais et la question de son efficacité, i.e. la covariance d'erreur finale moyennée sur toutes les observations Z, peut légitimement être posée. Cette question est semble-t-il pour UKF, comme pour les filtres EKF et d'ordre 2, toujours ouverte. Toujours est-il que les trois derniers filtres que nous venons d'aborder appartiennent donc à la même classe de filtres approchés « gaussiens ». Leur seule distinction se situe dans la gestion des développements de Taylor des opérateurs non-linéaires. Malgré la notion de particules introduites par le filtre UKF, leur nombre et leur positionnement font qu'on est très loin des approches particulaires (cf. Legland (2002)) et même de filtrage d'ensemble EnKF synthétisé dans Evensen (2007). Pour le premier, toute la distribution est estimée à partir de particules via des méthodes de Monte-Carlo. Pour le deuxième, souvent imaginé comme une version stochastique du filtre EKF, on reste dans la catégorie des filtres « gaussiens » qui ne cherchent que les deux premiers moments de la distribution de probabilité à partir de moyennes et covariances empiriques de particules (voir Blum et al (2008)). Cependant, pour ce filtre EnKF comme pour le filtre particulaire, on utilise un très grand nombre de particules alors qu'UKF sélectionne le nombre « juste » nécessaire à l'interpolation des opérateurs non-linéaires. Ce dernier est pour ces raisons fondamentalement déterministe par rapport aux deux autres.

1.5 Conclusion

Ce premier chapitre est donc un tour d'horizon des différentes techniques d'estimation en dimension finie. Il nous a permis d'introduire l'ensemble des concepts utiles pour structurer notre problématique d'estimation et de montrer comment en passant du point de vue probabiliste au point de vue déterministe on peut apporter de nouveaux éclairages. Notamment nous avons fait le lien entre d'un côté l'approche dite variationnelle consistant à minimiser un critère de coût relatif à l'erreur d'observation et d'a priori et de l'autre les approches séquentielles. Cependant, nous avons déjà commencé a observer que l'équivalence ne vaut que dans

le cadre LQ. Dès qu'on s'écarte de cette situation, on peut alors choisir de continuer à minimiser le critère ou de trouver un filtre efficace malgré tout. Reste maintenant à appliquer ce que nous avons vu en grande dimension avec des vecteurs aléatoires approchant des champs continus relatif à la modélisation de la contraction mécanique du coeur. Or en pratique, il existe de nombreuses limitations liées à la taille du système notamment pour les approches séquentielles comme déjà constaté avec HJB. Ainsi l'approche variationnelle est « souvent » préférée dans les disciplines historiques de l'estimation pour ces modèles distribués, i.e l'assimilation de données. Nous choisirons pourtant une approche séquentielle où nous allons proposer des filtres implémentables, y compris en non-linéaire car adaptés à la physique du problème considéré. Pour ce faire, il nous faut d'abord maîtriser le problème direct de la contraction cardiaque en profondeur, soit l'objectif du chapitre suivant.

CHAPITRE

2

Modélisation cardiaque

There are three things that every beginner in elasticity theory should know. The first is that « Kirchhoff » has two h's in it. The second is that Hooke's law will not be found as a basic axiom (it « really » means you are working with linearized theory). The third is that researchers in elasticity theory are very opinionated, even when they are wrong.

— Jerrold E. Marsden and Thomas J.R. Hugues, *preface of Marsden et Hughes (1983)*

Sommaire du chapitre

2.1 Le cœur **68**
2.2 Eléments de physiologie cardiaque **69**
2.2.1 Anatomie Cardiaque 69
2.2.2 Cycle cardiaque 71
2.2.3 Formation et conduction de l'excitation électrique 73
2.2.4 La fonction cardiaque 75
2.2.5 Le réseau vasculaire 78
2.2.6 Physiologie détaillé des tissus musculaires cardiaques 80
2.3 Modélisation électromécanique cardiaque **85**
2.3.1 Formulation mécanique en Lagrangien total 85
2.3.2 Loi de comportement mécanique 88
2.3.3 Modèles de circulation 100
2.3.4 Modèle cardiaque complet 105
2.3.5 Conditions aux limites 105
2.3.6 Eléments de modélisation électrique 107
2.4 Analyse numérique et simulations **112**
2.4.1 Modèles géométriques 112
2.4.2 Discrétisation 124
2.4.3 Calibration du modèle 130
2.4.4 Simulations numériques 131
2.5 Observations cardiaques **136**
2.5.1 IRM 136
2.5.2 Echographie 143
2.5.3 Mesures de pression 152
2.5.4 Mesures électriques 152
2.6 Modèles et observateurs en 1D **153**
2.7 Conclusion **156**

2.1 Le cœur

Le cœur est un organe à l'activité complexe. C'est une pompe musculaire commandée par des signaux électriques dont la fonction est de pulser le sang à chaque battement. Son architecture en plusieurs chambres d'admission et d'éjection ainsi que ses propriétés électro-mécaniques en font un moteur adapté à l'organisme. L'objectif de ce chapitre est de présenter une modélisation possible de cette activité, principalement dans ses aspects de mécanique du solide lors de la contraction de l'organe. Cette modélisation, fondée sur une approche multi-échelle allant de la fibre musculaire à l'organe tout entier, est le résultat d'un travail collectif à l'INRIA auxquelles les contributions de cette thèse sont principalement numériques. Et si nous présentons en détails dans ce chapitre l'effort de modélisation bien que l'objet de cette thèse soit le problème inverse, c'est que nous sommes convaincus que pour présenter et valider une méthodologie d'estimation il est absolument indispensable de bien comprendre les propriétés intrinsèques du modèle direct. Les propriétés strictement quantitatives sont, en revanche, secondaires dans cette démarche. Certes envisager une application clinique à plus long terme nécessitera une confrontation quantitative intensive à l'expérience. Cependant cette validation est partiellement découplée de la définition des méthodes d'assimilation de données si on peut assurer que celles-ci sont définies synthétiquement sur des modèles et des mesures numériques représentatifs des difficultés d'observabilité du problème. Ainsi, ce chapitre est plus un itinéraire à travers toutes les briques nécessaires et effectivement assemblées dans un simulateur de mécanique cardiaque avant (ou plutôt en parallèle de) la définition et la formalisation des estimateurs inverses.

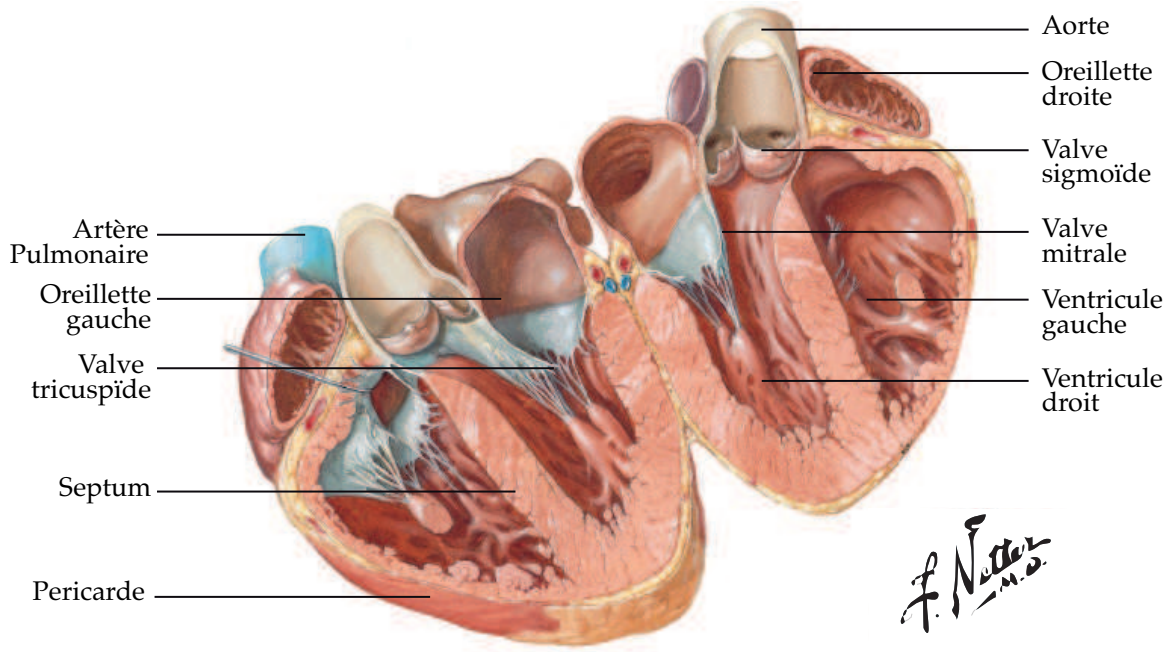


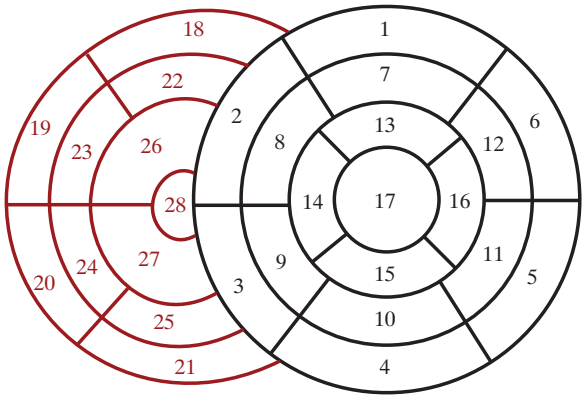
FIGURE 2.1: Coupe anatomique tirée de Netter (1969).

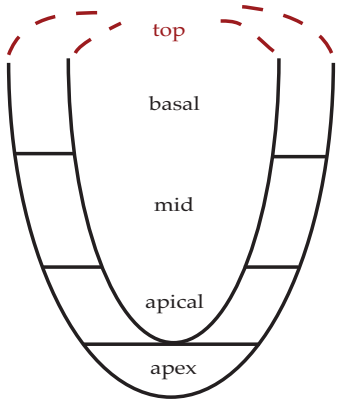
2.2 Eléments de physiologie cardiaque

2.2.1 Anatomie Cardiaque

L'anatomie cardiaque est adaptée à sa fonction principale de pompe cardiaque. Elle comprend quatre chambres. Les oreillettes droites et gauches sont en quelque sorte des chambres d'accès vers leur ventricule respectif, droit et gauche, eux-mêmes véritables chambres de propulsion vers respectivement les poumons (petite circulation), et respectivement le reste de l'organisme (la grande circulation) via l'aorte. Le mécanisme de remplissage-éjection est contrôlé par différentes valves qui assurent l'écoulement unidirectionnel du sang. Tout d'abord les valves *auriculo-ventriculaires* (*tricuspide* à droite et *mitrale* à gauche) remplissent deux fonctions : d'une part elles canalisent le sang de l'oreillette vers le ventricule ; d'autre part elles préviennent, lors de la contraction ventriculaire (*systole*), le reflux du sang vers l'oreillette. Comparativement, les valves dites *sigmoïdes* situées à la jonction entre l'un des ventricules et, soit l'aorte à gauche ou l'artère pulmonaire à droite, empêchent, pendant le remplissage, le reflux du sang des artères vers les ventricules.

Les deux ventricules, qui sont donc les forces motrices, correspondent chacun à un ellipsoïde épais et tronqué. Le gauche est approximativement un solide de révolution, alors que le droit vient s'encastrer dans le premier (cf. Schéma 2.4). Principalement constitué du myocarde, muscle s'étendant de l'endocarde délimitant l'intérieur des ventricules, à l'épicarde délimitant l'extérieur, les ventricules ont une épaisseur respective moyenne de $0.5\ cm$ pour le droit, et $1.5\ cm$ pour le gauche qui génère le plus gros effort de propulsion. L'endocarde et l'épicarde sont deux couches minces délimitant donc la région musculaire, et l'ensemble du cœur est lui-même contenu dans le sac péricardique.





1. basal anterior
2. basal anteroseptal
3. basal inferoseptal
4. basal inferior
5. basal inferolateral
6. basal anterolateral
7. mid anterior
8. mid anteroseptal
9. mid inferoseptal
10. mid inferior
11. mid inferolateral
12. mid anterolateral
1.3 apical anterior
14. apical septal
15. apical inferior
16. apical lateral
17. apex
18. basal anterior
19. basal anterolateral
20. basal inferolateral
21. basal inferior
22. mid anterior
23. mid anterolateral
24. mid inferolateral
25. mid inferior
26. basal anterior
27. basal inferior
28. apex

FIGURE 2.2: Régions AHA officielles pour le ventricule gauche et extension au ventricule droit.

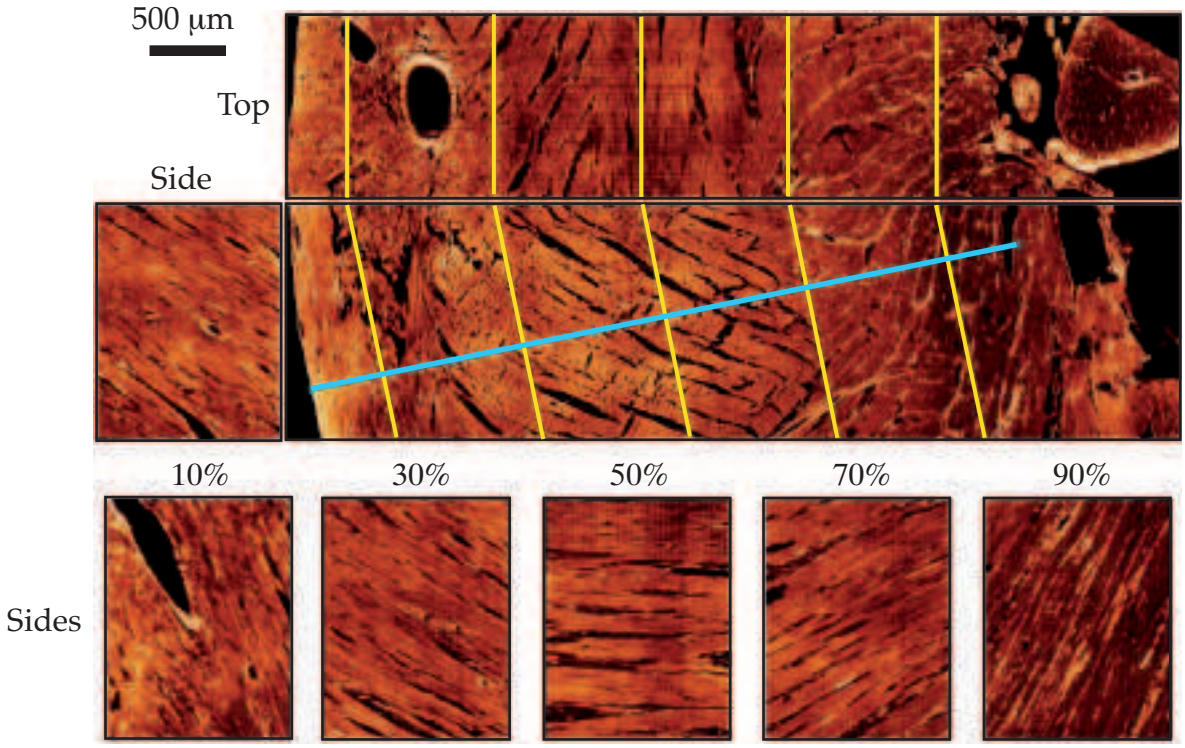


FIGURE 2.3: Exemple d'imagerie de tissu cardiaque de rat. Échantillon transmural complet (haut) et plusieurs coupes à différentes distances de l'épicarde (bas). L'image est tirée du cours Smith et al (2000).

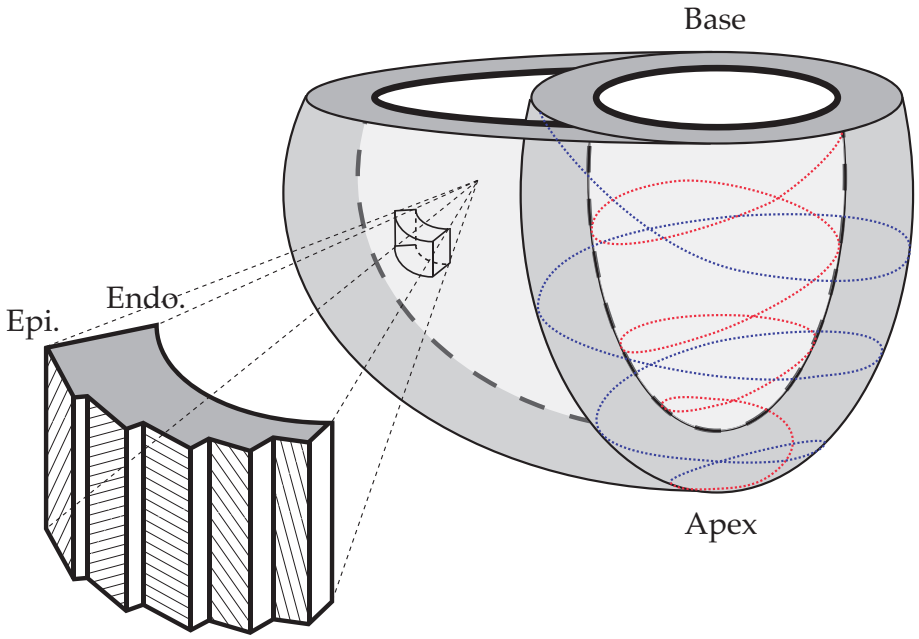


FIGURE 2.4: Schéma de la répartition transmurale des fibres (à gauche) et de l'enroulement dans le ventricule gauche (à droite).

L'orientation du cœur dans l'organisme peut varier sensiblement en fonction de la corpulence des individus, et on définit souvent son grand axe comme le plus petit axe principal d'inertie du ventricule gauche. La partie inférieure des ventricules, c'est-à-dire de plus petite élévation par rapport au grand axe, se nomme l'apex et la supérieure où viennent se greffer les oreillettes est appelée la base. L'American Heart Association, AHA/ACC/SNM (1992), propose de plus de régionaliser les ventricules en dix-sept régions pour le gauche, auxquelles nous en ajoutons onze pour le droit. Malgré d'assez fortes disparités ces régions correspondent sensiblement à des zones d'effet de coronaires principales et donc elles sont raisonnables à la fois du point de vue de la localisation spatiale et de la physiologie.

Au niveau du tissu maintenant, le myocarde est essentiellement constitué d'un enroulement de fibres, un peu à la manière d'une pelote de laine. De nombreuses études anatomiques ont été consacrées à décrire l'organisation géométrique des fibres source d'anisotropie structurelle pour le tissu (voir Mourad (2003) pour une présentation complète). Il a été montré notamment par Streeter (voir notamment Streeter (1979) et ses références) que les fibres courent sur un ensemble de surfaces emboîtées en passant de l'endocarde à l'épicarde via l'apex. Ainsi on distingue trois couches :

- une couche *superficielle ou sous-épicardique* où les fibres sont hélicoïdales d'angle compris entre -60 et -80 degrés ;
- une couche *moyenne* qui n'est présente que dans le ventricule gauche où les fibres sont orthoradiales par rapport au grand axe ;
- une couche *profonde ou sous-endocardique* où les fibres sont hélicoïdales d'angle compris entre -60 et -80 degrés.

La séparation entre ces trois couches est en réalité progressive, néanmoins on pourra la conserver conventionnellement pour sous échantillonner la « régionalisation AHA ».

REMARQUE 2.2.1
Plusieurs études (voir Smith et al (2000) et ses références ou Usyk et al (2000)) ont de plus montré l'existence d'une structure en feuillets orthogonaux à la fibre puisque composés en fait de plusieurs d'entre elles regroupées dans la direction radiale. Ainsi le myocarde présenterait non pas une mais deux directions privilégiées.

2.2.2 Cycle cardiaque

Le cycle cardiaque est présenté sur la figure 2.5. L'activité se décompose ainsi en quatre phases se succédant en moins d'une seconde en moyenne (la fréquence cardiaque étant généralement de 60-80/min). On observe ainsi successivement :

1. **la phase de contraction de la *systole*** (phase I, env. 50ms) :
   Les ventricules se contractent sous l'effet de l'onde de dépolarisation (voir le paragraphe 2.2.3). Toutes les valves étant fermées (la contraction est dite isovolumétrique), la pression monte alors très rapidement jusqu'à atteindre la pression dans l'aorte pour le ventricule gauche (resp. l'artère pulmonaire pour le ventricule droit) d'environ 80 mmHg (resp 10mmHg). Les valves sigmoïdes s'ouvrent alors.

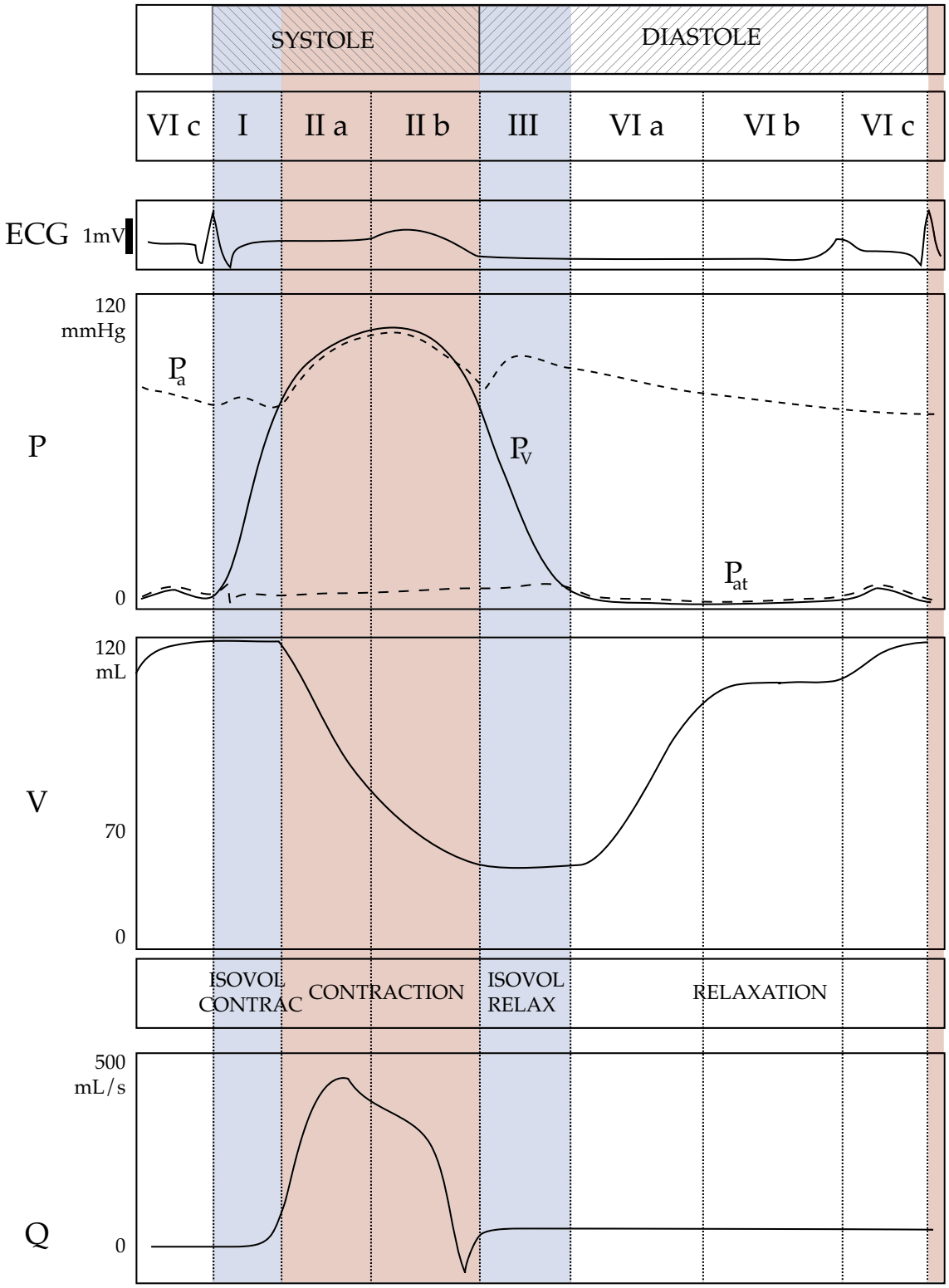


FIGURE 2.5: Principaux indicateurs cardiaques au cours du cycle (ECG, pression, volume, débit dans l'aorte)

2. **la phase d'éjection** (phase II, env. 210ms au repos) :
   La pression dans le ventricule gauche et dans l'aorte continuent de croître jusqu'à la *pression systolique* (phase II env. 120mmHg). En début de phase IIa, la plus grande partie du *volume systolique* V_s est expulsé rapidement et le débit aortique atteint alors son maximum. Par la suite, l'excitation du myocarde cesse et la pression ventriculaire commence à diminuer jusqu'à devenir inférieure à celle de l'aorte (resp. l'artère pulmonaire) entraînant la fermeture des valves. Le volume d'environ 40ml qui reste dans le ventricule gauche est appelé alors *volume télésystolique*.

3. **la phase de relaxation isovolumétrique** (phase III env.60ms) :
   Pendant ce temps, les oreillettes se sont à nouveau remplies. La pression ventriculaire chute alors brusquement tandis que la pression auriculaire augmente jusqu'à ouverture de la valve mitrales à gauche ou pulmonaire à droite.

4. **la phase de remplissage de la *diastole*** (phase IV env. 500ms) :
   Le sang s'écoule alors très rapidement des oreillettes vers les ventricules de sorte que ceux-ci se remplissent à environ 80% en seulement le quart de la durée de la diastole (phase de remplissage rapide (IVa)). Puis le remplissage se ralentit et les oreillettes se contractent permettant la fin du remplissage (soit pour environ 20 % du remplissage total).

Cette description qualitative correspond à une fréquence cardiaque normale au repos ; lorsque le rythme cardiaque augmente, le cycle cardiaque diminue aux dépens de la diastole. De ce fait, c'est alors la pression auriculaire qui participe quantitativement le plus au remplissage des ventricules.

2.2.3 Formation et conduction de l'excitation électrique

Bien que notre problématique soit essentiellement focalisée sur les aspects mécaniques de la contraction, il est fondamental de comprendre la nature de l'excitation électrique du myocarde comme origine de sa contraction.

La genèse de l'excitation siège dans l'organe lui-même, contrairement à ce qui peut se passer au niveau des muscles squelettiques : on parle alors de rythme spontané ou d'autonomie du cœur. Du point de vue électrique, le myocarde, auriculaire et ventriculaire, est un syncitium, c'est-à-dire que les cellules ne sont pas isolées les unes des autres mais reliées par des *gap-junctions*. Ainsi, une excitation naissant dans une des oreillettes se propage jusqu'à contraction complète de celles-ci, et de même pour les ventricules.

L'excitation du cœur naît au niveau du nœud sinusal qui constitue en quelque sorte le pacemaker. La propagation de l'excitation s'étend à partir de ce point sur les deux oreillettes et au *nœud atrioventriculaire* (nœud AV) où elle rejoint le myocarde ventriculaire par les deux branches du faisceau de His et par le réseau de Purkinje. Là, elle parcourt le myocarde de l'intérieur vers l'extérieur et de la pointe (*l'apex*) vers la base (voir Figure 2.6).

REMARQUE 2.2.2
Le cœur peut donc battre de manière autonome, sans innervation extérieure, et sa fréquence est le reflet de l'émission périodique et de la propagation de cette onde d'excitation. Cependant, l'innervation est

nécessaire à l'adaptation de la fonction cardiaque à la demande variable de l'organisme sur des échelles de temps allant de la dizaine de battements à plusieurs minutes. Dans ce cas, les trois caractéristiques suivantes peuvent être modifiées :

- **la fréquence** de formation des impulsions par le pacemaker, donc la fréquence des battements cardiaques (*effet chronotrope*),
- **la vitesse de conduction de l'excitation**, spécialement dans le nœud AV (*effet dromotrope*),
- **la force de contraction musculaire** (effet *inotrope*).

Ces modifications de l'activité sont provoquées par un certain nombre de neuro-transmetteurs tels que l'acetylcholine ou la noredraline .

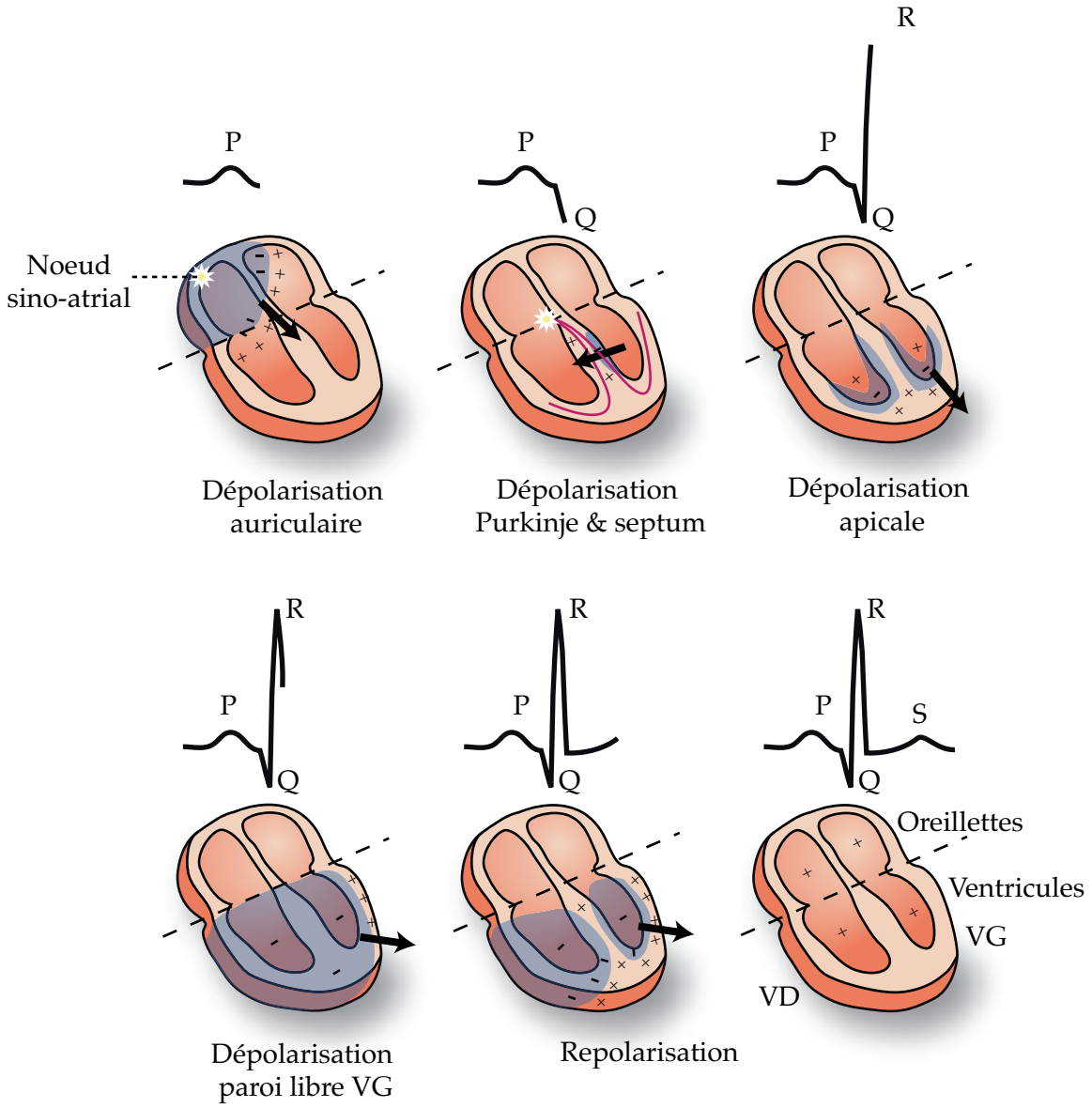


FIGURE 2.6: Propagation de l'excitation électrique cardiaque. Schéma inspiré de Netter (1969)

Le tracé des *électrocardiogrammes* (ECG) renseigne sur les différents instants de la formation et conduction de l'onde électrique en visualisant les différences de potentiels (quelques mV) entre différents points de la surface du thorax résultant de la diffusion de l'activité électrique cardiaque. En d'autres termes, l'ECG correspond au champ électrostatique « lointain » de l'état de la dépolarisation à la surface[1] du cœur. vis-à-vis de l'ECG cette dernière est d'ailleurs équivalente à un dipôle électrique qui varie tout au long du cycle tant en taille qu'en direction en fonction de la propagation. On réalise alors jusqu'à douze mesures de différence de potentiel, appelées dérivations et illustrées Figure 2.7, empiriquement suffisantes au thérapeute. Plus précisément, sur chacun des tracés on retrouve trois déflexions caractéristiques de la dépolarisation

- **l'*onde*** P correspondant à l'activation des oreillettes,
- **le *complexe*** QRS correspondant à celle des ventricules,
- **l'*onde*** T traduisant la repolarisation des cellules ventriculaires,

et c'est l'analyse de ces différents événements sur les dérivations qui permet le diagnostic.

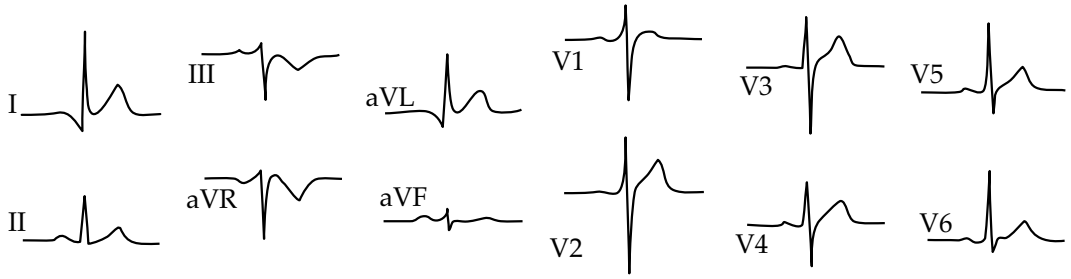


FIGURE 2.7: Illustration des signaux des 12 dérivations ECG

2.2.4 La fonction cardiaque

□ DIAGRAMME PRESSION-VOLUME ET TRAVAIL DU CŒUR – Du point de vue de l'organe *cœur*, la relation pression intra-ventriculaire - volume ventriculaire (conséquence de la relation longueur - étirement musculaire cf Section 2.2.6) exprime le travail de la pompe cardiaque. Dans cet examen, on privilégie le ventricule gauche qui est responsable de l'éjection du sang dans la grande circulation et constitue donc la principale fonction motrice de l'organe. La construction de ce diagramme nécessite tout d'abord 4 courbes d'appuis :

- **la courbe d'étirement au repos :** Elle indique les pressions qui se forment passivement pour divers volumes de remplissage,
- **la courbe des maxima isovolumétriques :** Correspond à la pression maximale obtenue à volume constant,
- **la courbe des maxima isotoniques :** Correspond au volume maximal obtenu à pression constante cette fois,

[1] simplement à la surface pour des raisons de champ équivalent

- **la courbe des maxima auxotoniques :** Durant la systole, la contraction isovolumétrique est suivie d'une phase d'éjection, dite *auxotonique*, où le volume diminue alors que la pression continue d'augmenter. Ce type mixte de contraction porte le nom de contraction à *postcharge*. Pour un volume de remplissage donné, son maximum change, celui-ci étant lié à la pression télédiastolique, mais tous ces maxima sont situés sur une courbe qui relie le point isovolumétrique et le maximum isotonique correspondant à un remplissage donné.

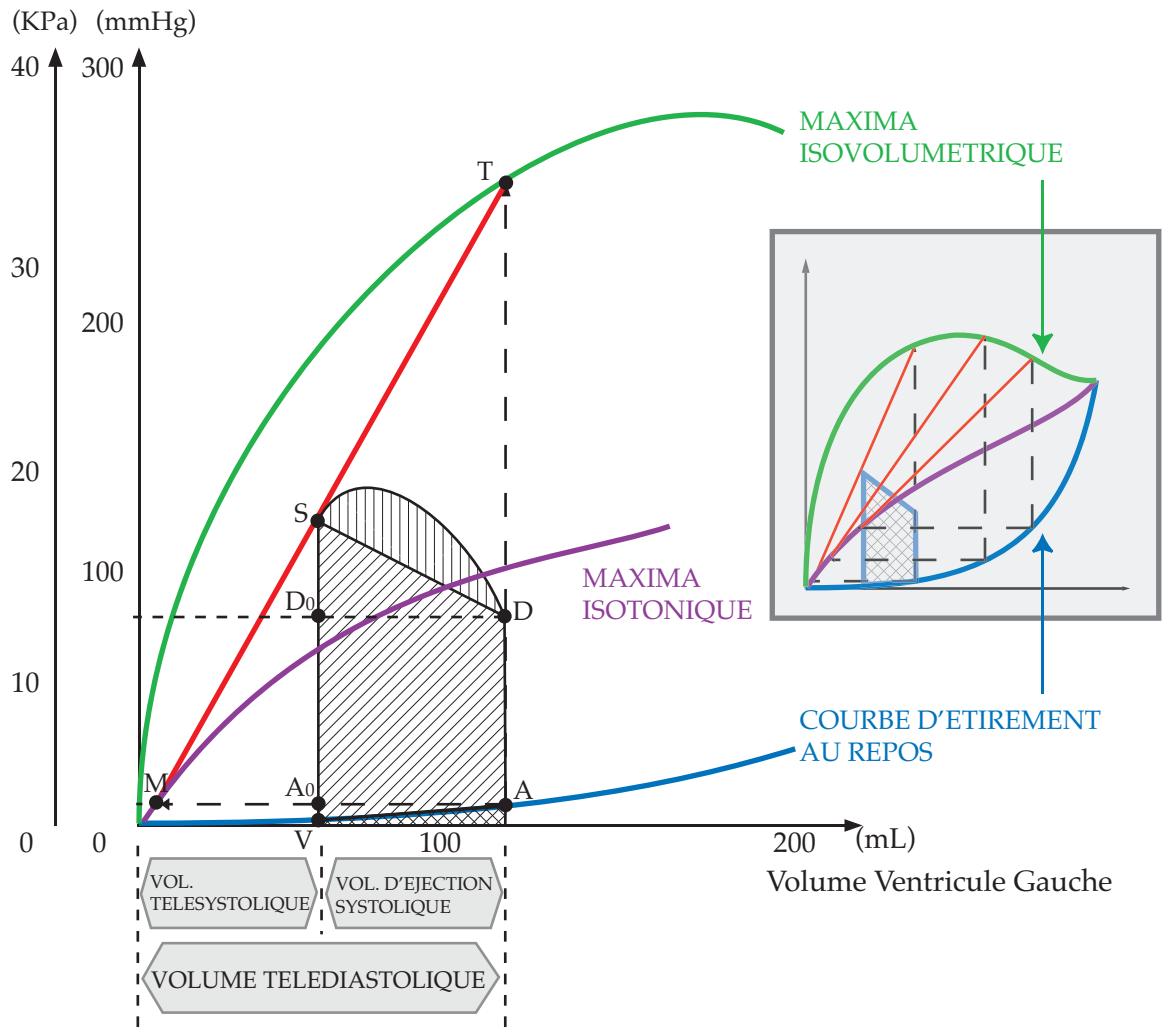


FIGURE 2.8: Diagramme PV de la pompe cardiaque

Le cycle de travail s'inscrit donc entre ces quatre courbes dont les valeurs caractéristiques ont déjà été données Section 2.2.2. Il faut noter que comme tous les cycles « thermodynamiques » de moteur, ce cycle se parcourt dans le sens trigonométrique et que le travail effectué par le ventricule gauche correspond à l'aire de la surface délimitée par le diagramme. Ainsi on obtient typiquement

$$\int P\,dV = 13300 \text{ Pa} \,.\, 80.10^{-6} \text{ m}^3 = 1.07 \text{ J [N.m]},$$

pour le ventricule gauche, et 0.16 J pour le ventricule droit. La plus grande partie du travail global P-V systolique est fournie directement par le myocarde lors de la contraction active ce qui correspond à l'aire $[A–D–D_0–A_0]$ (Figure 2.8). Une petite partie est fournie par la réaction passive et élastique à sa distension lors du remplissage (surface $[A_0–A–V]$ Figure 2.8). En fait ce travail est assuré à 80% indirectement par le myocarde ventriculaire et le reste directement par la contraction auriculaire, ainsi que par les muscles respiratoires et les muscles squelettiques qui participent ensemble au retour veineux.

Cependant, le travail systolique des deux ventricules ne représente pas tout le travail du cœur, car celui-ci doit fournir encore 20% de surplus à l'onde pulsative (distension des parois des vaisseaux), ce qui correspond à la surface $[D\frown S–D_0]$ Figure 2.8. On obtient ainsi comme travail total $W_{\text{TOTAL}} \simeq (1.07+0.16)\,.\,1.2 \simeq 1.45$ soit une puissance au repos, pour une fréquence de 70 pulsations par minute, de $P_W = 1.47 \cdot 1.17^{-1} = 1.71$ W.

□ RÉGULATION DU VOLUME D'ÉJECTION – Un phénomène très important dans la fonction cardiaque est l'ensemble des contrôles qui permettent son adaptation aux changements d'environnement. Or, autant l'adaptation de l'activité cardiaque aux besoins de l'organisme est sous contrôle *extrinsèque*, autant l'adaptation du volume d'éjection (V_{ES}) au volume de remplissage (liées à la position du corps, à la respiration...) et à la pression aortique dépendent directement de l'étirement télédiastolique[2] du myocarde. Ce mécanisme, connu sous le nom de *mécanisme de Frank-Starling*, se traduit par trois phénomènes :

1. Premièrement lorsqu'il y a **augmentation de remplissage** (*précharge* ou *preload* en anglais). Dans ce cas le début de la phase de mise en tension du muscle (point A) est déplacé vers la droite. De fait, le volume télédiastolique augmente de même que le volume d'ejection, ce qui conduit donc à une régulation. Le travail global lors de ce cycle est donc lui aussi plus important.

2. Deuxièmement lorsque **la pression aortique augmente** (*postcharge* ou *afterload*), la valve aortique ne s'ouvre que pour une pression aortique plus élevée. Ainsi pendant une phase transitoire le volume d'éjection est moindre.

3. Enfin, lorsque la modification de l'activité cardiaque est indépendante de la précharge ou de la postcharge, (on parle de **contractilité modifiée** ou *d'inotropie*), c'est que la contractilité est modifiée sous l'effet de l'augmentation de la fréquence cardiaque ou par l'action inotrope de la noredrénaline ou de l'adrénaline. Les maxima isovolumétriques sont alors déplacés vers le haut. Le cœur peut ainsi travailler avec une pression plus élevée et son travail augmente.

Une conséquence directe du mécanisme de Franck-Starling pour la régulation globale du système cardio-vasculaire est l'égalisation des volumes d'éjection dans les deux ventricules évitant ainsi toute stase dans le circuit pulmonaire et systémique.

[2] en fin de diastole

2.2.5 Le réseau vasculaire

a. Description générale

Le réseau vasculaire est constitué d'un ensemble de vaisseaux de diamètres et de propriétés différents comme résumé sur la figure 2.10 et le tableau 2.9. A partir du cœur le cheminement du sang dans la grande circulation est alors le suivant. Expulsé par le ventricule gauche dans l'aorte, le sang est ensuite réparti par les grosses artères dans le réseau périphérique constituée d'artérioles puis de capillaires. Ensuite, le sang « remonte » vers le cœur par les veines, les branches veineuses, les grosses veines et enfin les veines caves. La pression sanguine moyenne passe alors de 100 mmHg dans l'aorte à environ $2-4$ mmHg dans les veines caves. Ceci donne une perte de charge de 97 mmHg qui correspond à une résistance périphérique totale (RPT) de 18 mmHg.min.L^{-1} d'après la relation simple (équivalent de la loi d'Ohm électrique)

$$\Delta P = Q.R_p,$$

où Q est le débit sanguin. D'après la loi de Poiseuille, on sait que la résistance à l'écoulement dans un tube dépend de la longueur du tube, de la viscosité du liquide et de la puissance 4 du rayon sous la forme $R \propto 8\frac{l\eta}{\pi r^4}$, ce qui permet d'imaginer l'importance de chaque réseau de vaisseaux dans la RPT, notamment :

- les petites artères sont responsables de près de 50% de la RPT et la modification du rayon de ces vaisseaux a de fortes conséquences sur la RPT,
- les capillaires, bien qu'ayant un rayon plus petit participent tout de même à 27% de la RPT, car leur nombre est considérable

	Pression (mmHg)	Flux (cm/s)	Surface de la section (cm2)
Aorte		18,0	5,3
Grandes Artères	100	5,0	20
Branches Artérielles	40	1,5	20
Artériolles		0,02-0,2	500
Capillaires	25	1,0	3500
Veines	20		2700
Branches Veines		6,0	100
Grosses Veines			30
Veine Cave	2-4		10

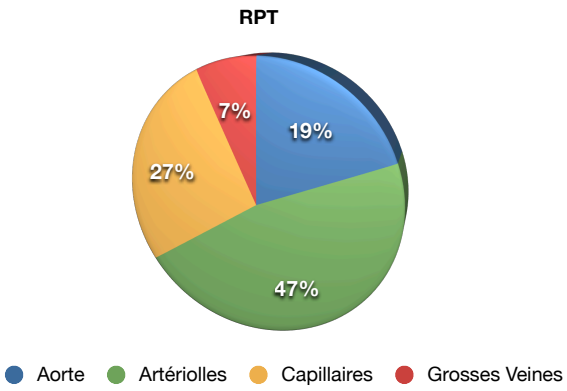


FIGURE 2.9: Pression moyenne, flux, et surfaces pour différentes catégories de vaisseaux.

b. Effet Windkessel

Au niveau des grandes artères, un phénomène vient s'ajouter à la simple perte de charge. Leur paroi se distend durant la phase d'éjection sous l'effet du débit créant une sorte de capacité

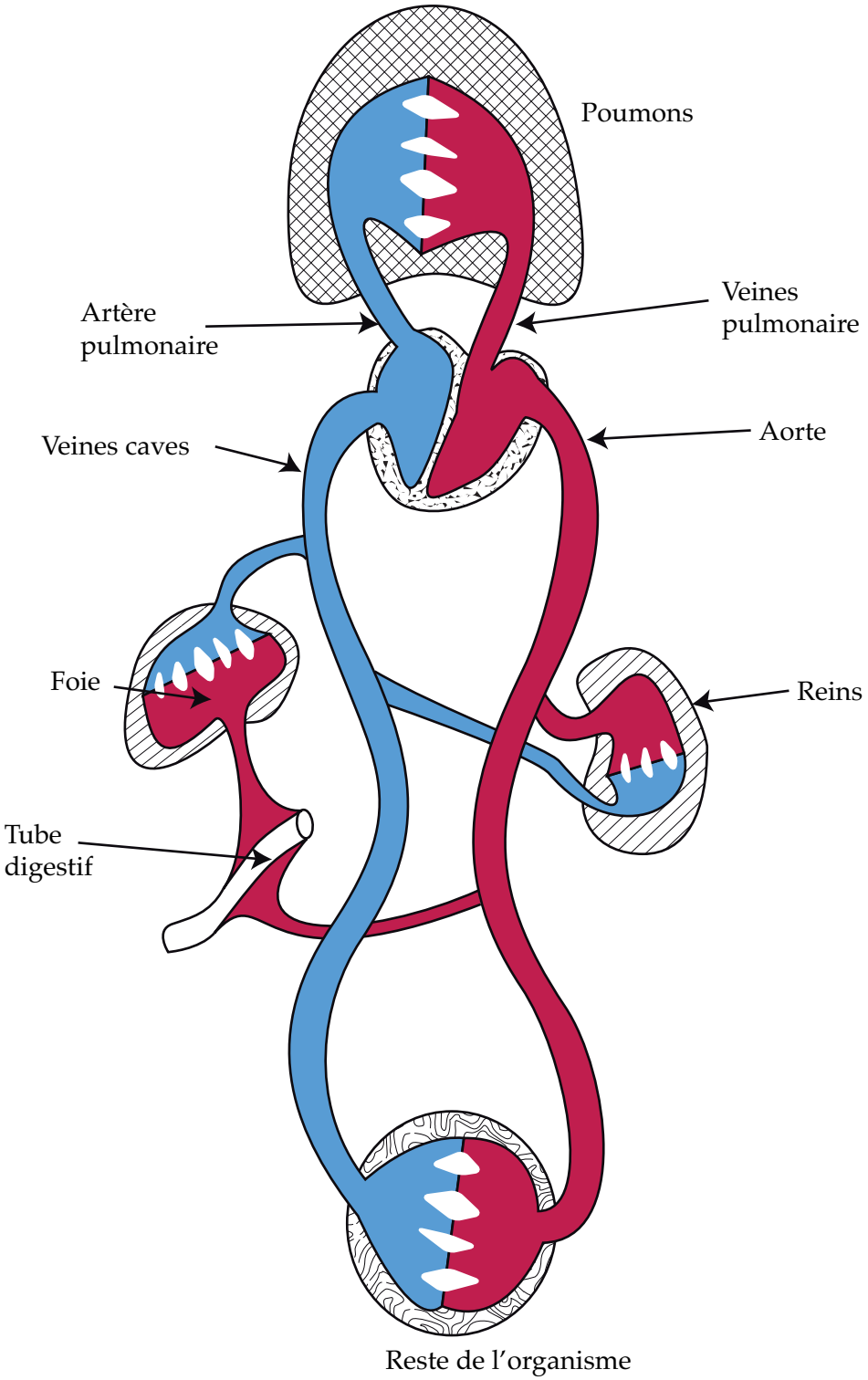


FIGURE 2.10: Réseau artériel et veineux desservi et desservant le cœur

(toujours par analogie électrique) puisqu'une partie du volume éjecté est « temporairement stockée » dans la lumière élargie du vaisseau. Ce phénomène est souvent appelé *effet Windkessel*.

c. Retour veineux

C'est le réseau veineux qui constitue finalement le réservoir à sang de l'organisme, d'autant que les veines sont bien plus extensibles que les artères (phénomène de rétention bien connu). Vu la faible pression régnant les veines, la force de propulsion du flux veineux est due à quatre actions :

- le résidu de pression artérielle maintenue au delà du réseau capillaire,
- le phénomène de succion dû à l'abaissement du plancher valvulaire lors de la contraction cardiaque,
- la pression exercée sur les parois des veines par les muscles squelettiques,
- la respiration qui provoque une hypertension dans l'abdomen et une dépression dans la cage thoracique.

2.2.6 Physiologie détaillé des tissus musculaires cardiaques

Nous venons de présenter rapidement les grandes lignes du fonctionnement du cœur à l'échelle de l'organe. Passons maintenant à une échelle plus petite dite mésoscopique de description de la fibre musculaire cardiaque avant de nous intéresser au comportement microscopique qui régit sa contraction. Le description qui suit s'applique à ce qu'on appelle les muscles striés, c'est-à-dire qu'elle est valide en grande partie à la fois pour les fibres cardiaques et les fibres squelettiques.

Les muscles striés, illustrés Figure 2.11, sont des structures « multi-échelles » dont l'élément principal est la cellule musculaire. Comme nous l'avons déjà mentionné, celle-ci est une fibre dont le diamètre moyen varie en moyenne de 10 à 100 μm et dont la longueur peut atteindre 15 cm pour les muscles squelettiques [3] ! La fibre musculaire est délimitée par une membrane cellulaire appelée *sarcolème* contenant le *sarcoplasme* (cytoplasme), des sarcosomes (mitochondries, cellules associées à la respiration), ainsi qu'une centaine de *myofibrilles*.

A l'échelle de la myofibrille, celle-ci est divisée en compartiments d'environ 2 μm de longueur, appelés sarcomères et limités par des disques Z. Lorsqu'on les observe au microscope, ils apparaissent comme une succession de bandes alternativement claires et sombres d'où leur dénomination de *muscle strié*. Ceci est dû en fait à la disposition de filaments épais de *myosine II*, et de filaments fins *d'actine*. On note, à l'intérieur de deux disques successifs, trois bandes :

- *la bande I* ne comportant que des filaments d'actine, (on en compte quelques milliers par sarcomère)

[3] les fibres de viandes reconnaissables à l'œil nu sont en fait des réseaux de 100-1000 μm formés d'une dizaine de fibres

- *la bande A* dans laquelle les deux types de filaments se chevauchent, et forment en son centre *le disque M*
- *la bande H* ne comportant que des filaments de myosine.

A l'échelle, cette fois, du filament de myosine, on observe que ce dernier est constitué d'un paquet d'environ 300 molécules de myosine II. Chaque molécule possède deux têtes globulaires reliées par *la partie cervicale* (analogie simplement compréhensible d'après la Figure 2.11) à *la partie caudale* filiforme du reste de la molécule. Chaque tête de myosine possède un *domaine moteur* avec une *poche à nucléotide* (ATP) et un *site de liaison à l'actine* et l'ensemble de la partie tête-cou-partie caudale peut faire basculer ces têtes de myosine lors de leur interaction avec l'actine (c'est le glissement des filaments).

Concernant l'actine, molécule globulaire, env. 400 molécules de celle-ci peuvent s'enchaîner à la manière d'un collier de perles pour former un polymère. En fait, ce sont deux molécules de ce types qui, associées à la molécule de tropomyosine, forment le filament d'actine comme illustré Figure 2.11. La molécule de *tropomyosine* qui possède une structure similaire s'enroule donc autour des deux polymères d'actine et tous les 40 nm environ vient se fixer une molécule de *troponine* qui sera responsable de l'interaction avec les têtes de myosine.

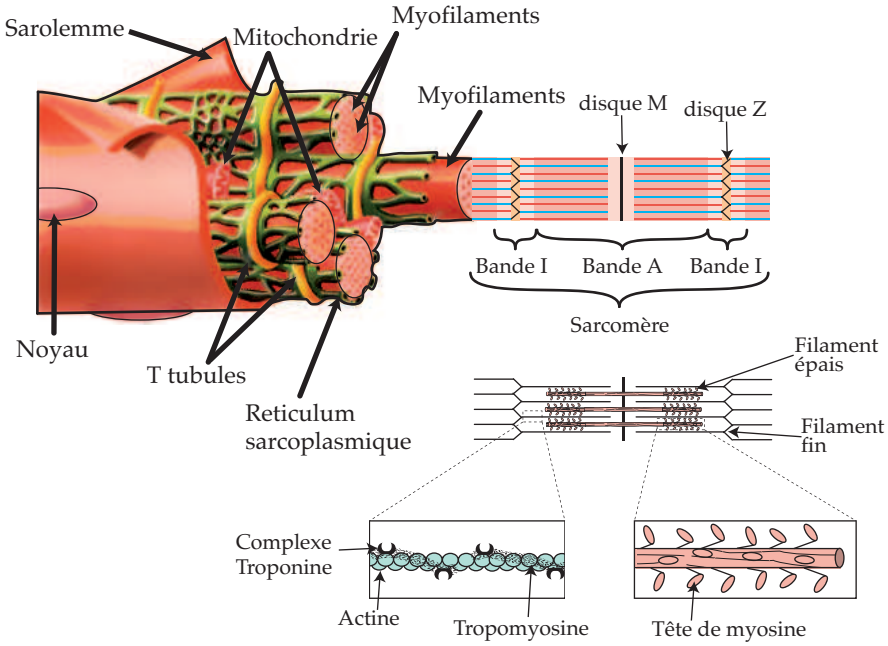


FIGURE 2.11: Description multi-échelles de la fibre musculaire

Le sarcomère comprend encore un système élastique de plus de 100 nm de long, la titine, qui lie les disques Z au disque M.

Enfin, on sait que toute les cellules comportent un *réticulum endoplasmique* qui est normalement responsable, au niveau du *cytoplasme*, des échanges protéolipidiques avec la cellule. Pour les cellules musculaires, ce reticulum est un peu différent et porte le nom de *reticulum sarcomérique*. Il est formé de compartiments fermés, parallèles aux myofibres appelés *tubes longitudinaux* qui sont essentiels dans le stockage des ions Ca^{2+}, cheville ouvrière de la propagation de

l'impulsion électrique dans le tissu.

a. Contraction de la fibre musculaire

□ EXCITATION DE LA FIBRE MUSCULAIRE – Au niveau électrique, la membrane cellulaire constitue une barrière pour les ions porteurs de charges entre le milieu intra- et extra-cellulaire. Néanmoins, certaines espèces et en particulier les ions Na^+, K^+ et Ca^{2+} initialement stockés dans le reticulum endoplasmique peuvent être échangés entre ces deux milieux au travers de pompes et échangeurs chimiques complexes. Les gradients ainsi formés notamment en ions Ca^{2+} entraînent une différence de potentiel transmembranaire qui se propage de cellule en cellule formant une vague qui s'entretient par l'activation des canaux ioniques. On parle alors de potentiel d'action décrit macroscopiquement Section 2.2.3. Lors de la relaxation, la cellule revient à l'équilibre électrique avec en particulier la recaptation des ions calcium dans le reticulum sous l'action d'ATP, carburant de ces mécanismes.

□ LE FILAMENT GLISSANT – Le mécanisme conduisant à la contraction à partir du gradient en ions calcium est connu sous le nom de théorie du filament glissant dont les différentes étapes sont présentées Figure 2.12. La myosine II et les filaments d'actine d'un sarcomère sont en effet disposés de telle manière qu'ils peuvent s'emboîter ensemble. Les têtes de myosine se fixent alors au filament fin en formant un angle donné. Suite à une modification structurale de la molécule de myosine II cet angle se modifie et la tête de myosine entraîne avec elle tout le filament fin dans son glissement. La tête se détache alors, se distend à nouveau, et est prête pour une nouvelle liaison dans le cycle suivant.

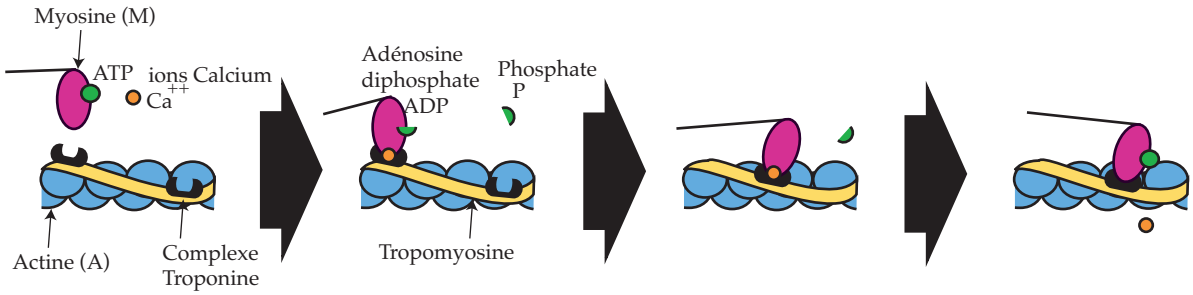


FIGURE 2.12: Etapes de la formation d'un pont actine-myosine

D'un point de vue du filament complet, ce mécanisme est la cause du rapprochement des deux disques Z par la réduction de la zone H. Quand les filaments épais se retrouvent en butée sur les disques Z, le muscle est alors en contraction maximum.

Du point de vue chimique, le mécanisme d'appariement de la tête de myosine puis la modification d'angulation sont des phénomènes complexes. La liaison initiale entre la tête de myosine et l'actine nécessite tout d'abord la fixation d'une molécule d'ATP. Le complexe ATP-myosine réalise alors un angle d'environ 90 degrés avec l'actine mais la liaison est faible. L'arrivée d'ions Ca^{2+} sur le complexe tropomyosine-troponine entraîne, par l'intermédiaire de l'actine, l'ATPase de la myosine, soit l'hydrolyse de la myosine. Il s'ensuit la formation d'un complexe $A-M-ADP-P$. Le détachement du P change alors la conformation du complexe et la liaison actine-myosine devient forte. Les têtes pivotent alors pour former un angle de

40 degrés en entraînant le filament de myosine. La libération de l'ADP amène alors les têtes dans leur position finale (45 degrés) dans une configuration extrêmement stable puisqu'en l'absence d'ATP pour faire remonter les têtes elle correspond à l'état observé pendant la raideur cadavérique. Quand, en revanche, l'ATP vient de nouveau se fixer sur les têtes, celles-ci se redressent. Parallèlement à cela la concentration d'ions calcium diminue pour passer en dessous des $1\,\mu$mol.l^{-1}. C'est alors que les deux molécules d'actine et myosine se détachent en attendant un nouveau potentiel d'action pour recommencer un cycle.

Ainsi le cycle complet de réaction conduisant à l'attachement-détachement de la tête de myosine sur l'actine s'écrit (voir définition des symboles Figure 2.12)

$$\begin{aligned} M.ATP &\xrightarrow{k_1} M.ADP.P \\ M.ADP.P + A &\xrightarrow{k_2} AM.ADP.P \\ AM.ADP.P &\xrightarrow{k_3} AM + ADP + P \\ AM + ATP &\xrightarrow{k_4} M.ADP + A \\ ADP + P &\xrightarrow{k_5} ATP \end{aligned} \tag{2.1}$$

A noter que l'équation 5 est présente uniquement pour fermer le cycle. De plus les ions Ca^{2+} n'apparaissent pas directement dans ce cycle.

Du point de vue un peu plus macroscopique, ce mécanisme au niveau des têtes de myosine n'est pas synchronisé sur tous les filaments donc la contraction s'opère par saccades. De plus, il n'opère pas dans toutes les configurations du tissu. En effet, quand l'étirement est maximum, il n'est plus possible de faire d'autres ponts, et inversement quand la longueur est trop faible, les ponts se recouvrent les uns les autres et ne contribuent plus. Ainsi la tension maximale est obtenue pour une longueur de repos « moyenne ».

b. Propriétés mécaniques du muscle

Le modèle du filament glissant définit le comportement mécanique actif de contraction cardiaque en l'associant directement au nombre de ponts actine - myosine formés. Cependant, le comportement mécanique des fibres musculaires ne se limite pas au sarcomère et s'exprime aussi de façon passive (indépendamment de l'activité électrique), notamment dans les phases de relaxation du tissu (lors de la repolarisation du tissu).

Du point de vue passif donc, un muscle au repos peut être étiré comme un élastique et la force résultante au début est très faible. Elle croit ensuite de manière exponentielle avec l'allongement : on obtient alors la courbe d'étirement de repos. Lors d'un allongement très important, le tissu de soutien résiste et empêche le démantèlement des fibres. Mais c'est avant tout la *titine* qui assure ce rôle. Cette molécule se situe à deux endroits. Au niveau de la bande A pour positionner le filament de myosine mais aussi sur la bande I en série avec les autres composantes.

Ces deux comportements, actifs et passifs, permettent de tracer les relations *Tension – Longueur* $T - L$ sur un diagramme, la contrainte totale pouvant être développée dans la fibre musculaire devenant la somme des deux contributions actives et passives. En régime mus-

culaire « normal », les différentes contractions possibles vont en fait s'inscrire à l'intérieur de ces courbes maximales avec en particulier deux contractions caractéristiques :

- la contraction *isométrique* s'opérant à longueur constante (correspondant à un déplacement vertical sur le diagramme $T-L$). Ce type de contraction n'est rendu possible que par la *composante élastique en série* correspondant aux bandes Z qui compense la modification de longueur provoquée par l'attachement des ponts.

- la contraction *isotonique* à pression constante (correspondant à un déplacement horizontal sur le diagramme $T-L$) lorsque la charge sur la fibre est nettement supérieure à ce qu'elle peut équilibrer.

Ces deux types de contractions sont, dans le cadre de la fibre cardiaque, corrélées respectivement au comportement isovolumique (puisque le myocarde correspond à une pelote de fibres enroulée sur elle-même), ou à la contraction isobare du myocarde complet. Ainsi, le diagramme (P, V) du cœur complet est similaire à celui du diagramme (T, L) de la fibre cardiaque (similarité des Figures 2.8 et 2.13). Mécaniquement, il y a donc une analogie possible entre le muscle tout entier et le comportement moyen de la fibre. Dans les autres situations, la contraction du muscle est dite *auxotonique* et correspond à une tension et une longueur variables.

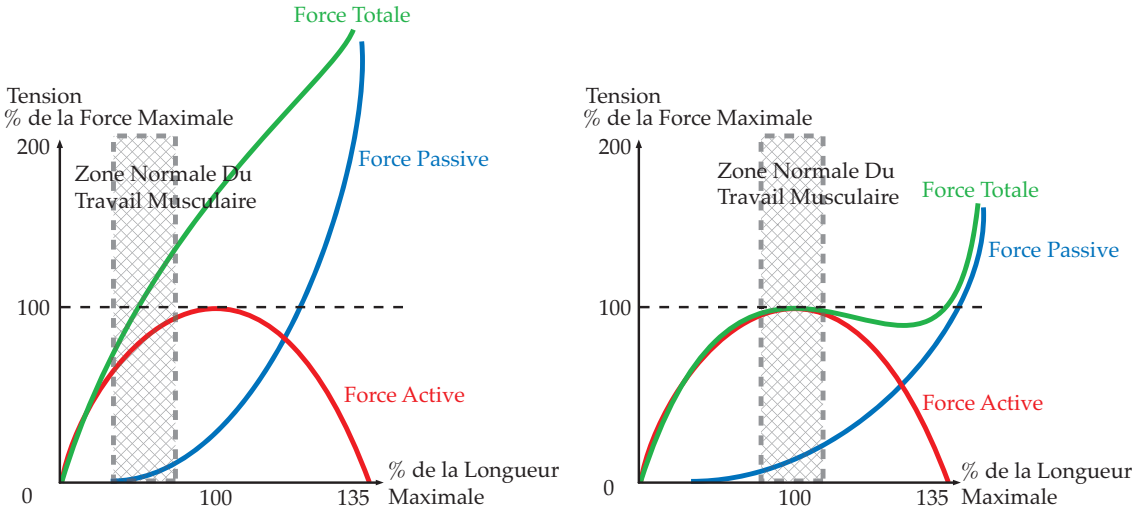


FIGURE 2.13: Diagramme Tension-Longueur pour le muscle cardiaque (à gauche) et à titre de comparaison pour le muscle squelettique (à droite).

REMARQUE 2.2.3
Il est intéressant de constater que, bien que très similaires dans leur principes, les muscles striés présentent des différences fonctionnelles essentielles suivant qu'ils sont squelettiques ou cardiaques.

- le muscle squelettique est plus extensible que le muscle cardiaque.

- le point de fonctionnement est différent pour les deux types de muscles et situé au niveau du maximum de la force active pour le muscle squelettique. Ceci laisse une zone d'activité supplémentaire pour le cœur, modifiable par la concentration en ions $[Ca^{2+}]$ et observable à travers l'effet Frank-Starling.

- une différence dans l'excitation puisque le plateau est plus long dans le muscle cardiaque et la contraction dépend de la longueur de ce plateau. De plus l'excitation s'étend à tout le cœur dès qu'elle a commencé.

2.3 Modélisation électromécanique cardiaque

Nous venons de parcourir l'essentiel de la physiologie cardiaque nécessaire à la formulation d'un modèle mécanique pour le cœur. Pour ce faire, nous commençons par quelques rappels indispensables de mécanique des milieux continus avant de présenter la loi de comportement utilisée pour la formulation de notre problème direct.

2.3.1 Formulation mécanique en Lagrangien total

a. Cinématique

On considère un solide déformable représenté à tout instant t par un ouvert $\Omega(t)$ (noté Ω quand il n y a pas d'ambiguïté). De plus, on suppose que $\partial\Omega(t)$ existe et est C^1 par morceaux (en particulier une normale est définie en tout point de $\partial\Omega(t)$). En configuration Lagrangienne, on se donne une configuration de référence $(\Omega_0, \partial\Omega_0)$ (pas nécessairement égale à $(\Omega(0), \partial\Omega(0))$) telle que la déformation est une application bijective $\underline{\phi}$ de la configuration de référence sur la configuration courante. La déformation donne la position de chaque point du solide à tout instant :

$$\underline{\phi} : \begin{cases} \overline{\Omega}_0 & \rightarrow & \mathbb{R}^3 \\ \underline{\xi} & \mapsto & \underline{\mathrm{x}} = \underline{\phi}(\underline{\xi}, t) \end{cases}$$

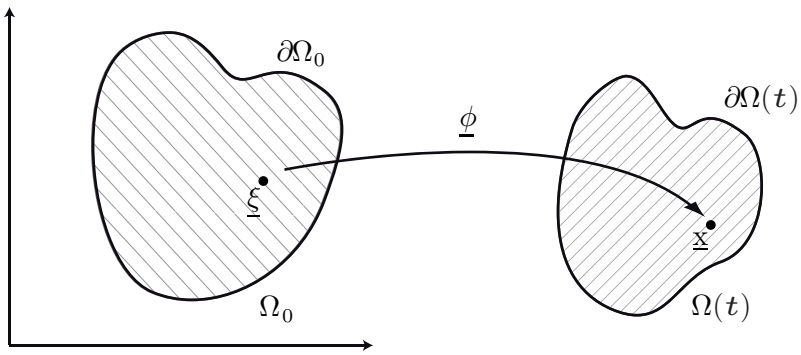


FIGURE 2.14: Configuration de référence, déformée et transformation

On note $\underline{y}$ le déplacement

$$\underline{y}(\underline{\xi}, t) = \underline{\mathrm{x}} - \underline{\xi} = \underline{\phi}(\underline{\xi}, t) - \underline{\xi},$$

$\underline{\underline{F}}$ est le gradient de transformation

$$\underline{\underline{F}}(\xi, t) = \underline{\underline{\nabla}}_{\underline{\xi}} \underline{\phi} = \underline{\underline{\mathbb{1}}} + \underline{\underline{\nabla}}_{\underline{\xi}} \underline{y},$$

tel que toute modification de volume est obtenue par $J d\Omega$ où $J = \det \underline{\underline{F}}$ et $d\Omega$ désigne la mesure de volume (ici en configuration de référence). De même, toute modification d'aire est donnée

par $JF^{-T}\underline{dS}$. De plus, $\underline{\underline{C}} = \underline{\underline{F}}^T\underline{\underline{F}}$ est le tenseur de déformation à droite de Cauchy-Green, ou tenseur de dilatation puisque

$$\forall \underline{\delta\xi}^1, \underline{\delta\xi}^2, \quad < \underline{\delta x}^1, \underline{\delta x}^2 > = < \underline{\underline{F}} \cdot \underline{\delta x}^1, \underline{\underline{F}} \cdot \underline{\delta x}^2 > = < \underline{\delta\xi}^1, \underline{\delta\xi}^2 >_{\underline{\underline{C}}} .$$

Le tenseur de Green-Lagrange est noté $\underline{\underline{e}}$,

$$\underline{\underline{e}} = \frac{1}{2}(\underline{\underline{C}} - \underline{\underline{\mathbb{1}}}) = \frac{1}{2}\Big(\underline{\underline{\nabla}}_{\underline{\xi}}\underline{y} + (\underline{\underline{\nabla}}_{\underline{\xi}}\underline{y})^T + (\underline{\underline{\nabla}}_{\underline{\xi}}\underline{y})^T \cdot \underline{\underline{\nabla}}_{\underline{\xi}}\underline{y}\Big),$$

et son linéarisé $\underline{\underline{\varepsilon}}$

$$\underline{\underline{\varepsilon}} = \frac{1}{2}\Big(\underline{\underline{\nabla}}_{\underline{\xi}}\underline{y} + (\underline{\underline{\nabla}}_{\underline{\xi}}\underline{y})^T\Big).$$

b. Principe fondamental de la dynamique

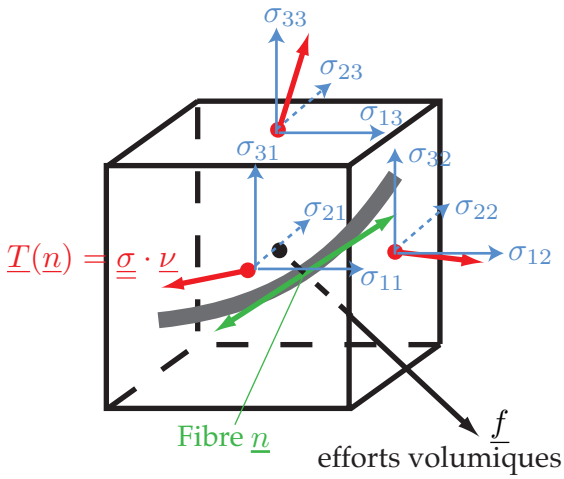


FIGURE 2.15: Cube de Cauchy pour le tissu cardiaque. Sont représentées les composantes du tenseur de Cauchy $\underline{\underline{\sigma}}$ et les efforts résultant sur chaque surface $\underline{T}$. La fibre apparaîtra dans la loi de comportement Section 2.3.2

.

En configuration déformée, le principe fondamental de la dynamique liant contraintes et forces extérieures se retrouve à partir du cube de Cauchy Figure 2.15,

$$\operatorname{div} \underline{\underline{\sigma}} + \rho(\underline{f} - \underline{\ddot{y}}) = 0, \text{ sur } \Omega$$

où $\underline{\underline{\sigma}}$ est le tenseur de contraintes de Cauchy, ρ la masse volumique du solide et $\underline{f}$ la résultante des efforts. En formulation faible sur l'espace V des déplacements admissibles, on obtient

$$\forall \underline{v} \in V(\Omega), \quad \int_\Omega \underline{\underline{\sigma}} : \nabla_{\underline{x}}\underline{v}\, d\Omega = \int_\Omega \rho(\underline{f} - \underline{\ddot{y}}) \cdot \underline{v}\, d\Omega.$$

Le tenseur de Cauchy étant symétrique, on peut symétriser aussi $\nabla_{\underline{x}}$ et ainsi faire apparaître $\underline{\underline{\varepsilon}}(\underline{v}) = \frac{1}{2}\big(\underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} + (\underline{\underline{\nabla}}_{\underline{\xi}}\underline{v})^T\big)$ le tenseur de déformation linéarisé en configuration déformée

$$\forall \underline{v} \in V(\Omega), \quad \int_\Omega \underline{\underline{\sigma}} : \underline{\underline{\varepsilon}}(\underline{v})\, d\Omega = \int_\Omega \rho(\underline{f} - \underline{\ddot{y}}) \cdot \underline{v}\, d\Omega,$$

Pour passer en Lagrangien total on effectue le changement de variable $\underline{x} \to \underline{\xi}$. Comme

$$\underline{\underline{\nabla}}_{\underline{x}}\underline{v} = \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot \underline{\underline{\nabla}}_{\underline{x}}\underline{\xi} = \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot (\underline{\underline{\nabla}}_{\underline{\xi}}\underline{x})^{-1} = \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot \underline{\underline{F}}^{-1},$$

on obtient

$$\forall \underline{v} \in V(\Omega_0), \quad \int_{\Omega_0} \underline{\underline{\sigma}} : \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot \underline{\underline{F}}^{-1} J \, d\Omega = \int_{\Omega_0} \rho_0(\underline{f} - \underline{\ddot{y}}) \cdot \underline{v} \, d\Omega.$$

En toute rigueur, seules les forces extérieures f dites conservatives (i.e. ne dépendant pas de la configuration) peuvent s'écrire simplement sur la configuration de référence ; les autres, comme par exemple les efforts de pression, sont à composer avec la déformation, mais nous ne le préciserons pas dans la formulation pour alléger l'écriture. En introduisant $\underline{\underline{T}} = J\underline{\underline{\sigma}} \cdot \underline{\underline{F}}^{-T}$ le *premier tenseur de Piola-Kirchhoff* (non symétrique) on obtient donc

$$\forall \underline{v} \in V(\Omega_0), \quad \int_{\Omega_0} \underline{\underline{T}} : \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot \, d\Omega = \int_{\Omega_0} \rho_0(\underline{f} - \underline{\ddot{y}}) \cdot \underline{v} \, d\Omega,$$

qui donne en formulation forte

$$\operatorname{div} \underline{\underline{T}} + \rho_0(\underline{f} - \underline{\ddot{y}}) = 0, \text{ sur } \Omega_0.$$

Pour conserver des tenseurs symétriques, on utilise le *second tenseur de Piola-Kirchhoff*

$$\underline{\underline{\Sigma}} = \underline{\underline{F}}^{-1} \cdot \underline{\underline{T}} = J\underline{\underline{F}}^{-1} \cdot \underline{\underline{\sigma}} \cdot \underline{\underline{F}}^{-T},$$

la formulation faible des équations du mouvement devenant

$$\forall \underline{v} \in V(\Omega_0), \quad \int_{\Omega_0} \underline{\underline{\Sigma}} : \underline{\underline{F}}^T \cdot \underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} \cdot \, d\Omega = \int_{\Omega_0} \rho_0(\underline{f} - \underline{\ddot{y}}) \cdot \underline{v} \, d\Omega.$$

Or $\underline{\underline{\nabla}}_{\underline{\xi}}\underline{v} = \mathrm{d}_{\underline{y}}\underline{\underline{F}} \cdot \underline{v}$, donc le symétrisé de $\underline{\underline{F}}^T \cdot \underline{\underline{\nabla}}_{\underline{\xi}}$ n'est autre que la différentielle du tenseur de déformation par rapport au déplacement, soit

$$\mathrm{d}_{\underline{y}}\underline{\underline{e}} \cdot \underline{v} = \frac{1}{2}\big((\mathrm{d}_{\underline{y}}\underline{\underline{F}} \cdot \underline{v})^T \cdot \underline{\underline{F}} + \underline{\underline{F}}^T \cdot \mathrm{d}_{\underline{y}}\underline{\underline{F}} \cdot \underline{v}\big),$$

conduisant à l'expression du principe de la dynamique en configuration de référence que nous utiliserons par la suite

$$\forall \underline{v} \in V(\Omega_0), \quad \int_{\Omega_0} \rho_0 \underline{\ddot{y}} \cdot \underline{v} \, d\Omega + \int_{\Omega_0} \underline{\underline{\Sigma}} : d_{\underline{y}}\underline{\underline{e}} \cdot \underline{v} \, d\Omega = \int_{\Omega_0} \rho_0 \underline{f} \cdot \underline{v} \, d\Omega. \tag{2.2}$$

A noter qu'on utilise souvent en mécanique la notation compactée $\underline{\underline{\delta e}} = d_{\underline{y}}\underline{\underline{e}} \cdot \underline{v}$.

REMARQUE 2.3.1
Les bilans énergétiques en mécanique sont obtenus classiquement en testant les formulations variationnelles sur les vitesses $\underline{v} = \underline{\dot{y}}$. Ainsi le tenseur de Cauchy $\underline{\underline{\sigma}}$ est le conjugué énergétique du tenseur de déformation linéarisée puisque $\underline{\underline{\varepsilon}}(\underline{\dot{y}}) = \underline{\underline{\dot{\varepsilon}}}$. De même, le premier tenseur de Piola-Kirchhoff $\underline{\underline{T}}$ est le conjugué énergétique du gradient de déformation $\underline{\underline{F}}$ car $\underline{\underline{\nabla}}_{\underline{\xi}}\underline{\dot{y}} = \underline{\underline{\dot{F}}}$ et le second tenseur de Piola-Kirchhoff $\underline{\underline{\Sigma}}$ est celui du tenseur de déformation Green-Lagrange $\underline{\underline{e}}$ car $d_{\underline{y}}\underline{\underline{e}} \cdot \underline{\dot{y}} = \underline{\underline{\dot{e}}}$.

2.3.2 Loi de comportement mécanique

Commençons par rappeler une décomposition classique du tenseur des contraintes en fonction du travail développé sur les déformations de type dilatation ou contraction isotrope. Soit

$$p = -\frac{1}{3J}\underline{\underline{\Sigma}} : \underline{\underline{C}} \tag{2.3}$$

et

$$\underline{\underline{\Sigma}}_d = \underline{\underline{\Sigma}} + pJ\underline{\underline{C}}^{-1} \tag{2.4}$$

Alors pour tout déplacement virtuel $\underline{v} = \lambda\underline{\mathrm{x}}$ où $\mathrm{d}_{\underline{y}}\underline{\underline{F}} \cdot \underline{v} = \lambda\underline{\underline{F}}$ et donc $\mathrm{d}_{\underline{y}}\underline{\underline{e}} \cdot \underline{v} = \lambda\underline{\underline{C}}$

$$\underline{\underline{\Sigma}}_d : \mathrm{d}_{\underline{y}}\underline{\underline{e}} \cdot \underline{v} = \lambda\underline{\underline{\Sigma}} : \underline{\underline{C}} + \lambda pJ\underline{\underline{C}}^{-1} : \underline{\underline{C}} = 0.$$

Donc $\underline{\underline{\Sigma}}_d$, appelée *partie déviatorique* du tenseur des contraintes, ne travaille pas sur les dilatations ou contractions isotropes.

De plus, la variable p s'interprète comme *la pression hydrostatique* du système. En effet, si le système est en équilibre avec une pression extérieure p_e,

$$\int_{\Omega_0} \underline{\underline{\Sigma}} : d_{\underline{y}}\underline{\underline{e}}\, d\Omega = \int_{\partial\Omega} -p_e\underline{\nu} \cdot \underline{v}\, d\Omega.$$

et en introduisant le volume

$$|\Omega| = \int_{\Omega_0} J\, d\Omega,$$

dont la différentielle est

$$d_{\underline{y}}\,|\Omega| \cdot \underline{v} = \int_{\Omega_0} J\underline{\underline{F}}^{-1} : d_{\underline{y}}\underline{\underline{F}} \cdot \underline{v}\, d\Omega = \int_{\Omega} \mathrm{div}_{\mathrm{x}}(\underline{v})\, d\Omega, \tag{2.5}$$

on obtient pour tout déplacement virtuel $\underline{v} = \lambda\underline{\mathrm{x}}$

$$p_e \int_{\partial\Omega} \underline{\nu} \cdot \underline{v}\, d\Omega = p_e d_{\underline{y}}\,|\Omega| \cdot \underline{v} = 3p\lambda\,|\Omega| \Rightarrow p = p_e.$$

a. Hyperélasticité

On définit un matériau hyperélastique comme un corps élastique homogène ne dissipant pas d'énergie pendant un cycle (voir Le Tallec (1994)). Ainsi pour toute transformation $\underline{\phi}(\underline{\xi}, t) = \underline{\underline{F}}(t).\underline{\xi} + c(t)$ périodique de période T,

$$\mathcal{W}_{\text{cycle}} = \int_0^T \int_{\Omega_0} \underline{\underline{T}}(\underline{\underline{F}}) : \dot{\underline{\underline{F}}}\, d\Omega dt = \int_{\underline{\underline{F}}(0)}^{\underline{\underline{F}}(T)=\underline{\underline{F}}(0)} T(\underline{\underline{F}}) : d\underline{\underline{F}} = 0 \tag{2.6}$$

D'après le théorème de Stokes, il existe donc une fonction W^e tel que $\underline{\underline{T}}(\underline{\underline{F}}) = \frac{\partial W^e}{\partial \underline{\underline{F}}}$. Ceci correspond physiquement au cas particulier de l'inégalité de Clausius-Duhem en l'absence de dissipations Truesdell et Noll (1965).

On peut aller plus loin dans la formulation de W^e puisqu'une rotation de corps rigide n'entraîne pas de modification énergétique. Mathématiquement on a donc,

$$\forall \underline{\underline{Q}} \in SO_3(\mathbb{R}), \quad W^e(\underline{\underline{F}}) = W^e(\underline{\underline{Q}} \cdot \underline{\underline{F}}). \tag{2.7}$$

En choisissant $\underline{\underline{Q}} = \underline{\underline{Q}} \cdot \sqrt{\underline{\underline{C}}} \cdot \underline{\underline{F}}^{-1}$, on obtient que W^e est en en fait une fonction de $\underline{\underline{C}}$, qui filtre toutes les rotations dans $\underline{\underline{F}}$. En outre, dans le cas isotrope, l'indépendance par rotation de la configuration initiale impose que

$$\forall \underline{\underline{Q}} \in SO_3(\mathbb{R}), \quad W^e(\underline{\underline{C}}) = W^e(\underline{\underline{Q}} \cdot \underline{\underline{C}} \cdot \underline{\underline{Q}}^T). \tag{2.8}$$

En choisissant pour $\underline{\underline{Q}}$ la matrice des vecteurs propres du tenseur (symétrique) de Cauchy-Green, on obtient finalement que W^e est une fonction uniquement des valeurs propres du tenseur. Autrement dit, elle s'exprime uniquement en fonction de ses invariants

$$I_1 = \mathrm{tr}(\underline{\underline{C}}), \quad I_2 = \tfrac{1}{2}(\mathrm{tr}(\underline{\underline{C}})^2 - \mathrm{tr}(\underline{\underline{C}}^2)) \quad I_3 = \det(\underline{\underline{C}}) = J^2 \tag{2.9}$$

En second tenseur de Piola-Kirchhoff cette relation se traduit en fonction de son conjugué énergétique sous la forme

$$\text{Il existe } W^e = W^e(I_1, I_2, I_3) \text{ tel que } \underline{\underline{\Sigma}}(\underline{\underline{e}}) = \frac{\partial W^e}{\partial \underline{\underline{e}}}(I_1, I_2, I_3). \tag{2.10}$$

En pratique, le tenseur des contraintes est donc obtenu directement à partir de la dérivée de W^e par rapport aux invariants

$$\underline{\underline{\Sigma}} = 2 \sum_{1 \le i \le 3} \frac{\partial W^e}{\partial I_i} \cdot \frac{\partial I_i}{\partial \underline{\underline{C}}},$$

avec les identités classiques

$$\frac{\partial I_1}{\partial \underline{\underline{C}}} = \mathbb{1}, \quad \frac{\partial I_2}{\partial \underline{\underline{C}}} = (I_1 \mathbb{1} - \underline{\underline{C}}), \quad \frac{\partial I_3}{\partial \underline{\underline{C}}} = I_3 \underline{\underline{C}}^{-1}. \tag{2.11}$$

Du point de vue énergétique, on remarque alors que le travail des forces intérieures $\mathcal{W}^{int}$

$$\mathcal{W}^{int}(t) = \int_0^t \int_{\Omega_0} \frac{\partial W^e}{\partial \underline{\underline{e}}} : d_y \underline{\underline{e}} \cdot \dot{y} \, d\Omega \, dt = \int_0^t \frac{\mathrm{d}}{\mathrm{dt}} \int_{\Omega_0} W^e(\underline{\underline{e}}) \, d\Omega \, dt = \mathcal{E}_m(t) - \mathcal{E}_m(0),$$

où $\mathcal{E}_m(t)$ devient l'énergie élastique du système dont le principe des travaux virtuels dérive directement en statique

$$\min_{\underline{y}} (\mathcal{E}_m - \mathcal{W}^{\text{ext}}).$$

□ EXEMPLES – Il existe de nombreux exemples de lois hyperélastiques, par exemple les énergie de type « néo-Hookéen » $W^e = \kappa_1(I_1 - 3)$ ou « Mooney-Rivlin » $W^e = \kappa_1(I_1 - 3) + \kappa_2(I_2 - 3)$. Une énergie plus riche et que nous utiliserons par la suite fut proposée par Ciarlet et Geymonat (1982)

$$W^e(\underline{\underline{e}}) = \kappa_1(I_1 - 3) + \kappa_2(I_2 - 3) + a(I_3 - 1) - (\kappa_1 + 2\kappa_2 + a)\ln(I_3). \tag{2.12}$$

La loi la plus simple dans cette classe correspond au matériau de Saint-Venant-Kirchhoff caractérisée par l'énergie

$$W^e(\underline{\underline{e}}) = \frac{\lambda}{2}(\mathrm{tr}(\underline{\underline{e}}))^2 + \mu \, \mathrm{tr}(\underline{\underline{e}}^2), \tag{2.13}$$

qui elle aussi peut se réécrire en fonction des invariants. Cependant, à la différence de l'énergie proposée par Ciarlet-Geymonat, elle est infiniment compressible à énergie finie et ne vérifie

pas les hypothèses de polyconvexité utilisées dans les théorèmes d'existence de solution (voir Ciarlet (1988)). En revanche, cette énergie reste utilisée en pratique car elle est l'extension non-linéaire directe de la loi de Hooke pour les matériaux élastiques. En fait tous les matériaux hyperélastiques isotropes se ramènent à cette loi de Hooke dans le cas des petites déformations, où l'on a

$$\Sigma(\underline{\underline{\varepsilon}}) = \lambda \operatorname{tr}(\underline{\underline{\varepsilon}}) Id + 2\mu \underline{\underline{\varepsilon}}, \tag{2.14}$$

avec λ,μ les coefficients de Lamé équivalents.

□ INVARIANTS RÉDUITS – On définit souvent des invariants réduits qui se prêtent plus naturellement à la décomposition pression-tenseur déviatorique. A cette fin, on considère de nouveau les déformations virtuelles de type contraction-dilatation isotrope $\underline{v} = \lambda \underline{\mathrm{x}}$, $\mathrm{d}_{\underline{y}} \underline{\underline{e}} \cdot \underline{v} = \lambda \underline{\underline{C}}$. Or

$$\frac{\partial I_1}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = I_1, \quad \frac{\partial I_2}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = 2I_2, \quad \frac{\partial I_3}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = 3I_3,$$

si bien qu'en choisissant

$$J_1 = I_1 I_3^{-\frac{1}{3}}, \quad J_2 = I_2 I_3^{-\frac{2}{3}}, \quad J_3 = I_3^{\frac{1}{2}} = J. \tag{2.15}$$

ceux-ci vérifient

$$\frac{\partial J_1}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = 0, \quad \frac{\partial J_2}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = 0, \quad \frac{\partial J_3}{\partial \underline{\underline{C}}} : \underline{\underline{C}} = \frac{3}{2} J.$$

si bien que

$$\underline{\underline{\Sigma}}_d = 2 \sum_{i=1,2} \frac{\partial W^e}{\partial J_i} \cdot \frac{\partial J_i}{\partial \underline{\underline{C}}}, \quad p = -\frac{\partial W^e}{\partial J},$$

avec les identités

$$\frac{\partial J_1}{\partial \underline{\underline{C}}} = I_3^{-\frac{1}{3}} (\mathbb{1} - \frac{1}{3} I_1 \underline{\underline{C}}^{-1}), \ \frac{\partial I_2}{\partial \underline{\underline{C}}} = I_3^{-\frac{2}{3}} (I_1 \mathbb{1} - \underline{\underline{C}} - \frac{2}{3} I_2 \underline{\underline{C}}^{-1}), \ \frac{\partial I_3}{\partial \underline{\underline{C}}} = \frac{1}{2} I_3^{\frac{1}{2}} \underline{\underline{C}}^{-1}. \tag{2.16}$$

□ INCOMPRESSIBILITÉ – Pour les matériaux incompressibles, toute déformation préserve le volume donc les seuls transformations admissibles sont telles que $J = \det F = 1$. On remarque que, dans ce cas, la démonstration de la dépendance de l'énergie en fonction des invariants reste valable mais n'est plus qu'une fonction des deux premiers invariants $W^e = W^e(J_1, J_2)$ puisque $J_3 = 1$. De plus, la contrainte cinématique introduit un multiplicateur de Lagrange dans la formulation variationnelle (2.2) puisque

$$\min_{\underline{y}, J=1} (\mathcal{E}_m - \mathcal{W}^{\text{ext}}) = \min_{\underline{y}} \max_{p} \left(\int_{\Omega_0} W^e(\underline{\underline{e}}) + p(1-J)\, d\Omega - \mathcal{W}^{\text{ext}} \right).$$

Ainsi le tenseur des contraintes devient

$$\underline{\underline{\Sigma}} = \frac{\partial W^e}{\partial \underline{\underline{e}}} - p \frac{\partial J}{\partial \underline{\underline{e}}} = \frac{\partial W^e}{\partial \underline{\underline{e}}} (J_1, J_2) - pJ \underline{\underline{C}}^{-1},$$

et on retrouve ici que le multiplicateur p n'est autre que la pression hydrostatique alors que $\frac{\partial W^e}{\partial \underline{\underline{e}}}(J_1, J_2)$ est la partie déviatorique.

Dans le cas de la quasi-incompressibilité on transforme la formulation mixte en formulation pénalisée du type $W^e(\underline{\underline{e}}) = W^e(J_1, J_2) + \frac{\kappa}{2}(1-J)^2$ avec κ grand pour que la formulation

variationnelle statique redevienne $\min_{\underline{y}}(\mathcal{U}^i)$. Dans ce cas, on obtient pour la pression hydrostatique

$$p = \kappa(1 - J).$$

De même, l'énergie de type Ciarlet-Geymonat, réécrite sur les invariant réduits sous la forme

$$W^e(\underline{\underline{e}}) = \kappa_1(J_1 - 3) + \kappa_2(J_2 - 3) + \kappa(J - 1) - \kappa \ln(J),$$

peut conduire aussi à une formulation pénalisée où la parabole $(1 - J)^2$ est remplacée par $\kappa(J - 1) - \kappa \ln(J)$. Dans cette expression que nous utiliserons finalement pour nos matériaux hyperélastiques on a alors

$$p = \kappa \frac{1 - J}{J}.$$

REMARQUE 2.3.2
Certains matériaux ne sont pas isotropes et dans ce cas, on peut montrer (mais c'est nettement plus technique, voir Raoult (2009)) que d'autres invariants interviennent. Par exemple pour un matériau dit isotrope transverse, c'est-à-dire avec une direction privilégiée en tout point $\underline{n}(\mathrm{x})$ on peut alors décomposer l'énergie suivant deux autres invariants

$$I_4 = \underline{n} \cdot \underline{\underline{C}} \cdot \underline{n}, \quad I_5 = \underline{n} \cdot \underline{\underline{C}}^2 \cdot \underline{n},$$

tels que $W^e = W^e(I_1, I_2, I_4, I_5)$ et avec ces invariants[4]

$$\underline{\underline{\Sigma}} = 2\left(\left(\frac{\partial W^e}{\partial I_1} - I_1 \frac{\partial W^e}{\partial I_2}\right)\underline{\underline{\mathbb{1}}} - \frac{\partial W^e}{\partial I_2}\underline{\underline{C}} + I_3 \frac{\partial W^e}{\partial I_3}\underline{\underline{C}}^{-1} + \frac{\partial W^e}{\partial I_4}\underline{n} \otimes \underline{n} + \frac{\partial W^e}{\partial I_5}(\underline{\underline{C}} \cdot \underline{n} \otimes \underline{n} + \underline{n} \otimes \underline{\underline{C}} \cdot \underline{n})\right). \quad (2.17)$$

Pour des matériaux complètement anisotropes appelés matériaux orthotropes, on fait cette fois apparaître 8 invariants où I_6 et I_7 sont les équivalents de I_4 et I_5 sur la nouvelle direction $\underline{n}'$ et $I_8 = \underline{n} \cdot C \cdot \underline{n}'$. Cependant, indépendamment de la complexité que représente la quantification numérique de tous ces invariants, il y a manifestement une surdétermination puisque, $\underline{\underline{C}}$ étant symétrique, seuls 6 termes sont nécessaires à sa définition.

b. Viscoélasticité

En présence de viscosité, l'égalité (2.6) devient inégalité de Clausius-Duhem, soit

$$\forall \underline{\underline{e}}, \dot{\underline{\underline{e}}}, \quad \left(\frac{\partial W^e}{\partial \underline{\underline{e}}} - \underline{\underline{\Sigma}}\right) : \dot{\underline{\underline{e}}} \leq 0.$$

Cette inégalité est notamment respectée en introduisant une fonctionelle $W^v(\underline{\underline{e}}, \dot{\underline{\underline{e}}})$ convexe en $\dot{\underline{\underline{e}}}$ avec $\frac{\partial W^v}{\partial \dot{e}}(\underline{\underline{e}}, \underline{\underline{0}}) = \underline{\underline{0}}$ telle que

$$\underline{\underline{\Sigma}} = \frac{\partial W^e}{\partial \underline{\underline{e}}} + \frac{\partial W^v}{\partial \dot{\underline{\underline{e}}}}.$$

Il est cependant très difficile de quantifier expérimentalement ces fonctionnelles et certains travaux ont conduit à la définition de lois relativement complexes pour les tissus biologiques

[4] on peut aussi faire le même type de formulation avec des invariants réduits compatibles avec la notion de tenseur déviatorique

Pioletti et al (1998). En l'absence de véritable consensus sur le sujet nous nous limiterons à des éléments visqueux relativement simples de la forme

$$W^v = \frac{\eta}{2}\,\mathrm{tr}(\dot{e}^2) \Rightarrow \underline{\underline{\Sigma}}^v = \frac{\partial W^v}{\partial \underline{\underline{\dot{e}}}} = \eta \underline{\underline{\dot{e}}}$$

c. Modélisation microscopique des myofibrilles

Nous venons de présenter des lois de comportements mécaniques classiques pour traduire un comportement élastique passif d'un matériau. Mais nous avons vu que la mécanique cardiaque est principalement fondée sur l'existence d'un comportement de contraction actif, c'est-à-dire commandé par l'activité électrique, au niveau des sarcomères. C'est ce comportement qu'il nous faut à présent modéliser.

□ MODÈLE DE CONTRACTION DE HUXLEY – La mécanique d'attachement et de détachement des ponts est représentée par la Figure 2.16. Huxley (1957) postule que l'attachement d'une tête de myosine n'est qu'une fonction de $\mathrm{x} \in [0, h]$ avec h un seuil d'élongation. On introduit la variable $s = \frac{\mathrm{x}}{h}$ (c'est-à-dire une forme de variable angulaire traduisant le fait que la tête « pivote » pour s'attacher). Soit n la fraction (« statistique ») de ponts existant à cette élongation.

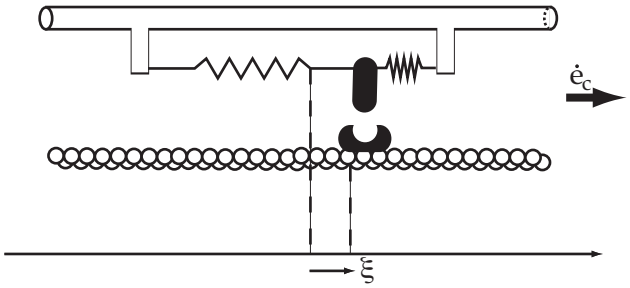


FIGURE 2.16: Schéma mécanique associé au mécanisme du filament glissant

Le modèle de Huxley détermine la dynamique de création des ponts sous la forme

$$\frac{dn}{dt}(s,t) = \frac{\partial n}{\partial t} + \dot{e}_c \frac{\partial n}{\partial s} = (1-n)f - ng, \tag{2.18}$$

avec $\dot{e}_c = \frac{\partial s}{\partial t}$, en supposant qu'une fois la myosine attachée, elle se déplace à la vitesse de l'ensemble. On désigne par f la fréquence de création d'un pont et g la fréquence de détachement. En particulier si f et g sont nulles le nombre de ponts est constant, sinon f (resp. g) s'applique sur le nombre de pont potentiels a créer $(1-n)$ (resp. détruire n). L'effort de modélisation réside ainsi dans le choix judicieux de f et g et nous résumons ici les travaux issus de Bestel (2000).

Lorsque le cycle chimique (2.1) d'attachement a lieu, on peut comparer la concentration $[AM]$ d'actine myosine liées à n. La loi d'action de masse écrite pour le cycle donne alors après simplifications :

$$\begin{aligned} \frac{d}{dt}[AM] &= k_3[AM.ADP.P] - k_4[AM][ATP] \\ &= k_2[A][M.ADP.P] - k_4[AM][ATP] \\ &= k_1[M.ATP] - k_4[AM][ATP]. \end{aligned} \tag{2.19}$$

Notons alors μ_A et μ_M les quantités d'actine et de myosine. Elles sont constantes et égales et donnent, en considérant l'une ou l'autre des espèces comme limitant, la concentration totale en paires possibles.

$$\begin{aligned}\mu = \mu_A &= [AM] + [A] + [AM.ADP.P],\\ &\text{ou}\\ &= \mu_M = [AM] + [M.ATP] + [AM.ADP.P] + [M.ADP.P].\end{aligned}$$

Huxley (1957) choisit de prendre $\mu = \mu_M$ tel que $n = \frac{[A]}{\mu}$. On obtient alors grâce à la loi d'action de masse

$$\mu = [AM] + \left(1 + \frac{k_1}{k_3} + k_1 k_2 [A]\right)[M.ATP]$$

, puis avec (2.19)

$$\frac{d[AM]}{dt} = \frac{k_1}{1 + \frac{k_1}{k_3} + \frac{k_1}{k_2[A]}}(\mu - [AM]) - k_4[ATP][AM]$$

. Un certain nombre d'études ont pu donner des ordres de grandeur des vitesses de réactions $k_1 \gg k_2 \gg k_3$ et on notera $k_{ATP} = k_3$ caractérisant l'activité du cycle (2.1) en présence d'ATP. On obtient finalement

$$\frac{d}{dt}\left(\frac{[A]}{\mu}\right) = k_{ATP}\left(1 - \frac{[A]}{\mu}\right) - k_4[ATP]\frac{[A]}{\mu}. \tag{2.20}$$

Ainsi nous obtenons pendant le cycle chimique $f = k_{ATP}$ supposé constante car l'ATP est majoritaire et $g = k_4[ATP]$. Enfin, les considérations numériques (voir encore Bestel (2000)) permettent de poser $k_4[ATP] = k_3$ (typiquement $k_3 \simeq 75s^{-1}$, $k_4 = 750.10^3 L.\text{mol}^{-1}.s^{-1}$ et $[ATP] = 100\mu\text{mol}.L^{-1}$) donc $g = k_{ATP}$ également. Enfin le cycle n'a lieu qu'en présence d'ions calcium et toujours pour $s \in [0,1]$ donc

$$f(s,t) = k_{ATP}\mathbb{1}_{s\in[0,1]}\mathbb{1}_{[Ca^{2+}]>C},$$

où C est la concentration seuil d'enclenchement de la réaction 2.1.

Pour la fréquence de détachement, les phénomènes sont plus complexes. Pendant la relaxation lorsque $[Ca^{2+}] < C$, des ponts sont détruits par l'aspiration des ions calcium par les pompes sarcoplasmiques. De plus, en contraction comme en relaxation, certains ponts sont détruits par des phénomènes purement mécaniques quand la vitesse relative des filaments est trop grande. Finalement la fonction g est définie sur toute la période de contraction-relaxation par

$$g(s,t) = \alpha\,|\dot{e}_c| + k_{ATP}\mathbb{1}_{s\notin[0,1]}\mathbb{1}_{[Ca^{2+}]>C} + k_{RS}\mathbb{1}_{[Ca^{2+}]<C}.$$

Il est possible de définir une seule variable de commande u pour représenter ces deux fonctions f et g.

$$\begin{aligned}u(t) = |u(t)|_+ - |u(t)|_- \text{ avec } &\begin{cases}|u(t)|_+ = k_{ATP}\mathbb{1}_{[Ca^{2+}]>C}\\ |u(t)|_- = k_{RS}\mathbb{1}_{[Ca^{2+}]<C}\end{cases}\\ f(s,t) &= |u(t)|_+\,\mathbb{1}_{s\in[0,1]},\\ g(s,t) &= |u(t)| + \alpha\,|\dot{e}_c| - f(s,t).\end{aligned} \tag{2.21}$$

Cette commande u est fonction de la concentration en ions calcium et pourra donc être dérivée « indirectement » (i.e. à k_{ATP} et k_{RS} près) des modèles d'activation électrique.

□ EQUATIONS DES MOMENTS – Le passage à l'échelle « mésoscopique » du sarcomère à partir du modèle de Huxley utilise des équations de moments sur la variable « statistique » s inspirées de Zahalac (1981). On définit ainsi les moments successifs

$$\mu_p = \int_{\mathbb{R}} s^p n(s,t)\, ds.$$

Le choix judicieux des fonctions f et g, notamment tel que la somme $f+g$ est indépendante de s, permet de déterminer une relation entre un moment d'ordre p et **seulement** ses prédécesseurs.

$$\begin{aligned}\dot{\mu}_p &= \int_{\mathbb{R}} s^p (f - (f+g)n - \dot{e}_c \frac{\partial n}{\partial s})\, ds \\ &= f_p - (f+g)\mu_p + p\dot{e}_c \mu_{p-1},\end{aligned}$$

où $f_p = \int_{\mathbb{R}} s^p f\, ds$ est le moment d'ordre p de f.

De ces moments, on fait alors apparaître des variables macroscopiques naturelles en considérant le sarcomère comme une collection de ressorts statistiques.

1. Moment d'ordre 0 : soit k_c la raideur associée au moment d'ordre 0 traduisant la rigidité du sarcomère en fonction du nombre de ponts attachés

$$k_c = k_0 \int_{\mathbb{R}} n(s,t)\, ds\ ;$$

2. Moment d'ordre 1 : soit τ_c la contrainte associée à cet ensemble de « ressorts » que l'on considère à l'équilibre en $s = -s_0 < 0$ (i.e. en $s = 0$ le ressort est sous tension)

$$\tau_c = k_0 \int_{\mathbb{R}} (s + s_0) n(s,t)\, ds.$$

Le système d'équations différentielles sur les moments conduit à un modèle de loi de comportement que nous appellerons par la suite *modèle « Bestel-Clément-Sorine »* Bestel et al (2001)

$$\begin{cases} \dot{k_c} = -(|u| + \alpha\, |\dot{e_c}|) k_c + k_0\, |u|_+ \\ \dot{\tau_c} = -(|u| + \alpha\, |\dot{e_c}|) \tau_c + \dot{e_c} k_c + \sigma_0\, |u|_+ \end{cases} \tag{2.22}$$

où tous calculs faits on a posé $\sigma_0 = (\frac{1}{2} + s_0) k_0$.

A la contrainte τ_c nous devons ajouter deux éléments pour obtenir la contrainte finale. Tout d'abord, il y a un certain nombre d'éléments visqueux dans le processus d'attachement-détachement. Ensuite, souhaitant prendre en compte l'effet Starling qui stipule que plus la déformation e_c est importante (i.e. au cycle précédent le cœur s'est beaucoup rempli) plus la contrainte doit être forte pour permettre une éjection on ajoute une fonction $d(e_c)$ en facteur de τ_c et on profite de l'intégration de cette fonction pour prendre en compte un deuxième phénomène physiologique qui est le démantèlement possible des fibres lors de déformations trop fortes. Ainsi d a la forme présentée figure 2.17. Finalement la contrainte totale dans le sarcomère est donc

$$\sigma_c = d(e_c)\tau_c + \mu_c \dot{e}_c. \tag{2.23}$$

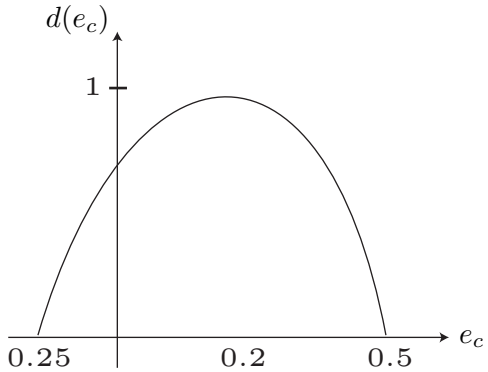


FIGURE 2.17: Fonction d associé à l'effet Starling et au démantèlement des fibres

□ BILAN ENERGÉTIQUE DANS LE SARCOMÈRE – Encore une fois, ce sont les moments (ici d'ordre 2) qui nous fournissent la variable interprétable macroscopiquement. En effet, l'énergie élastique contenue dans un ressort microscopique est $\frac{k_0}{2}(s+s_0)^2$ donc on pose comme énergie « linéique » moyennée sur la collection de ponts

$$U_c = \frac{k_0}{2}\int_{\mathbb{R}} (s+s_0)^2 n(s,t)\, ds.$$

D'après l'équation différentielle sur les moments on obtient

$$\begin{aligned} \dot{U}_c &= \frac{k_0}{2}\int_{\mathbb{R}} (s+s_0)^2 f\, ds - (|u| + \alpha\,|\dot{e}_c|)U_c + \dot{e}_c\tau_c \\ &= -(|u| + \alpha\,|\dot{e}_c|)U_c + \dot{e}_c\tau_c + U_0\,|u|_+\,, \end{aligned} \tag{2.24}$$

avec $U_0 = \frac{k_0}{2}(s_0^2+s_0+\frac{1}{3}) > 0$. Cependant, on aimerait pouvoir définir cette énergie directement par rapport aux variables macroscopiques car cette grandeur nous aurait donné une quantité naturelle à « contrôler » pour définir un schéma en temps. En fait, l'énergie la plus naturelle à définir à partir de τ_c et k_c est l'énergie élastique « linéique » suivante

$$\Psi_c = \frac{1}{2k_c}\tau_c^2, \tag{2.25}$$

dont la dynamique est

$$\begin{aligned} \dot{\Psi}_c &= \frac{\tau_c}{k_c}\dot{\tau}_c - \frac{1}{2k_c}\tau_c^2\frac{\dot{k}_c}{k_c} \\ &= -(|u| + \alpha\,|\dot{e}_c|)\Psi_c + \dot{e}_c\tau_c + k_0\left(\frac{\tau_c}{k_c}(s_0 + \frac{1}{2} - \frac{1}{2}\frac{\tau_c}{k_c})\right)|u|_+\,. \end{aligned} \tag{2.26}$$

Cette dynamique est très proche de celle de l'énergie micro à la différence de la constante positive U_0 remplacée par l'expression (dynamique) $\frac{\tau_c}{k_c}(s_0 + \frac{1}{2} - \frac{1}{2}\frac{\tau_c}{k_c})$. On sait que $\frac{\tau_c}{k_c}$ est une forme de déformation moyenne dont la dynamique est d'ailleurs

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}}\frac{\tau_c}{k_c} &= \frac{\dot{\tau}_c k_c - \dot{k}_c\tau_c}{k_c^2} \\ &= \dot{e}_c + \frac{k_0}{k_c}(\frac{1}{2} + s_0 - \frac{\tau_c}{k_c})\,|u|_+\,. \end{aligned}$$

Autrement dit la variable $\eta_c = \frac{\tau_c}{k_c} - (\frac{1}{2} + s_0)$ vérifie

$$\dot{\eta}_c = -\frac{k_0}{k_c} |u|_+ \eta_c + \dot{e}_c,$$

qui se stabilise que si $\dot{e}_c = 0$ (i.e $\frac{\tau_c}{k_c} \to \frac{1}{2} + s_0$ dans ce cas). Par conséquent il est possible de contrôler Ψ_c à partir des seules variables macroscopiques τ_c et k_c, mais $\frac{\tau_c}{k_c}$ ne correspond pas exactement à l'énergie emmagasinée microscopiquement dans les ressorts élémentaires pour le choix de f effectué.

d. Modèles rhéologiques

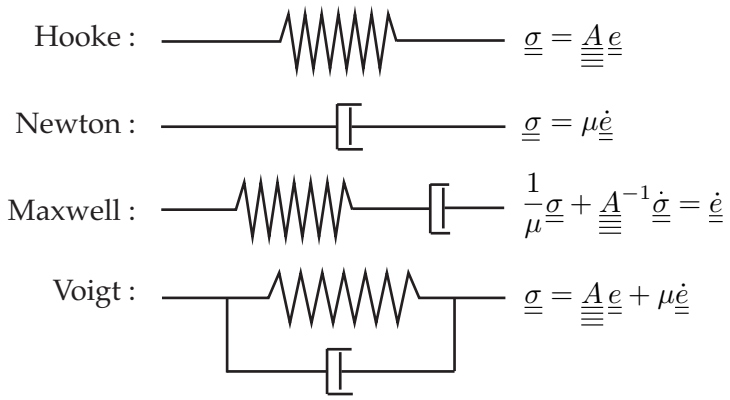


FIGURE 2.18: Principaux montages rhéologiques linéaires

Des comportements mécaniques complexes, notamment pour les tissus biologiques, peuvent être formalisés par des modèles rhéologiques (voir notamment Le Tallec (1994)). L'idée sous-jacente est de décrire le comportement macroscopique global comportant plusieurs éléments constitutifs par analogie avec un ensemble de ressorts et amortisseurs en série et parallèles. Les éléments les plus utilisés sont décrits en Figure 2.18.

Dans le cas de petits déplacements, on démontre sans peine que l'association de différents éléments suit les lois naturelles des montages en série et parallèle.

En non-linéaire cependant, la loi de composition série est plus complexe. En effet, soit un élément 1D (qui nous le verrons nous suffira par la suite) correspondant à deux matériaux élastiques en série. Alors on a

$$\phi = \phi_2 \circ \phi_1 \Rightarrow \nabla\, \phi = \nabla\, \phi_2 \cdot \nabla\, \phi_1,$$

d'où

$$1 + 2e = (1 + 2e_1)(1 + 2e_2). \tag{2.27}$$

Pour les contraintes, les énergies libres s'additionnent $W(e) = W_1(e_1) + W_2(e_2)$ sous la contrainte cinématique (2.27). La minimisation de l'énergie (en statique) permet aussi de définir e_1 et e_2 à l'équilibre

$$(e_1, e_2) = \underset{1+2e=(1+2e_1)(1+2e_2)}{\operatorname{argmin}} W(e),$$

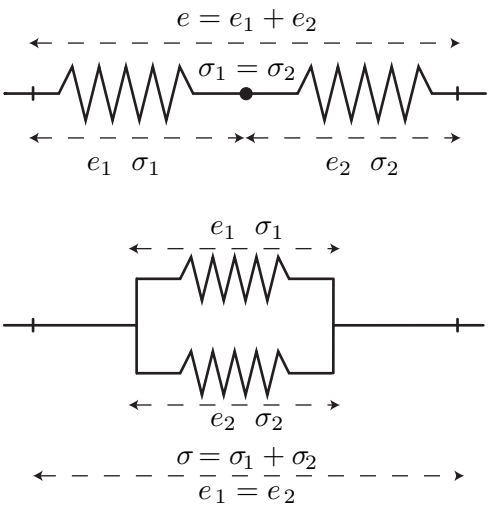


FIGURE 2.19: Schéma rhéologique série et parallèle et règles d'association linéaires

d'où on tire les relations

$$\begin{cases} W_1' \, \mathrm{d}e_1 + W_2' \, \mathrm{d}e_2 = 0 \\ (1 + 2e_2) \, \mathrm{d}e_1 + (1 + 2e_1) \, \mathrm{d}e_2 = 0 \end{cases}$$

(avec W_1', resp W_2', la dérivée de W_1, resp W_2, par rapport à la déformation scalaire e_1, resp e_2). On obtient alors

$$\frac{W_1'}{1 + 2e_2} = \frac{W_2'}{1 + 2e_1}.$$

Enfin la contrainte totale dans l'élément $\sigma = \frac{\partial W}{\partial e}$ est donnée par

$$\begin{aligned} \mathrm{d}W = \sigma \, \mathrm{d}e &= W_1' \, \mathrm{d}e_1 + W_2' \, \mathrm{d}e_2 \\ &= \frac{W_1'}{1 + 2e_2}((1 + 2e_2) \, \mathrm{d}e_1 + (1 + 2e_1) \, \mathrm{d}e_2) \\ &= \frac{W_1'}{1 + 2e_2} \, \mathrm{d}e, \end{aligned}$$

et de même en intervertissant 1 et 2. Ainsi, la loi rhéologique est finalement modifiée sous la forme

$$\sigma = \frac{\sigma_1}{1 + 2e_2} = \frac{\sigma_2}{1 + 2e_1}. \tag{2.28}$$

e. Loi de comportement cardiaque

Reste maintenant l'assemblage final de Chapelle et al (2001) constitué des éléments passifs et actifs décrits précédemment suivant le schéma rhéologique étendu ici en non-linéaire. Ce principe de composition rhéologique avait d'ailleurs était initialement exprimé pour le cœur par Hill (1938). Nous avons d'après le schéma final Figure 2.20 :

- l'élément actif décrivant le comportement actif du sarcomère d'après le modèle Bestel-Clément-Sorine

$$\sigma_c = d(e_c)\tau_c + \mu_c \dot{e}_c.$$

- en série, un élément résistif correspondant aux différent disques Z. Cet élément assure la propriété d'isométricité locale de la fibre cardiaque puisque l'élément contractile peut se contracter $(e_c \neq 0)$ sans déformation totale de la fibre $e = 0$. En notant e_s, σ_s les déformations et contraintes monodimensionnelles de cet élément, on a

$$\sigma_s = E_s e_s,$$

et le modèle rhéologique fournit

$$\sigma_c = d(e_c)\tau_c + \mu \dot{e_c} = E_s \frac{e_{1\mathrm{D}} - e_c}{(1+2e_c)^3}(1+2e_{1\mathrm{D}}), \tag{2.29}$$

où $e_{1\mathrm{D}} = \underline{n} \cdot \underline{\underline{e}} \cdot \underline{n}$ est la déformation dans la direction de la fibre.

- En parallèle, les éléments passifs décrivant la matrice de collagène et d'élastine entourant la fibre. Cet élément passif est 3D et contient les lois hyperélastique et visqueuse présentées.

$$\begin{cases} \underline{\underline{\Sigma}}^p = \underline{\underline{\Sigma}}^e + \underline{\underline{\Sigma}}^v = \dfrac{\partial W^e}{\partial \underline{\underline{e}}}(J_1, J_2) + \dfrac{\partial W^v}{\partial \underline{\underline{\dot{e}}}} + \kappa(J-1)\underline{\underline{C}}^{-1} \\ W^e = \kappa_1(J_1 - 3) + \kappa_2(J_2 - 3) \\ W^\nu = \dfrac{\eta}{2}\,\mathrm{tr}(\underline{\underline{\dot{e}}}^2) \end{cases} \tag{2.30}$$

L'ensemble de ces éléments définissent le tenseur des contraintes total

$$\underline{\underline{\Sigma}} = \underline{\underline{\Sigma}}^p + \sigma_{1\mathrm{D}} \underline{n} \otimes \underline{n}, \tag{2.31}$$

avec pour la contrainte spécifique du sarcomère

$$\sigma_{1\mathrm{D}} = \frac{\sigma_c}{1+2e_s} = \frac{\sigma_s}{1+2e_c}.$$

Dans le cas où $d(e_c) = 1$, la relation rhéologique ci-dessus conduit (elle a en fait été écrite pour cela cf. Section 2.3.2.d.) à partir de l'équation (2.31) à l'identité

$$\underline{\underline{\Sigma}} : \mathrm{d}_{\underline{y}} \underline{\underline{e}}.v = \underline{\underline{\Sigma}}^p : \mathrm{d}_{\underline{y}} \underline{\underline{e}}.\underline{v} + \sigma_s\, \mathrm{d}e_s.v' + \sigma_c\, \mathrm{d}e_c.v'. \tag{2.32}$$

Testée sur la vitesse, elle donne

$$\underline{\underline{\Sigma}} : \underline{\underline{\dot{e}}} = \underline{\underline{\Sigma}}^p : \underline{\underline{\dot{e}}} + \sigma_s \dot{e}_s + \sigma_c \dot{e}_c,$$

permettant l'écriture du **bilan d'énergie total**

$$\frac{d}{dt}\left(\mathcal{E}_c + \mathcal{E}_m + \frac{1}{2}\int_{\Omega_0} E_s e_s^2\, d\Omega + \int_{\Omega_0} U_c\, d\Omega\right) = \\ \mathcal{P}_{\text{ext}} + \int_{\Omega_0} U_0\, |u|_+\, d\Omega - \int_{\Omega_0} (|u| + \alpha\, |\dot{e}_c|) U_c\, d\Omega - \int_{\Omega_0} \frac{\partial W^v}{\partial \underline{\underline{\dot{e}}}} : \underline{\underline{\dot{e}}}\, d\Omega, \tag{2.33}$$

où, en résumé,

- $\mathcal{E}_c$ est l'énergie cinétique

- $\mathcal{E}_m$ est l'énergie hyperelastique isotrope
- $\frac{1}{2}\int_{\Omega_0} E_s e_s^2 \, d\Omega$ est l'énergie de l'élément passif série sur la fibre On remarque que cette formulation est compatible avec les matériaux isotropes-transverse de la remarque 2.3.2
- U_c est l'énergie élastique microscopique des ponts actine-myosine décrite en (2.24)

les termes sources

- $\mathcal{P}_{\text{ext}}$ est la puissance des efforts extérieurs
- $\int_{\Omega_0} U_0 \, |u|_+$ est la puissance positive fournie par le *moteur actine-myosine*

et les termes dissipatifs

- $\int_{\Omega_0} (|u| + \alpha \, |\dot{e}_c|) U_c \, d\Omega$ est l'énergie dissipée par les mécanismes de création et destruction des ponts
- $\int_{\Omega_0} \eta \frac{\partial W^v}{\partial \dot{\underline{\underline{e}}}} : \dot{\underline{\underline{e}}}$ est l'énergie dissipé par la viscosité passive

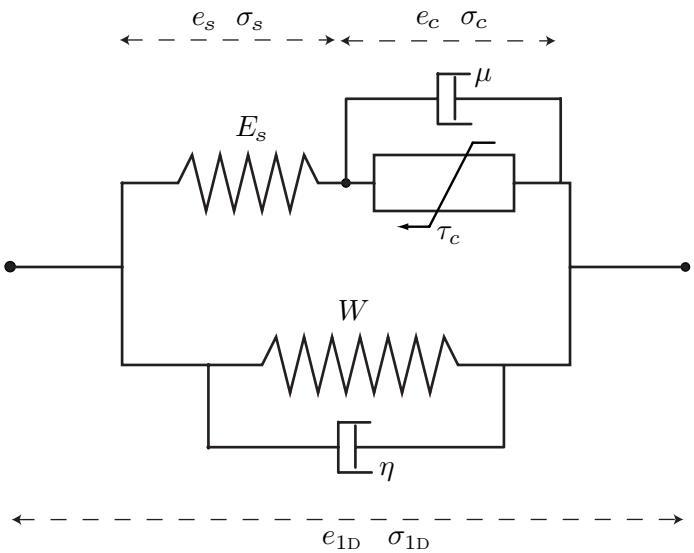


FIGURE 2.20: Schéma rhéologique complet pour le modèle de fibre active (dans la direction $\underline{n}$).

Il existe beaucoup d'autres modèles de contraction cardiaques par exemple ceux présentés dans Schmid et Hunter (2006) ; Hunter et al (1998) ; Sachse (2004). Ces modèles sont tous fondés sur une loi de comportement comprenant un composé actif et un passif. La plupart sont issus d'interpolation de données expérimentales qui les légitiment. Cependant soit ils ne vérifient pas de propriété énergétique globale, soit ils utilisent des fonctionnelles « singulières ». Le modèle le plus utilisé historiquement est certainement celui de Hunter rappelé dans Schmid et Hunter (2006). A noter aussi les tentatives d'obtention des lois macroscopiques par des procédés d'homogénéisation par exemple Caillerie et al (2003) qui est cependant restreinte à la seule partie passive. Encore une fois, l'objectif n'est certainement pas de comparer ces modèles entre

eux « qui ont tous des vertus ». Le modèle présenté ici est récent et cette synthèse s'attache à montrer ses qualités énergétiques, de contrôle par l'activation, d'interprétation physiologique pour un ensemble relativement réduit de paramètres, qualités qui, nous allons le voir, sont fondamentales pour son analyse numérique et sa simulation mais aussi pour structurer les problèmes d'assimilation de données.

2.3.3 Modèles de circulation

Les efforts *extérieurs* sur myocarde sont la résultante de la pression P_V exercée par le sang contenu dans les cavités cardiaques sur leurs paroi. Autrement dit, la puissance virtuelle des efforts extérieurs s'écrit

$$\forall v \in V, \quad \mathcal{P}_{\text{ext}}(\underline{v}) = -\int_{\Gamma_{\text{endo}}} P_V \underline{\nu} \cdot \underline{\underline{F}}^{-1} \cdot \underline{v} J \, dS.$$

où ν le vecteur normal et $\Gamma_{\text{endo}} = \cup_{i=g,d} -_i$ représente l'ensemble des deux endocarde gauches et droits[5] $\mathcal{C}_i$, $i = g, d$. Dans le cas où l'une des valves cardiaques est ouverte, alors P_V est imposée par la pression dans les autres compartiments circulatoires (oreillettes ou artères) ; mais dans le cas où toutes les valves sont fermées, c'est-à-dire dans les phases isovolumiques du cycle cardiaque décrit précédemment, la pression interne n'est plus physiquement qu'un multiplicateur de Lagrange associé à la contrainte d'incompressibilité du fluide (donc de volume constant pour les cavités fermées). Les différents changements de phase du cycle correspondent in fine à la résultante des efforts de pressions sur les valves ; ces dernières sont ouvertes (resp fermées) si la pression dans les cavités concernées sont supérieures (resp. inférieures) aux pressions correspondantes dans les artères ou oreillettes. Ces différentes remarques montrent combien il est indispensable de modéliser ce que nous appellerons par la suite la *circulation extérieure* pour rendre compte du cycle cardiaque complet. En d'autres termes, à la différence d'autres systèmes mécaniques, la pression intracavitaire est une variable d'état du système et pas une donnée indépendante. Il existe depuis longtemps des modèles distribués de circulation sanguine dans les cavités et les artères ainsi que, plus récemment, des modèles de valves complet comme ceux de Diniz Dos Santos et al (2008) ; Astorino et al (2008), mais, pour des raisons évidentes de complexité, on a recours à un modèle de circulation simplifié considérant P^v uniforme dans les cavités. Cette hypothèse se justifie partiellement dans la modélisation fine du tissu mécanique si on garde à l'esprit le fait que l'origine et la plus grande contribution pour les déformations locales du tissu est avant tout l'onde électrique parcourant le myocarde, soit le terme de puissance virtuelle des efforts intérieurs contrôlé par l'activation $u(t, \mathrm{x})$ et la pression moyenne dans la cavité plus que ses fluctuations locales.

a. Modèle de valves

Le modèle de valve gère les différentes phases du modèle définissant la pression extérieure ou le volume intracavitaire comme des contraintes sur le système. En effet la version simplifiée du diagramme (P, V) de la pompe cardiaque serait un rectangle qui signifiant que deux phases sont contrôlées par un volume constant, les fameuses phases isovolumiques, quand

[5] Dans certains cas seule la cavité gauche peut-être modélisée

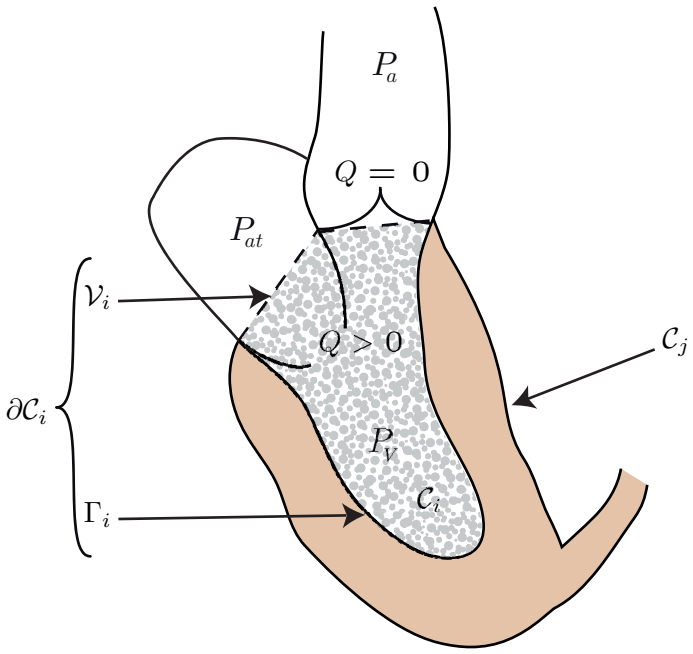


FIGURE 2.21: Configuration de la cavité cardiaque. $\partial\mathcal{C}_i = \Gamma_i \cup \mathcal{V}_i$ correspond à la cavité fermée par la valve idéalisée, c'est-à-dire la frontière de $\mathcal{C}_i$.

deux autres le sont par la pression extérieure imposée par la circulation au moment où les valves sont ouvertes. Le modèle de valve permet de traduire successivement en une seule relation ces deux types de phases. Une analogie intéressante peut être le problème du contact. En effet considérant deux solides S_1 et S_2, la distance d entre eux et la force de contact F sont liées. Si les deux solides se touchent alors $d = 0$ et $F > 0$ est inconnu et sinon c'est $d > 0$ qui devient une inconnue et $F = 0$. On traduit ces deux états par la relation toujours vraie $Fd = 0$ ($F \geq 0$, $d \geq 0$). Pour notre problème, la relation est un peu plus sophistiquée car les deux phases contrôlée en pression correspondent à l'ouverture de valves différentes : d'un côté l'artère pendant la systole, de l'autre l'oreillette en fin de diastole. Si nous considérons par exemple, le flux sanguin Q éjecté par le ventricule gauche ou droit dans l'artère correspondante, nous obtenons $Q\,|P^v - P^{ar}|_- = 0$. De même au niveau de l'oreillette $Q\,\left|P^v - P^{at}\right|_+ = 0$. Ces deux relations traduisent

$$\begin{cases} Q \geq 0 \text{ si } P^v = P^{ar} \\ Q = 0 \text{ si } P^{at} < P^v < P^{ar} \\ Q \leq 0 \text{ si } P^v = P^{at} \end{cases} \tag{2.34}$$

conduisant à la représentation de la figure 2.21. En fait la prise en compte de la perte de charge dans l'artère ou l'oreillette permet d'obtenir une version régularisée qui sera d'ailleurs plus simple à numériser. On obtient alors une relation globale du type

$$-\dot{V} = Q = f(P^v, P^{ar}, P^{at}) = \begin{cases} 0 \text{ si } P^{at} \leq P^v \leq P^{ar} \\ K_{ar}(P^v - P^{ar}) \text{ si } P^v \geq P^{ar} \\ K_{at}(P^v - P^{at}) \text{ si } P^v \leq P^{at} \end{cases} \tag{2.35}$$

où $K_{ar} \propto \frac{\pi R^2}{\rho c}$ où R est la rayon de la valve, ρ la masse volumique du sang et c sa vitesse.

Pour des raisons numériques, nous serons aussi amenés à prendre en compte une version régularisée de cette fonction autour de $Q = 0$ définie comme suit

$$-\dot{V} = Q = f_\delta(P^v, P^{ar}, P^{at}) = \begin{cases} K_{iso}(P^v - P^{at}) \text{ si } P^{at} \leq P^v \leq P^{ar} \\ K_{ar}(P^v - P^{ar}) + K_{iso}(P^{ar} - P^{at}) \text{ si } P^v \geq P^{ar} \\ K_{at}(P^v - P^{at}) \text{ si } P^v \leq P^{at} \end{cases} \tag{2.36}$$

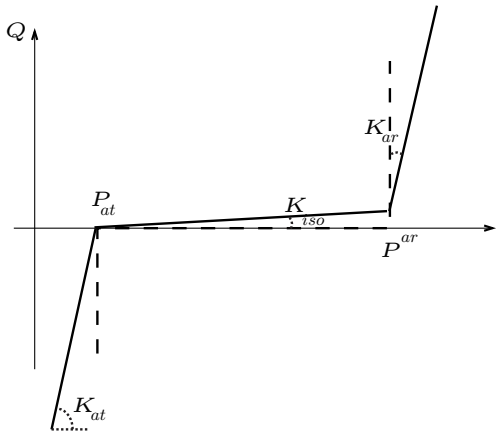


FIGURE 2.22: Fonctions de valve. En pointillés, la fonction idéale (2.34) **et en traits pleins la fonction** f_δ **définie en** (2.36)

La dérivée en temps du volume, $\dot{V}$, est donc le débit sanguin éjecté par l'un ou l'autre des ventricules et est donné variationnellement à partir de l'expression du volume V de sang contenu dans le ventricule. A valve ouverte, cette notion révèle une certaine ambiguïté pour délimiter la frontière fictive correspondant à l'intérieur des ventricules au niveau des valves. Cependant, cette ambiguïté est en partie levée par une certaine rigidité des anneaux valvulaires permettant de définir une sorte de disque valvulaire. Ainsi par Stokes on obtient directement les expressions des volumes V et débit $\dot{V}$ de chaque cavité à partir des déplacements de l'épicarde. En effet

$$V = \int_{\mathcal{C}_i} d\Omega = \frac{1}{3} \int_{\mathcal{C}_i} \operatorname{div}(\underline{x})\, d\Omega,$$

donc

$$V = \frac{1}{3} \int_{\partial\mathcal{C}_i} \underline{x} \cdot \underline{\nu}\, dS \tag{2.37}$$

et, utilisant la formule (2.5) de dérivation du volume déjà utilisée en introduction de Section 2.3.2 et après simplifications,

$$\dot{V} = \int_{\partial\mathcal{C}_i} \underline{\dot{y}} \cdot \underline{\nu}\, dS. \tag{2.38}$$

REMARQUE 2.3.3 (ONDE DICROTE)
Certains ont imaginé des modèles de contact plus complexes, introduisant des fonctions non-linéaires pour prendre en compte des phénomènes physiologiques plus subtile tels que la prise en compte de l'onde dicrote. Cette onde de surpression est visible juste après la fermeture de la valve aortique mais son interprétation est encore sujette à débat. Une hypothèse est que le flux sanguin en sortie de ventricule permet le remplissage des coronaires qui fonctionnent alors comme une réserve de sang. Cette

réserve se vide un peu après la fermeture entraînant un effet de suppression. Pour modéliser cet effet, on peut imaginé une fonction de valve non-linéaire de la forme :

$$\frac{1}{R_{av}} = \frac{1}{\rho_{av}}(1 + \tanh(K_3(P^v - P^{ar} - P_{av}))),$$

où P_{av} est une pression seuil, R_{av} et ρ_{av} des constantes supplémentaires.

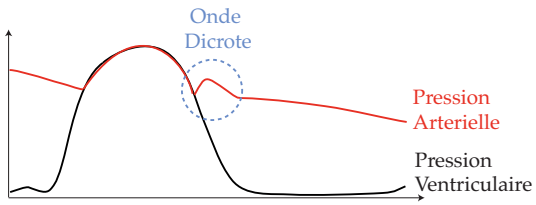


FIGURE 2.23: Onde dicrote schématisée.

b. Modèle de Windkessel

Afin d'obtenir un diagramme (P, V) réaliste, il est nécessaire de modéliser la circulation extérieure et donc que P^{ar} varie au long du cycle cardiaque et en particulier pendant l'éjection. Là encore plusieurs niveaux de modélisation en espace sont possibles et nous nous limitons à une modélisation « 0D ». La modélisation de type Windkessel pour l'écoulement artériel consiste en une analogie électrique où l'intensité représente le flux et la tension la pression. L'idée fut initialement proposée par S. Hales en 1733 puis développée par O. Frank en 1899. Nous considérons ici un modèle de Windkessel à 3 éléments équivalent à schéma RLC électrique. Typiquement la capacitance C_w représente l'énergie stockée par le vaisseau grâce à sa déformation, R_w représente la résistance équivalente du système circulatoire complet,

$$\frac{L_w C_w}{R_w}\ddot{P}^{ar} + C_w \dot{P}^{ar} + \frac{P^{ar} - P_{sv}}{R_w} = (1 + \frac{R_w}{R_q})Q + C_q R_w \dot{Q} + L_q C_q \ddot{Q}, \tag{2.39}$$

avec $Q = \dot{V}$ le débit sanguin. Le second membre et formulé ici de façon très général comme dans Stergiopulos et al (1999). Cette expression nécessite cependant la calibration des différents paramètres ce qui représente un véritable enjeux. En effet Stergiopulos et al (1999) calibre ce modèle à Q donné alors que celui-ci est normalement couplé par ailleurs à la pression via la dynamique cardiaque.

Afin d'interpréter ce modèle nous lui préférons donc une version premier ordre sous une forme où P_d représente une pression plus éloignée dans le bras artériel appelée *pression distale*,

$$\begin{cases} C_p \dot{P}^{ar} + \dfrac{P^{ar} - P^d}{R_p} = Q, \\ C_d \dot{P}^d + \dfrac{P^d - P^{ar}}{R_p} = \dfrac{P_{sv} - P_d}{R_d}, \end{cases} \tag{2.40}$$

Dans ces lois, R_p et C_p sont les paramètres proximaux et R_d, C_d sont les paramètres distaux modélisant le reste du système artériel. Enfin P_{sv}, la pression dans le réseau veineux, est supposée constante et faible. Pour retrouver le système du second ordre, on a tous calculs faits,

$$L_w^{-1} = \frac{1}{R_p}\left(\frac{1}{R_p C_d} + \frac{1}{R_d C_d} + \frac{1}{R_p C_p}\right), \quad C_w = C_p\left(1 + \frac{R_p}{R_d} + \frac{C_d}{C_p}\right), \quad R_w = R_p,$$

et

$$R_q = R_d, \quad C_q = C_d, \quad L_q = 0.$$

Du point de vue énergétique (ce qui ensuite nous servira pour le choix du schéma en temps), on a immédiatement

$$\frac{\mathrm{d}}{\mathrm{dt}}\left(\frac{C_p}{2}(P^{ar})^2 + \frac{C_d}{2}(P^d)^2\right) = -\frac{(P^{ar}-P^d)^2}{R_p} - \frac{(P^d-P_{sv})^2}{R_d} + QP^{ar} + \frac{P_{sv}^2}{R_d}.$$

A noter que l'analogie électrique convient pour le modèle de la valve que nous avons simplement représenté par des diodes (ou diodes non-linéaires dans certains cas). Ainsi tout le modèle de circulation peut être représenté par un schéma complet présenté Figure 2.24.

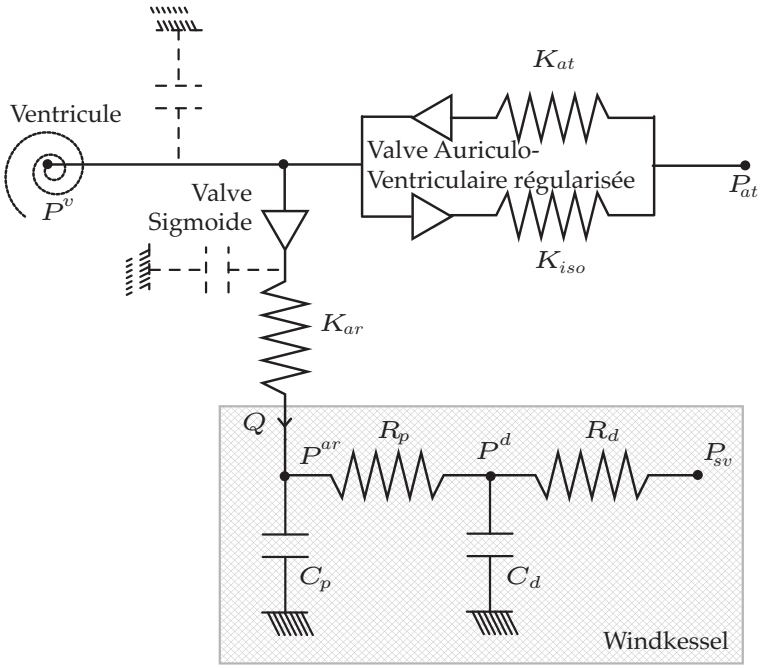


FIGURE 2.24: Schéma de circulation complet comprenant les valves sous forme de diode et le montage artériel Windkessel

2.3.4 Modèle cardiaque complet

Nous déduisons des sections précédentes le modèle de contraction mécanique cardiaque complet de Sainte-Marie et al (2006) et utilisé dans les simulations numériques,

$$\begin{cases} \displaystyle\int_{\Omega_0} \rho \ddot{\underline{y}} \cdot \underline{v}\, d\Omega + \int_{\Omega_0} \underline{\underline{\Sigma}}(e, e_c) : \mathrm{d}_{\underline{y}} \underline{\underline{e}}.\underline{v}\, d\Omega + \sum_{i=g,d} \int_{\partial \mathcal{C}_i} P_i^v \underline{\nu} \cdot \underline{\underline{F}}^{-1} \cdot \underline{v} J dS = 0, \quad \forall \underline{v} \in V \\ \underline{\underline{\Sigma}} = -pJ\underline{\underline{C}}^{-1} + \dfrac{\partial W^e}{\partial \underline{\underline{e}}} + \dfrac{\partial W^v}{\partial \dot{\underline{\underline{e}}}} + \sigma_{1\mathrm{D}}(e_{1D}, e_c) \underline{n} \otimes \underline{n}, \quad \sigma_{1\mathrm{D}} = \dfrac{\sigma_c(1+2e_c)}{1+2e_{1\mathrm{D}}} \\ \sigma_c = d(e_c)\tau_c + \mu \dot{e_c} = E_s \dfrac{e_{1\mathrm{D}} - e_c}{(1+2e_c)^3}(1+2e_{1\mathrm{D}}), \quad e_{1\mathrm{D}} = \underline{n} \cdot \underline{\underline{e}} \cdot \underline{n} \\ \dot{k_c} = -(|u| + \alpha\, |\dot{e_c}|)k_c + k_0\, |u|_+ \\ \dot{\tau_c} = -(|u| + \alpha\, |\dot{e_c}|)\tau_c + \dot{e_c} k_c + \sigma_0\, |u|_+ \\ Q_i = \dot{V_i} = \displaystyle\int_{\partial \mathcal{C}_i} \underline{\nu}.\underline{\underline{F}}^{-1}.\delta \underline{v} J dS = -f_\delta(P_i^v, P_i^{ar}, P_i^{at}), \quad i = g,d \\ L_w \ddot{P}_i^{ar} + C_w \dot{P}_i^{ar} + \dfrac{P_i^{ar} - P_{sv}}{R_w} = (1 + \dfrac{R_w}{R_q})Q_i + C_q R_w \dot{Q}_i, \quad i = g,d \end{cases} \tag{2.41}$$

2.3.5 Conditions aux limites

La question des conditions aux limites pour le système mécanique cardiaque est sans doute une des plus difficiles à résoudre dans le cadre de la modélisation. En effet, dans l'organisme le cœur est maintenu « en place » par l'ensemble des organes qui l'entourent. En particulier, il est en contact avec le diaphragme, le sternum, les poumons et est sans doute partiellement retenu par l'aorte et l'artère pulmonaire. Ainsi, du point de vue de l'observation cinématique le cœur semble se contracter en s'enfonçant dans le jour de la cavité cardiaque au niveau du diaphragme de telle sorte que l'apex paraît très peu se déplacer au contraire de la base.

Il est difficile de tenir compte de l'ensemble des organes encadrant le cœur pour définir des conditions aux limites « exactes », surtout s'il faut résoudre pour chacun un problème de contact. L'objectif est donc de les remplacer par des conditions aux limites plus simples avec un certain nombre de précautions.

□ CONDITIONS AUX LIMITES SUR LES VALVES – Le premier choix est soit de bloquer (conditions aux limites de Dirichlet), soit de contraindre (condition aux limites de type Robin) un certain nombre de degrés de liberté de la base afin, au minimum, de filtrer les modes de corps rigide. Il est aisé, par la suite, de rendre fixe l'apex par post-traitement alors qu'il eût été difficile de bloquer au moins les six degrés de liberté de corps rigide en appliquant des conditions uniquement autour de celui-ci.

Une autre remarque s'impose sur les conditions aux limites relativement au couplage modèle 3D de déformation avec le modèle 0D de valve+Windkessel. En effet, à valve ouverte par exemple, si on suppose une déformation de la valve nous avons vu que il était difficile de définir ce que signifiait le volume de la cavité. De plus l'échange d'énergie entre le ventricule et l'artère ne se réalise plus simplement à travers le fluide mais aussi à travers le déplacement solide de l'interface (ici l'anneau valvulaire). Or cette énergie mécanique échangée n'est pas intégrée au niveau du Windkessel sauf à imaginer (et rien ne dit que ça soit possible) de modifier

les constantes C et R moralement dépendante de la géométrie en fonction du déplacement... Autrement dit, du point de vue énergétique, l'échange énergétique entre la première équation de (2.41) (testée sur les vitesses) et la dernière (testé sur les pressions) est exacte à valves fixes. C'est en effet le seul cas où l'évolution de $\partial\mathcal{C}_i$ au cours du temps est exactement celle de $\mathcal{C}_i$ puisque l'anneau est fixe. Sinon il y a un petit apport énergétique associé aux déplacements des valves.

Ces conditions aux limites ne sont pas idéales et elle impliquent que le comportement mécanique de l'organe doit être jugé loin du lieu de leur prescription. En effet, *Le principe de Saint-Venant* souvent mentionné en mécanique des poutres dit en substance que *si on remplace, sur une zone petite, les conditions aux limites en efforts par des conditions aux limites ayant une résultante et un moment identiques, les deux solutions diffèrent peu, loin de cette zone*. Reste finalement que la notion de « loin de cette zone » serait de l'ordre de la dimension du cœur et qu'il est donc indispensable de pouvoir traiter au moins les conditions sur la base le plus loin possible au niveau des artères...

□ CONDITIONS AUX LIMITES DE SUPPORT ÉLASTIQUE – Un modèle de conditions aux limites un peu plus réalistes consiste à prendre en compte l'environnement de l'organe par des conditions de support viscoélastiques équivalentes dans les régions où prévalent les attachements principaux. Sous forme variationnelle, elles s'écrivent sous la forme d'une puissance d'efforts extérieurs supplémentaires

$$\mathcal{P}^{\text{ext}} = -\int_{\mathcal{B}} (a(x)\underline{y} + b(x)\underline{\dot{y}}) \cdot \underline{v}\, dS, \tag{2.42}$$

sur la surface $\mathcal{B}$ avec a et b les modules d'élasticité et de viscosité correspondants.

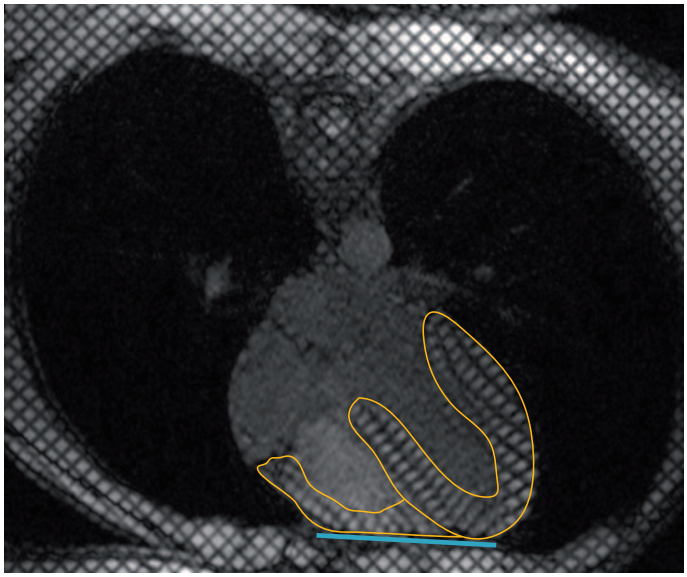


FIGURE 2.25: IRM taggé + segmentation : coupe 4 cavités de cœur en systole (contraction). En bleu la zone d'attachement à l'apex (remerciements à Radomir Chabiniok, Macs, INRIA et IKEM, Prague).

Typiquement, des clichés d'IRM (par exemple Figure 2.25) tendent à définir $\mathcal{B}$ comme la zone de contact avec le sternum près de l'apex. Cependant, reste alors à calibrer ce « modèle

équivalent » en choisissant des fonctions a et b adéquates. Cette calibration peut d'ailleurs être à terme complétées par une phase d'assimilation de données. Cette question de modélisation des conditions aux limites, fondamentale dans tous les cas, s'exprime ici différemment, car mathématiquement ces dernières sont paramétrées. Elles donnent donc lieu à un problème d'estimation similaire à celui de l'identification des paramètres de la loi de comportement.

2.3.6 Eléments de modélisation électrique

Encore une fois, l'entrée du modèle mécanique cardiaque, représentée par la variable u, est directement liée à la concentration en ions calcium variable au cours de la dépolarisation. Dans la volonté d'offrir une synthèse complète, nous souhaitons donner quelques éléments de modélisation des courants ioniques circulant dans le tissu cardiaque lors de dépolarisation cardiaque. Cette courte section résume ainsi succinctement les travaux de l'équipe Reo de l'INRIA (publiés dans Boulakia et al (2007, 2008)) sur la modélisation électrique de la contraction cardiaque actuellement **utilisée en entrée** du simulateur cardiaque développé côté mécanique. Les résultats issus de ce couplage électromécanique faible, au sens où on considère que la propagation électrique n'est pas affecté par la mécanique,[6] seront présentés conjointement entre les équipes Reo et Macs dans l'article en préparation Chapelle et al (2008).

□ MODÈLE BIDOMAINE – Le modèle d'activité électrique mise en oeuvre est un modèle bidomaine dont on trouvera une description complète dans Sundnes et al (2006). Le milieu cardiaque est constitué de deux domaines dit intra-cellulaire Ω_i et extra-cellulaire Ω_e ayant chacun leur densité de courant respectives $\underline{j}_i$ et $\underline{j}_e$ et potentiels u_i et u_e tel que la cellule cardiaque soit représentable par le schéma Figure 2.26.

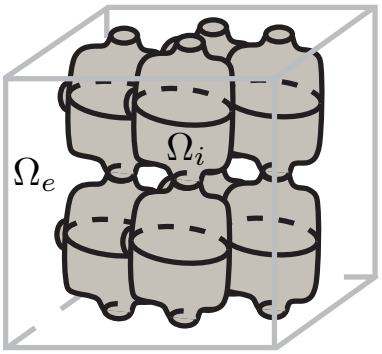


FIGURE 2.26: Echantillon de tissu avec les milieux intra-cellulaire et extra-cellulaire. L'homogénéisation consiste à faire croître le nombre de cellules.

La loi homogénéisée de conservation de la charge s'écrit

$$\mathrm{div}(\underline{j}_i + \underline{j}_e) = 0,$$

En outre, la relation tension-intensité pour le potentiel transmembranaire $V_m = u_i - u_e$ est

$$A_m \left(C_m \frac{\partial V_m}{\partial t} + I_{\text{ion}}(V_m, w) \right) + \mathrm{div}(\underline{j}_i) = I_{\text{app}}$$

[6] le couplage fort appelé effet mécano-électrique est un phénomène dont la communauté de physiologie mathématique cherche encore à quantifier l'importance.

où A_m est la densité moyenne de surface cellulaire par unité de volume, I_{ion} est une fonction représentant les échanges ioniques au niveau de la cellule à définir précisément plus loin, I_{app} est un courant de stimulation et C_m la capacité moyenne de la membrane cellulaire (voir Figure 2.29). Cette équation traduit pour l'instant la cinématique de la propagation électrique mais, dépend de la façon dont on modélise I_{ion} en fonction de la dynamique de w représentant l'ouverture des échangeurs ioniques entre les milieux intra- et extra-cellulaire

$$\frac{\partial w}{\partial t} + g(V_m, w) = 0.$$

Il existe de nombreux modèles de courant ioniques allant de ceux représentant toutes les espèces pour tous les types de cellules cardiaques (voir Luo et Rudy (1994)) ou au contraire des modèles plus simplifiés. Pour les cellules courantes dans le ventricule le modèle le plus complet actuel est peut-être ten Tusscher et al (2004) et comporte 17 variables d'état. Un modèle plus phénoménologique à deux variables est choisi par Boulakia et al (2008) pour ses propriétés vis-à-vis de l'ECG, celui de Mitchell et Schaeffer (2003) :

$$I_{\text{ion}}(V_m, w) = \frac{w}{\tau_{\text{in}}}(1 - V_m)V_m^2 + \frac{V_m}{\tau_{\text{out}}},$$

et

$$g(V_m, w) = \begin{cases} \dfrac{w-1}{\tau_{\text{open}}} & \text{si } V_m \leq V_{gate} \\ \dfrac{w-1}{\tau_{\text{close}}} & \text{si } V_m \geq V_{gate} \end{cases}$$

Ce modèle comporte 2 variables d'état représentant un courant d'entrée et de sortie. Un autre modèle utilisé et limitant le nombre de variables d'états est celui proposé par Fenton et Karma (1998) et comporte 3 variables représentant les courants ioniques du calcium, sodium, potatium. Enfin un modèle permettant l'extension à d'autres cellules cardiaques (oreillettes, réseau de conduction rapide voir Figure 2.27) a aussi été implémenté par cette équipe au cours d'une collaboration avec l'équipe Sysiphe de l'INRIA comme présenté dans Djabella et Sorine (2005) comporte 8 variables d'état semble être un bon compromis en terme de complexité.

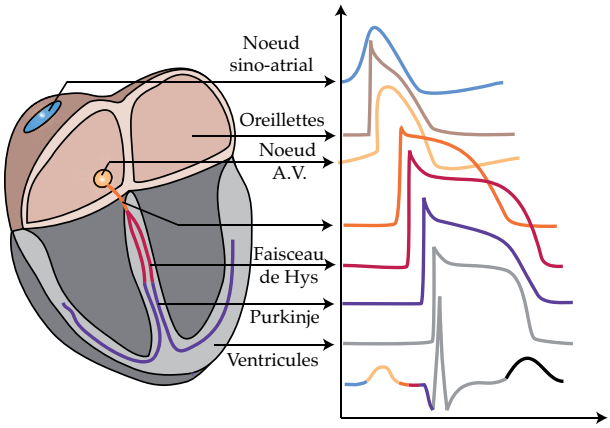


FIGURE 2.27: Réponse électrique de chaque type de cellule cardiaque. Un type de modèle ionique suffit pour le ventricule gauche qui contribut au complexe QRS de l'ECG. (l'activité électrique se situe entre -20 et 80 mV). Schéma inspiré de Netter (1969).

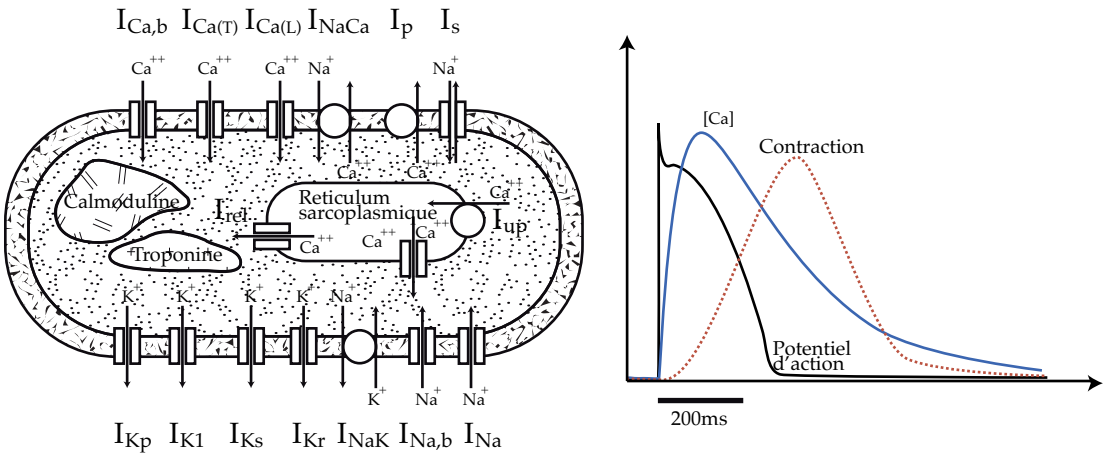


FIGURE 2.28: Exemple de modèle de courants ioniques pour la cellule cardiaque

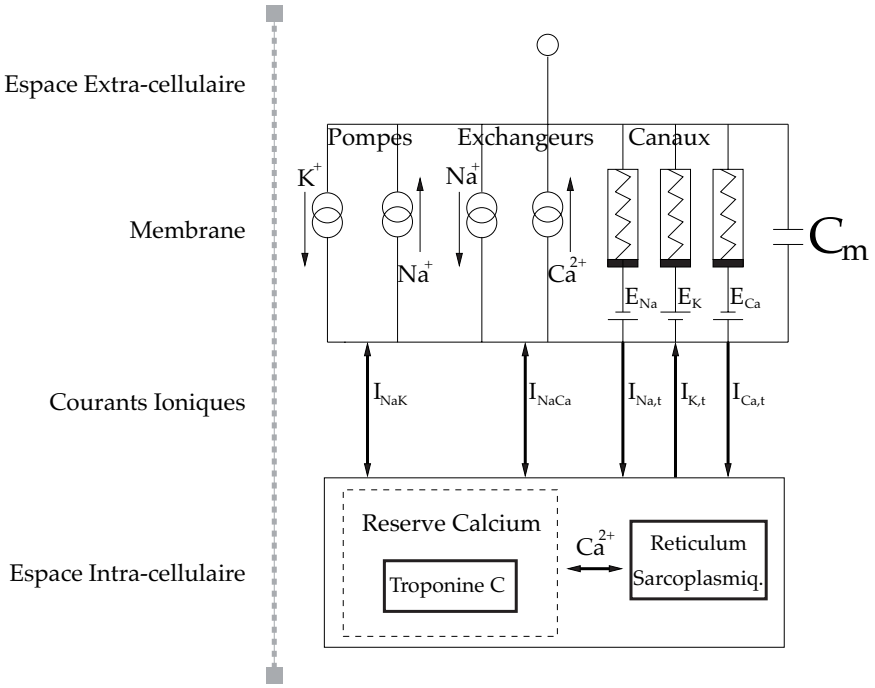


FIGURE 2.29: Traduction sous forme de circuit d'un modèle de courants ioniques

A ces équations, on ajoute classiquement la loi d'Ohm,

$$\underline{j}_i = \underline{\underline{\sigma}}_i \cdot \underline{\nabla} u_i, \quad \underline{j}_e = \underline{\underline{\sigma}}_e \cdot \underline{\nabla} u_e$$

afin d'exprimer en fonction des potentiels les dynamiques complètes de l'activité électrique

$$\begin{cases} A_m \left(C_m \dfrac{\partial V_m}{\partial t} + I_{\text{ion}}(V_m, w) \right) - \operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m) = \operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} u_e) + I_{\text{app}}, \\ -\operatorname{div}((\underline{\underline{\sigma}}_i + \underline{\underline{\sigma}}_e) \cdot \nabla u_e) = \operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m), \\ \dfrac{\partial w}{\partial t} + g(V_m, w) = 0 \end{cases} \quad \text{sur } \Omega \tag{2.43}$$

De plus, il faut ajouter des conditions aux limites traduisant

- que $\underline{j}_i$ ne circule pas en dehors du cœur, donc

$$(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} u_i) \cdot \underline{n} = 0, \quad \text{sur } \partial\Omega.$$

- la continuité des potentiels, en supposant que le péricarde est un conducteur parfait,

$$\begin{cases} u_T = u_e, \quad \text{sur } \partial\Omega \\ (\underline{\underline{\sigma}}_e \cdot \underline{\nabla} u_e) \cdot \underline{n} = (\underline{\underline{\sigma}}_T \cdot \underline{\nabla} u_T) \cdot \underline{n}, \quad \text{sur } \partial\Omega \end{cases}$$

  où u_T est le potentiel électrique dans le reste du thorax conducteur (sinon on n'aurait pas d'ECG)

Ce potentiel u_T dans le thorax Ω_T suit une simple équation de Poisson classique en électrostatique dans un milieu conducteur en l'absence de charges supplémentaires d'où

$$\operatorname{div}(\underline{\underline{\sigma}}_T \cdot \underline{\nabla} u_T) = 0, \quad \text{sur } \Omega_T.$$

avec comme condition aux limites pour le courant à la surface du thorax

$$(\underline{\underline{\sigma}}_T \cdot \underline{\nabla} u_T) \cdot \underline{n} = 0, \quad \text{sur } \Gamma_{\text{ext}}.$$

Or c'est ce sont des différences de potentiel u_T mesurés en des point spécifiques de la surface extérieure Γ_{ext} du thorax qui produisent les différentes dérivations de l'ECG.

□ MODÈLE MONODOMAINE POUR CŒUR ISOLÉ – Le modèle bidomaine peut être simplifié en une équation dite monodomaine (voir Sundnes et al (2006)) en considérant que les deux tenseurs de conductivité sont proportionnels

$$\underline{\underline{\sigma}}_i = \lambda \underline{\underline{\sigma}}_e$$

pour un paramètre λ de l'ordre de l'unité. Dans ce cas dans (2.43), la deuxième équation donne

$$-\operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} u_e) = \frac{1}{1+\lambda} \operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m), \quad \text{sur } \Omega, \tag{2.44}$$

impliquant que la première équation n'est plus fonction que de V_m

$$A_m \left(C_m \frac{\partial V_m}{\partial t} + I_{\text{ion}}(V_m, w) \right) + \frac{\lambda}{1+\lambda} \operatorname{div}(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m) = I_{\text{app}}, \quad \text{sur } \Omega. \tag{2.45}$$

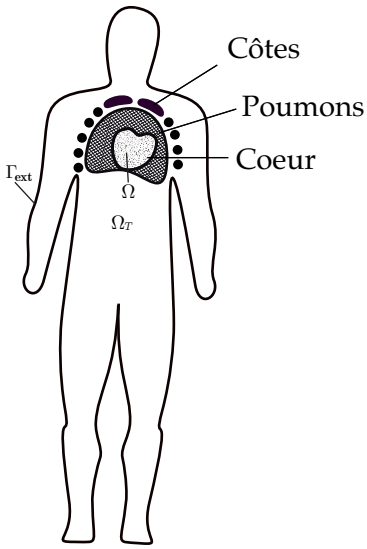


FIGURE 2.30: Schéma présentant le domaine électrique. La propagation de la dépolarisation a lieu dans le tissu cardiaque mais l'ECG est calculé via une équation de Poisson sur tout le domaine où les conductivités dépendent du type de tissu.

Pour résoudre uniquement cette équation sur V_m, on ajoute alors l'hypothèse supplémentaire d'isolation du cœur par rapport au thorax. La condition aux limites pour (2.45) se résume alors à

$$(\underline{\underline{\sigma}}_i \cdot \underline{\nabla} V_m) \cdot \underline{n} = 0, \quad \text{sur } \partial\Omega \tag{2.46}$$

Pour le torse, on résout d'abord l'équation (2.44) pour obtenir u_e étant donné V_m, puis

$$\begin{cases} \operatorname{div}(\underline{\underline{\sigma}}_T \cdot \underline{\nabla} u_T) = 0, & \text{sur } \Omega_T \\ u_T = u_e & \text{sur } \partial\Omega \\ (\underline{\underline{\sigma}}_T \cdot \underline{\nabla} u_T) \cdot \underline{n} = 0 & \text{sur } \Gamma_{\text{ext}} \end{cases} \tag{2.47}$$

où $u_T = u_e$ est la seule condition aux limites conservée dans le cas d'un cœur isolé.

Finalement, vis-à-vis du couplage mécanique, on simplifie l'interdépendance de l'entrée électrique du modèle Bestel-Clément-Sorine avec les courant ioniques et donc le potentiel transmembranaire par une relation affine du type

$$u = \alpha V_m + \beta,$$

où α et β sont des constantes de couplage électromécanique globale qui pourrait (devrait) être précisées si on peut modéliser finement la dépendance de la commande u introduite en fonction des espèces du modèle ionique considéré. C'est finalement ce modèle qui sert d'entrée au modèle mécanique présenté.

□ MATRICE DE TRANSFERT – La condition d'isolation du cœur permet l'obtention des différentes potentiels à la surface du thorax (donc les dérivations ECG) revient à résoudre un problème elliptique (statique) linéaire. Par conséquent, on peut remplacer cette résolution pas une matrice de transfert (calculée une fois pour toutes) afin d'accélérer les calculs. En pratique, cette matrice de transfert comporte en ligne la solution du problème où comme condition aux limites 1 pour un élément de la surface et 0 ailleurs.

2.4 Analyse numérique et simulations

2.4.1 Modèles géométriques

Afin de réaliser des simulations cardiaques réalistes un pré-requis indispensable est de définir une géométrie discrète adaptée à la simulation. Autrement dit la première étape de la discrétisation n'est pas celle des équations mais bien celle de la géométrie. Or celle-ci est souvent jugée acquise alors qu'elle représente un effort humain considérable. On dit souvent que dans l'industrie, plus de la moitié du temps de simulation est passé à la définition de la géométrie. C'est pourquoi les paragraphes qui suivent (jusqu'à la Section 2.4.2) présentent les grandes étapes de la définition géométrique du modèle que sont :

- l'obtention de la géométrie
- la définition des zones AHA
- la définition du champ de fibres

Ils sont essentiellement opérationnels et techniques et ne sont donc pas dans le même registre que le reste du chapitre. Pour autant il participe à la constitution du simulateur direct et à son transfert...

a. Modèles anatomiques

Dans toute la suite nous utiliserons préférentiellement deux géométries.

□ VG ANALYTIQUE – Une première dite *analytique* représentant un ventricule gauche défini à partir de deux ellipsoïdes tronqués emboîtés d'équations respectives

$$\begin{cases} \text{Epicarde : } \dfrac{\mathrm{x}^2}{a_L^2} + \dfrac{\mathrm{y}^2}{b_L^2} + \dfrac{\mathrm{z}^2}{c_L^2} = 1 \\ \text{Endocarde : } \dfrac{\mathrm{x}^2}{(a_L - e_L)^2} + \dfrac{\mathrm{y}^2}{(b_L - e_L)^2} + \dfrac{\mathrm{z}^2}{(c_L - e_L)^2} = 1 \end{cases}$$

où e_L représente l'épaisseur du ventricule et numériquement

$$a_L = 40\,\mathrm{mm},\ b_L = 45\,\mathrm{mm},\ c_L = 80\,\mathrm{mm},\ e = 16\,\mathrm{mm}.$$

La troncature intervient pour $z > h$ avec $h = 20\,\mathrm{mm}$.

Le grand axe est défini sur $0\mathrm{z}$ et ces ellipses sont donc tronquées pour $\mathrm{z} > h$. Cette géométrie donne lieu alors à plusieurs maillages surfaciques puis volumiques respectivement à partir des codes `Yams` et `GHS3D`[7] (voir George et al (1990) ; Frey (2001) et plus généralement Frey et George (2008) et les références incluses). Les différentes discrétisations utilisées sont résumées sur le graphe suivant.

Sur cette géométrie, on peut définir aisément les régions AHA à partir de secteurs angulaires dans le plan $(O\mathrm{x}, O\mathrm{y})$. On obtient alors les maillages référencées présentées notamment Figure 2.31.

[7] distribués commercialement par Distène

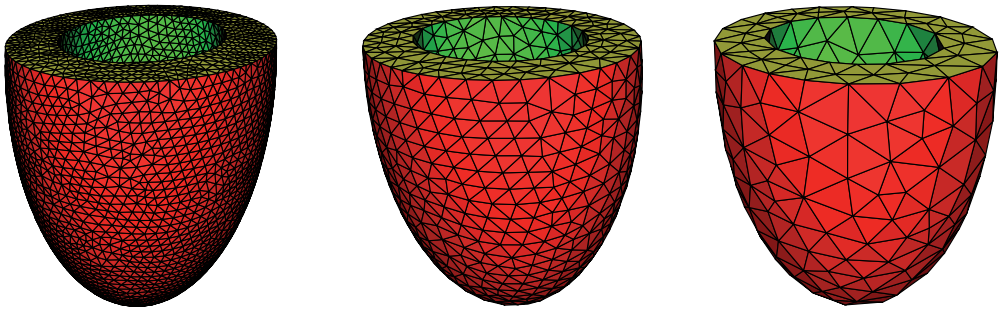


FIGURE 2.31: 3 Exemples de discrétisations possibles sur la géométrie analytique

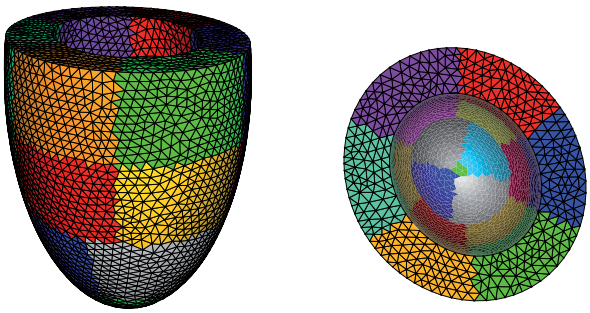


FIGURE 2.32: Régions AHA fixées sur la géométrie analytique

On peut aussi ajouter à ce ventricule gauche un ventricule droit par intersection avec un deuxième couple d'ellipsoïdes ($a_R = 80\,mm$, $b_R = 45\,mm$, $c_R = 75\,mm$, $e_R = 8\,mm$,) plus grandes et tronquées cette fois en $x > 0$ et $z > h$.

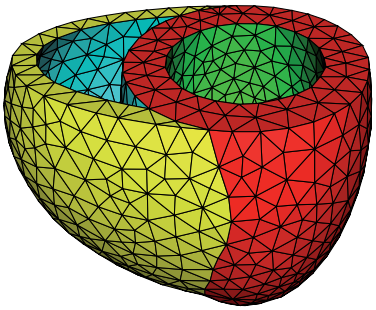


FIGURE 2.33: Exemple de géométrie bi-ventriculaire analytique

□ MAILLAGE 3D SCIENCE – L'action d'envergure Cardiosense3D a fait l'acquisition auprès de *Zygote Media Group* de l'ensemble des maillages surfaciques représentant le système cardiovasculaire du projet 3D Science. Cette géométrie provient vraisemblablement de post-traitements à partir de données CT (Computed Tomography) hautes résolutions. L'obtention de maillages de calcul à partir de ces éléments se confronte à deux difficultés majeures. Tout d'abord ces maillages sont avant tout des maillages de visualisation et pas de calculs éléments finis, donc il faut les « corriger » afin qu'il ne soient plus auto-intersectant et que les éléments soient « de bonne qualité » au sens de l'analyse numérique. De plus, l'acquisition n'a pas pu avoir lieu in-vivo sans représenter une surdose en radiation pour l'individu, donc cette géomé-

trie, malgré le travail dû à Zygote Media Group, correspond encore à un cœur non irrigué donc affaissé. Ainsi nous avons en partie redonner une forme adaptée à l'organe par application de surpressions sur les cavités en configuration quasi-statique sur la partie passive hyperelastique de la loi de comportement cardiaque. L'ensemble des opérations de nettoyage de ce maillage a été effectué par l'ingénieur associée E. Phé avec le logiciel *3-matic* de la société *Materialise*.

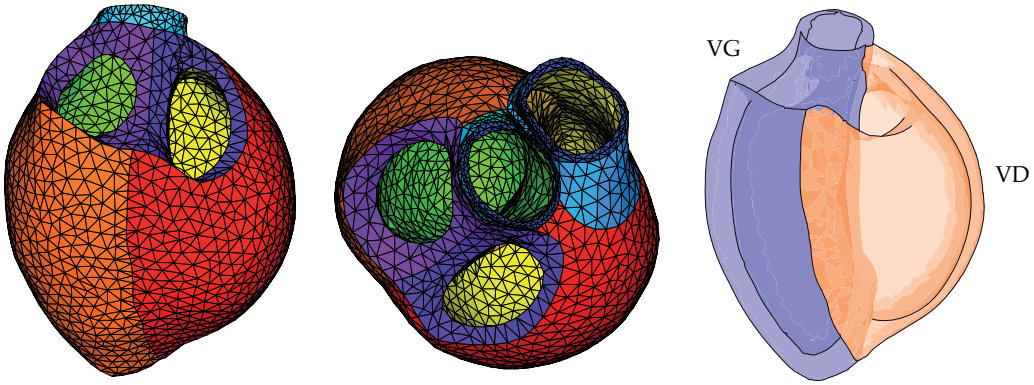


FIGURE 2.34: Un exemple de maillage de calcul mécanique obtenue à partir de la géométrie Zygote.

Afin de générer les régions AHA, nous avons choisit de recaler (i.e. ici appliquer une transformation affine) le cœur tel que le grand axe soit selon Oz. Pour cela, nous avons défini le grand axe comme l'axe principal d'inertie du ventricule gauche en dessous d'un certain plan de coupe délimitant les valves du reste du tissu, là où celui est finalement « proche » (homéomorphe en tout cas) d'un ellipsoïde tronqué. Notons par $\tilde{\Omega}$ (resp. $\tilde{\Gamma}$) les restrictions du domaine volumique (resp. surfacique) en dessous de ce plan :

$$\tilde{\Omega} = \{\mathrm{x} \in \Omega \quad | \quad \mathrm{sign}(\underline{A} - \Pi_P(\underline{A})) \cdot (\underline{\mathrm{x}} - \Pi_P \underline{\mathrm{x}})) = 1\},$$

où $\underline{A}$ désigne l'apex. On définira aussi le vecteur normal $\underline{n}$ au plan opposé à l'apex par

$$\begin{cases} \underline{n}.\underline{v} = 0, \forall \underline{v} \in P \\ (\underline{A} - \Pi_P(\underline{A})) \cdot \underline{n}_P = -1 \end{cases}$$

Une fois ce plan de coupe défini, on peut automatiquement obtenir le grand axe comme l'axe principal d'inertie de plus grande valeur propre pour l'endocarde gauche. Soit O le barycentre de l'endocarde gauche coupé défini à partir des coordonnées $(x_i)_{1,\leq i \leq 3}$ correspondant à (x, y, z) :

$$\underline{O} = \frac{1}{\int_{\tilde{\Gamma}_g} dS} \left(\int_{\tilde{\Gamma}_g} \mathrm{x}_i dS \right)^T_{1 \leq i,j \leq 3}.$$

Le tenseur d'inertie J est alors donné par

$$\underline{\underline{J}} = \left(\int_{\tilde{\Gamma}_g} (\mathrm{x}_i - \mathrm{x}_{Oi})(\mathrm{x}_j - \mathrm{x}_{Oj}) dS \right)_{1 \leq i,j \leq 3}.$$

La direction du grand axe sera alors notée (par convention) $\underline{z}$ et correspond au vecteur normalisé vérifiant

$$\begin{cases} \underline{\underline{J}}.\underline{z} = \min(\mathrm{eig}(\underline{\underline{J}}))\underline{z} \\ \mathrm{sign}(\underline{z}.\underline{n_P}) = 1 \end{cases}$$

Le grand axe Oz est alors l'axe de direction z passant pas O.

Une fois ce grand axe défini, pour toute élévation z on détecte l'intersection entre ventricule droit et gauche afin que les régions 2-3, 8-9 et 14 de la Figure 2.2 correspondent bien au septum.

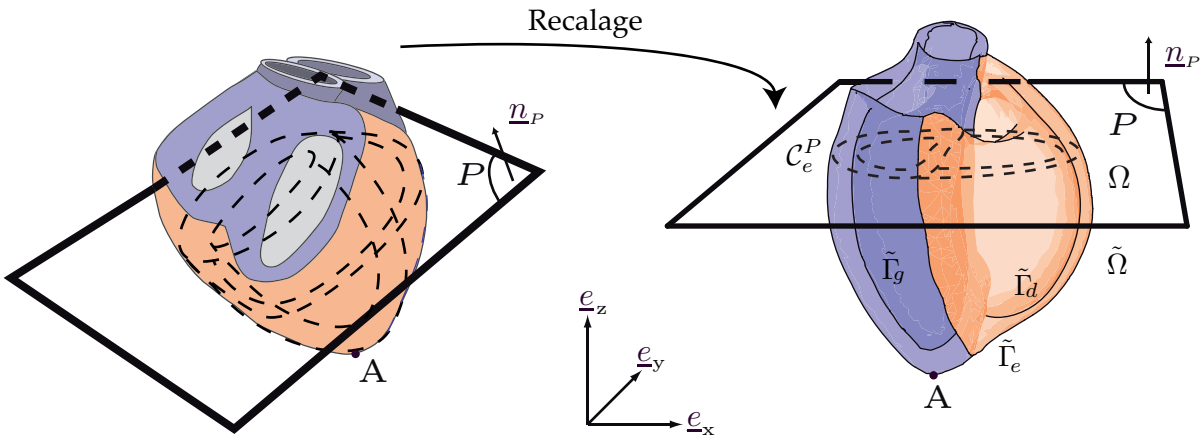


FIGURE 2.35: recalage d'axe et notations géométriques nécessaires pour le modèle de fibres défini ci-dessous

REMARQUE 2.4.1
D'autres géométries sont employées au sein de l'équipe-projet Macs et de Cardiosense3D, notamment à partir de données CT sur des cochons obtenues à l'hôpital Henry Mondor. Dans ce cas, la géométrie a été obtenue après segmentation semi-automatique utilisant le logiciel yav de l'équipe projet Asclepios. Le maillage est alors obtenu par application d'un algorithme de *marching-cube* (introduit par Lorensen et Cline (1987)) sur l'image segmentée.

b. Modèle de fibres

Sur ces deux géométries il nous faut définir un champ de fibres cardiaques. Normalement les fibres cardiaques peuvent être obtenues, soit par dissection, soit par un procédé d'imagerie par résonance magnétique IRM appelé imagerie de tenseur de diffusion. Dans tout milieu, l'agitation thermique entraîne la diffusion brownienne des molécules d'eau. Cette diffusion peut être anisotrope en fonction de directions préférentielles du tissu, et on définit donc un tenseur de diffusion qui caractérise en fait l'anisotropie du tissu. L'IRM de tenseur de diffusion vise à mesurer ce tenseur en bénéficiant de la réponse magnétique différente des molécules d'eau en fonction de leur mouvement (voir aussi Section 2.5.1 pour les principes généraux de l'imagerie magnétique). Cette technique fut initialement mise au point pour le suivi des axones dans le cerveau et a donné lieu à de nombreux travaux, par exemple Fillard et al (2007), sur leur manipulation et traitement. En fait, elle permet aussi la détection des fibres cardiaques et des feuillets (voir par exemple Peyrat et al (2007)) avec typiquement des rapports de proportionnalité entre les valeurs principales du tenseur de l'ordre de

$$\underline{\underline{\mathrm{DTI}}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.65 & 0 \\ 0 & 0 & 0.55 \end{pmatrix}.$$

Cependant les mouvements de l'organe limitent considérablement la mesure dont les temps d'acquisition sont normalement très grands devant le battement cardiaque. Ainsi cette technique n'est pas utilisée à l'heure actuelle en imagerie cardiaque pour des cœur in-vivo. Il faut donc imaginer un modèle de fibres basés sur les connaissances physiologiques à notre disposition (cf. Section 2.2.1) pour obtenir le champ de fibres nécessaire à la simulation numérique du modèle de cœur présenté plus haut.

□ BASE LOCALE – Afin de caractériser géométriquement le champ de fibres, nous utilisons une base locale adaptée à leur orientation. En effet nous avons vu que les fibres étaient définies dans *un plan transverse à l'épaisseur* alors que les feuillets étaient des nappes dont l'orientation se faisaient au travers de l'épaisseur. Mathématiquement, définir ces notions impose la définition de surfaces isoépaisseurs. Ces surfaces sont définies intuitivement dans le cas de notre ventricule gauche analytique puisqu'elles correspondent à une équation du type

$$\forall 0 \leq e \leq e_L, \quad \mathcal{S}(e) = \left\{ (\mathrm{x}, \mathrm{y}, \mathrm{z}) | \frac{\mathrm{x}^2}{(a_L - e)^2} + \frac{\mathrm{y}^2}{(b_L - e)^2} + \frac{\mathrm{z}^2}{(c_L - e)^2} = 1 \right\}.$$

On définit alors une fonction $u \in [-1, 1]$ caractérisant le positionnement dans l'épaisseur par

$$u = 1 - 2e = \begin{cases} -1 \text{ si } \underline{\mathrm{x}} \in \Gamma_g, \Gamma_d \\ 1 \text{ si } \underline{\mathrm{x}} \in \Gamma_e \end{cases}$$

Mais pour une géométrie quelconque comprenant notamment deux ventricules, la situation est plus complexe. On décide alors de calculer la distance d'un point à chacune des surfaces suivantes : épicarde, endocarde gauche, endocarde droit. A partir de ces distances il est possible de définir la séparation entre les deux ventricules. Ainsi le ventricule droit est défini par

$$\begin{aligned} \Omega_d = \Big\{ \underline{x} \in \Omega \,|\, & d(\underline{\mathrm{x}}, \Gamma_d) < d(\underline{\mathrm{x}}, \Gamma_g), \\ & d(\underline{\mathrm{x}}, \Gamma_e) < d(\underline{\mathrm{x}}, \Gamma_g) \\ & \|\underline{n}(\underline{\mathrm{x}}, \Gamma_d) \wedge \underline{n}(\underline{\mathrm{x}}, \Gamma_e)\| < 1/2 \\ & \underline{n}(\underline{\mathrm{x}}, \Gamma_d) \cdot \underline{n}(\underline{\mathrm{x}}, \Gamma_e)) < 0 \Big\}. \end{aligned}$$

où $\underline{n}(.\Gamma_i)$ désigne la direction de la fonction distance à Γ_i. La définition est délicate du fait de la jonction entre les deux ventricules et les différentes situations rencontrées sont résumées Figure 2.36. Bien sûr, on a ensuite

$$\Omega_g = \Omega \backslash \Omega_d,$$

et on peut alors proposer une fonction épaisseur $u \in [-1, 1]$ définie séparément dans chaque ventricule

$$u = 1 - 2\frac{\underline{d}_i^{\text{ext}} \cdot (\underline{d}_i^{\text{ext}} - \underline{d}_i^{\text{int}})}{\|\underline{d}_i^{\text{ext}} - \underline{d}_i^{\text{int}}\|^2}, \quad i = g,\, d$$

où $\underline{d}_i^{\text{ext}}$, resp. $\underline{d}_i^{\text{int}}$, est la distance à l'extérieur, resp. l'intérieur de $\partial\Omega_g$ (ou $\partial\Omega_d$).

Numériquement on utilise la fonction distance à la surface d'un maillage proposée par Baerentzen et Aanaes (2005). Leur approche permet en effet de définir une distance signée telle que **la propriété de signe soit exactement vérifiée sur tout maillage triangulaire ($\mathbb{P}_1$) de surface**.

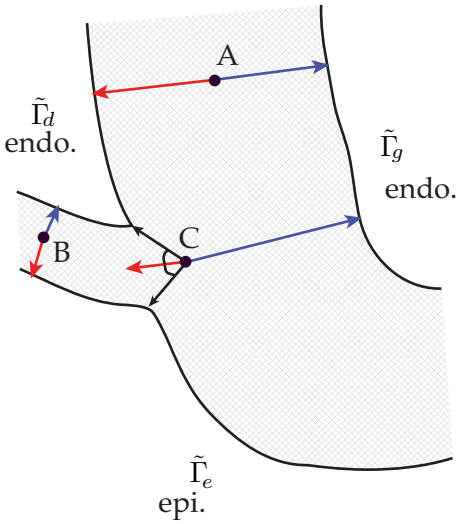


FIGURE 2.36: Différentes configurations dans le calcul des distances à chacune des surface $\tilde{\Gamma}_i$.. Cas particulier notamment du point C à la jonction entre les ventricules permettant de définir une séparation.

On peut définir des surfaces d'isovaleur de cette fonction de vecteur normal associé au point $\underline{\nu}(\underline{x})$. Notons $\underline{u}_\xi$ ce premier vecteur de la base locale. On la complète par deux vecteurs du plan tangent à l'isosurface définies par

$$\begin{cases} \underline{u}_\eta = \dfrac{\Pi_T(\underline{z})}{\|\Pi_T(\underline{z}))\|} \\ \underline{u}_\phi = \underline{u_\xi} \wedge \underline{u_\eta} \end{cases}$$

où le grand axe est recalé sur l'axe $OZ = z$.

Cette base est notée ainsi $(\underline{u}_\xi, \underline{u}_\eta, \underline{u}_\phi)$ par analogie avec la base prolate-sphérique elle-même extension pour les géométries elliptiques de la base associées aux coordonnées sphériques (voir Figure 2.37). En effet ces coordonnées sont souvent utilisées dans d'autres modèles de géométrie cardiaque (notamment par Schmid et Hunter (2006),ou d'autres comme Vallet et al (2006)) car bien adaptées à la forme des ventricules. Dans la suite, on notera U la matrice de changement de base entre cette base « pseudo-prolate » et la base cartésienne $U : (\underline{u}_\xi, \underline{u}_\eta, \underline{u}_\phi) \mapsto (\underline{e}_{\mathrm{x}}, \underline{e}_{\mathrm{y}}, \underline{e}_{\mathrm{z}})$:

$$U(\underline{x}) = \begin{pmatrix} u_{\xi,\mathrm{x}} & u_{\eta,\mathrm{x}} & u_{\phi,\mathrm{x}} \\ u_{\xi,\mathrm{y}} & u_{\eta,\mathrm{y}} & u_{\phi,\mathrm{y}} \\ u_{\xi,\mathrm{z}} & u_{\eta,\mathrm{z}} & u_{\phi,\mathrm{z}} \end{pmatrix}.$$

Le changement de base $(\underline{e}_{\mathrm{x}}, \underline{e}_{\mathrm{y}}, \underline{e}_{\mathrm{z}}) \mapsto (\underline{u}_\xi, \underline{u}_\eta, \underline{u}_\phi)$ est alors évidement donné par U^T.

□ TENSEUR DE DIFFUSION – Nous avons vu que physiologiquement, la fibre cardiaque était située dans le plan tangent à l'isosurface épaisseur avec un angle élévation qui varie entre -60° et $+60^o$ à travers l'épaisseur (voir Figure 2.4 de la Section 2.2.1).

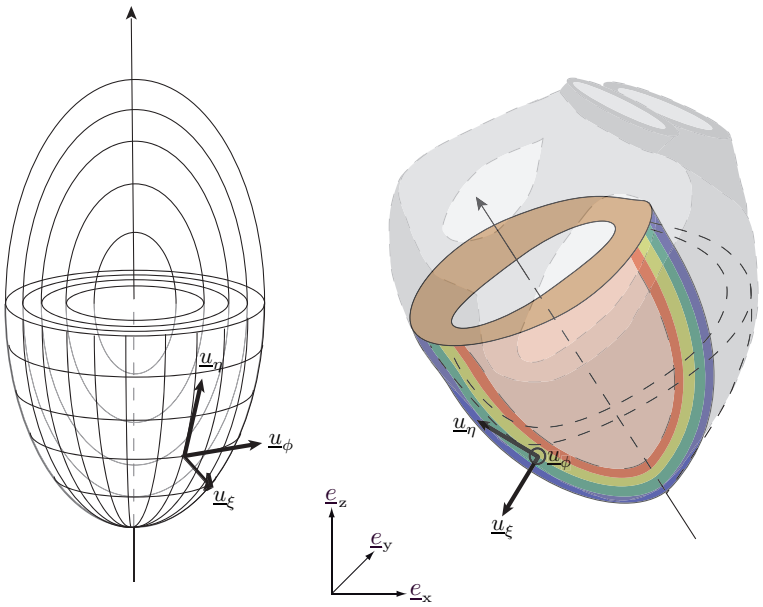


FIGURE 2.37: Coordonnées *prolates* permettant de définir des géométries ellipsoïdales et leur « extension » sur une géométrie cardiaque quelconque à partir d'une carte de distance sur chaque ventricule (3 isovaleurs représentées).

Ainsi, pour tout $\underline{x}$, il existe une matrice $R(\underline{x}) \in \mathrm{SO}(\mathbb{R})$ de rotation dans le plan $(\underline{x}_\xi, \underline{x}_\eta)$

$$R(\underline{x}) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

telle que le tenseur de diffusion $\underline{\underline{T}}$ soit définit en tout point à partir du tenseur diagonal $\underline{\underline{\mathrm{DTI}}}$ par

$$\underline{\underline{T}}(\underline{x}) = U(\underline{x}) \cdot R(\underline{x}) \cdot \underline{\underline{\mathrm{DTI}}} \cdot R(\underline{x})^T \cdot U(\underline{x})^T. \tag{2.48}$$

L'angle $\theta = \theta(u)$ est lui-même défini en fonction de la position dans l'épaisseur par

$$\theta(u) = f(u)\theta_{\text{max}} + (1 - f(u))\theta_{\text{min}}, \tag{2.49}$$

où f est une fonction d'interpolation typiquement

- linéaire : $f(u) = \frac{u+1}{2}, -1 \leq d_e \leq 1$
- sigmoïde : $f(u) = \frac{ath(u)+1}{2}, -1 \leq u \leq 1$

et $\theta_{\text{min}}, \theta_{\text{max}}$ sont des constantes de l'ordre de 60° correspondant à l'angle d'inclinaison respectivement sur l'endocarde et l'épicarde (ou le septum côté ventricule droit). Plus précisément le champ de fibres est donc donné en tout point par

$$\underline{n} = \cos(\theta_{\text{min}} + f(u)(\theta_{\text{max}} - \theta_{\text{min}}))\underline{u_\xi} + \sin(\theta_{\text{min}} + f(u)(\theta_{\text{max}} - \theta_{\text{min}}))\underline{u_\eta}. \tag{2.50}$$

□ CONTRAINTES AU NIVEAU DES VALVES – Cette définition du tenseur de diffusion est largement valide (et d'ailleurs les autres groupes de modélisation cardiaque procèdent de manière analogue) sur $\tilde{\Omega}$ comme le montrent les premiers résultats ci-dessous. Cependant, cette

définition n'est valide que dans les régions globalement homéomorphes à des ellipsoïdes. Près des valves, la nature plus complexe de la géométrie nécessite une approche différente. Ainsi nous choisissons de nous appuyer pour ces régions sur les connaissances physiologiques (celles proposées par Sachse et al (1999) et ses références) qui permettent d'imposer l'orientation au voisinage des valves cardiaques. On dénote par $\mathcal{B}_j,\ 1 \leq j \leq N_{\mathcal{B}}$ les différentes régions autour des valves où on fixe a priori une orientation « soit radiale, soit orthoradiale » au tissu. Cette orientation imposée ainsi que le modèle à 4 paramètres présenté plus haut ($\theta_{\text{min}}, \theta_{\text{max}}$ pour chaque portion de ventricule $\tilde{\Omega}_i$) sont alors suffisantes pour déterminer l'orientation dans la région restante de $\Omega \backslash \tilde{\Omega}$ par **interpolation**.

Cette dernière doit cependant être réalisée sous contraintes pour respecter globalement l'orientation transverse des fibres dans tout le tissu. Ainsi notre interpolation suit deux étapes successives

- Une **interpolation géodésique** sur la surface. On commence donc par interpoler l'orientation des fibres sur les différentes régions manquantes des endocardes et epicardes. On définit $\mathcal{C}_i^P$ comme l'intersection de chacune des surfaces Γ_i avec le plan de coupe puis Γ_i^P une surface étroite (de largeur ϵ) autour de $\mathcal{C}_i^P$ par

$$\forall i, \quad \Gamma_i^P = \{x \in \Gamma_i \mid , d(x, P) < \epsilon\}$$

et, pour tout point $x \in \Gamma_i \backslash \tilde{\Gamma}_i$, on va définir une interpolation surfacique entre Γ^P et les $\mathcal{B}_j$ à partir d'une distance géodésique calculée sur Γ_i entre x et les surfaces où le tenseur de diffusion nous convient. Autrement dit

$$\underline{\underline{T}}(x) = \alpha_i d_G(x, \Gamma_i^P) \underline{\underline{T}}_{\Gamma_i^P} + \sum_j \beta_{i,j}, d_G(x, \mathcal{B}_j) \underline{\underline{T}}_{\mathcal{B}_j}$$

où les $\underline{\underline{T}}_{\Gamma_i^P}$, resp. $\underline{\underline{T}}_{\mathcal{B}_j}$, sont les tenseurs du point le plus proche (au sens de d_G) de Γ_i^P, resp. $\mathcal{B}_j$. Afin d'interpoler correctement sur la surface, d_G est distance géodésique sur les surfaces $\Gamma_i \backslash \tilde{\Gamma}_i$ définie donc par

$$\begin{cases} d_G(x, \Gamma) = \min_s (d_G(x, s, \Gamma)) \\ d_G(x, s, \Gamma) = \min_{\gamma \in \mathcal{C}(x,s,\Gamma)} \int_0^1 \|\gamma'(t)\| dt, \\ \mathcal{C}(x, s, \Gamma) = \{\gamma \in C([0,1]) \mid \gamma(t) \in \Gamma,\ \gamma(0) = x,\ \gamma(1) = s\} \end{cases} \tag{2.51}$$

Les coefficients de l'interpolation sont normalisés. En plus de l'interpolation classique, nous avons aussi utilisé une interpolation associée à une métrique logarithmique. Cette dernière est classique en interpolation de champs de tenseurs symétriques définis positifs (cf. Fillard et al (2007) et références). Cette interpolation, consiste simplement à interpoler les logarithmes matriciels des tenseurs. Du point de vue numérique, la distance géodésique est approchée sur le maillage par une méthode classique de *Fast-Marching*. Celle-ci se fonde sur le fait que la distance géodésique vérifie l'équation *Eikonale* suivante

$$\|\underline{\nabla_x} d_G(x, s, \Gamma)\| = 1,$$

et nous utilisons les techniques et algorithmes de Peyré et Cohen (2005) et mise à disposition par Peyré pour sa résolution approchée dans le cas d'une donnée de type maillage.

- Une **interpolation volumique** dans l'épaisseur. On interpole dans le volume $\Omega \backslash \tilde{\Omega}$ en cherchant un point représentatif sur chaque surface (deux en tout) qui entourent un point $x \in \Omega \backslash \tilde{\Omega}$ afin d'y retrouver les valeurs θ_{min} et θ_{max} désormais spécifiques. Pour ce faire, on pourrait simplement calculer les points les plus proches mais nous préférons utiliser la fonction u précédemment calculée pour remonter sur les surfaces. Autrement dit on résout successivement les deux équations de transport

$$\begin{cases} \dfrac{d\underline{X}}{dt} = \pm \underline{\nabla} u(X), \\ X(0) = x \end{cases} \tag{2.52}$$

jusqu'à intersection avec les surfaces Γ_i en x_{Γ_i}. Une fois ces deux points définis, on considère toujours que dans l'épaisseur la fibre tourne entre les deux valeurs extrêmes via la rotation définie plus haut. Numériquement, vu la dimension caractéristique de l'épaisseur près des valves, on se satisfait de

$$\begin{cases} \dfrac{d\underline{X}}{dt} = \pm \underline{\nabla} u(x), \\ X(0) = x \end{cases} \tag{2.53}$$

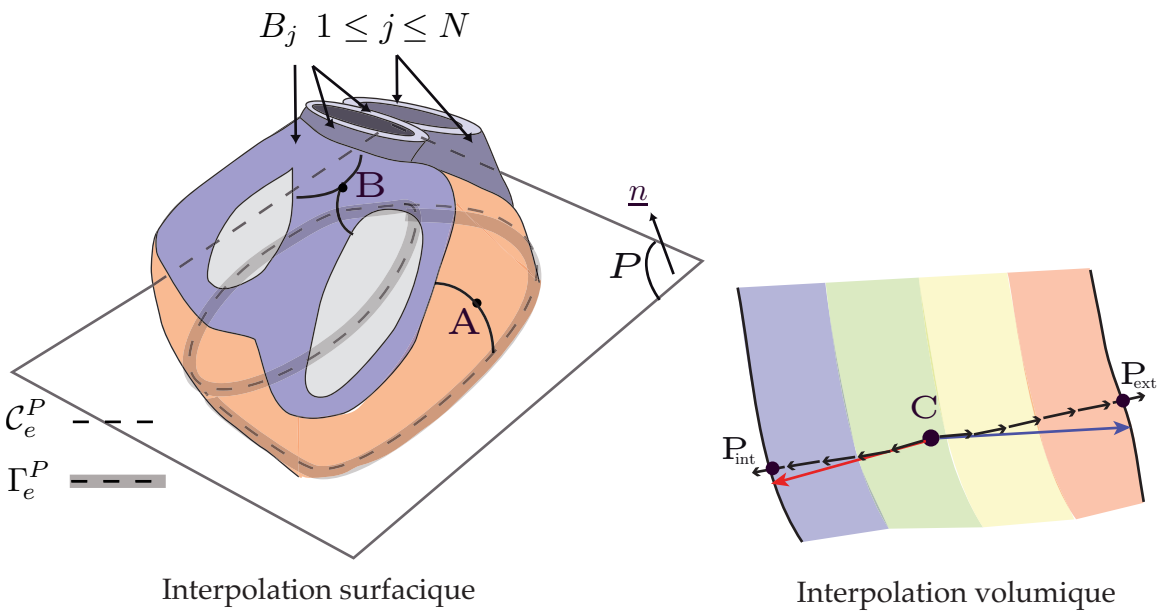


FIGURE 2.38: Les deux types d'interpolation. D'abord l'interpolation sur chaque surface à partir des régions connues. Puis à l'intérieur de l'épaisseur à partir des valeurs surfaciques.

□ RÉSULTATS – Nous présentons l'ensemble des étapes sur le maillage « Zygote » pour une discrétisation de l'ordre de 5000 noeuds. Premièrement, la carte de distance aux endocardes-épicardes permettant la définition de la base locale est présentée Figure 2.39. Cette base locale nous permet d'appliquer sur toute la géométrie le modèle à 4 paramètres. Il apparaît en bleu

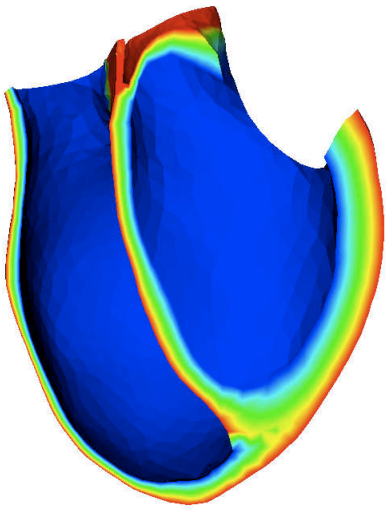


FIGURE 2.39: Distance numérique calculée sur le maillage « Zygote »

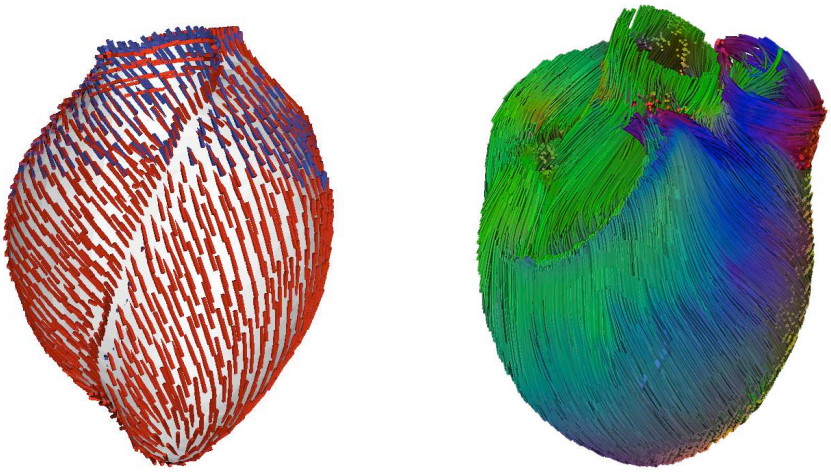


FIGURE 2.40: Modèle de fibres à 4 paramètres (bleu). A gauche, Version avec contraintes sur les valves et reinterpolation (rouge). A droite, visualisation à partir du logiciel MedINRIA (remerciements à l'Equipe-Projet Asclepios).

Figure 2.40. A celui-ci nous superposons en rouge les fibres finales où nous avons en plus imposé l'orientation aux valves et réinterpolé au dessus du plan de coupe P. Nous utilisons pour cela des cartes de distances géodésiques aux régions $\mathcal{B}$ et Γ^P comme celles de la Figure 2.41.

Comme première validation de ce modèle de fibres, nous avons choisi de le comparer, Figure 2.43, à un atlas de fibres de chiens de Peyrat et al (2007) afin de déterminer s'il dispose de suffisamment de paramètres. Nous disposons de la géométrie et des fibres de l'atlas et nous calculons les fibres pour $\theta_g = 60^o$ et $\theta_d = 40^o$. Pour mémoire l'angle à droite est plus faible car il n'est pas calculé par rapport à l'axe d'inertie du ventricule droit mais par rapport à celui de gauche qui lui est incliné d'une dizaine de degrés. La comparaison est ensuite réalisée en projetant la fibre sur les 3 axes ou en comparant les normes de la différence des champs de vecteurs. On pourrait sans doute aller plus loin en choisissant une norme d'erreur mieux adaptée à la comparaison de vecteurs unitaires ou de tenseur à l'image ce qui est fait dans

Peyrat et al (2007).

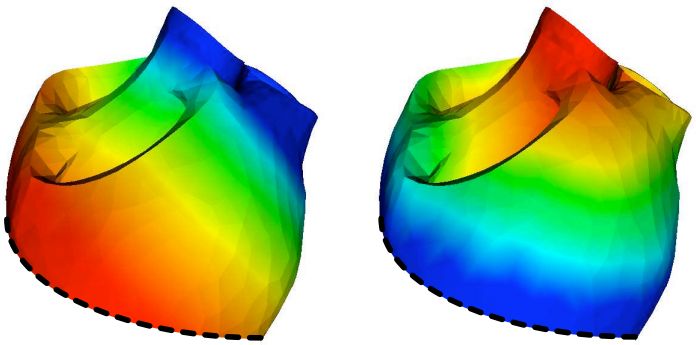


FIGURE 2.41: Exemple de calcul de distance géodésique à la valve sigmoïde (gauche) et à la bande Γ^P (droite)

REMARQUE 2.4.2
Toujours sur le modèle animal[8] de l'hôpital Henri Mondor, il a, en revanche, été possible d'effectuer une acquisition de tenseur de diffusion voir Figure 2.42. Le traitement et la comparaison avec le modèle paramétrique défini ici sont en cours au sein de l'équipe projet Macs. Les valeurs angulaire pour ce modèle ont été choisies en fonction du comportement mécanique global lors de la contraction comparé aux images IRM disponibles sur ce modèle animal. Les valeurs numériques choisies sont $\theta_g = 60$ et $\theta_d = 40$ et ce sont les valeurs de référence que nous avons garder pour la géométrie Zygote.

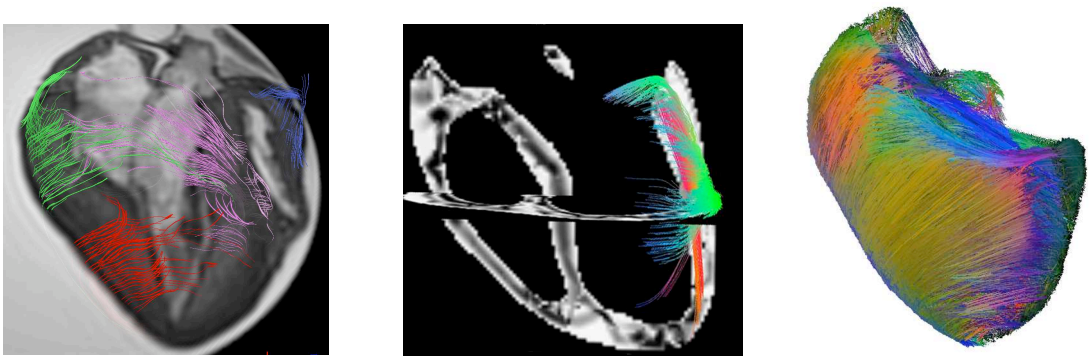


FIGURE 2.42: Exemple de fibres obtenues par DTI sur un cœur de cochon (gauche). Même type de visualisation à partir des DTI synthétiques (centre). Rendu complet du DTI synthétique (droite). (Remerciements à Radomir Chabiniok, Macs, INRIA et Hôpital Henri-Mondor, Créteil)

[8]un modèle animal dans le contexte médical signifie une expérimentation animale. On imagine ici la source de quiproquo potentiel à parler de modèle entre un physicien-numéricien et un clinicien

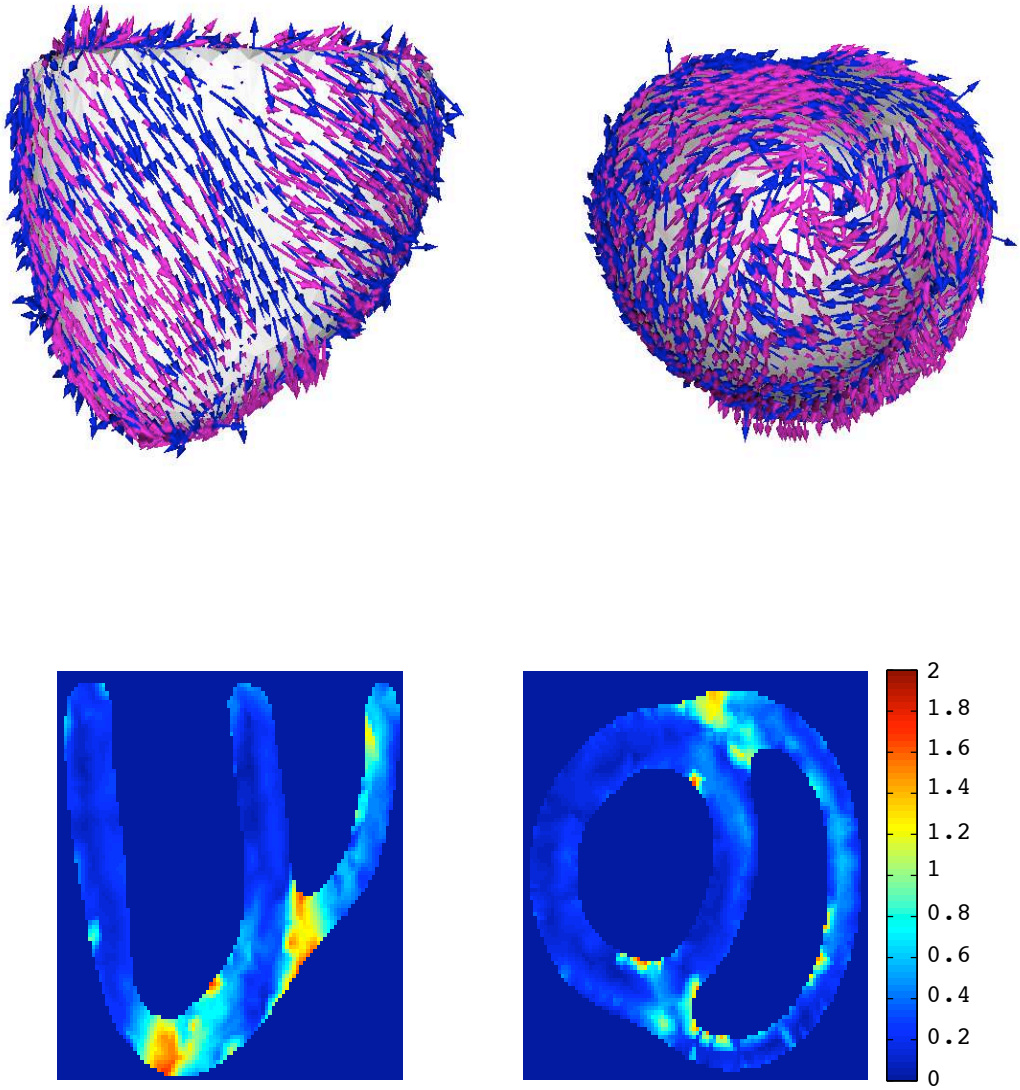


FIGURE 2.43: Comparaison entre le champ de fibres moyen issue de Peyrat et al (2007) et le champ de fibres issu du modèle. Figure du bas : norme d'erreur entre les deux champs. Les erreurs sont concentrées à la jonction VG-VD et près de l'apex où les champs sont très irréguliers.

2.4.2 Discrétisation

a. Discrétisation spatiale

La discrétisation spatiale est très classique puisque les variables de déplacement sont approchées par une formulation éléments finis $\mathbb{P}_1$ (voir Bathe (1996)). Ce choix n'est pas seulement motivé par les facilités d'implémentation qu'il procure. En effet, il est à l'heure actuelle presque impossible de bénéficier de maillages de calcul hexaédriques directement à partir de données réelles donc il faut sans doute écarter les éléments $\mathbb{Q}_p$. On pourrait cependant souhaiter monter en degré sur les éléments triangulaires, notamment pour mieux traiter l'incompressibilité. En effet, notre discrétisation $\mathbb{P}_1$ nous oblige à ne pas dépasser des valeurs de $\nu = 0.49$ sous peine de risques de verrouillage numérique (voir de nombreux exemples dans Bathe (1996)). Cependant la question de l'incompressibilité (exacte) du tissu cardiaque est encore largement ouverte, et il semble qu'en fait elle soit due beaucoup plus au sang contenu dans le tissu qu'au tissu lui même et donc que cette difficulté sera contournée par le modèle de perfusion du tissu. De plus, quand bien même on souhaiterait un tissu très incompressible, l'erreur de 5% faite en se contentant de $\nu = 0.49$ [9] est largement compensée par les autres erreurs de modélisation au point où nous en sommes aujourd'hui.

On note $\underline{y}_h$ le champ approché dans l'espace d'approximation V_h qui prend en compte les conditions aux limites essentielles. La vitesse est elle aussi discrétisée dans V_h et on note $x_h = (\underline{y}_h \ \underline{\dot{y}}_h)^T$. Du point de vue numérique, le nombre de degrés de liberté pour les variables déplacements est N_d correspondant à trois fois le nombre de points moins les degrés de liberté fixés par les conditions aux limites de Dirichlet. On note respectivement Y et $\dot{Y}$ les vecteurs de degrés de liberté correspondant aux champs approchés et $X = (Y \ \dot{Y}) \in R^N$ avec $N = 2N_d$. L'opérateur d'interpolation définissant le champ approché $\underline{y}_h$ à partir du vecteur de degrés de libertés est noté $\mathcal{I}_h$ et, par extension, on utilisera aussi $x_h = \mathcal{I}_h X$ sachant que c'est normalement $\mathrm{diag}([\mathcal{I}_h, \mathcal{I}_h])$.

En élastodynamique linéaire isotrope, la dynamique du mouvement s'exprime alors très simplement en fonction des matrices de masse M, raideur K et viscosité C.

$$\begin{cases} M\ddot{Y} + C\dot{Y} + KY = F_e(t) \\ Y(0),\ \dot{Y}(0) \end{cases} \tag{2.54}$$

où

$$\forall U, V \in R_{N_d}, \quad U^T M V = \int_{\Omega_0} \rho \underline{u}_h \cdot \underline{v}_h \, d\Omega,$$

$$\forall V \in R_{N_d}, \quad (F^e)^T V = \int_{\Omega_0} \rho \underline{f}^e \cdot \underline{v}_h \, d\Omega,$$

$$\forall U, V \in R_{N_d}, \quad U^T K V = \int_{\Omega_0} \underline{\underline{\varepsilon}}(\underline{u}_h) : \underline{\underline{\underline{\underline{A}}}} : \underline{\underline{\varepsilon}}(\underline{v}_h) \, d\Omega,$$

où

$$\underline{\underline{\underline{\underline{A}}}} : \underline{\underline{\varepsilon}}(u_h) = \lambda \, \mathrm{tr}(\underline{\underline{\varepsilon}}(\underline{u}_h)) \underline{\underline{\mathbb{1}}} + 2\mu \underline{\underline{\varepsilon}}(\underline{u}_h),$$

[9] l'erreur sur la solution est linéaire en ν en élasticité linéaire

et pour la viscosité dans les différents cas qui nous intéressent

$$\forall U, V \in R_{N_d}, \quad U^T C V = \int_{\Omega_0} \underline{\underline{\varepsilon}}(\underline{u}_h) : \underline{\underline{\underline{\underline{A^v}}}} : \underline{\underline{\varepsilon}}(\underline{v}_h)\, d\Omega.$$

En linéaire, ceci nous conduira dans la plupart des études à choisir finalement $C = \eta K$ qui est un cas particulier de l'amortissement de Rayleygh $C = \alpha M + \beta K$ si souvent utilisé pour les simplifications qu'il engendre dans les études modales malgré une définition un peu *ad hoc*. La dynamique (2.54) du deuxième ordre peut se réécrire simplement sous forme du premier ordre. On obtient alors le système de Cauchy sur R^N,

$$\begin{cases} \dot{X} = AX + R \\ X(0) \end{cases} \tag{2.55}$$

la matrice A et le second membre R étant définis par

$$A = \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{pmatrix} \quad R = \begin{pmatrix} 0 \\ M^{-1}F \end{pmatrix}.$$

En élastodynamique non-linéaire, les efforts intérieurs seront définis par l'opérateur

$$\forall U, V \in R_{N_d}, \quad F^i(U)^T V = \int_{\Omega_0} \underline{\underline{\Sigma}} : d_{\underline{y}} \underline{\underline{e}}(\underline{v}_h) d\mathrm{x},$$

tels que la dynamique devienne

$$\begin{cases} M\ddot{Y} + C(Y)\dot{Y} + F^i(Y) = F^e(t) \\ Y(0),\ \dot{Y}(0) \end{cases} \tag{2.56}$$

Et on notera $K(Y) = \mathrm{d}_Y F^i(Y)$, et $K = \mathrm{d}_Y F^i(0)$ c'est-à-dire le linéarisé dans la configuration de référence. Sous forme du premier ordre, on notera sans plus de précision

$$\begin{cases} \dot{X} = A(X, t) \\ X(0) \end{cases} \tag{2.57}$$

Pour le modèle de contraction cardiaque complet, il faut ajouter à la dynamique classique la résolution du modèle de contraction active ainsi que la dynamique des variables de pressions dans les efforts extérieurs. Ainsi, en chaque point d'intégration où il faut évaluer les efforts intérieurs, les dynamiques de τ_c, k_c et σ_c de (2.41) sont résolues conduisant à autant de valeurs de e_c que de points d'intégrations (fixés à 4 par éléments, $N_{\text{pi}} = 4N_{\text{elt}}$). Pour la dynamique des variables de pressions, on doit notamment calculer les volumes des cavités et débits sanguins comme des termes d'intégrale sur chaque cavité interne fermée. Pour ce faire nous définissons la fermeture de ces cavités au niveau des valves à l'aide du barycentre de chaque contour. Ces points sont donc contraints à rester toujours les barycentres de leur contours respectifs définissant ainsi une condition aux limites en contrainte linéaire. On définit ensuite l'opérateur S calculant la pression constante P_V sur la cavité ouverte sous la forme

$$P_V S(Y)^T V = P_V \int_{\partial \mathcal{C}_i} \underline{\nu} \cdot \underline{\underline{F}}^{-1} \cdot \underline{v} J dS,$$

dont la différentielle est donnée par le calcul de la différentielle du volume. De même on aura $S(Y)^T\dot{Y}$ pour le calcul du débit si les valves sont fixées par les conditions aux limites. Si les valves bougent, en revanche, il faut prendre en compte toute la cavité fermée $\partial\mathcal{C}_i$ et on note alors $\bar{S}(Y)^T\dot{Y}$ cette variation de volume au cours du temps prenant en compte la contrainte barycentrique. Evidemment, ce calcul est mené pour chaque cavité $1 \leq i \leq N_{\text{cav}}$

Le système complet a donc finalement pour variable d'état

$$X = (Y, \dot{Y}, e_c^{\text{pi}}, \tau_c^{\text{pi}}, k_c^{\text{pi}}, \{P_i^v, P_i^{ar}\}_{1\leq i\leq N_{\text{cav}}})^T.$$

Là encore on note $A(.)$ la dynamique globale non-linéaire pour ce vecteur d'état que nous allons préciser après discrétisation en temps.

b. Discretisation en temps

Reste à définir la discrétisation en temps pour pouvoir implémenter le modèle. On suppose que l'intervalle $[0, T]$ est subdivisé en N_T intervalles de même longueur Δt. On note $X_n = \begin{pmatrix} Y_n \\ \dot{Y}_n \end{pmatrix}$ la variable discrète[10] approchant $X(n\Delta t)$. Nous avons choisi un schéma en temps de type Newmark centré utilisé en mécanique pour ses propriétés de conservativité dans les modèles linéaires (cf. Bathe (1996)).

□ SCHÉMA DE NEWMARK CENTRÉ EN ÉLASTODYNAMIQUE LINÉAIRE – On définit donc

$$\begin{cases} \ddot{Y}_{n+1/2} \overset{def}{=} \dfrac{\ddot{Y}_{n+1} + \ddot{Y}_n}{2} = \dfrac{\dot{Y}_{n+1} - \dot{Y}_n}{\Delta t} \\ \dot{Y}_{n+1/2} \overset{def}{=} \dfrac{\dot{Y}_{n+1} + \dot{Y}_n}{2} = \dfrac{Y_{n+1} - Y_n}{\Delta t} \end{cases} \tag{2.58}$$

Ce choix réduit à chaque pas de temps la résolution de (2.54) à

$$\begin{cases} \dfrac{Y_{n+1} - Y_n}{\Delta t} = \dfrac{\dot{Y}_{n+1} + \dot{Y}_n}{2} \\ M\dfrac{\dot{Y}_{n+1} - \dot{Y}_n}{\Delta t} + C\dfrac{\dot{Y}_{n+1} + \dot{Y}_n}{2} + K\dfrac{Y_{n+1} + Y_n}{2} = F^e_{n+\frac{1}{2}} \\ Y_{n=0} \quad \dot{Y}_{n=0} \end{cases} \tag{2.59}$$

où $F^e_{n+\frac{1}{2}} = \dfrac{F^e((n+1)\Delta t) + F^e(n\Delta t)}{2}$ est la demi-somme des forces suivant la règle des trapèzes. On définit

$$A_1 = \begin{pmatrix} \frac{1}{\Delta t} & -\frac{1}{2} \\ \frac{1}{2}K & \frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}, \quad A_0 = \begin{pmatrix} \frac{1}{\Delta t} & \frac{1}{2} \\ -\frac{1}{2}K & -\frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}, \quad \tilde{R}_{n+1} = \begin{pmatrix} 0 \\ F^e_{n+\frac{1}{2}} \end{pmatrix},$$

permettant de préciser l'opérateur de transition $A_{n+1|n} = A_1^{-1}A_0$ et $R_{n+1} = A_1^{-1}\tilde{R}_{n+1}$ tel que

$$A_1 X_{n+1} = A_0 X_n + \tilde{R}.$$

Du point de vue algorithmique on résout plutôt le système suivant (à Y_n et $\dot{Y}_n$ connus)

$$\begin{cases} \dot{Y}_{n+1} = \frac{2}{\Delta t}(Y_{n+1} - Y_n) - \dot{Y}_n \\ \left(\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M\right) Y_{n+1} = \left(-\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M\right) Y_n + \frac{2}{\Delta t}M\dot{Y}_n + F^e_{n+\frac{1}{2}} \\ Y_{n=0},\ \dot{Y}_{n=0} \end{cases} \tag{2.60}$$

[10] $\dot{Y}_n$,$\ddot{Y}_n$ est une notation pour l'approximation de $\dot{Y}(n\Delta t)$,$\ddot{Y}(n\Delta t)$ et ne correspond pas à une dérivée en temps

La première équation du système ci-dessus ne fait intervenir que des coefficients scalaires et donc $\dot{Y}_{n+1}$ s'obtient directement à partir de Y_{n+1} lui-même solution d'un système à N_d degrés de liberté. Donc on ne doit inverser qu'un système linéaire creux de taille N_d. Nous utilisons pour ce faire des solveurs directs soit *open-source* (`MUMPS`, `UMFPACK`) soit libre d'accès dans le cadre universitaire (`Pardiso`). Ces solveurs peuvent être d'algorithmique et d'implémentation assez différentes mais nous garderons en tête la complexité usuelle de l'inversion de tels systèmes en $N_d^2.L$ où L est la largeur de bande de la matrice creuse $(\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M)$.

L'énergie discrète du système est ainsi exactement dissipative en l'absence de forces extérieures puisque, en multipliant la seconde équation de (2.59) par $\frac{1}{2}(\dot{Y}_{n+1} + \dot{Y}_n)^T$, on obtient en effet

$$\mathcal{E}_{n+1} - \mathcal{E}_n = \Delta t \Big(\frac{\dot{Y}_{n+1} + \dot{Y}_n}{2} \Big)^T C \frac{\dot{Y}_{n+1} + \dot{Y}_n}{2} > 0,$$

avec

$$\mathcal{E}_n = \tfrac{1}{2}\, \dot{Y}_n^T M \dot{Y}_n + \tfrac{1}{2}\, Y_n^T K Y_n.$$

Ce schéma vérifie donc une propriété de stabilité L_2.

□ SCHÉMA EN ÉLASTODYNAMIQUE NON-LINÉAIRE – En grand déplacements, les travaux de Hauret et Le Tallec (2006) montrent comment le schéma classique de Newmark doit être adapté pour respecter une conservativité exacte. Par exemple pour les contraintes, on peut choisir en principe entre plusieurs stratégies pour représenter ce que nous définirons par

$$F^i_{n+\frac{1}{2}}(U)^T V = \int_{\Omega_0} \underline{\underline{\Sigma}}_{n+\frac{1}{2}}(\underline{u}_h) : \mathrm{d}_{\underline{y}} \underline{\underline{e}}_{n+\frac{1}{2}} \cdot \underline{v}_h \, d\mathrm{x}$$

Les choix possibles sont :

- le point milieu des déformations conduisant à

$$F^i_{n+\frac{1}{2}} \stackrel{def}{=} F^i(Y_{n+\frac{1}{2}}),$$

- le point milieu des contraintes

$$F^i_{n+\frac{1}{2}} \stackrel{def}{=} \frac{1}{2} \left(F^i(Y_{n+1}) + F^i(Y_n) \right),$$

- toute combinaison linéaire entre ces alternatives dont on sait d'après les formules de Taylor qu'une est exactement conservative (voir Hauret et Le Tallec (2006) et références pour plus de détails)

$$F^i_{n+\frac{1}{2}} \stackrel{def}{=} \alpha F^i(\beta Y_{n+1} + (1-\beta) Y_n) + (1-\alpha) F^i((1-\beta) Y_{n+1} + \beta Y_n).$$

Nous avons décidé d'implémenter le premier pour les simplifications qu'elles induisent dans la mise en oeuvre tout en limitant l'erreur de conservativité de l'énergie mécanique en $\mathcal{O}(\Delta t^2)$. Ce choix est raisonnable tant que les efforts d'accélération ne sont pas dominants ou que le pas de temps pour la résolution n'est pas trop grand. Or, l'activation étant très rapide, nous verrons que pour capter ce phénomène nous serons contraints de choisir un pas de temps relativement bas ($1ms$).

Une attention particulière a cependant été accordée à la viscosité pour garder un terme exactement dissipatif à partir du choix simple de pseudo-potentiel visqueux présenté Section 2.3.2.b. En effet, pour $\underline{\underline{\Sigma}}^v = \eta \underline{\underline{\dot{e}}}$, on doit vérifier que testée sur une fonction-test correspondant à la vitesse au point milieu, la contribution de la formulation variationnelle correspondant à la viscosité passive est strictement négative. On doit donc définir $\underline{\underline{\dot{e}}}_{n+\frac{1}{2}}$ et $\mathrm{d}_{\underline{y}} e_{n+\frac{1}{2}}$ tels que

$$\int_\Omega \eta \underline{\underline{\dot{e}}}_{n+\frac{1}{2}} : \mathrm{d}_{\underline{y}} \underline{\underline{e}}_{n+\frac{1}{2}} \cdot \frac{\underline{y}_{h,n+1} - \underline{y}_{h,n}}{\Delta t} \, d\Omega > 0.$$

En remarquant que $\underline{\underline{\dot{e}}} = \mathrm{d}_{\underline{y}} \underline{\underline{e}} \cdot \underline{\dot{y}}$, on voit bien qu'il suffit de considérer simplement le même $\mathrm{d}_{\underline{y}} \underline{\underline{e}}$ de chaque côté du produit de contraction et toujours $\underline{\dot{y}} = \frac{\underline{y}_{h,n+1} - \underline{y}_{h,n}}{\Delta t}$. De plus, une simplification importante intervient lorsque la différentielle du tenseur des déformations est évaluée au déplacement milieu. En effet, on vérifie

$$\begin{aligned}
\mathrm{d}_{\underline{y}} \underline{\underline{e}}_{n+\frac{1}{2}} \cdot \underline{\dot{y}}_{n+\frac{1}{2}} &= \frac{1}{2}(\underline{\underline{F}}(\underline{y}_{n+\frac{1}{2}})^T \cdot \mathrm{d}\underline{\underline{F}} \cdot \underline{\dot{y}}_{n+\frac{1}{2}} + (\mathrm{d}\underline{\underline{F}} \cdot \underline{\dot{y}}_{n+\frac{1}{2}})^T \cdot \underline{\underline{F}}(\underline{y}_{n+\frac{1}{2}})) \\
&= \frac{1}{2}\left(\left(\frac{\underline{\underline{F}}_{n+1} + \underline{\underline{F}}_n}{2} \right)^T \cdot \frac{\underline{\underline{F}}_{n+1} - \underline{\underline{F}}_n}{\Delta t} + \left(\frac{\underline{\underline{F}}_{n+1} - \underline{\underline{F}}_n}{\Delta t} \right)^T \cdot \frac{\underline{\underline{F}}_{n+1} + \underline{\underline{F}}_n}{2} \right) \\
&= \frac{1}{\Delta t}(\underline{\underline{F}}_{n+1}^T \cdot \underline{\underline{F}}_{n+1} - \underline{\underline{F}}_n^T \cdot \underline{\underline{F}}_n) = \frac{\underline{\underline{e}}_{n+1} - \underline{\underline{e}}_n}{\Delta t},
\end{aligned} \tag{2.61}$$

permettant d'exprimer la viscosité simplement en fonction des déplacement $\underline{y}_{h,n}$ et $\underline{y}_{h,n+1}$.

De manière générale en élastodynamique, ces différents choix conduisent à un schéma en temps complet non-linéaire

$$\begin{cases} \mathcal{F}(Y_{n+1}, Y_n, n) = 0 \\ \dot{Y}_{n+1} = \frac{2}{\Delta t}(Y_{n+1} - Y_n) - \dot{Y}_n, \end{cases}$$

résolu par une méthode de Newton (cf. Figure 1.3 du premier chapitre)

$$\mathrm{d}_{Y_{n+1}} \mathcal{F}(Y_{n+1}^{(i)}, Y_n, n) \cdot (Y_{n+1}^{(i+1)} - Y_{n+1}^{(i)}) = -\mathcal{F}(Y_{n+1}^{(i+1)}, Y_n, n).$$

On notera que ce système n'est pas symétrique donc nécessite d'utiliser des solveurs qui ne sont pas limités sur les opérateurs symétriques.

□ MODÈLE DE CONTRACTION – Ce que nous venons de présenter vaut en toute généralité mais, dans le cadre du modèle cardiaque présenté, nous avons vu qu'il fallait résoudre un problème local donné par un système différentielle sur k_c et τ_c pour obtenir la loi de comportement aux points d'intégration. Pour ce faire, on peut soit suivre une règle des trapèzes classiques, soit utiliser l'énergie linéique macroscopique définie en (2.25) (de dynamique (2.26)) pour contrôler k_c et τ_c. On discrétise alors les variables $\frac{\tau_c}{\sqrt{k_c}}$ et $\sqrt{k_c}$ par un schéma point milieu

$$\begin{aligned}
\left(\frac{\tau_c}{\sqrt{k_c}}\right)_{n+1} - \left(\frac{\tau_c}{\sqrt{k_c}}\right)_n &= (\sqrt{k_c})_{n+1/2}((e_c)_{n+1} - (e_c)_n) + \frac{\sigma_0 \Delta t (|u|_+)_{n+1/2}}{(\sqrt{k_c})_{n+1/2}} \\
&- \left(\frac{k_0 \Delta t (|u|_+)_{n+1/2}}{(k_c)_{n+1/2}} + \Delta t(|u|_{n+1/2} + \alpha |(e_c)_{n+1} - (e_c)_n|) \right) \left(\frac{\tau_c}{2\sqrt{k_c}} \right)_{n+1/2}
\end{aligned} \tag{2.62}$$

$$(\sqrt{k_c})_{n+1} - (\sqrt{k_c})_n = \frac{k_0 \Delta t (|u|_+)_{n+1/2}}{2(\sqrt{k_c})_{n+1/2}} - (\Delta t |u|_{n+1/2} + \alpha |(e_c)_{n+1} - (e_c)_n|) \frac{(\sqrt{k_c})_{n+1/2}}{2} \tag{2.63}$$

Comme expliqué précédemment, ce système est résolu pour chaque point d'intégration. Ainsi pour résoudre un problème d'élastodynamique avec cette loi de contraction, il faut résoudre par une méthode de Newton un système non-linéaire de la forme

$$\mathcal{F}^{BS}(Y_{n+1}, Y_n, (e_c)_{n+1}, (e_c)_n, n) = \begin{pmatrix} \mathcal{F}(Y_{n+1}, Y_n, (e_c)_{n+\frac{1}{2}}), n) \\ \mathcal{S}(e_{n+1}, e_n, (e_c)_{n+1}, (e_c)_n, (k_c)_{n+1}, (k_c)_n, (\tau_c)_{n+1}, (\tau_c)_n, n) \end{pmatrix} = 0,$$

où $\mathcal{F}$ représente le résidu issu de la formulation de la dynamique et $\mathcal{S}$ est le résidu issu du couplage micro-macro de la dynamique avec le modèle de contraction « Bestel-Clément-Sorine »

$$\mathcal{S} = d((e_c)_{n+\frac{1}{2}})(\tau_c)_{n+\frac{1}{2}} + \mu \dot{e}_{c_{n+\frac{1}{2}}} - E_s \frac{e_{n+\frac{1}{2}} - (e_c)_{n+\frac{1}{2}}}{(1 + 2(e_c)_{n+\frac{1}{2}})^3}(1 + 2e_{n+\frac{1}{2}}).$$

Après substitution de 2.62 et 2.63, le système tangent s'écrit alors sous la forme

$$\begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} Y_{n+1}^{(i+1)} - Y_{n+1}^{(i)} \\ (e_c)_{n+1}^{(i+1)} - (e_c)_{n+1}^{(i)} \end{pmatrix} = \begin{pmatrix} \mathcal{F} \\ \mathcal{S} \end{pmatrix},$$

avec $K_{22} \in \mathbb{M}_{N_{\text{pi}}}(\mathbb{R})$ diagonale puisque le modèle microscopique est obtenu par résolution d'une ODE et pas une EDP. Par conséquent K_{22} s'inverse simplement et on résout en fait un système uniquement défini sur les déplacements grâce à un complément de Schur

$$(K_{11} - K_{12}K_{22}^{-1}K_{21})(Y_{n+1}^{(i+1)} - Y_{n+1}^{(i)}) = \mathcal{F} - K_{12}K_{22}^{-1}\mathcal{S},$$

suivi de

$$(e_c)_{n+1}^{(i+1)} - (e_c)_{n+1}^{(i)} = K_{22}^{-1}(\mathcal{S} - K_{21}(Y_{n+1}^{(i+1)} - Y_{n+1}^{(i)})).$$

□ MODÈLE COMPLET – Reste enfin à traiter les variables de pression pour obtenir le système cardiaque complet. Tout d'abord la relation algébrique associée au modèle de valve est discrétisée sous la forme

$$\dot{V}_{n+\frac{1}{2}} = \frac{1}{2}\left(f_\delta(P_{n+1}^v, P_{n+1}^{ar}, P_{n+1}^{at}) + f_\delta(P_n^v, P_n^{ar}), P_n^{at}\right),$$

mais comme cette loi exprimant le passage entre différents régimes est très peu régulière on choisit de résoudre le Newton en fixant l'une ou l'autre des phases en procédant à une adaptation du pas de temps si le système ne peut converger dans aucune d'entre elles. Le choix de la discrétisation de $\dot{V}_{n+\frac{1}{2}}$ n'est pas non plus anodin, car elle traduit les échanges énergétiques entre la mécanique et le système artériel lorsque la valve est ouverte. On veut qu'au moins en l'absence de mouvement au niveau des valves les échanges soient exacts au niveau discret. Si on se limite à la dynamique couplée au Windkessel artériel, on a (en se limitant au premier terme de débit) lorsque la valve est ouverte (à modèle de valve parfait)

$$\begin{cases} M\ddot{Y} + C(Y)\dot{Y} + F^i(Y) = -P^{ar}S(Y) \\ \dfrac{L_w C_w}{R_w}\ddot{P}^{ar} + C_w\dot{P}^{ar} + \dfrac{1}{R_w}(P^{ar} - P_{sv}) = S(Y)^T\dot{Y} \end{cases}$$

si bien qu'en multipliant la première par $\dot{Y}$ et la deuxième par P_{ar}, les termes en $PS(Y)^T\dot{Y}$ s'annulent. Pour discrétiser en temps le Windkessel, on choisit donc en cohérence avec le reste

du schéma une règle des trapèzes pour la pression (un Newmark 0D plus exactement) et les termes contenant $S(Y)$ tels que l'énergie échangée discrète soit (après multiplication par le conjugué énergétique discret)

$$P^{ar}_{n+\frac{1}{2}} S(Y_{n+\frac{1}{2}})^T \dot{Y}_{n+\frac{1}{2}} = \Big(\frac{P^{ar}_{n+1} + P^{ar}_n}{2}\Big) S\Big(\frac{Y_{n+1} + Y_n}{2}\Big)^T \Big(\frac{\dot{Y}_{n+1} + \dot{Y}_n}{2}\Big).$$

2.4.3 Calibration du modèle

□ LOI DE COMPORTEMENT MÉCANIQUE – Lors de la phase de remplissage, et plus précisément lors de la contraction des oreillettes, seul le comportement passif joue ce qui nous permet de le calibrer en premier. La raideur passive est alors réglée pour que l'augmentation de volume corresponde à l'accroissement de pression et la viscosité en fonction de la constante de temps du remplissage.

Une fois l'élément passif calibré, on peut ajuster la contractilité en systole (éjection). Pour cela, on fixe une pression d'éjection constante à une valeur physiologique nominale ($12000Pa$ pour le ventricule gauche typiquement et deux fois moins pour le droit). La contractilité permet alors de varier le volume d'éjection pour obtenir la fraction d'éjection souhaitée (de l'ordre de 60%). La loi de Laplace indique en effet que la pression équilibre la tension active moins la compression passive.

Nous ne nous attardons pas sur la calibration indispensable, mais largement découplée, des autres éléments de la loi de contraction Bestel-Clément-Sorine qui procède d'une analyse au niveau de la fibre en 1D (cf. Section 2.6 et les indications de calibration de Chapelle et al (2001) ; Sainte-Marie et al (2006)).

□ WINDKESSEL – Commençons par introduire les constantes temporelles de ce système

$$\tau_p = R_p C_p = \frac{1}{\delta_p}, \quad \tau_d = R_d C_d = \frac{1}{\delta_d}.$$

Pour la circulation artérielle, on a alors en pratique

$$\tau_p \ll \tau_d \quad \text{et} \quad R_p < R_d, \quad C_p < C_d.$$

Ces ordres de grandeurs nous permettent d'analyser le comportement du double Windkessel en fonction du petit paramètre

$$\epsilon_w = \frac{\tau_p}{\tau_d},$$

pour extraire deux dynamiques principales : une lente et une rapide. Soit $\alpha = \frac{C_p}{C_d} < 1$, les valeurs et vecteurs propres du système (2.40) sont au premier ordre en ϵ_w

$$\lambda_w^{\text{slow}} = -\frac{\epsilon_w}{1+\alpha}\delta_p = -\frac{1}{R_d(C_d + C_p)}, \quad P^{ar} - P^d = 0,$$

et

$$\lambda_w^{\text{fast}} = -(1+\alpha)\delta_p = -\frac{1}{R_p C_d} - \frac{1}{R_p C_p}, \quad \alpha P^{ar} + P^d = 0.$$

Le premier mode correspond simplement à l'évolution en temps long où les deux capacitances sont en parallèle et la résistance R_p est transparente. Le second correspond lui à la dynamique

rapide où les deux tronçons artériels proximal et distal rétablissent leur écart de quantité de fluide stockée.

Pour la calibration, on commence par la dynamique lente. La constante de temps $\tau_d(1+\alpha)$ est fixée pour que la décroissance de la pression aortique après fermeture des valves soit compatible avec la périodicité. Puis en maintenant cette constante de temps, on règle R_d vis-à-vis de l'augmentation de pression distale pendant l'éjection. On obtient alors la pression de fermeture de valve aortique recherchée. On passe alors à la partie rapide du windkessel pour laquelle le réglage de R_p permet une pression aortique maximale physiologique en systole à constante de temps de l'ordre de quelques millisecondes.

2.4.4 Simulations numériques

Nous pouvons maintenant présenter une simulation complète du battement cardiaque préparée pour l'article en cours de rédaction Chapelle et al (2008). La simulation commence avec la contraction de l'oreillette, non modélisée électriquement et correspondant simplement à une surpression imposée de P_{at}. L'activation électrique ventriculaire débute après $60ms$ et est initiée sur les deux endocarde. Elle est représenté par la carte de couleur sur les Figures 2.47 et suivantes avec un échelle comprise entre $-80\ mV$ et $20\ mV$. Elle correspond de plus à l'ECG Figure 2.44 réalisé à partir de la matrice de transfert initialement calculée sur le maillage de thorax présenté à droite de la figure. La discrétisation et les valeurs numériques utilisées pour cette simulation électrique sont développées dans Boulakia et al (2008). Les résultats méca-

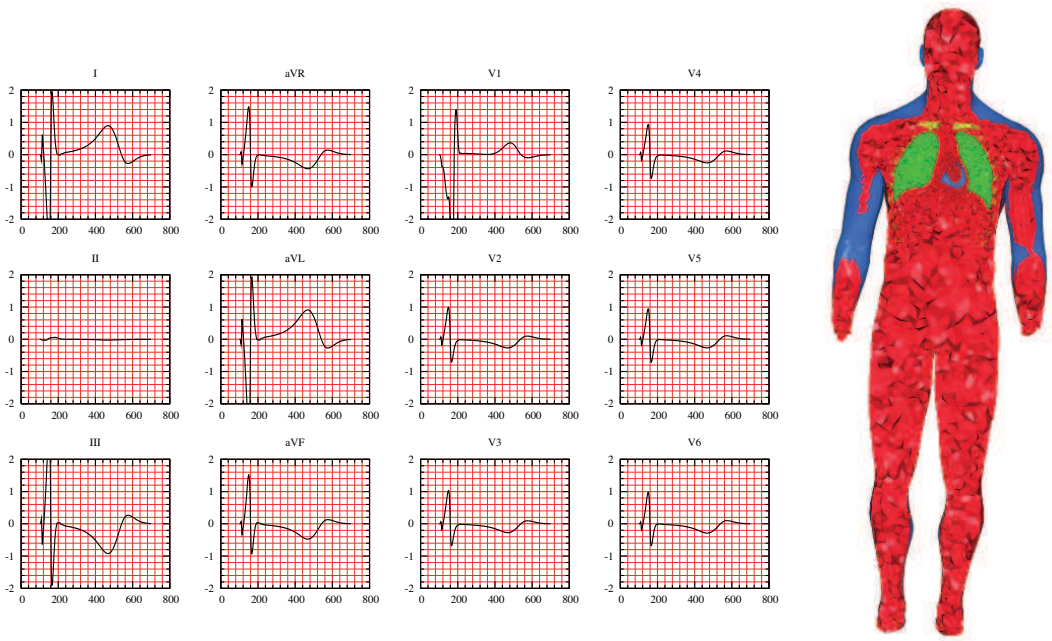


FIGURE 2.44: ECG correspondant à l'activation électrique permettant la contraction cardiaque

niques sont synthétisés sur les Figure 2.47 et pour des coupes petit axe et grand axe Figure 2.49 et Figure 2.48. On constate sur ces dernières deux caractéristiques classiques du battement cardiaque, le phénomène de piston sur la coupe grand axe et le phénomène de torsion de quelques degrés (« *twist* ») sur la coupe petit axe. Sur chaque vignette l'ombre de la configuration de référence est laissée en noir. Elle permet de constater où sont situées, un peu au dessus de l'apex, les conditions aux limites élastiques. Ces simulations correspondent aux indicateurs physiologiques, volumes et pressions, de la Figure 2.45 qui sont évidemment à comparer à ceux de l'illustration Figure 2.5. La volume du ventricule droit est un peu plus grand mais la fraction d'éjection est comparable à celle du gauche comme attendu. On remarque enfin le rôle différent des deux ventricules sur la Figure 2.46 où le diagramme PV prouve que le ventricule gauche est véritablement la pompe principale. Cette dernière figure est elle aussi à comparer avec le schéma Figure 2.8.

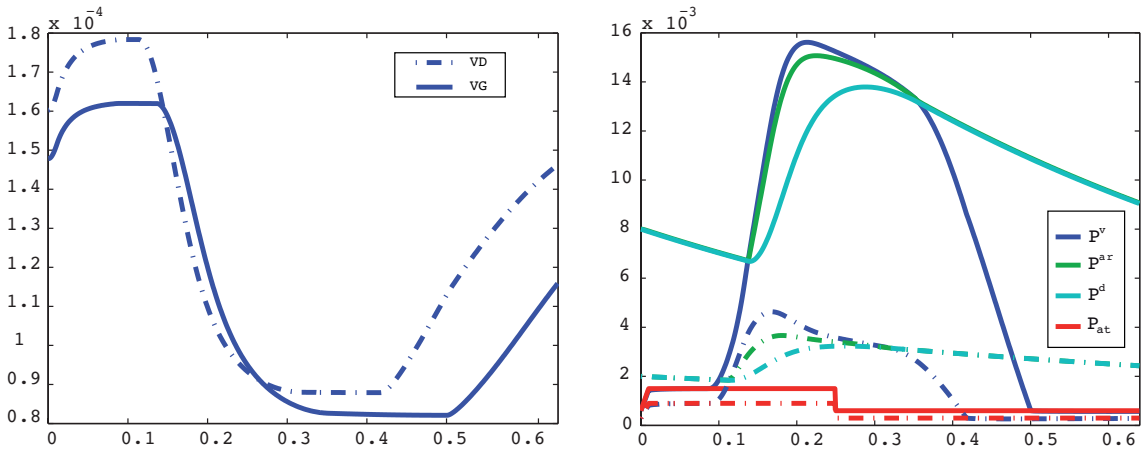


FIGURE 2.45: Courbes de volumes et de pressions pour la simulation cardiaque

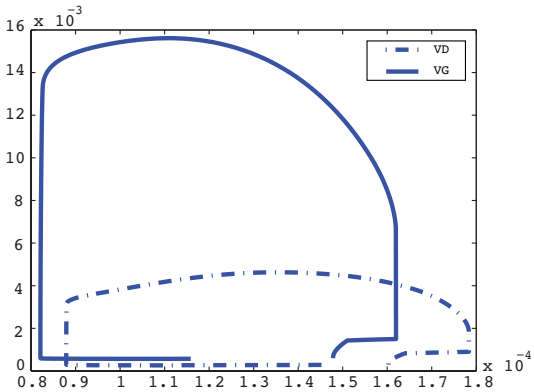


FIGURE 2.46: Diagramme P - V pour la simulation cardiaque

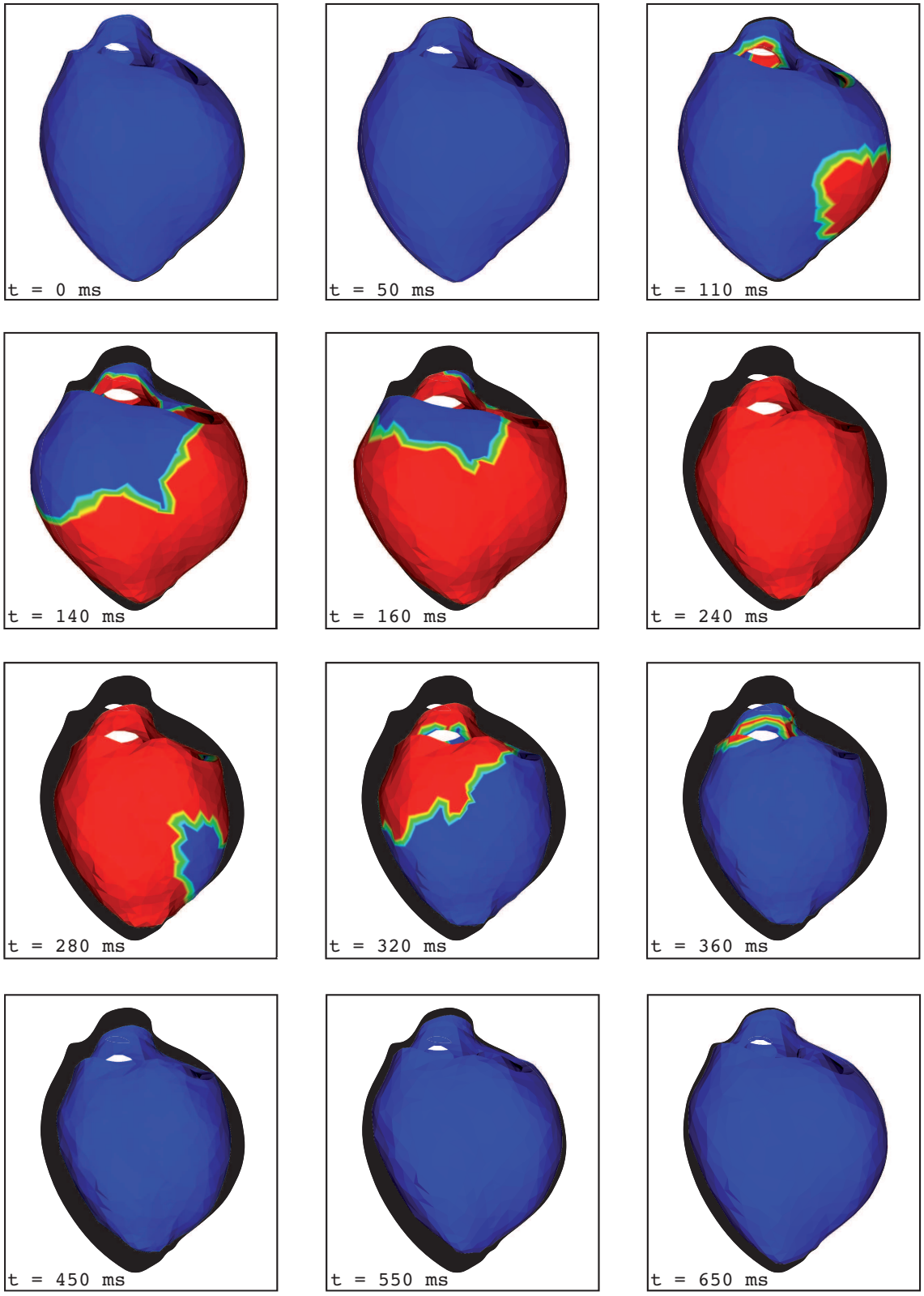


FIGURE 2.47: Simulation cardiaque globale. L'activation électrique est représentée en couleur (bleu : -80 mV, rouge : 20 mV)

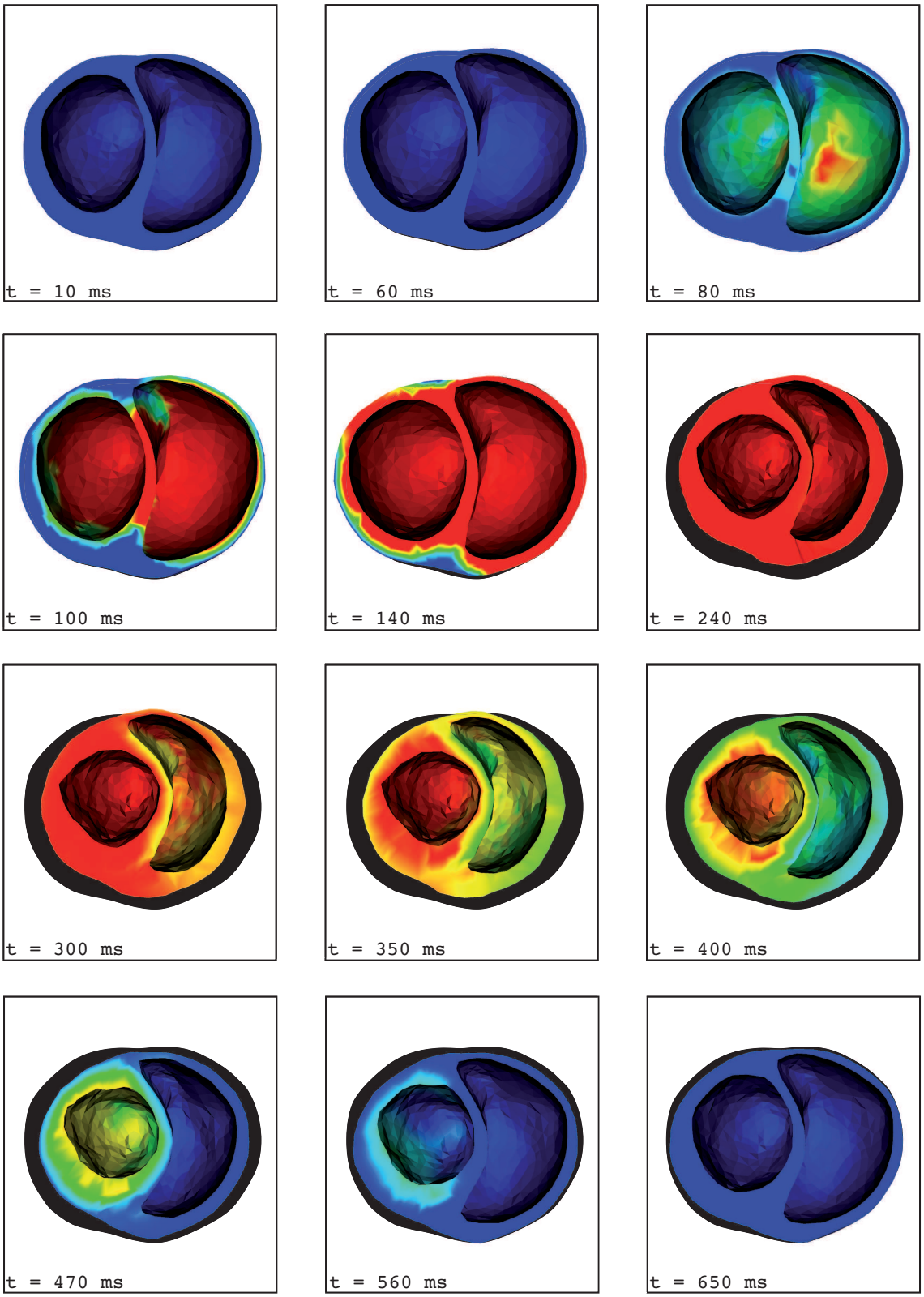


FIGURE 2.48: Coupe petit axe de la simulation cardiaque

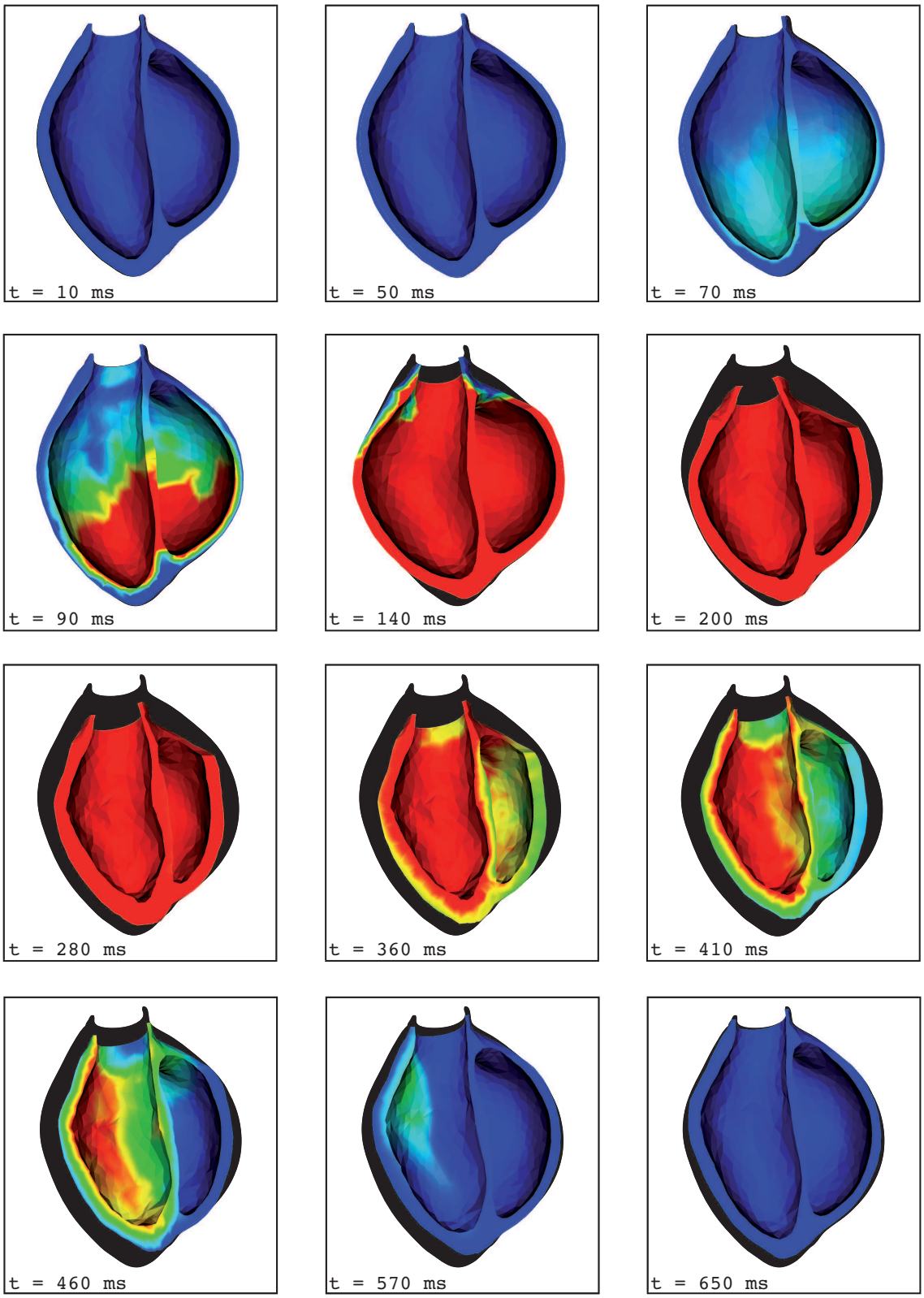


FIGURE 2.49: Coupe grand axe de la simulation cardiaque

2.5 Observations cardiaques

Nous venons de présenter l'ensemble des éléments nécessaires à la définition et la simulation de l'activité mécanique cardiaque. Cependant, celle-ci n'est jamais observée dans sa « configuration complète » mais au travers de mesures. Ces dernières sont donc des sorties potentielles du modèle direct afin de le représenter dans un espace d'observation usuel pour les cliniciens. Elle seront à l'inverse les entrées des problèmes d'estimation cardiaque des chapitres suivants. L'objectif de cette section est de formaliser les mesures actuellement disponibles pour le cœur afin de définir

- dans quels espaces mathématiques celles-ci se situent,
- quels sont leur liens au cours du temps avec la configuration mécanique x du système simulé,
- quelles sont les bruits caractéristiques qui les composent.

Suivant le formalisme présenté au chapitre précédent, on considère donc un ensemble (nécessairement fini) de mesures sur le système, représentées par un vecteur Z et un opérateur d'observation[11] $\mathcal{H}$ tel que

$$Z = \mathcal{H}(x) + \chi,$$

où χ est un bruit additif constitutif de la mesure considérée et à caractériser. L'opérateur d'observation n'est donc rien d'autre qu'un modèle d'instrument de mesures. Plus celui-ci projette la configuration x dans un espace de structure très différente ou d'indentifiabilité faible, moins il faudra compter sur ces mesures pour résoudre un problème d'estimation. La question de l'identifiabilité semble assez naturelle mais la complexité l'est tout autant. Par exemple, les données de volumes, d'épaississement des parois, de twist issus de post-traitements des données d'imagerie, très utilisées en milieu clinique, sont peu envisageables comme modèle d'observation alors même qu'elles sont apparemment discriminantes sur la santé d'un patient pour le médecin qui les consulte. Nous leur préférerons la donnée d'imagerie brute dans différentes modalités notamment celles non-invasives que constituent l'IRM ou l'échographie. A noter enfin que les mesures cardiaques disponibles à travers les images sont presque exclusivement cinématiques. Il n'y a pas d'informations sur les contraintes dans le tissu sauf peut-être à travers certains marqueurs de contraste ou en élastographie mais celle-ci est encore peu développée. De même on peut espérer bénéficier de données de pression seulement de façon exceptionnelle, lorsque le patient est cathéterisé.

2.5.1 IRM

a. Principes physiques

□ CINÉ-IRM – L'IRM est une technique d'imagerie datant du début des années 80 fondée sur l'observation du retour à l'équilibre de moments magnétiques de protons perturbés par l'application d'un champ radio-fréquence. La mesure repose sur le principe de la *RMN*[12] au-

[11] Nous utilisons la notation $\mathcal{H}$ pour insister sur le fait que l'opérateur agit sur les champs de déplacements et vitesses

[12] décrit par Felix Bloch et Edward Mills Purcell en 1946, Prix Nobel de physique en 1952.

quel on ajoute une localisation en espace. En effet, dans un tissu biologique et en l'absence de champ magnétique, les spins des protons sont aléatoirement distribués. *A contrario*, sous l'effet d'un champ magnétique extérieur B_0, les spins s'alignent globalement sur ce champ et majoritairement dans le même sens conduisant à la création d'un moment magnétique M_0 effectuant un mouvement de précession autour de B_0 à une fréquence appelée fréquence de Larmor ω. Sous une perturbation de ce premier champ magnétique par un champ B_1 perpendiculaire à B_0 suffisamment forte (pour cela B_1 est lui-même en précession à la fréquence de résonance ω), on observe une transition d'une partie des atomes qui s'alignent sur le champ B_1 et créent un moment M_1. Est alors mesuré le temps de relaxation des noyaux excités par B_1 : T_1 est le temps de relaxation de la composante du moment sur M_0, et T_2 celle sur le plan contenant M_1, tous deux étant dépendant du type de tissu. Pour la localisation spatiale, le principe de l'IRM est de construire des champs dépendant de la position

$$B_z(x) = B_0 + G_x.x.$$

Dans ce cas, la précession de Larmor sera dépendante de ce gradient sous la forme

$$w(x) = w_0 + \gamma G_x.x.$$

Ainsi pour chaque atome, les fréquences de précession dépendent de la position et des temps de relaxation caractéristiques du type de tissu. On reconstruit alors une image bidimensionnelle ou tridimensionnelle du tissu et, plus l'intensité des champs est grande, plus le contraste de celle-ci est fort. En pratique, la plupart des IRM cliniques sont de $1,5$ ou 3 Tesla, mais il existe des IRM entre 10 et 20 Tesla à des fin de recherche ou pour le petit animal.

Cette mesure IRM peut être reproduite au cours du temps et produire une séquence de volumes d'acquisition pendant tout un cycle cardiaque. En fait le phénomène de relaxation implique que les mesures ne sont pas instantanées. Il faut donc jouer sur plusieurs périodes cardiaques pour acquérir tout un cycle. En ajustant le déclenchement (*trigger*) de la mesure sur la mesure ECG effectuée simultanément, on obtient des images correspondant à des instants différents d'une sorte de cycle moyenné sur une à quelques dizaines de battements.

Plus spécifiquement en imagerie cardiaque, l'acquisition de mesure IRM présente des difficultés inhérentes à l'observation d'un organe en mouvement. De plus, la respiration est un facteur de bruit important et la plupart des mesures sont prises en apnée limitant ainsi le nombre de cycles.

Ainsi la résolution finale espace-temps en IRM cardiaque est conditionnée par l'ensemble des aspects physiques et pratiques du protocole de mesure mentionné, et l'ordre de grandeur actuel (en évolution constante) est aujourd'hui approximativement

$$(\Delta \mathrm{x}, \Delta \mathrm{y}, \Delta \mathrm{z}, \Delta T) = (1.5, 1.5, 1.5, 50)\ mm^3 \times ms$$

REMARQUE 2.5.1 (SIMULATEUR D'IRM)
Il est possible de simuler le processus complet d'acquisition de mesures IRM à partir des sorties du modèle grâce par exemple au simulateur *SIMRI* de Benoit-Cattin et al (2005), fondé sur les principes physiques présentés. En imagerie cardiaque, ce simulateur a été utilisé par Haddad et al (2007) où on trouvera les différents ordres de grandeurs des paramètres concernés. Une question particulièrement délicate pour les IRM cardiaques est la prise en compte des mouvements du cœur au cours de l'acquisition, y compris dans les types de bruits induits par ces mouvements. Du point de vu pratique, on utilise

à partir des maillages déformés du modèle une technique dite de rasterisation (voir l'algorithme complet dans Sermesant (2003)) pour générer des images (fantômes) binaires correspondant aux maillages comme illustré Figure 2.50. Ces images sont alors utilisées en entrée du simulateur d'IRM comme nous l'avons fait pour notre part en échographie Section 2.5.2.

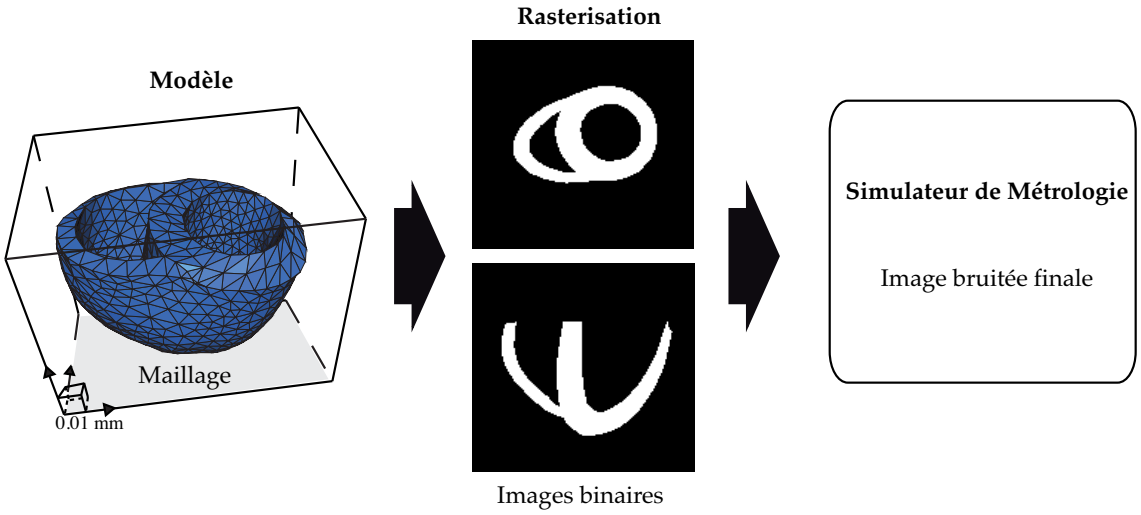


FIGURE 2.50: Principe d'un générateur de mesures réalistes utilisé en sortie d'un simulateur éléments finis via une procédure de rasterisation des maillages en images.

REMARQUE 2.5.2 (IMAGES CT)
A titre de comparaison la tomographie axiale calculée ou imagerie CT offre des résolutions, notamment spatiales, bien plus intéressantes que l'IRM. Cependant, comme toutes les techniques de scanner, cette résolution est conditionnée par la quantité de rayons X, voir d'éléments de contraste absorbés. Elle est donc bien plus invasive que les techniques de résonance magnétique et la plupart des images haute résolutions disponibles sont soit statiques sur des coeurs morts, soit sur des cœurs animaux.

□ IRM-TAGGÉS – A priori, les images issues d'un IRM classique appelé aussi ciné-IRM sont faiblement texturées et il est peu envisageable de suivre des points matériels du tissu sauf certains éléments précis de la géométrie (*landmark*). Il existe cependant un type de séquence IRM appelé IRM-taggés (ou marqués)[13] qui vise à combler cette lacune. Pour ce faire, on utilise un gradient sinusoïdal suivant deux directions permettant d'inscrire une grille magnétique sur le tissu. Les points matériels dont le spin est ainsi opposé aux autres sont véritablement marqués magnétiquement et ne vont plus apparaître tout au long de la séquence. Ainsi la grille régulière initiale se déforme avec le tissu au cours du battement. Les limitations principale de ce type de mesure sont :

- le contraste du marquage est exploitable pendant une durée de l'ordre de T_1 soit 400 ms.
- Seules deux directions de plans orthogonaux sont utilisables dans une même séquence. Ainsi sur une séquence d'acquisition 3D, on n'obtient qu'une projection en 2D du déplacement matériel. On utilise alors plusieurs coupes combinées *grand axe* et *petit axe* pour

[13] tagged-MRI en anglais

analyser le mouvement d'ensemble. L'obtention de grilles magnétiques véritablement 3D à l'image de Pan et al (2005) est un enjeu indéniable pour le développement de cette technique.

b. Segmentation pour l'IRM et recalage de tags

Il est a priori possible d'utiliser directement les données images dans un éventuel opérateur d'observation. Cependant, nous avons fait le choix de considérer les observations comme le résultat d'un chaîne de traitement sur la donnée image pour en extraire l'information qu'elle visait à représenter. Dans le cadre du ciné-IRM, on considère donc que la donnée pertinente issue de l'image est la trace des contours du cœur pendant sa déformation. Cette trace est obtenue par segmentation et donne lieu à de très nombreux développements en traitement d'image avec différentes techniques (morphologie mathématique, contours actifs ou modèles déformables, level-sets, recalage non rigide ...). C'est d'ailleurs dans la conjonction de techniques multiples que s'inscrivent aujourd'hui les dernières réussites opérationnelles (voir par exemple Peters et al (2007)). Classiquement, le sang étant fortement contrasté dans les IRM, la segmentation la plus simple est celle de l'endocarde du ventricule gauche et dans une moindre mesure du ventricule droit. Inversement, l'épicarde pour le ventricule droit est très difficile à segmenter.

Pour l'IRM-taggé, l'objectif de la mesure est de représenter le déplacement des points matériels à partir de la grille initiale. Comme nous venons de le voir, la donnée consiste en plusieurs séquences d'intersection de plans dans seulement deux directions, donc retrouver le déplacement en recalant ces plans entre eux nécessite des outils adaptés (voir en particulier Declerck et al (1999) ; Axel et al (2005)) et constitue encore une difficulté majeure dans l'exploitation des tags.

REMARQUE 2.5.3
L'utilisation d'un simulateur de protocole de mesure combiné à un modèle de contraction cardiaque peut fournir des données synthétiques pertinentes pour l'évaluation des différentes techniques de traitement de la mesure, notamment pour le recalage de plan de tags et l'extraction des déplacement 3D.

c. Opérateur d'observation en déplacement

Commençons par l'IRM-taggé qui fournit l'opérateur le plus simple, une fois la mesure supposée traitée. En effet, on dispose alors de déplacements de points de la grille par rapport à une configuration de référence. Autrement dit, à chaque pas de temps d'échantillonage de la mesure, on dispose du déplacement « de points du tissu » en configuration Lagrangienne. Considérons donc une séquence de $N_{\Delta T}$ tags, échantillonnée sur le battement tous les ΔT, alors

$$Z_k = \mathcal{H}x = (\mathcal{H}^d\ 0)(\underline{y}(k\Delta T)\ \underline{\dot{y}}(k\Delta T))^T, \quad 1 \leq k \leq N_{\Delta T},$$

où $\mathcal{H}^d$ est linéaire et travaille uniquement sur la partie déplacement de l'état x et correspond grosso-modo à un opérateur de sélection. Plus précisément, la mesure a en réalité une extension spatiale correspondant à l'épaisseur de la grille. Les déplacements, s'ils sont correctement extraits des tags, correspondent donc à des moyennes pondérées sur des cellules d'extension

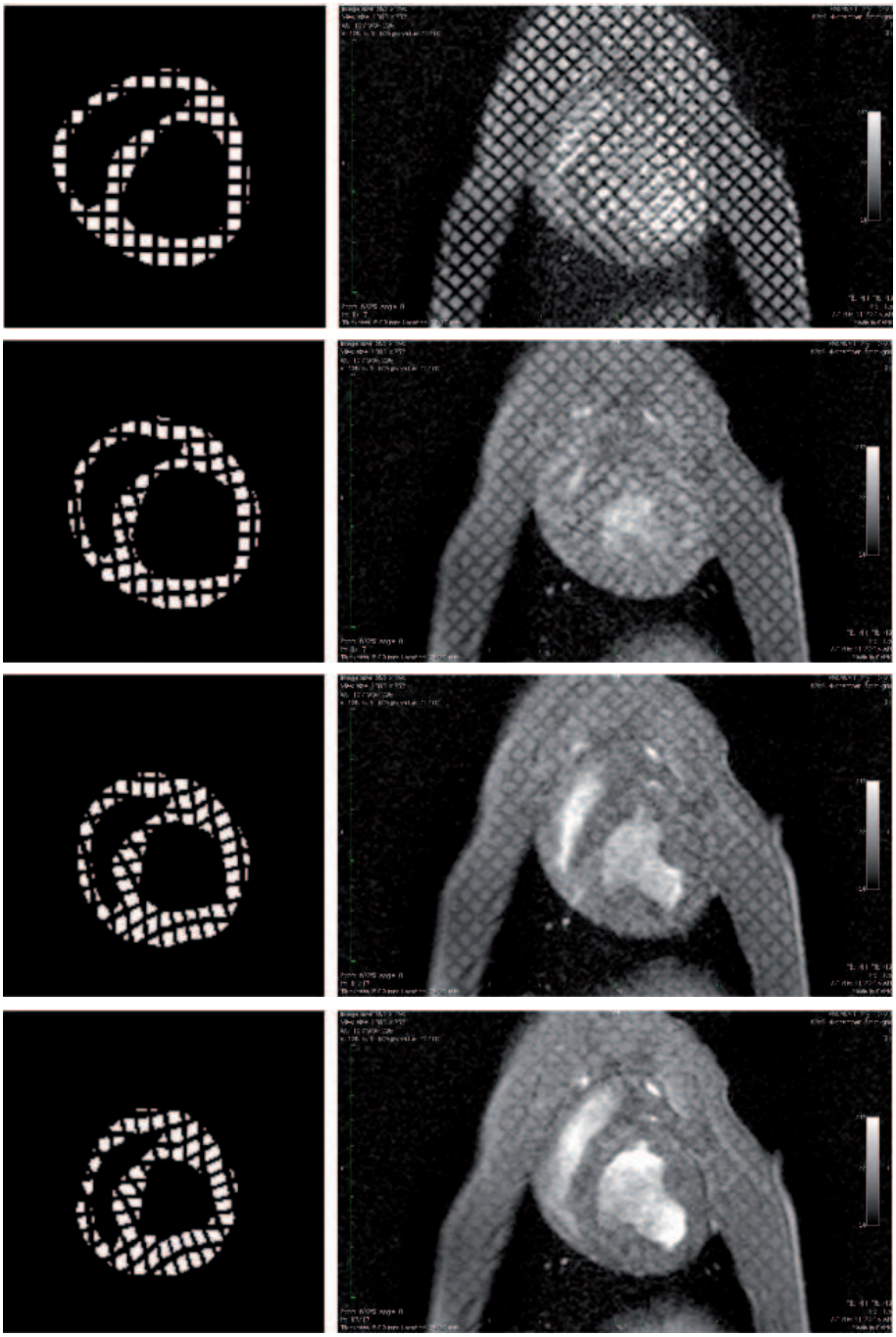


FIGURE 2.51: Exemple de tags synthétiques (gauche) et réels (droit) pour un coeur de cochon. On observe nettement la déformation au cours de la systole. Sur les tags réels, le marquage magnétique s'estompe au cours de l'acquisition. Remerciements à Radomir Chabiniok, Macs, INRIA

spatiale finie, les fonctions de pondération (gaussienne ou sinusoïdale par exemple, voir Figure 2.52) étant liées à la mesure. On note q le nombre de déplacements tridimensionnels contenus à partir de q cellules de mesures non recouvrantes de la zone de mesure globale noté Ω_m. Les fonctions de pondération sont notées $(s_i)_{i=1}^q$. On choisit de définir chaque fonction sur tout Ω_m, leur support étant d'intersection vide. Elles sont enfin normalisées $\|s_i\|_{L^2\Omega_m} = 1$. Ainsi, chaque composante des Z_k est donnée par

$$(Z_k)_i = \int_{\Omega_m} s_i \underline{y}\, d\Omega, \quad 1 \leq i \leq q. \tag{2.64}$$

De plus on supposera que chaque mesure est perturbée par un bruit gaussien indépendant.

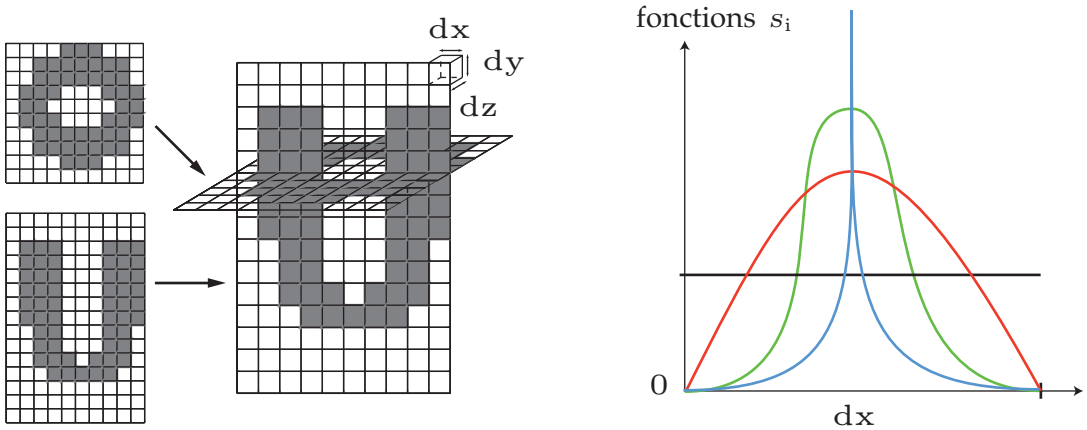


FIGURE 2.52: Principes de cellules de mesures issues des tags. Les déplacements sur les cellules sont moyennés par des fonction poids s_i dont on donne quelques exemples potentiels.

Dans le cas des ciné-IRM la définition de l'opérateur d'observation est bien plus complexe. En effet, nous avons vu que l'information pertinente contenue dans la donnée correspond à des

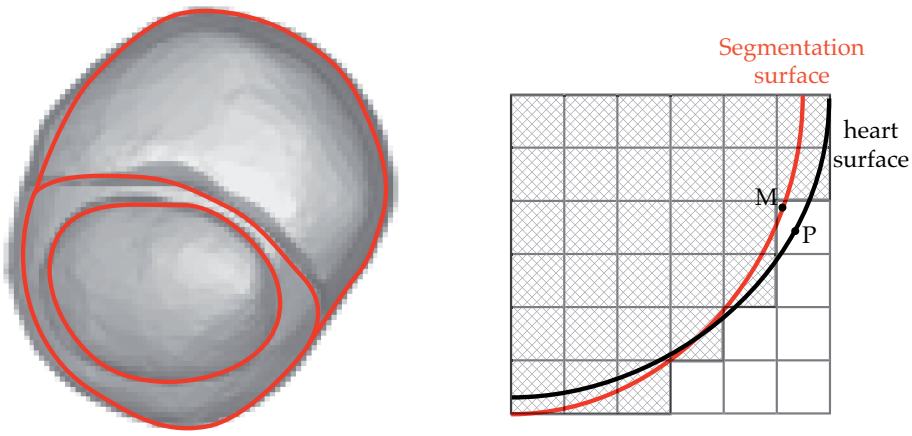


FIGURE 2.53: Schéma d'une segmentation. Le bruit est défini en configuration déformée comme une distance entre les deux contours.

surfaces intérieures (et dans certains cas extérieures) de l'objet. On considère qu'on dispose d'images segmentées sous forme de surfaces régulières maillées (à partir d'un algorithme de marching-cube puis de régularisation, ou directement à partir de la segmentation suivant Pons et al (2007)). Par exemple, au temps $t = k\Delta T$, on note S_k cette surface maillée. On pourrait donc dire que l'ensemble des noeuds de S_k correspond à la mesure Z sous forme de position en configuration Eulérienne. Cependant ces positions sont peu exploitables car l'opérateur passant des déplacements Lagrangien à ces postions Eulériennes est défini uniquement en l'absence de bruit sous une forme un peu artificielle du type

$$(Z_k)_i = \underline{\phi}(\underline{\mathrm{x}}_i - \underline{y}(\phi^{-1}(\underline{\mathrm{x}}_i))), \quad i \in \mathcal{T}(S_k).$$

En effet, la transformation Φ n'est définie que sur Ω_m, or il n'y a jamais correspondance exacte entre un point du cœur et la surface extraite comme illustré Figure 2.53. Ce qui, dans la mesure, est exploitable correspond plutôt à une correspondance sous forme de distance minimum entre un point de l'objet mécanique et la surface issue de la segmentation. Ainsi, la donnée n'est pas vraiment du type $Z = \mathcal{H}^d(\underline{y}) + \chi$, mais plutôt

$$\mathrm{dist}(\underline{\xi} + \underline{y}, S_k) = \chi, \quad \forall \underline{\xi} \in \partial\Omega_m,$$

noté génériquement $D(Z, X) = 0$ comme au chapitre 1, et extension non-linéaire de « $Z - H^d(\underline{y}) = \chi$ » pour une donnée initialement fournie sous la forme $(S_k)_{k=1}^{N_{\Delta T}}$.

REMARQUE 2.5.4 (OPÉRATEUR D'OBSERVATION EULERIEN POUR LES TAGS)
Nous avons vu comment nous pouvions utiliser des tags « recalés » comme opérateur d'observation. Une question légitime est de savoir s'il n'existe pas un moyen de les utiliser dans une forme plus brute. Certaines équipes de traitement d'image essaient notamment d'extraire les déplacements des tags à partir de modèles déformables et on pourrait légitimement se demander s'il est possible de fusionner ces deux étapes (1. Extraction des tags à partir des modèles déformables - 2. Estimation sur le modèle à partir des déplacements). Par exemple, à l'image de ce que nous avons dit sur la segmentation on pourrait imaginer un opérateur d'observation sous la forme $D(Z, X) = 0$ en calculant la distance au plans de tags les plus proches de chaque point ce qui éviterait de devoir recaler entre eux les plans issues de deux jeux de mesures.

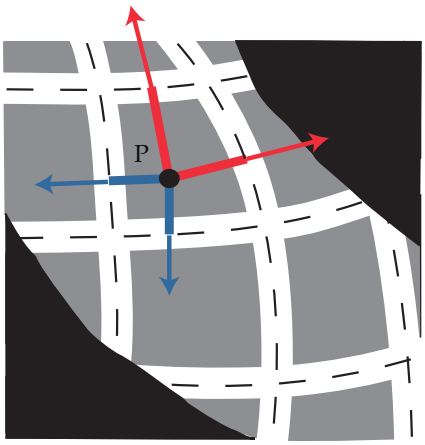


FIGURE 2.54: Schéma de calcul de distance au plan de tags qui pourrait permettre d'éviter le recalage des plans entre eux.

2.5.2 Echographie

a. Principes physiques et simulations directes

Une onde ultrason est une onde mécanique longitudinale se propageant dans le milieu à une fréquence $\equiv 20kHz$. D'après Shung (1993), on peut supposer que les déplacements causés par la propagation de l'onde sont petits et vérifient l'équation d'onde

$$\rho\ddot{u} + \nu\dot{u} - \nu'\Delta\dot{u} - (\lambda + \mu)\Delta u = 0,$$

où u est le déplacement longitudinal lié à la pression exercée par l'onde dans le milieu par $u = \frac{p}{Z}$ à travers l'impédance caractéristique du tissu parcouru. De plus, lorsque l'onde traverse un milieu hétérogène, elle est soit diffusée soit réfléchie-transmise suivant les règles classiques en électromagnétisme en fonction de la taille caractéristique de l'hétérogénéité. On parle en effet de

- **diffusion** si la longueur d'onde $\Lambda > L$: l'onde réfléchit dans des directions arbitraires une faible partie de son énergie. On parle alors de phénomène de *scattering* causé tout au long du trajet de l'onde par les hétérogénéités locales du tissu parcouru. Sur l'image résultante, cette diffusion produit des schémas d'interférence appelés *speckle* caractéristiques de chaque tissu et du système d'échographie utilisé. Ces derniers sont la source de l'effet granulaire et donc bruité de l'image échographique. Cependant ils sont en pratique recherchés par le praticien car ils fournissent une information de localisation relative au tissu. En effet, pour un même appareil de mesure, un élément de tissu fixe émettra le même type de speckle sur l'image échographique et on peut espérer qu'il en sera de même pour des mouvements faibles.
- **réflexion spéculaire** si la longueur d'onde $\Lambda < L$, l'obstacle est localement considéré comme un demi-espace d'indice de reflexion différent et les règles de l'optique géométrique s'appliquent. Ce type de réflexion intervient principalement à l'interface entre deux types de tissus radicalement différents par exemple sang-muscle, air-os etc...

De plus, il existe un certain nombre d'éléments d'atténuation de l'onde. Tout d'abord la diffusion au sens du scattering est clairement limitante pour l'énergie finale captée par l'échographe. Il y a aussi la dissipation classique liée au milieu caractérisé dans l'équation par ν et ν'. Enfin il y a une perte d'énergie entre l'écho émis et celui détecté par l'appareil liée à la dilatation du front (voir Figure 2.55) d'onde qui n'est pas rigoureusement planaire. On note que l'atténuation globale exprimée en décibel est dépendante de la fréquence du signal. Ainsi une haute fréquence, donc une faible longueur d'onde, doit améliorer la résolution de l'image finale mais est plus fortement atténuée.

La plupart des échographies actuelles possèdent une ligne d'émetteurs-capteurs piezo-éléctriques émettant une onde planaire dans le milieu et visualisant l'écho ultrason aussi appelé signal radio-fréquence (RF). Dans ce cas, l'image reconstruite est une image 2D correspondant au tissu observé dans le plan de parcours. Il existe désormais des dispositifs d'acquisition directement en 3D où les capteurs sont soit disposés sur une grille soit déplacés pour balayer un plan d'émission. Dans les deux cas, l'image reconstruite correspond directement à un volume 3D de tissu observé. La plupart des dispositifs ont des spécifications propres dans la

manière dont ils émettent mais surtout traitent le signal pour former l'image, notamment dans la façon dont ils gèrent le phénomène de « speckle ». La machine Volumetrics, développée à *Duke University* est sans doute le seul dispositif à caractéristiques ouvertes et dont on peut essayer de reproduire numériquement le fonctionnement.

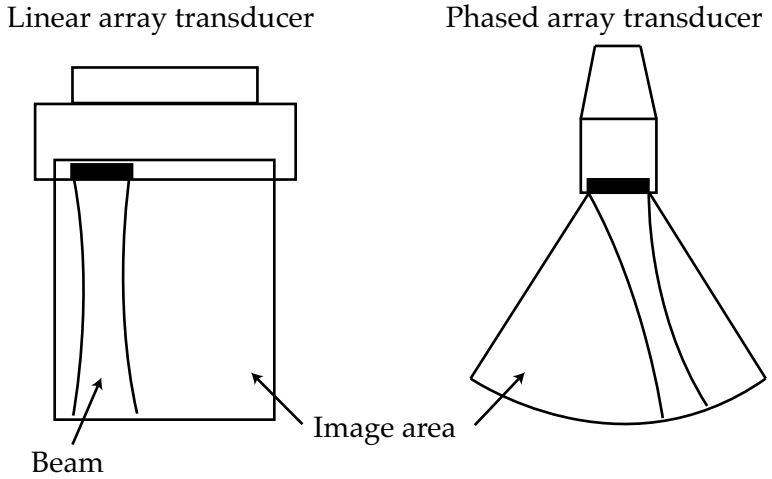


FIGURE 2.55: Exemple de dispositifs d'échographie 3D. A droite le schéma de fonctionnement de la machine Volumetrics où l'acquisition est donc en coordonnées sphériques.

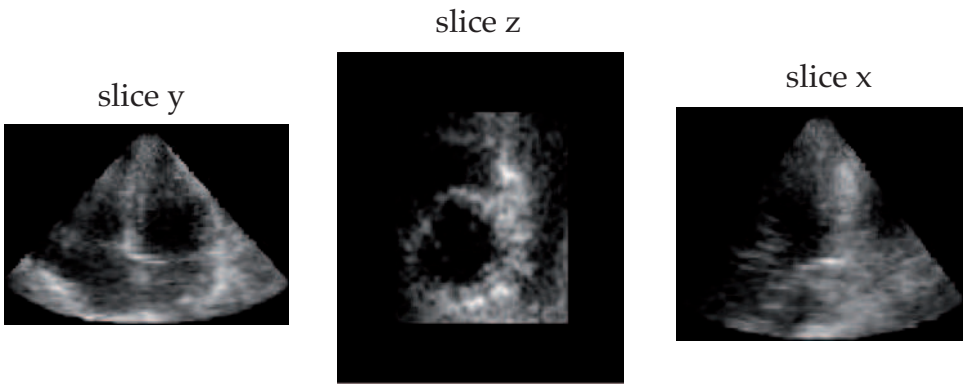


FIGURE 2.56: Exemple de données échographiques 3D. Remerciement à E. Angelini, Telecom ParisTech.

Nous avons voulu simuler l'acquisition de mesures échographiques en sortie du modèle à la fois pour fournir un outil pédagogique mais surtout pour analyser le type de données qu'on pouvait en extraire et la validité des algorithmes de traitement d'images fondés sur ce type d'acquisition (voir section suivante). L'approche suivie a été principalement guidée par le traitement de signal et pas nécessairement par la modélisation physique fine de la propagation, simplement par volonté d'efficacité sur un phénomène de scattering avant tout statistique. Le modèle est inspiré principalement des travaux de Meunier et Bertrand (1995) et présenté dans Duan et al (2007). Il a donné lieu à des extensions présentées dans Talcoth (2007) pour prendre en compte la géométrie d'acquisition de la machine Volumetrics ainsi que le comportement du sang et les effets de réflexion spéculaire. Dans ces modèles, le signal radio-fréquence (RF)

est représenté en bout de chaîne comme une convolution en tout point du volume de tissu observé Ω_m

$$\forall \underline{x} = (\mathrm{x}, \mathrm{y}, \mathrm{z}) \in \Omega_m, \quad RF(\mathrm{x}, \mathrm{y}, \mathrm{z}) = H(\mathrm{x}, \mathrm{y}, \mathrm{z}) * T(\mathrm{x}, \mathrm{y}, \mathrm{z})$$

entre

- **Une « *point spread function* »** H qui correspond à la réponse d'un point matériel de matière. Basiquement, cette fonction néglige l'atténuation, ne prend pas en compte les zones de réflexion spéculaire, et la propagation (ici z pour fixer les idées) n'est prise en compte qu'à travers l'anisotropie de la réponse

$$H(\mathrm{x}, \mathrm{y}, \mathrm{z}) = \exp\left(-\frac{1}{2}\left(\frac{\mathrm{x}^2}{s_{\mathrm{x}}^2} + \frac{\mathrm{y}^2}{s_{\mathrm{y}}^2} + \frac{\mathrm{z}^2}{s_{\mathrm{z}}^2}\right)\right) \cos(2\pi \frac{\mathrm{z}}{\lambda}).$$

  où $s_{\mathrm{z}} >> s_{\mathrm{x}}, s_{\mathrm{y}}$

- **la réponse impulsionnelle du tissu** T modélisée (voir Meunier et Bertrand (1995)) comme la dérivée seconde d'une impédance acoustique $Z_{3D} = \sum_n Z_n$, où Z_n est l'impédance de chacun des éléments diffusifs du tissu appelés *scatterrers*. Ces derniers, en nombre bien plus important que la résolution de l'image formée, sont regroupés par « paquet » en chaque point $\underline{x}_n = (\mathrm{x}_n, \mathrm{y}_n, \mathrm{z}_n)$ du tissu. L'*écogéneité* du tissu considéré en chaque point est directement proportionnel au cardinal de scatterrers dans chaque paquet.

$$T(\mathrm{x}, \mathrm{y}, \mathrm{z}) = \frac{\partial^2}{\partial z^2} \sum a_n C(\mathrm{x} - \mathrm{x}_n, \mathrm{y} - \mathrm{y}_n, \mathrm{z} - \mathrm{z}_n),$$

  où

$$C(\mathrm{x}, \mathrm{y}, \mathrm{z}) = \exp\left(-\frac{1}{2}\left(\frac{\mathrm{x}^2}{\sigma_{\mathrm{x}}^2} + \frac{\mathrm{y}^2}{\sigma_{\mathrm{y}}^2} + \frac{\mathrm{z}^2}{\sigma_{\mathrm{z}}^2}\right)\right).$$

Enfin le dispositif physique d'acquisition effectue pour former l'image une détection d'enveloppe à partir du signal RF correspondant à une transformée de Hilbert. Ainsi, l'image finale est

$$I(\mathrm{x}, \mathrm{y}, \mathrm{z}) = |RF(\mathrm{x}, \mathrm{y}, \mathrm{z}) + j\mathrm{Hilbert}_z(RF(\mathrm{x}, \mathrm{y}, \mathrm{z}))|$$

Les paramètres sont choisis et optimisés en fonction des données physiques relatives au tissu considéré mais aussi à partir de bases de données d'images issues des machines Volumetrics. La Figure 2.59 fournit des exemples de simulations d'échographie en sortie du modèle. Figure 2.58, on rappelle d'autres exemples de simulation échographiques antérieures dans des configurations saines et pathologique correspondant aux résultats des travaux publiés dans Duan et al (2007).

b. Segmentation et flot optique

Une première donnée extraite des images échographiques est strictement l'équivalent de ce que nous avons vu pour les IRM, c'est-à-dire une segmentation des contours de l'objet. Bien évidemment, les méthodes de segmentation sont différentes puisque largement dépendant de

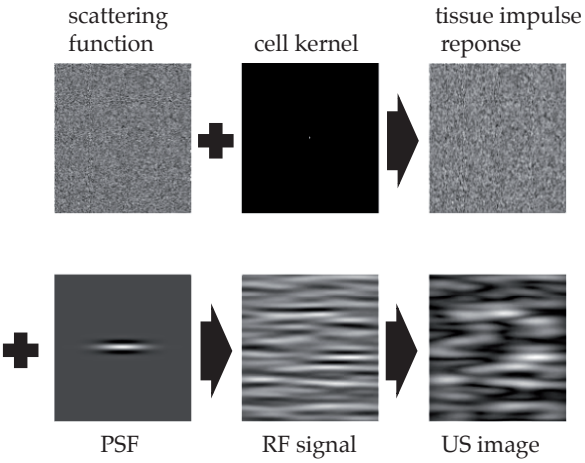


FIGURE 2.57: Différents échantillons d'images successifs permettant la formation de l'image échographique.

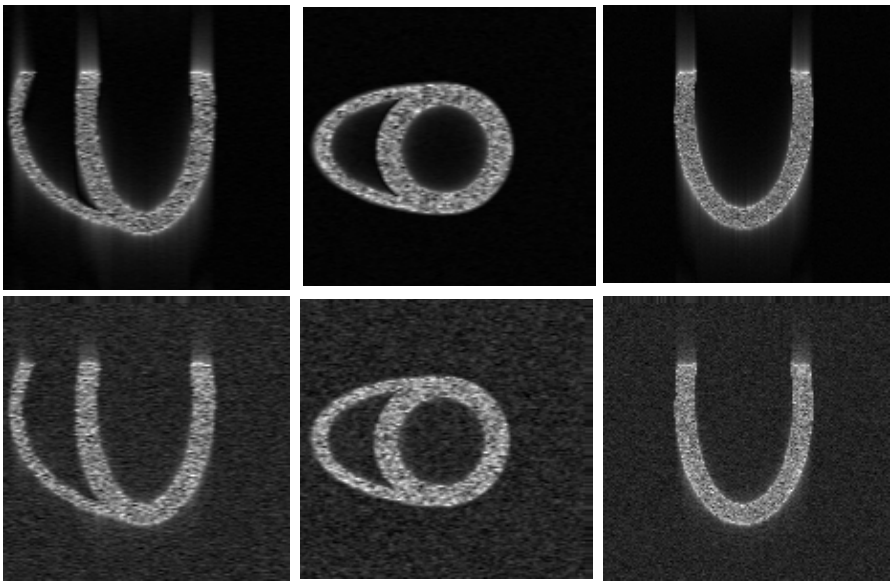


FIGURE 2.58: Exemple de simulation d'ultrasons issue de la collaboration Duan et al (2007)

leur stratégie de la modalité, et côté cardiaque se concentre le plus souvent sur le ventricule gauche (voir Vallet et al (2006) pour un exemple).

Sur la structure des images échographiques et en bénéficiant de sa haute résolution temporelle de l'ordre de 60 à 100 images par secondes, d'autres traitements cinématiques sont envisageables, notamment l'utilisation de *flot optique*. Le flot optique consiste à déterminer le déplacement de points entre deux images par recalage d'intensité (et de texture). Si une image a suffisamment d'hétérogénéités, et si ces hétérogéneités se transportent au cours du temps, alors en recherchant les similarités d'intensité entre deux images on espère retrouver une information Lagrangienne au sens de la configuration de l'objet (déplacement de points matériels) dans la donnée d'image profondément Eulerienne. Les deux conditions ci-dessus sont en partie remplies par l'échographie comme nous venons de le voir en considérant que les différents speckles se transportent au cours du temps et diffusent dans les mêmes conditions.

Du point de vue technique, Angelini et Gérard (2006) rappelle que deux types de méthodes permettent le calcul du flot optique. La première est une technique fondée sur le calcul des dérivées spatio-temporelle de la carte d'intensité. Si $I(\underline{\mathrm{x}}, t)$ est l'image au cours du temps, la dérivée particulaire de cette image est supposé nulle au cours du temps puique l'intensité est conservée

$$\frac{\partial I}{\partial t} + \underline{\dot{y}}(\underline{\mathrm{x}}) \cdot \nabla I = 0.$$

où $\underline{\dot{y}}(\underline{\mathrm{x}})$ représente la vitesse eulérienne du point matériel en $\underline{\mathrm{x}}$. Cette contrainte scalaire n'est pas suffisante pour déterminer $\underline{y} \in \mathbb{R}^3$ donc on ajoute d'autres contraintes, soit par exemple la valeur de l'intensité, soit sur le type d'espace dans lequel on attend le déplacement. Dans ce cas, on retrouve un problème de minimisation de fonctionnelle quadratique avec un terme d'a priori exprimé dans une norme de Sobolev bien choisie... La deuxième technique consiste véritablement à faire du recalage locale de petits motifs de l'image. Elle est aussi appelée speckle tracking et a été choisie pour l'étude. Soit $I(\underline{\mathrm{x}}, t)$ et $I(\underline{\mathrm{x}}, t + \Delta t)$, les deux images et $\mathcal{O}$ une fenêtre de recherche. On maximise alors la covariance normalisée

$$\Delta \mathrm{x}(\mathrm{x}) = \mathrm{argmax}_{\Delta \mathrm{x}} \rho(\mathrm{x}, \Delta \mathrm{x}) = \frac{\sum_{\mathrm{x} \in \mathcal{O}} \mathbb{E}(I(\mathrm{x}, t) I(\mathrm{x} + \Delta \mathrm{x}, t + \Delta t))}{\left(\sum_{\mathrm{x} \in \mathcal{O}} \mathbb{E}(I(\mathrm{x}, t)^2) \sum_{\mathrm{x} \in \mathcal{O}} \mathbb{E}(I(\mathrm{x} + \Delta x, t + \Delta t)^2)\right)^{\frac{1}{2}}}.$$

Au vu de la résolution temporelle, on peut obtenir une estimation de la vitesse finalement par

$$\underline{\dot{y}}(\underline{\mathrm{x}}) = \frac{\Delta \mathrm{x}(\mathrm{x})}{\Delta t}.$$

Les premiers résultats de flot optique des échographies synthétiques à partir du simulateur cardiaque utilisent la deuxième méthode déjà mise en oeuvre dans Duan et al (2006). Ils ont été publiés dans Duan et al (2007) et présentent des déplacements cohérents (voir un exemple Figure 2.60).

REMARQUE 2.5.5 (IMAGERIE DOPPLER ET CHAMP DE VITESSES)
Si l'échographie et le flot optique seuls semblent limitées pour l'obtention de champs denses de vitesses à tout instant de la séquence d'images, Chapelle et al (2008) propose de coupler l'échographie à l'imagerie Doppler (TDI). Cette dernière utilise le fait que la fréquence de l'onde échographique est modifiée par l'effet Doppler lorsque les tissus sont en mouvement. Cette modification de fréquence permet de reconstituer une champs de vitesses dans la direction de l'onde échographique alors que le flot optique était plus performant dans la direction orthogonale à la source.

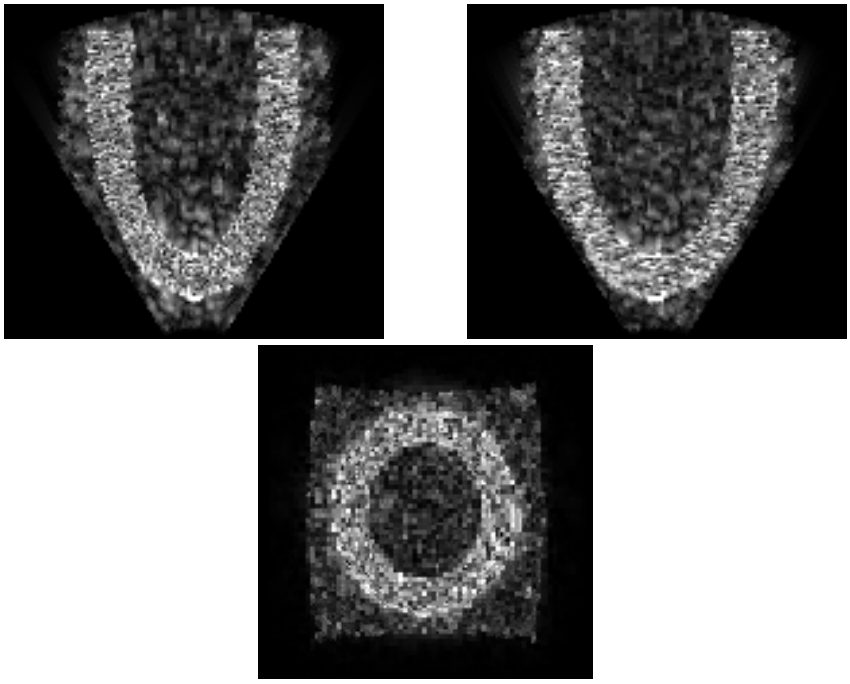


FIGURE 2.59: Exemple de simulations réalisées par Oscar Talcoth en fin de stage dans l'équipe en partenariat avec E. Angelini de l'ENST. Sont pris en compte la grille sphérique, l'échogeneité des tissus, les reflexions spéculaires (au niveau de l'apex).

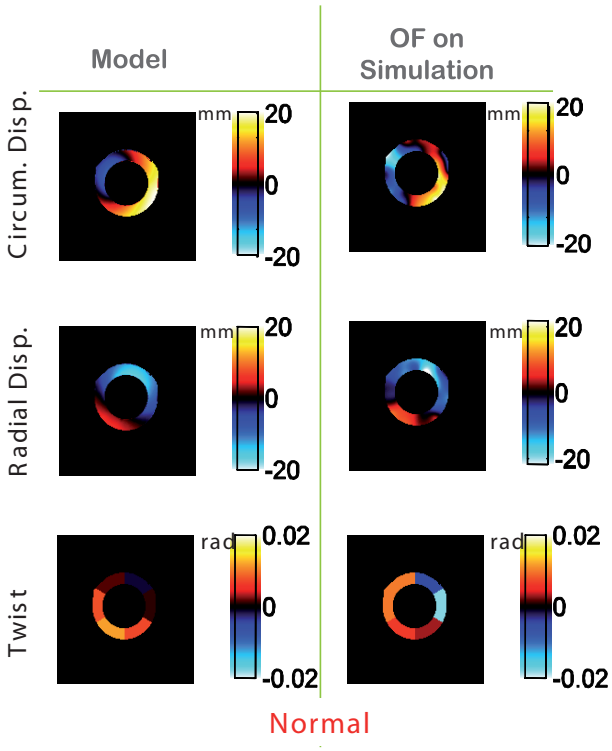


FIGURE 2.60: Exemple de reconstruction de mouvement à partir de flot optique sur des images petit axe de la simulation d'échographie synthétique

c. Opérateur d'observation

Quelle que soit la modalité utilisée pour l'obtention du champ de vitesses, celui reste fondamentalement Eulérien. Pour, déterminer un déplacement Lagrangien il faut ensuite intégrer ces mesures au cours du temps, c'est-à-dire successivement suivre un point matériel depuis l'image initiale jusqu'à l'image courante. Dans ce cas, on obtient idéalement un opérateur d'observation fournissant à tout instant le déplacement ou la vitesse (résolution temporelle suffisante) des points suivis. Concernant les points suivis, le principe du flot optique est de suivre des éléments de textures caractéristiques. Ceci s'applique aux speckles suivant les hypothèses formulées plus haut, mais doit aussi être favorisé par la présence de contours dans l'image. Autrement dit, on peut considérer que ces déplacement seront les moins sujets aux erreurs et donc que l'opérateur d'observation est un opérateur frontière du type

$$\mathcal{H}y = y|_{\partial\Omega},$$

qui comme pour les tags doit en toute rigueur être représenté à partir de cellules frontières corrélées aux cellules à partir desquelles le flot optique a été calculé. De plus, il reste à déterminer quel type de bruit est généré dans ce traitement de mesure et c'est tout l'avantage de pouvoir tester ces algorithmes sur des données synthétiques créées par un fantôme numérique de battement cardiaque. D'après ce que nous avons vu depuis le début ceci est équivalent en formulation « déterministe » à déterminer dans quel espace et pour quelle norme se situent les erreurs en déplacements-vitesses issus du flot optique. Ce travail est toujours en cours avec Elsa Angelini après des premiers résultats d'estimation d'erreurs peu concluants.

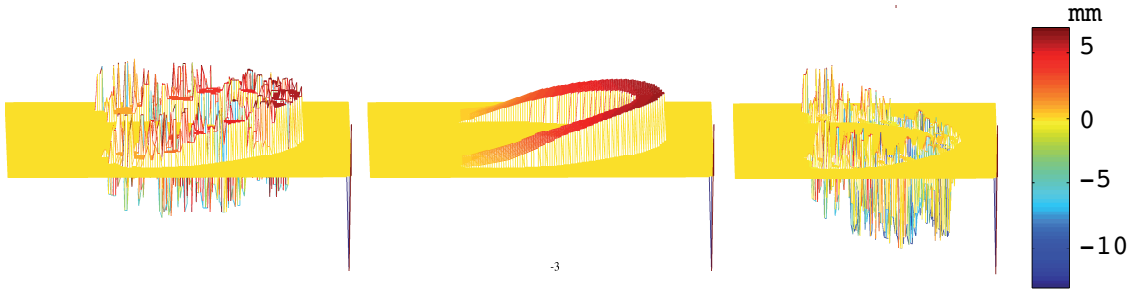


FIGURE 2.61: A gauche résultat du flot optique sur un mouvement de ventricule gauche. Au centre, le déplacement ayant généré les échographies. A droite l'estimation d'erreur L^2. Celle-ci est de l'ordre de grandeur du déplacement, alors qu'« à l'œil », le flot optique a capté le mouvement d'ensemble.

En effet, ce ne sont finalement pas nécessairement les éléments du bord qui sont le plus facilement détectés car les déformations sont telles que les éléments de textures qu'ils représentent n'apparaissent pas conservés au cours du déplacement au sens où le suppose le flot optique type « speckle tracking ». En effet ce dernier ne cherche que des déplacements de type translation pour l'analyse de ces cellules de texture et pas les configurations de rotation, cisaillement possibles. A l'inverse, dans le tissu le caractère conservatif de l'emission des speckle laisse penser qu'on peut suivre des petits cubes de matière à l'intérieur du tissu. Dans ce cas on se rapproche alors d'un observateur de type de celui développé pour les IRM taggés sauf que cette propriété intérieure « au tissu échographique » est sans doute idéalisée par rapport

à la situation réelle ou les speckles n'émettent pas identiquement au cours de la déformation. Pour notre simulateur on doit donc se limiter à de faibles déplacements et finalement plutôt au cœur du tissu. Dans ce cas, on constate que les déplacements ne sont pas nécessairement bien évalué en tout point ce qui introduit des hautes fréquences dans l'erreur. Une hypothèse est alors de considérer l'erreur dans une norme H^s simple à calculer sur une image à partir de sa transformée de Fourier $\mathcal{F}$

$$\|\tilde{I}\|_{H^s} = \left(\int (1+|\xi|^2)^s \left| \mathcal{F}\tilde{I}(\xi) \right|^2 \right).$$

Cependant, les performances actuelles du flot optique sont insuffisantes encore pour ces normes et les erreurs s'analysent plutôt par le modèle suivant : Soit $\underline{y}_{of}$ le déplacement obtenu par le flot optique et $\underline{y}$ le déplacement de référence, alors

$$\underline{y}_{of} = \begin{cases} \underline{y} + \epsilon, & \text{avec probabilité } p_1 \\ \chi, & \text{avec probabilité } p_0 = 1 - p_1 \end{cases}$$

avec ϵ l'erreur liée à la résolution de l'image échographique et χ l'erreur quand le flot optique ne trouve pas le bon point cible dans son analyse. Par conséquent, ϵ est centré et de l'ordre de grandeur de la résolution spatiale de l'image. Par ailleurs χ est centré mais de la taille de la fenêtre de la recherche $\mathcal{O}$ car le flot optique fournit un résultat arbitraire dans cet fenêtre quand il n'a pas pu trouver de correspondance réelle. Notons d_b cette distance, elle est donc du même ordre de grandeur que $|\underline{y}|$. Soit e_{rel} l'erreur relative L^2

$$e_{rel} = \frac{\|\underline{y}_{of} - \underline{y}\|_{l^2}}{\|\underline{y}\|_{L^2}}.$$

On a

$$\mathbb{E}(\|\underline{y}_{of} - \underline{y}\|^2) = p_1 \mathbb{E}(\epsilon^2) + p_0(d_b^2 + |\underline{y}|^2).$$

Donc sur toute une fenêtre de flot optique

$$\sum_{\mathcal{O}} \mathbb{E}(\|\underline{y}_{of} - \underline{y}\|^2) = p_1 \mathbb{E}(\epsilon^2)\,|\mathcal{O}| + p_0 d_b^2\,|\mathcal{O}| + p_0 \|\underline{y}\|^2_{L^2(\mathcal{O})}.$$

Ainsi l'erreur relative est toujours

$$e_{rel} > \sqrt{2p_0}.$$

Et les calculs seraient similaires en normes H^s ou en $p_0^{\frac{1}{p}}$ pour une norme L^p. Ainsi, avant le choix de la norme, il faut pouvoir améliorer la qualité de détection du flot optique « speckle tracking » ou lui associer un indice de confiance. Ce dernier servirait pour détecter les cas où le flot optique n'a aucune idée du bon déplacement et choisit donc « au hasard ». Afin d'illustrer cette idée, nous montrons par exemple Figure 2.63 une configuration où on ne garde que les points à indice de confiance a posteriori supérieurs à un seuil.

Ainsi, la chaîne complète, battement simulé, échographie synthétique, flot optique, calculs d'erreurs est un outil essentiel pour l'amélioration des techniques de traitements d'images à supposer que la simulation mécanique permette des mouvements « réalistes » et la mesure

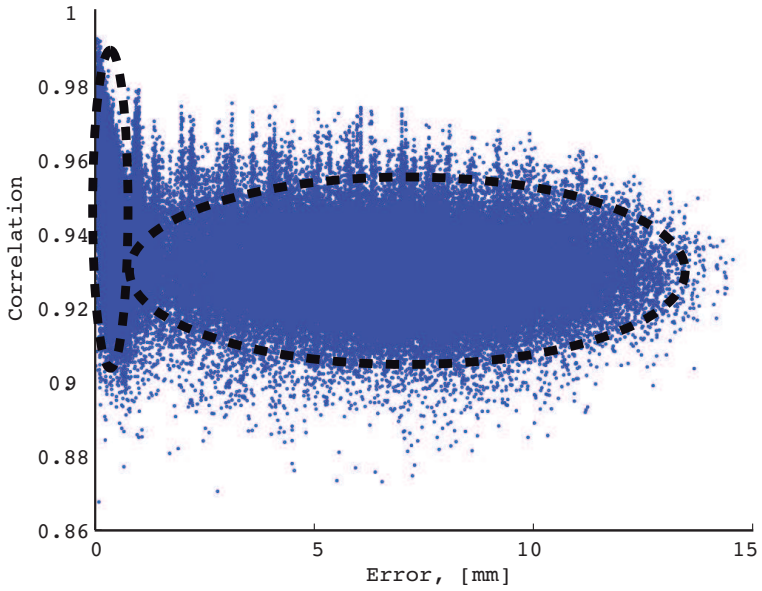


FIGURE 2.62: Relation a posteriori entre la corrélation et les erreurs. On constate bien deux zones principales. Au dessus d'un seuil de corrélation on ne garde plus que des points bien détectés.

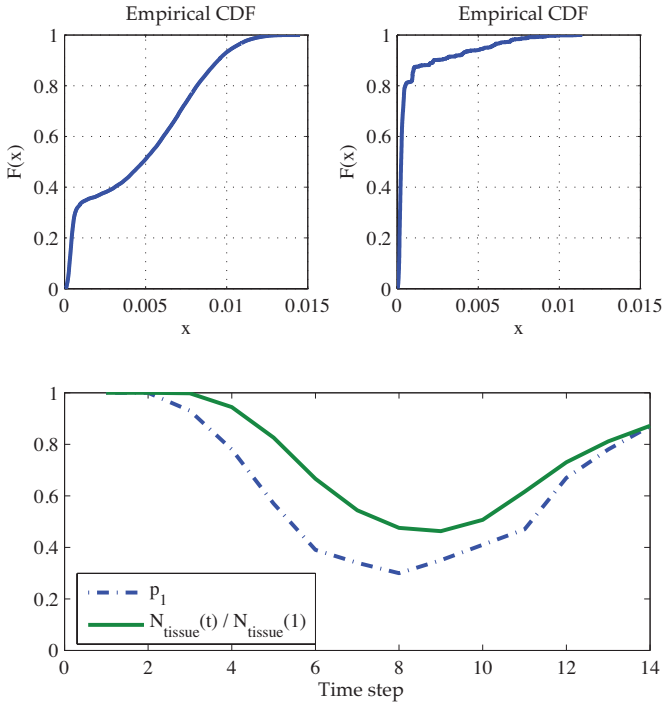


FIGURE 2.63: Piste d'amélioration du simulateur échographique relatives au travail d'Oscar Talcoth. En haut, amélioration des normes d'erreur de déplacement quand on ne conserve (a posteriori) plus que les points de corrélation élevée dans la détermination du flot optique. En bas, phénomène de perte de speckles (en vert) lié à un échantillonnage spatial incompatible avec la contraction du tissu.

synthétique ne favorise (ou défavorise) pas les outils de post-traitements. De nombreux efforts sont encore à réaliser sur ce point par le simulateur échographique. En particulier, dans la gestion des speckles, car nous avons observé que certains d'entre eux étaient arbitrairement perdus au cours de l'échantillonnage spatiale.

Ce travail est encore une fois en cours et sert d'introduction devant cette nécessité fondamentale d'analyser en détails les opérateurs d'observations et les mesures en tant que sorties du modèle. Car ces sorties se transforment en entrées des problèmes inverses d'estimation que nous allons maintenant aborder.

REMARQUE 2.5.6 (OPÉRATEUR D'OBSERVATION EULERIEN DE VITESSES SUR LE BORD)
Là aussi, comme pour les données de segmentation IRM ou les tags, on pourrait concevoir un observateur sur les vitesses directement en configuration eulérienne au travers d'un opérateur $D(Z, X) = 0$ à imaginer ... Par exemple, dans le cas de vitesses sur le bord, comparer la vitesse du point du contour déformé à celle de son projeté sur l'image.

2.5.3 Mesures de pression

Le système cardiaque ne se résume pas aux seules variables cinématiques, déplacements et vitesses, mais il faut aussi prendre en compte les variables de pression. Celles-ci sont d'autant plus importantes qu'elles sont l'unique reflet des contraintes à l'intérieur du système. Or, pour l'identifiabilité globale du système, on imagine déjà qu'il sera difficile d'identifier de façon absolue des raideurs ou contractilités si on ne dispose pas d'un étalon général de la contrainte dans le tissu.

Hélas, les mesures de pression ne sont pas simplement accessibles. Le type de mesure le plus employé est invasif et correspond à l'introduction d'un cathéter dans les ventricules. La pression dans le ventricule droit est dans ce cas plus simple à mesurer, car le cathéter est remonté à l'intérieur du système veineux basse pression jusqu'à l'oreillette et le ventricule droit. On dispose alors de mesures directes d'une variable d'état du système cardiaque.

Il n'existe pas de mesures non-invasives en milieu clinique des variables de pression. Cependant, Crépeau et Sorine (2007) ; Laleg et al (2007) ont imaginé une méthodologie pour identifier la pression en sortie du ventricule à partir de mesures de pressions au doigt P^f. Il existe en effet des dispositifs, appelés FINAPRESS (brevet américain déposé par Wesseling (1985)), qui permettent des mesures de pression au doigt précises au cours du temps, c'est-à-dire à l'extrémité du réseau artériel. On peut alors imaginer utiliser cette pression soit

- comme mesure $P^f_{\mathfrak{m}}$ sur le modèle existant via un opérateur d'observation à définir à partir de la fonction de transfert $(P^{ar}, P^d) \to P^f_{\mathfrak{m}}$;
- comme étage supplémentaire de la modélisation conduisant alors à une nouvelle mesure directe d'une variable d'état P^f.

2.5.4 Mesures électriques

Même si l'activité électrique n'est pas le sujet de cette thèse, nous pouvons donner rapidement quelques éléments autour des mesures de l'activité électrique en entrée du système. À

terme, il sera de toute façon indispensable de coupler assimilations de données mécaniques et électriques pour une mise en place effective dans un contexte clinique (voir les perspectives méthodologiques en fin de document Section 4.5). Nous renvoyons notamment à Moreau-Villégier (2005) pour une revue détaillée des différentes modalités de mesure.

L'ECG est le premier type de mesure. Non-invasif, il comporte jusqu'à 12 dérivations mais, dans certaines expérimentations, des vestes comportant quelques 250 électrodes peuvent être utilisées. Cette mesure est une observation distante de l'activité cardiaque, et la question de son utilisation en assimilation de données se pose à nouveau, soit avec un opérateur d'observation complexe à partir des composantes électriques à la surface du myocarde, soit en incluant, au reste du modèle, le modèle de diffusion dans le thorax. Ici, cette dualité est moins importante que pour les variables de pression car le phénomène physique est statique (en supposant le cœur isolé) mais, en revanche, on sait que le problème de l'inversion de la matrice de transfert coeur – thorax est mal posé.

Pour éviter ces difficultés théoriques, il faut alors passer à des mesures plus directes, donc invasives. Nous pouvons citer des exemples d'expérimentations sur des modèles animaux. Tout d'abord McVeigh et al (2001) a mis au point un protocole d'expérimentation consistant à positionner une chaussette d'électrodes à la surface d'un coeur de chien pour en tirer directement des mesures de potentiels surfaciques. Une autre stratégie, employée par M. et al (2007), utilise l'imagerie par fluorescence pour obtenir des « cartes de l'état de dépolarisation ». Ces cartes doivent encore être étudiées pour déterminer précisément l'opérateur d'observation passant du potentiel transmembranaire à celles-ci. Dans le pire des cas, si elle rendent simplement compte du passage de la dépolarisation, elles pourront être utilisées comme un masque au même titre que les images segmentées sont un masque de la déformation mécanique. Cette méthodologie se limite pour l'instant à des expérimentations et des coeurs ex-vivo tant que l'imagerie fluorescente ne permet pas une acquisition à travers le thorax. Le dernier exemple, présenté dans Sermesant et al (2004), est lui aussi invasif mais néanmoins mis en oeuvre en milieu clinique au King's College de Londres. Le principe est d'utiliser un cathéter disposant d'un ballon flexible d'électrodes. Ce ballon est introduit dans la cavité cardiaque et n'entre pas en contact avec la surface du tissu afin de ne pas en modifier l'activité électrique. Les courants détectés sont alors extrapolés sur la surface de la cavité pour reconstituer le potentiel endocardique.

2.6 Modèles et observateurs en 1D

Nous finissons ce chapitre orienté modélisation, y compris côté observations, par une version 1D des modèles présentés. Il y a deux objectifs à cette formulation. Tout d'abord ce modèle 1D est par définition plus simple en terme d'implémentation mais aussi de complexité algorithmique. Donc, il va nous permettre de tester des méthodes d'estimation sur des configurations moins demandeuses en terme de puissance de calcul. A l'autre bout de la chaîne, c'est-à-dire du point de vue théorique, le modèle 1D offre de nombreuses simplifications voir même spécificités (les fonctions H^1 sont continues etc...) ; en particulier l'élastodynamique redevient un système scalaire semblable structurellement à l'équation des ondes. Enfin, cette modélisation n'est pas inutile du point de vue pratique car le modèle de contraction cardiaque que nous venons de formuler provient de l'analyse physiologique du comportement mécanique de la

fibre cardiaque. Or celle-ci est fondamentalement unidimensionnelle et le paramétrage de ce modèle peut-être évalué à partir d'expériences directes sur celles-ci. Nous l'avions rapidement évoqué mais il existe plusieurs types d'expériences sur des éléments de fibres musculaires de l'échelle macroscopique à l'échelle microscopique.

- A l'échelle macroscopique, les expériences menés par Huxley ou Hill (1975) font encore partie du paysage physiologique. Elles consistent à accrocher un élément de muscle caridiaque (plus présisément muscle papillaire) de rat entre une butée et une masse et de soumettre l'ensemble à une sollicitation électrique comme présenté Figure 2.64. Elles permettent de décomposer le comportement de la fibre musculaire en plusieurs modes de fonctionnement : isotonique, isotonique etc...

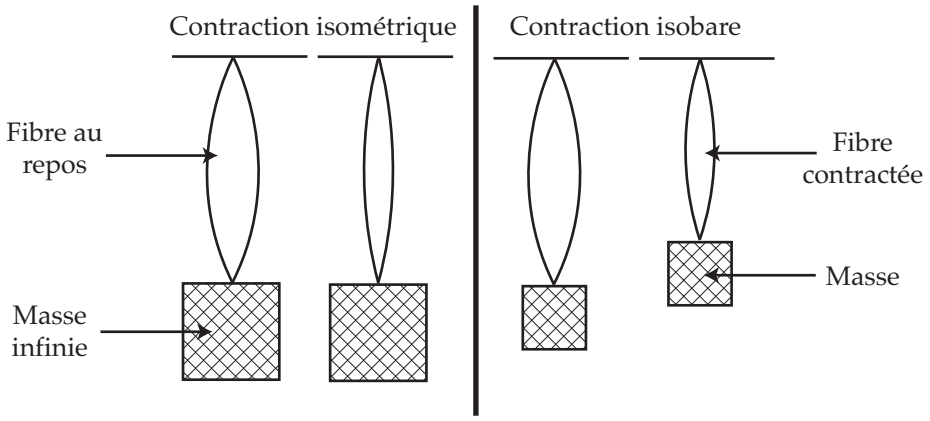


FIGURE 2.64: Expérience de contraction isométrique ou isotonique sur des échantillons de muscle cardiaques

- A une échelle inférieure, on peut citer par exemple les travaux de Pustoc'h et al (2005) où sont enregistrés par vidéomicroscopie les contractions de cardiomyocytes de rat en culture. Il est alors possible de suivre la propagation de l'onde de raccourcissement longitudinal qui traverse la cellule pendant sa contraction. De plus l'onde de dépolarisation peut être suivie par marquage fluorescent post-traité par des techniques de flot optique.

- A l'échelle microscopique, il est enfin possible de réaliser le même type d'expériences que Huxley ou Hill sur une seule cellule cardiaque (voir Cazorla et al (2003)). Par exemple, l'utilisation de pinces optiques permet de « tenir » par deux lasers les extrémités d'une cellule cardiaque et de la soumettre au même type de sollicitations et de déformations isotoniques, isotoniques....

Ces différentes expériences peuvent évidemment donner lieu à des modélisations fines 3D, comme dans Cottet et Maitre (2006), mais motivent l'utilisation de modèles unidimensionnels dans des perspectives de paramétrage, de calibration et pourquoi pas d'estimation.

A partir de l'ensemble de toutes ces considérations nous présentons ici le petit jeu mathématique qui consiste, par exemple, à définir un modèle 1D de cavités cardiaques à partir du

modèle 3D dont on restreint la cinématique.

$$2\pi R_0 r_0 \left(\int_0^L \rho_0 \frac{\partial^2 y}{\partial t^2} \cdot v \, \mathrm{dx} + \int_0^L \Sigma_{\mathrm{xx}} \, \mathrm{d}_y e \cdot v dx \right) = P \pi R_0^2 v(L,t), \tag{2.65}$$

où $\mathrm{d}_y e \cdot v = (1 + \frac{\partial y}{\partial \mathrm{x}}) \frac{\partial v}{\partial \mathrm{x}}$ est la variation du tenseur de Green-Lagrange et $\pi R_0^2 v(L,t)$ est la puissance virtuelle des efforts de pression à l'extrémité du piston.

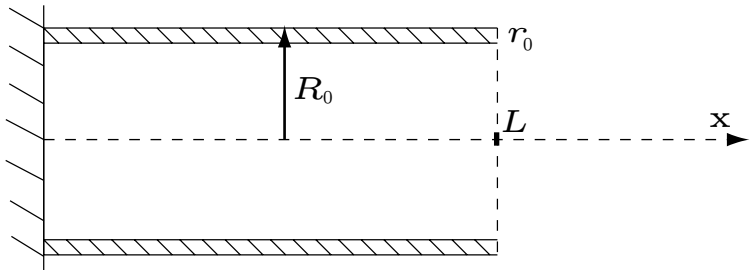


FIGURE 2.65: Exemple de cavité 1D

Bien évidemment on pourrait aussi obtenir un modèle absolument équivalent en considérant un piston plein qui peut monter en pression à longueur constante.

Du point de vue de la formulation forte de ces équations, la structure scalaire monodimensionnelle nous rapproche de l'équation des ondes. Pour se fixer les idées, écrivons-la dans l'hypothèse « des petits déplacements » qui implique que $\mathrm{d}_y e \cdot v = \frac{\partial v}{\partial \mathrm{x}}$ et $\Sigma_{\mathrm{xx}} = E \frac{\partial y}{\partial \mathrm{x}} + \eta \frac{\partial y}{\partial \mathrm{x} \partial t} + \sigma_c$, où σ_c est obtenu par les lois rhéologiques linéaires sur le modèle présenté Section 2.3.2.e.. On obtient alors l'EDP

$$\begin{cases} \rho_0 \dfrac{\partial^2 y}{\partial t^2} - \eta \dfrac{\partial^3 y}{\partial \mathrm{x}^2 \partial t} - E \dfrac{\partial^2 y}{\partial \mathrm{x}^2} - \dfrac{\partial \sigma_c}{\partial \mathrm{x}} = 0 \\ E \dfrac{\partial y}{\partial \mathrm{x}}(L,t) + \eta \dfrac{\partial y}{\partial \mathrm{x} \partial t}(L,t) + \sigma_c(L,t) = \dfrac{P R_0}{2 r_0} \end{cases} \tag{2.66}$$

En l'absence d'activation ($\sigma_c = 0$) on retrouve bien l'équation des ondes amorties. De même, si on suppose que σ_c n'est pas fonction de la déformation au premier ordre mais simplement de l'activation. La contrainte active σ_c devient alors une **précontrainte** active et on retrouve l'équation des ondes amorties en présence de forçage.

Pour ce qui est des observations, les différents types d'expériences présentées plus haut conduisent à imaginer les pendants 1D des observateurs présentés pour le coeur. En particulier on peut supposer qu'on dispose soit de mesures en différents points de la fibre soit seulement à l'extrémité de celle-ci :

$$\mathcal{H} y = y(L,t).$$

Dans ce cas, notre modélisation nous rapproche de la structure d'un problème très largement étudié en contrôle qui est l'équation des ondes contrôlée par le bord, sauf que dans notre cas nous allons considérer le problème de son observation frontière et non de son contrôle. Cette réduction nous est donc très utile pour formuler, tester, justifier et analyser la problématique d'assimilation de données dans le paysage scientifique actuel, et donc nous aider à formuler des stratégies originales. C'est tout l'objectif des chapitres suivants.

2.7 Conclusion

Nous venons de présenter l'ensemble des briques nécessaires à la formulation d'un modèle complet de contraction cardiaque. Nous avons souhaité cette présentation la plus exhaustive possible sur l'ensemble des travaux de modélisation et d'analyse numérique de Bestel (2000) puis Sainte-Marie et al (2006) jusqu'à aujourd'hui avec Chapelle et al (2008). C'est pour nous l'occasion de présenter dans son ensemble ce travail de synthèse effectué notamment depuis 3 ans avec la perspective nouvelle côté mécanique de structurer la modélisation et les aspects numériques dans l'optique de l'assimilation de données. Cependant, malgré son aspect un peu encyclopédique, cette présentation n'a pas vocation à imposer le modèle présenté par rapports aux autres. Cela nécessiterait un travail très important de validation quantitative qui n'était pas l'objectif de cette thèse. Encore une fois, notre besoin était de définir et structurer le modèle direct le plus précisément possible comme pré-requis à son inversion. En particulier, sa structure énergétique qui a guidé son développement est un avantage considérable en ce sens. Cependant ce modèle doit encore évoluer et être confronté intensivement à l'expérience pour représenter une solution alternative pertinente à des modèles mécaniques cardiaques historiques. Ce travail est d'ailleurs en cours et je remercie encore une fois R. Chabiniok, médecin en thèse au sein de l'équipe sur la validation à partir de données réelles 1D et 3D, pour les illustrations des différents résultats qu'il a obtenu sur la base du modèle et du code développés.

En assimilation de données, les objectifs de la simulation directe ne se limitent pas au phénomène physique étudié mais doivent aussi s'attacher à bien représenter les différents types de mesures utilisés par la suite en entrée des problèmes d'estimation. Nous avons donc voulu montrer quels étaient les enjeux que l'observation du système cardiaque implique, en particulier en échographie avec la collaboration avec E. Angelini, O. Talcoth, Q. Duan et A. Laine. De nombreux problèmes restent ouverts et impliquent des développements à la fois en modélisation de la mesure mais aussi en traitement d'images et des signaux. De plus, une fois toutes le modalités de mesure maîtrisées individuellement, il reste une difficulté majeure peu mentionnée dans notre exposé, avec la fusion des différentes données que ce soit spatialement (sur le même domaine géométrique) ou du point de vue temporel (point de départ du battement). Ce n'est qu'une fois toutes ces étapes réalisées qu'on peut traiter **en pratique** le problème de l'assimilation de données.

CHAPITRE

3

Estimation d'état pour une classe de systèmes hyperboliques

Tout est dans le contrôle

— Michel Platini

Sommaire du chapitre

3.1 Introduction **158**

3.2 Estimation d'état moindres carrés **161**

3.2.1 Formulation variationnelle spécifique 161

3.2.2 Approche séquentielle et complexité 165

3.2.3 Problématique 167

3.3 Filtrage asymptotique de Luenberger **168**

3.3.1 Système d'erreur 168

3.3.2 Fonctions de Liapounov 171

3.3.3 Stabilisation et contrôle 175

3.4 Estimation d'état en élastodynamique à partir de vitesses **177**

3.4.1 Rappels sur la position du problème 177

3.4.2 Définition de l'estimateur DVF 182

3.4.3 Estimations en dimension infinie issues du contrôle 183

3.4.4 Problème aux valeurs propres quadratiques 187

3.4.5 Estimation d'erreur en fonction des bruits 197

3.4.6 Observateur frontière 206

3.4.7 Extension en dynamique non-linéaire 212

3.5 Estimation d'état en élastodynamique à partir de déplacements **215**

3.5.1 Définition de l'estimateur 216

3.5.2 Analyse de l'estimateur 219

3.5.3 Schéma en temps 224

3.5.4 Résultats d'estimation 227

3.5.5 Cas d'observations eulériennes 227

3.5.6 Résultats 233

3.5.7 Extension en dynamique non-linéaire 233

3.5.8 Résultats d'estimation d'état pour le cœur 237

3.6 Conclusion et perspectives **243**

3.1 Introduction

L'Ensemble des éléments présentés jusqu'à présent nous permettent de préciser la problématique qui nous occupe réellement. En toute rigueur notre objectif est de simuler le comportement cardiaque d'un patient donné. A supposer que le travail de modélisation du chapitre 2 soit satisfaisant et intensivement validé, on peut considérer qu'on dispose, relativement à une application donnée (par exemple le diagnostic de troubles mécaniques de la contraction), de modèles satisfaisants à l'image de ce que l'on a l'habitude de considérer en sciences de l'ingénieur. Ainsi le système réel cardiaque est, du point de vue mécanique, représenté par ses champs de déplacements et vitesses $(\underline{y}, \underline{\dot{y}})$ vérifiant sous forme variationnelle les équations aux dérivées partielles présentées au chapitre 2. Ces formulations peuvent être écrites sous forme de système dynamique (par l'intermédiaire d'un semi-groupe typiquement en linéaire)

$$\begin{cases} \dot{x} = \mathcal{A}(x,t) \\ x(0) \end{cases} \tag{3.1}$$

Classiquement, dans un objectif de diagnostic ou de pronostic, cette dynamique est simulée numériquement via une discrétisation en espace et en temps du même type que celles présentées au chapitre 2 dont on peut analyser la convergence : c'est l'objectif de l'analyse numérique. La discrétisation en espace donne lieu à un système dynamique défini sur $\mathbb{R}^N$ du type

$$\begin{cases} \dot{X} = A(X,t) \\ X(0) \end{cases} \tag{3.2}$$

Par exemple dans le cas d'une méthode de Galerkin type éléments finis, on définit un champ $x_h = \mathcal{I}_h X$ où $\mathcal{I}_h$ est l'opérateur d'interpolation correspondant à l'ordre d'approximation considéré (ici $\mathbb{P}_1$). Les résultats d'analyse numérique fournissent alors une estimation d'erreur du type

$$\|x - x_h\|_{\mathcal{E}} = \mathcal{O}(h^p).$$

Par la suite, la discrétisation en temps approche $X(n\Delta t)$ par X_n en définissant un opérateur de transition

$$\begin{cases} X_{n+1} = A_{n+1|n}(X) \\ X_{n=0} \end{cases} \tag{3.3}$$

Cette fois l'analyse numérique fournit une estimation du type

$$\forall n, \quad \|X(n\Delta t) - X_n\|_{\mathcal{E}} = \mathcal{O}(\Delta t^q).$$

Par inégalité triangulaire x est ainsi approché en espace et en temps par X_n. Cette succession d'étapes assure le rôle, la nécessité et le quotidien de l'analyse numérique, mais procède d'un parti pris (en sus de l'adéquation du modèle aux phénomènes observés) : le fait que l'ingénieur dispose des entrées nécessaires à la simulation. Or, si nous supposons les conditions aux limites comme partie intégrante de la modélisation, supposée « juste » dans ce chapitre, la seule entrée indispensable est la condition initiale du système. La situation que nous étudions dans ce chapitre correspond alors au cas où cette connaissance n'est pas suffisamment précise par rapport aux ordres d'approximation numérique choisis. A une translation dans le temps près, cela signifie qu'à aucun moment on ne connaît suffisamment précisément l'état exact du

système. A la limite, ce qu'on espère de l'ingénieur est qu'il connaisse, par expérience, l'état du système « en moyenne » sur l'ensemble des réalisations du phénomène qu'il connaît. Nous notons x_0 cette connaissance si bien que

$$x(0) = x_0 + \zeta^x, \quad \mathbb{E}(\zeta^x) = 0.$$

Alors, par discrétisation en espace, on peut supposer de même qu'il existe ζ^X tel que

$$X(0) = X_0 + \zeta^X, \quad \mathbb{E}(\zeta^X) = 0.$$

Cette erreur initiale sur l'état limite la prédiction du modèle sur les intervalles futurs, *a fortiori* si la stabilité du système étudié est très dépendante de la condition initiale. En tout état de cause, même pour des systèmes stables, une incertitude de cette nature rend caduc un résultat de convergence d'analyse numérique au moins pendant le régime transitoire de la stabilisation de l'erreur. Or en mécanique cardiaque, il n'est pas clair que la dynamique du système nous assure une stabilisation de cette connaissance imparfaite sur la fenêtre d'analyse que nous nous sommes fixée, comprise dans le battement cardiaque. A défaut de cette connaissance, et comme dans les autres disciplines d'assimilation de données que sont la météorologie, la climatologie, la géophysique etc., on bénéficie en mécanique cardiaque de l'apport supplémentaire que représentent les mesures notamment l'imagerie cardiaque. Ces mesures contribuent à une connaissance partielle de l'état du système à différents instants. Nous l'avons déjà noté en introduction

$$Z = \mathcal{H}(x) + \chi,$$

et précisé ces opérateurs et bruits en fonction de la modalité. Ce chapitre souhaite donc apporter une réponse à la problématique suivante : **est-il possible de bénéficier des données d'observations pour contre-balancer notre connaissance imparfaite de la condition initiale ?** Dans un langage plus d'automaticien, les observations sont normalement des sorties du problème direct alors que la condition initiale est une entrée, donc l'objectif est de bénéficier de ces sorties pour retrouver une donnée d'entrée inconnue. Encore une fois, cette notion de donnée d'entrée et de sortie caractérise un problème inverse mais l'objectif final est un résultat d'analyse numérique.

Pour obtenir ce résultat, nous nous ramenons d'abord au système discret. Par exemple, pour la discrétisation en espace, soit $\hat{X}$ un estimateur de X. Alors on peut déterminer un vecteur de champ $\hat{x}_h = \mathcal{I}_h \hat{X}$ défini sur l'espace de x. On a alors, pour toute norme adaptée à x,

$$\|x - \hat{x}_h\| \leq \|x - x_h\| + \|X - \hat{X}\|,$$

sachant que

$$Z = H(X) + \chi + \epsilon_h,$$

où $\epsilon_h = \mathcal{H}(x) - \mathcal{H}(x_h)$ et $H = \mathcal{H} \circ \mathcal{I}_h$. Ainsi il suffit bien d'analyser la performance de l'estimateur discret $\hat{X}$ de $X \in R^N$, **uniformément par rapport à la discrétisation**. Le même type de manipulation peut ensuite être opéré pour exprimer Z en fonction de X_n avec un $\epsilon_{\Delta t}$ ou de considérer $\hat{X}_n$ comme une discrétisation en temps de $\hat{X}$. Enfin notons qu'en toute rigueur les données Z sont, comme nous l'avons déjà dit, fournies en nombre discret, notamment en temps tous les $\Delta T >> \Delta t$, mais allons voir comment on peut les considérer disponibles continûment en temps par simple interpolation en introduisant un autre $\epsilon_{\Delta T}$.

□ INTERPOLATION DES MESURES – Jusqu'à présent, nous avons toujours défini notre estimation à partir de données d'observations fournies par l'instrument de mesure *en continu* ou en tout cas à tous les pas de temps de discrétisation sachant que celui-ci a vocation à tendre vers 0 ! Cette approche n'est certes pas réaliste mais elle trouve sa justification dans les remarques suivantes.

Supposons que nous disposions seulement de $Z_k = Z(k\Delta T),\ 1 \leq k \leq N_{\Delta T}$. Par exemple pour les tags nous avons vu au chapitre précédent que $N_{\Delta T} = 20$ sur une seconde de battement donc $\Delta T \simeq 0.05s$ à comparer à $\Delta t < 1e-03$ pour capter les phases isovolumiques. Définissons alors une mesure continue Z_{int} reconstruite par interpolation. En l'absence de connaissance plus précise, on procède par interpolation linéaire :

$$Z_{\text{int}} = \alpha_k(t)Z_k + (1-\alpha_k(t))Z_{k+1},\ \text{avec}\ \alpha_k(t) = \tfrac{(k+1)\Delta t - t}{\Delta t}.$$

Dans ce cas on substitue Z_{int} à Z dans l'équation de l'estimateur et donc de l'erreur

$$\dot{\tilde{X}} = (A - K_X H)\tilde{X} - K_X(\chi + \epsilon_h + (Z_{\text{int}} - Z)).$$

Or $(Z_{\text{int}} - Z) = \mathcal{O}(\Delta T)$ ce qui est tout à fait attendu et ne va pas à l'encontre de la consistance de l'estimateur défini en introduction du premier chapitre. De plus, on peut espérer d'autres formes d'interpolation, fondées notamment sur une connaissance du système, telles que l'erreur d'interpolation soit, par exemple, quadratique. Rien n'empêche d'imaginer des interpolateurs s'appuyant en partie sur la physique du phénomène même si, dans ce cas, on peut légitimement s'interroger sur la redondance de cette étape vis-à-vis de l'estimateur. Nous avons déjà vu ces notions lors de l'analyse des observateurs et c'est une question ouverte de pouvoir justifier de l'intérêt, tour à tour, des différentes phases d'interpolation, de traitement ou d'estimation sur l'ensemble du processus d'assimilation. Le juge de paix reste une fois de plus la convergence de l'erreur, elle-même entièrement conditionnée par les questions d'observabilité. Ainsi toute interpolation en temps ou en espace est un cas particulier d'estimation qui vise simplement à utiliser un certain modèle comme une régularisation. Cette approche est donc utile tant que la régularisation ne modifie pas l'observabilité et donc ne filtre pas une information cruciale pour l'estimation finale. Si on pouvait assurer que l'interpolateur en temps « physique » est quadratique par rapport à la vraie donnée, il faudrait absolument l'utiliser, mais en pratique c'est rarement le cas.

Enfin, examinons le cas du bruit en plus de l'interpolation en temps des données. En effet, rien n'empêche de compter $(Z_{\text{int}} - Z)$ dans le bruit de mesure. Toutefois, nous avons beaucoup parlé de d'indépendance des bruits au cours du temps pour motiver les filtres sur les considérations probabilistes. Or, autant il facile d'imaginer ce que signifie alors un bruit blanc entre les différentes mesures Z_k :

$$(\chi_k)_k \text{ est une suite de bruits blancs en temps si et seulement si } E(\chi_i\chi_j) = \delta_{ij}W,$$

autant il est plus difficile de le manipuler en temps continu (cf. Section 1.3.3). Ainsi, continuer à supposer le bruit blanc en temps est artificiel à deux niveaux : artificiel car c'est avant tout un outil théorique, et artificiel car en réalité le bruit blanc est associé à la discrétisation de la mesure tous les ΔT. Côté mise en oeuvre numérique, interpoler la mesure mais continuer à jouer le jeu du bruit blanc aura une incidence sur le choix de W. On avait vu au chapitre 1

(cf. équation (1.68)) que le bruit blanc continu sous-jacent à un échantillonnage de mesure ΔT était

$$W = \Delta T\, W_{\Delta T},$$

avec $W_{\Delta T}$ la covariance de bruit de la mesure échantillonnée (Z_k). Identiquement, la discrétisation en temps de ce bruit blanc continu, consistante avec celle de la dynamique, serait naturellement (cf. équation (1.66) issue de la discrétisation du critère optimal)

$$\chi_n = \frac{1}{\Delta t}\int_t^{t+\Delta t} \chi,$$

de covariance

$$W_n = \frac{1}{\Delta t} W = \frac{\Delta T}{\Delta t} W_{\Delta T}.$$

Nous insistons sur le fait que l'instrument de mesure livrerait naturellement $W_{\Delta T}$ et que le W continu sous-jacent est largement artificiel.

3.2 Estimation d'état moindres carrés

3.2.1 Formulation variationnelle spécifique

a. Dérivation en élastodynamique de la minimisation

Nous commençons ce chapitre en mettant en oeuvre dans un cas d'élastodynamique linéaire ce que nous avons vu au premier chapitre en estimation moindres carrés. Nous développons la démarche classique d'estimation d'état par l'approche variationnelle en assimilation de données (voir Blum et al (2008)) avant d'expliquer dans la section suivante pourquoi il n'existe pas de résultats d'estimation d'état par filtrage de Kalman d'un système mécanique distribué (i.e. régi par un modèle de mécanique des milieux continus). A titre introductif nous nous limitons à l'estimation sur le système discret et négligeons dans un premier temps les erreurs de discrétisation. Seule la mise en oeuvre nous importe ici et nous servira de référence pour la suite. Le système discret est donc régi par les équations présentées au chapitre modélisation :

$$\begin{cases} M\ddot{Y} + C\dot{Y} + KY = F(t) \\ Y(0) = Y_0 + \zeta^Y \\ \dot{Y}(0) = \dot{Y}_0 + \zeta^{\dot{Y}} \end{cases} \tag{3.4}$$

réécrit sous la forme du premier ordre comme présenté au chapitre précédent

$$\begin{cases} \dot{X} = AX + R \\ X(0) = X_0 + \zeta^X \end{cases} \tag{3.5}$$

où

$$A = \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ M^{-1}F \end{pmatrix},$$

avec des observations du type

$$Z = HX + \chi = \begin{pmatrix} H_d & 0 \\ 0 & H_v \end{pmatrix} \begin{pmatrix} Y \\ \dot{Y} \end{pmatrix} + \begin{pmatrix} \chi^d \\ \chi^v \end{pmatrix}, \quad W = \begin{pmatrix} W^d & 0 \\ 0 & W^v \end{pmatrix},$$

et pour simplifier, H^d, H^v et donc W^d,W^v sont supposés indépendant du temps. Le critère de coût est celui du chapitre 1

$$J(\xi) = \frac{1}{2}\|\xi\|^2_{P_0^{-1}} + \frac{1}{2}\int_0^T \|Z - HX_\xi\|^2_{W^{-1}}, \tag{3.6}$$

dont la différentielle est toujours

$$\mathrm{d}J \cdot \delta\xi = \xi^T P_0^{-1}\delta\xi - p(0)^T\delta\xi, \tag{3.7}$$

avec p vérifiant l'équation adjointe

$$\begin{cases} \dot{p} + A^T p = -H^T W^{-1}(Z - HX) \\ p(T) = 0 \end{cases}$$

Nous pouvons désormais préciser cette dynamique adjointe en fonction de l'opérateur A. Notons[1] $p = \left(\begin{smallmatrix} p^d \\ p^v \end{smallmatrix}\right)$: On a

$$\begin{cases} \dot{p} + A^T_{\text{skew}} N^{-1} p = -H^T W^{-1}(Z - HX) \\ p(T) = 0 \end{cases}$$

avec

$$A_{\text{skew}} = \begin{pmatrix} 0 & K \\ -K & -C \end{pmatrix}, \quad N = \begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix}$$

Donc en posant $q = \left(\begin{smallmatrix} q^d \\ q^v \end{smallmatrix}\right) = N^{-1}p$, on obtient

$$\begin{cases} \dot{q}^d = q^v - K^{-1}H^{d^T}(W^d)^{-1}(Z^d - H^d Y) \\ M\dot{q}^v + Kq^d - Cq^v = -H^{v^T}(W^v)^{-1}(Z^v - H^v\dot{Y}) \end{cases}$$

Cette équation peut être reformulée sous forme d'une équation du second ordre si nous supposons la mesure suffisamment régulière (et donc c'est avant tout une réécriture formelle)

$$\begin{cases} M\ddot{q}^d - C\dot{q}^d + Kq^d = -H^{v^T}(W^v)^{-1}(Z^v - H^v\dot{Y}) \\ \qquad\qquad -MK^{-1}H^{d^T}(W^d)^{-1}(\dot{Z}^d - H^d\dot{Y}) + CK^{-1}H^{d^T}(W^d)^{-1}(Z^d - H^d Y) \\ q^d(T) = 0 \\ \dot{q}^d(T) = K^{-1}H^{d^T}(W^d)^{-1}(Z^d(T) - H^d Y^d(T)) \end{cases}$$

Donc q vérifie bien une équation de la dynamique « **rétrograde** ». La présence du $-C$ permet de conserver l'aspect dissipatif en sens rétrograde. On remonte alors aisément à p sous la forme

$$\begin{cases} p^d(0) = Kq^d(0) \\ p^v(0) = Mq^v(0) = M(\dot{q}^d(0) + K^{-1}H^{d^T}(Z^d(0) - H^d Y(0))) \end{cases}$$

Une fois p simulé (nous verrons plus bas comment en pratique), on obtient avec $p(0)$ le gradient de J que l'on peut minimiser jusqu'à convergence par un algorithme de type descente de gradient (méthode de descente classique ou avec « line-search », méthode de Gauss, méthode de gradient-conjugué, voir Bonnans et al (2006)).

Ces équations nous permettent de faire quelques remarques générales :

[1] Comme déjà fait implicitement, nous utiliserons couramment l'exposant « d » pour caractériser des variables ou des opérateurs associés au déplacements et « v » pour les vitesses

- La première est classique dans les approches variationnelles. Nous venons de préciser la différence entre une équation adjointe et un retournement temporel. En effet la première fait apparaître un $-C$ et est aussi stable que la dynamique directe alors que le second serait instable.
- La deuxième est plus spécifique à la nature des observations utilisées ici. On remarque que les observations de vitesses et déplacements ne sont pas intégrées de la même façon dans l'équation adjointe. En particulier cette dernière est largement simplifiée si $H^d = 0$ et ne nécessite pas la dérivation de la mesure. Nous verrons plus loin les causes profondes de cette structure.

b. Schéma en temps

Intéressons nous maintenant à la discrétisation en temps nécessaire à la mise en oeuvre numérique. Nous avons déjà vu que le schéma de Newmark présentait les bonnes propriétés (consistance quadratique en temps et stabilité énergétique) pour le problème direct. En pratique, on a vu que, en définissant

$$A_1 = \begin{pmatrix} \frac{1}{\Delta t} & -\frac{1}{2} \\ \frac{1}{2}K & \frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}, \quad A_0 = \begin{pmatrix} \frac{1}{\Delta t} & \frac{1}{2} \\ -\frac{1}{2}K & -\frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}, \quad \tilde{R}_{n+1} = \begin{pmatrix} 0 \\ F^e_{n+\frac{1}{2}} \end{pmatrix},$$

l'opérateur de transition s'écrivait $A_{n+1|n} = A_1^{-1}A_0$ et $R_{n+1} = A_1^{-1}\tilde{R}_{n+1}$ tel que

$$\begin{cases} A_1 X_{n+1} = A_0 X_n + \tilde{R}_{n+1} \\ X_{n=0} = X_0 + \zeta^X \end{cases}$$

La question légitime à ce stade est : que faut-il faire pour le problème adjoint ? A cette question, deux réponses sont possibles **a priori** :

- soit dériver un schéma à partir de l'équation adjointe continue,
- soit se donner un critère variationnel discret et en déduire la dynamique adjointe discrète associée.

La réponse n'est jamais absolue sachant que le juge de paix est normalement la convergence de la procédure variationnelle (itérations m dans la direction du gradient) couplée au schéma en temps (itérations k du schéma). Classiquement, il est conseillé (notamment Chavent (2008)) de choisir la deuxième pratique pour obtenir au final cette convergence. Ainsi, on reprend la version discrète introduit Section 1.3.1 du critère (3.6)

$$\mathrm{d}J_n(\xi) \cdot \delta\xi = \xi^T P_0^{-1} \delta\xi - \sum_{k=0}^{n} (Z_k - HX_k)^T W_{\Delta t}^{-1} H d_\xi X_k \cdot \delta\xi, \tag{3.8}$$

avec $W_k = \frac{W}{\Delta t} = W_{\Delta t}$. Rappelons alors l'équation adjointe

$$\begin{cases} p_k - A_{k+1|k}^T p_{k+1} = H^T W_{\Delta t}^{-1} (Z_k - HX_k) \\ p_{n+1} = 0 \end{cases} \tag{3.9}$$

et la dérivée du critère

$$\mathrm{d}J_n(\xi)\cdot\delta\xi = \xi^T P_0^{-1}\delta\xi - p_0^T\delta\xi.$$

Définissant alors la variable $q_k = A_1^{-T}p_k$, on a

$$A_1^T q_k = A_0^T q_{k+1} + H^T W_{\Delta t}^{-1}(Z_k - HX_k).$$

En posant $S_k = \left(\begin{smallmatrix} S_k^d \\ S_k^v \end{smallmatrix}\right) = H^T W_{\Delta t}^{-1}(Z_k - HX_k)$, on obtient

$$\begin{pmatrix} \frac{1}{\Delta t} & \frac{1}{2}K \\ -\frac{1}{2} & \frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}\begin{pmatrix} q_k^d \\ q_k^v \end{pmatrix} = \begin{pmatrix} \frac{1}{\Delta t} & -\frac{1}{2}K \\ \frac{1}{2} & -\frac{1}{2}C + \frac{1}{\Delta t}M \end{pmatrix}\begin{pmatrix} q_{k+1}^d \\ q_{k+1}^v \end{pmatrix} + \begin{pmatrix} S_k^d \\ S_k^v \end{pmatrix}$$

Ainsi on a

$$\begin{cases} K\dfrac{q_{k+1}^v + q_k^v}{2} = \dfrac{q_{k+1}^d - q_k^d}{\Delta t} + S_k^d \\ M\dfrac{q_{k+1}^v - q_k^v}{\Delta t} - C\dfrac{q_{k+1}^v + q_k^v}{2} + \dfrac{q_{k+1}^d + q_k^d}{2} = -S_k^v \end{cases}$$

Donc en posant $K\tilde{q}_k^d = q_k^d, \quad 1 \le k \le n+1$ on obtient

$$\begin{cases} \dfrac{q_{k+1}^v + q_k^v}{2} = \dfrac{\tilde{q}_{k+1}^d - \tilde{q}_k^d}{\Delta t} + K^{-1}S_k^d \\ M\dfrac{q_{k+1}^v - q_k^v}{\Delta t} - C\dfrac{q_{k+1}^v + q_k^v}{2} + K\dfrac{\tilde{q}_{k+1}^d + \tilde{q}_k^d}{2} = -S_k^v \end{cases}$$

On retrouve alors, au second membre près, un schéma de type Newmark (voir (2.59)) pour une dynamique rétrograde (on cherche k à partir de $k+1$) avec un terme de viscosité $-C$ attendu. Autrement dit, cette équation sur q_k est exactement la version discrète de celle sur q en temps continu. Ainsi, dans cette configuration linéaire le dilemme associé à la stratégie de discrétisation évoqué plus haut se résout de lui-même ! Du point de vue algorithmique on obtient finalement

$$\begin{cases} q_k^v = \frac{2}{\Delta t}(\tilde{q}_{n+1}^d - \tilde{q}_n^d) + 2K^{-1}S_n^d - q_{n+1}^v \\ \left(\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M\right)\tilde{q}_n^d = \left(-\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M\right)\tilde{q}_{n+1}^d \\ \qquad\qquad\qquad\qquad -\frac{2}{\Delta t}Mq_{n+1}^v + \frac{2}{\Delta t}MK^{-1}S_n^d + CK^{-1}S_n^d - S_n^v \\ q_{N+1}^v = 0 \\ \tilde{q}_{N+1}^d = 0 \end{cases}$$

à comparer à l'algorithmique du schéma de Newmark (2.60). Reste alors comme en continu à calculer $p_n = A_1^T q_n$,

$$\begin{cases} p_n^d = \frac{1}{\Delta t}q_n^d + \frac{1}{2}Kq_n^v = K(\frac{1}{\Delta t}\tilde{q}_n^d + \frac{1}{2}q_n^v) \\ p_n^v = -\frac{1}{2}q_n^d + \left(\frac{1}{2}C + \frac{1}{\Delta t}M\right)q_n^v = -\frac{1}{2}K\tilde{q}_n^d + \left(\frac{1}{2}C + \frac{1}{\Delta t}M\right)q_n^v \end{cases}$$

pour obtenir une expression du gradient du critère et permettre, par un algorithme de descente, la minimisation.

3.2.2 Approche séquentielle et complexité

a. Obtention du filtre de Kalman en élastodynamique

Profitons de cette introduction pour préciser aussi comment le filtre de Kalman opère potentiellement en estimation d'état pour l'élastodynamique. Là encore, les équations générales ont déjà été présentées au chapitre 1 :

$$\begin{cases} \dot{\hat{X}} = A(t)\hat{X} + R(t) + K(Z - H\hat{X}) \\ \hat{X}(0) = X_0 \end{cases} \tag{3.10}$$

où

$$K_{\text{kal}} = PH^T W^{-1} \tag{3.11}$$

Comme $H = \left(\begin{smallmatrix} H^d & 0 \\ 0 & H^v \end{smallmatrix}\right)$, la matrice de covariance est en toute généralité

$$P = \begin{pmatrix} P^{dd} & P^{dv} \\ P^{vd} & P^{vv} \end{pmatrix},$$

et donc

$$K_{\text{kal}} = \begin{pmatrix} P^{dd}(H^d)^T(W^d)^{-1} & P^{dv}(H^v)^T(W^v)^{-1} \\ P^{vd}(H^d)^T(W^d)^{-1} & P^{vv}(H^v)^T(W^v)^{-1} \end{pmatrix}.$$

Par conséquent, la première équation du système de l'estimateur représentée par le premier bloc de A (par exemple avec $A = \left(\begin{smallmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{smallmatrix}\right)$) n'exprime plus la vitesse comme dérivée du déplacement. Si $\hat{X} = \left(\begin{smallmatrix} \hat{X}^d \\ \hat{X}^v \end{smallmatrix}\right)$, alors on a désormais

$$\dot{\hat{X}}^d = \hat{X}^v + P^{dd}(H^d)^T(W^d)^{-1}(Z^d - H^d\hat{X}^d) + P^{dv}(H^v)^T(W^v)^{-1}(Z^v - H^v\hat{X}^v). \tag{3.12}$$

Ainsi on ne peut plus considérer l'état sous la forme $\hat{X} = \left(\begin{smallmatrix} \hat{Y} \\ \dot{\hat{Y}} \end{smallmatrix}\right)$ sans confusion sur la nature de la notation $\dot{\hat{Y}}$ (i.e est-ce une notation ou une dérivée en temps).

Pour le schéma en temps associé à cet estimateur, on pourrait utiliser une de règle des trapèzes (à l'image du schéma de Newmark) sur la partie filtrée $K_{\text{kal}}(Z - H\hat{X})$, c'est-à-dire discrétiser suivant $K_n(Z_{n+\frac{1}{2}} - H\frac{\hat{X}_n + \hat{X}_{n+1}}{2})$. Cependant, le filtre de Kalman discret et son équivalence avec le critère variationnel discret fournit immédiatement un schéma en temps sous la forme « one-step » (1.44). Volontairement, nous n'avons pas encore présenté les propriétés précises de stabilité de ce schéma, en dehors de son lien avec la fonctionnelle de coût discrétisée, mais de nombreuses précisions vont être apportées au cours des sections suivantes.

On rappelle enfin que le filtrage de Kalman ne donne pas accès directement à ζ^X mais simplement à $\hat{X}(T)$. Pour retrouver la condition initiale, il faut procéder à une inversion en temps de la dynamique qui peut donner lieu à des problèmes de stabilité. Mais dans le cas de l'estimation d'état pour un battement cardiaque, on cherche à poursuivre une trajectoire donc éventuellement à rattraper une erreur de condition initiale plutôt qu'à l'identifier. De plus l'aspect quasi-périodique du battement cardiaque permet de relier directement $\hat{X}(T)$ au ζ^X du battement suivant.

b. Complexité

Il est donc envisageable *a priori* de construire le même estimateur, soit sous forme variationnelle, soit sous forme séquentielle avec, pour chacun, un schéma en temps satisfaisant. Reste, dans la mise en oeuvre, à regarder le question de la complexité algorithmique. Cette complexité est de deux natures :

- Temps de calcul noté σ_T,
- Espace mémoire nécessaire noté σ_M.

Notons `linearDyn`, l'algorithme direct de calcul de la dynamique linéaire, `VariaLQ` l'algorithme variationnel associé au problème LQ (donc à la dynamique linéaire) et `Kalman` celui associé au filtrage de Kalman. Pour information, dans le calcul direct, on n'assemble et ne calcule les cofacteurs (LU, Cholesky ou autre) de la matrice **creuse** $(\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M)$ qu'une seule fois. Au cours des itérations en temps, on ne procède alors plus qu'à des assemblages de second membre et à des descentes-remontées de systèmes linéaires triangulaires. Rappelons de nouveau que l'inversion de cette matrice creuse est directe (en général, on utilise peu les solveurs itératifs en mécanique du solide) et sa complexité en fonction de la dimension du système est $\sigma_T = \mathcal{O}(NL^2)$ où L est la largeur de bande de la matrice (cf. Bernadou et al (1988)).

Pour le problème `VariaLQ`, en supposant disposer d'un bon algorithme de descente du gradient, on peut espérer, si le problème est bien posé, converger en une dizaine d'itérations de calcul de gradient de J. Or pour calculer ce gradient, on résout un problème direct et un problème adjoint. Donc

$$\sigma_T(\texttt{VariaLQ}) \simeq 20\sigma_T(\texttt{linearDyn}),$$

si on peut stocker en mémoire, à tous les pas de temps du problème direct, l'état du système nécessaire dans l'adjoint pour calculer $Z_n - HX_n(\xi)$ (on ne stocke en pratique que les déplacements car les vitesses sont directement déductibles de la loi des trapèzes). Dans ce cas, la complexité mémoire est

$$\sigma_M(\texttt{VariaLQ}) = \mathcal{O}(N.N_{\Delta t}) + \sigma_M(\texttt{linearDyn}).$$

Sur un battement cardiaque de $1s$, le pas de temps est en partie contraint par les phases rapides isovolumiques liées à la dépolarisation du tissu donc $\Delta t = 10^{-3} \Rightarrow N_{\Delta t} = 10^3$. En considérant les maillages présentés où le nombre de degrés de liberté N est compris entre 10^3 et 10^5 on obtient en occupation mémoire

$$8\,Mo \lesssim \sigma_M(\texttt{VariaLQ}) \lesssim 800\,Mo,$$

ce qui est envisageable sur les machines actuelles. De plus, en pratique, rien n'empêche de sauvegarder cet état en mémoire disque avec pour conséquence l'allongement du temps de calcul par les accès disques.

Pour l'estimation par `Kalman`, le temps de calcul correspond à la simulation directe plus éventuellement une simulation en temps rétrograde pour retrouver la condition initiale. Cependant, à chaque itération il faut calculer le filtre à partir d'inversions des matrices de covariances. Or celle-ci sont pleines (cf. paragraphe suivant pour les aspects mémoires) et donc

l'inversion de tels matrices et en $\mathcal{O}(N^3)$. Donc

$$\sigma_T(\texttt{Kalman}) \simeq \sigma_T(\texttt{linearDyn}) + \mathcal{O}(N^3.N_{\Delta t}).$$

De plus le stockage constitue aussi un enjeu puisque les matrices symétriques **pleines** (en toute généralité) de covariances impose

$$\sigma_M(\texttt{Kalman}) = \mathcal{O}(2N^2) + \sigma_M(\texttt{linearDyn}),$$

soit en pratique

$$15\,Mo \lesssim \sigma_M(\texttt{Kalman}) \lesssim 150\,Go.$$

et ce stockage est nécessairement en espace mémoire central. Ainsi la complexité du filtrage de Kalman est, encore aujourd'hui, critique pour les machines actuelles.

3.2.3 Problématique

Ce que nous venons de voir sous forme de « travaux pratiques » nous permet de clarifier la problématique évoquée en introduction de ce chapitre. En effet, l'approche contrôle optimal semble répondre à notre problème d'estimation et nous fournir potentiellement un schéma numérique en temps. Cependant, l'approche séquentielle à la Kalman, au vu de sa complexité, paraît largement hors de portée sauf à tenter de réduire la dimension de la covariance. Ces remarques sont tout à fait générales dans les autres domaines d'assimilation de données. Notamment en météorologie, la plupart des acteurs opérationnels français et européens se sont orientés vers la stratégie variationnelle pour sa complexité au regard du filtrage de Kalman. Seuls quelques acteurs opérationnels aux États-Unis ont choisi l'approche Kalmanienne mais en ayant recours à ce qu'on appelle le filtrage réduit, c'est-à-dire une situation où on force la covariance à ne représenter qu'un sous-espace de rang faible de l'état[2]. Cependant, l'approche séquentielle a ceci de séduisant qu'elle permet

- de bénéficier de nouvelles mesures sans reprendre toute l'estimation,
- d'éviter la mise en place d'un algorithme d'optimisation efficace et ses itérations potentielles, à supposer qu'on soit capable de calculer rapidement le filtre.

Notre problématique est la suivante : **peut-on imaginer des algorithmes séquentiels de filtrage, pas nécessairement fondés sur la notion de contrôle optimal mais permettant l'estimation efficace du système réel ?** Autrement dit nous souhaitons trouver un opérateur K, différent de celui fourni par le contrôle optimal, tel que le système

$$\begin{cases} \dot{\hat{X}} = \hat{A}X + R + K_X(Z - H\hat{X}) \\ \hat{X}(0) = X_0 \end{cases} \tag{3.13}$$

ayant pour condition initiale la seule donnée disponible X_0, estime efficacement le système réel. Or cette estimation efficace signifie simplement comme indiqué en introduction qu'on

[2]Nous verrons nous aussi, au chapitre 4, ces méthodes de filtrage réduit même si nous ne l'appliquerons pas du tout pour réduire la covariance d'état.

peut trouver une norme pertinente telle que l'erreur entre le système réel et l'estimateur numérique converge vers 0, quand T et les mesures augmentent et quand h et Δt tendent vers 0. Cet angle d'approche, concentré sur la maîtrise de l'erreur d'estimation, va nous permettre de définir des nouveaux filtres, notamment grâce à l'analyse des propriétés physiques (et dynamiques) des systèmes observés.

3.3 Filtrage asymptotique de Luenberger

3.3.1 Système d'erreur

Dans ce qui suit, nous nous focaliserons principalement sur un système mécanique classique au sens où les seuls champs que comporte l'état sont le déplacement et la vitesse. Nous n'insisterons pas sur le traitement d'autres variables d'état, par exemple des variables de pression que nous avons vu apparaître en modélisation cardiaque. On rappelle donc les notations, $x = \left(\begin{smallmatrix} y \\ \dot{y} \end{smallmatrix}\right)$ est l'état du système réel et $x_h = \mathcal{I}_h X$, avec $X = \left(\begin{smallmatrix} Y \\ \dot{Y} \end{smallmatrix}\right) \in \mathbb{R}^N$, est son approximation éléments finis. De plus, $\hat{X}$ est l'estimateur par filtrage de X, définissant $\hat{x}_h$. On note l'erreur d'estimation

$$\tilde{X} = X - \hat{X}.$$

On a pour toute norme adaptée à x

$$\|x - \hat{x}_h\| \leq \|x - x_h\| + \|\tilde{X}\|. \tag{3.14}$$

Ainsi il suffit d'analyser la norme de l'erreur $\tilde{X}$, uniformément par rapport à la discrétisation. Précisons pour ce faire la dynamique de l'erreur en fonction de celle de x.

a. Cadre linéaire : équation autonome

Dans le cas de l'élastodynamique linéaire, nous avons déjà vu la dynamique de X en (3.5) et celle de l'estimateur en tout généralité en (3.13). Supposant de plus l'opérateur d'observation linéaire de la forme

$$Z = HX + \chi + \epsilon_h,$$

la dynamique de l'erreur est alors

$$\begin{cases} \dot{\tilde{X}} = (A - K_X H)\tilde{X} + K_X(\chi + \epsilon_h) \\ \tilde{X}(0) = \zeta^X \end{cases} \tag{3.15}$$

Ce système est autonome et 0 est une position d'équilibre pour le système homogène associé

$$\dot{\tilde{X}} = (A - K_X H)\tilde{X} = \tilde{A}\tilde{X}. \tag{3.16}$$

Déterminer si l'estimateur d'état converge vers la cible revient donc ici à analyser la stabilité de ce système homogène autour de sa position d'équilibre dont la définition est rappelée ci-dessous.

DÉFINITION 3.3.1 (STABILITÉ)
Soit un système dynamique $\dot{X} = A(X)$ de dimension quelconque, un point d'équilibre $X = 0$ est dit

- ***stable** si*
$$\forall \epsilon > 0,\ \exists \delta > 0,\ \|X(0)\| \leq \delta \Rightarrow \|X(t)\| < \epsilon,\ \forall t,$$
- ***instable** sinon,*
- ***asymptotiquement stable** s'il est (uniformément) stable et que δ peut être choisi tel que*
$$\|X(0)\| \leq \delta \Rightarrow \lim_{t\to\infty} \|X(t)\| = 0,$$
- ***exponentiellement stable** s'il est asymptotiquement stable et qu'il existe C et γ tels que*
$$\exists t_0 \,|\, t > t_0, \Rightarrow \|X(t)\| \leq Ce^{-\gamma t}.$$

Ces notions de stabilité apparaissent déjà dans le cadre du problème direct. Cependant, elles nous intéressent particulièrement ici pour le système de l'erreur. Afin de lever toute ambiguïté sur le système pour lequel nous imaginons utiliser les résultats qui suivent, nous utiliserons donc la notation $\tilde{X}$, $\tilde{A}$ etc. Pour cette section, la stabilité du système linéaire autonome (3.16) est assurée par le théorème suivant (voir Hahn (1967)).

THÉORÈME 3.3.2 (STABILITÉ DES SYSTÈMES LINÉAIRES)
Soit un système dynamique autonome linéaire sur $\mathbb{R}^N$ régi par
$$\dot{\tilde{X}} = \tilde{A}\tilde{X}.$$
Notons (λ_i) l'ensemble des valeurs propres de $\tilde{A}$. Alors l'équilibre $X = 0$ est stable si et seulement si toutes les valeurs propres de $\tilde{A}$ satisfont $\Re\lambda_i \leq 0$ et si toutes les valeurs propres nulles sont associées à un bloc de Jordan d'ordre[3] 1. Il est asymptotiquement stable si toutes les valeurs propres vérifient $\Re\lambda_i < 0$.

■ ***Démonstration :*** *La preuve réside essentiellement dans l'écriture de la résolvante. On a vu au premier chapitre que*
$$\mathcal{R}(t, 0) = \exp(\tilde{A}t).$$
Or toute matrice $\tilde{A} \in \mathbb{M}_N(\mathbb{R})$ est jordanisable, c'est-à-dire qu'il existe une matrice $U \in \mathbb{G}L_n(\mathbb{C})$ telle que
$$U^{-1}\tilde{A}U = J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_r \end{pmatrix},$$
où J_i est soit diagonal, soit du type
$$J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in \mathbb{M}_{m_i}(\mathbb{C}),$$

[3] l'ordre d'un bloc de Jordan est la multiplicité de la valeur propre dans le polynome minimal de $\tilde{A}$

et m_i est l'ordre de λ_i. Dans ce cas

$$\exp(\tilde{A}t)V = U\exp(Jt)U^{-1}V = \sum_{i=1}^{r}\sum_{k=1}^{m_i}(V_{ik}^T \cdot V)\frac{t^{(j-k)}}{(j-k)!}\exp(\lambda_i t)V_{ik}.$$

où les V_{ik} forment une base orthonormée des sous-espaces caractéristiques de $\tilde{A}$. Le théorème se déduit alors directement de cette décomposition d'algèbre linéaire en dimension finie. ■

Un élément intéressant dans la démonstration de ce théorème est que qu'elle fournit dans le même temps une quantification de la « performance » de la stabilisation du système. En effet on obtient directement

$$\|\exp(\tilde{A}t)\| \leq Ce^{\delta t},$$

avec

$$\forall\delta \mid \sup_{n\geq 1}(\Re(\lambda_n)) < \delta < 0.$$

Les définitions et théorèmes précédents permettent d'envisager tout une nouvelle classe de filtre. Il « suffit » en effet d'imaginer K_X tel que les valeurs propres de l'opérateur $A - K_X H$ intervenant dans la dynamique de l'erreur (3.16) est asymptotiquement voir exponentiellement stable. Ce type d'observateur est identifié depuis les années 60 dans Luenberger (1963) et porte le nom *d'observateur asymptotique de Luenberger*. Voyons maintenant comment ces conditions formulées pour une dynamique linéaire s'appliquent dans un cadre général.

b. Cadre non-linéaire : équation non autonome

En non-linéaire, le système de l'erreur n'est plus autonome et nous devons le considérer couplé à la dynamique soit de l'estimateur, soit du système cible. Nous choisissons la dynamique réelle sur laquelle on peut a priori supposer plus de régularité. On étudie donc le système

$$\begin{cases} \dot{X} = A(X,t) \\ \dot{\tilde{X}} = (A(X,t) - A(X-\tilde{X},t)) - K_X(H(X) - H(X-\tilde{X})) + K_X(\chi + \epsilon_h) \\ X(0) = X_0 + \zeta^X \\ \tilde{X}(0) = \zeta^X \end{cases} \tag{3.17}$$

Dans ce cas, l'analyse du linéarisé fournit une condition suffisante (mais pas nécessaire) de stabilité (voir théorème 3.3.3). Cependant il y a deux choix de linéarisé possibles. Le premier correspond à une dynamique couplée sur l'ensemble des deux systèmes, c'est-à-dire en considérant la variable $\binom{X}{\tilde{X}}$. Dans ce cas, le point d'équilibre est $X = 0$ et $\tilde{X} = 0$ donc on analyse la stabilité autour de la trajectoire cible (statique) à l'équilibre. L'autre approche consiste à ne considérer que l'équation sur $\tilde{X}$ étant donnée la trajectoire de X. Ainsi, on étudie la stabilité autour uniquement de $\tilde{X} = 0$ pour un système non autonome dont la dépendance en temps est liée à la trajectoire du système cible. Le tangent de ce système est alors

$$\begin{cases} \dot{\widetilde{\delta X}} = \left(\frac{\partial A}{\partial X}(X(t),t) - K_X\frac{\partial H}{\partial X}(X(t),t)\right)\widetilde{\delta X} + K_X(\chi + \epsilon_h) \\ \widetilde{\delta X}(0) = \zeta^X \end{cases} \tag{3.18}$$

Les résultats de stabilité du linéarisé existent aussi pour des dynamiques non autonomes et s'écrivent de manière très générale. Nous en donnerons une justification au paragraphe suivant et nous contentons de l'énoncé (voir aussi Khalil (1992) pour des énoncés similaires) ici

THÉORÈME 3.3.3 (STABILITÉ DU LINÉARISÉ)
Soit un système dynamique sur $\mathbb{R}^N$ régi par

$$\dot{\tilde{X}} = \tilde{A}(\tilde{X}, t).$$

Si le système tangent défini par

$$\dot{\widetilde{\delta X}} = \frac{\partial \tilde{A}}{\partial X}(0, t) \cdot \widetilde{\delta X}$$

est (resp. asymptotiquement) stable, alors il existe un voisinage de 0 tel que le système non-linéaire est (resp. asymptotiquement) stable.

Donc la stabilité de l'erreur linéarisée permet d'assurer qu'il existe un voisinage pour ζ^X dans lequel la trajectoire de l'estimateur converge en temps vers celle du système cible. En revanche, l'instabilité du tangent ne prouve rien car certaines non-linéarités peuvent avoir un effet stabilisant sur le système. A titre d'exemple, on pourra consulter les nombreuses configurations présentées en dimension finie et infinie par Coron (2007). Sur le principe, nous devons retenir que l'analyse du linéarisé est un outil puissant et direct mais que les non-linéarités sont souvent un atout non-négligeable pour la stabilité du système.

3.3.2 Fonctions de Liapounov

En dehors de l'étude du spectre de l'opérateur qui, nous le verrons, a ses limites en dimension infinie, il existe d'autres approches pour prouver la stabilité du système de l'erreur (qui, combiné à une uniformité par rapport à la discrétisation, reste notre objectif). En particulier le théorème d'invariance de LaSalle (voir Hahn (1967)) et les fonctions de Liapounov sont des outils puissants.

a. Cadre continu en temps

Commençons par quelques définitions générales pour tout système dynamique, autonome ou non, de la forme $\dot{\tilde{X}} = \tilde{A}(\tilde{X}, t)$ défini sur R^N **dont** 0 **est un point d'équilibre**. On rappelle que, pour nos études futures, ce système est typiquement celui de l'erreur d'estimation.

DÉFINITION 3.3.4 (FONCTION DE LIAPOUNOV)
$V : B(0, R) \times \mathbb{R}^+$ *est* une fonction de Liapounov *si elle vérifie les propriétés suivantes*

- *V est de classe C^1 et $V(0, t) = 0$,*

- *il existe $W : B(0, R) \to \mathbb{R}^+$ continue telle que*

$$\forall t,\ \forall X, \quad W(X) \leq V(X, t),$$

- $W(X) = 0$ *si et seulement si* $X = 0$,

- *sur* $B(0, R) \times \mathbb{R}^+$
$$\frac{\mathrm{d}V}{\mathrm{dt}} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial X}\tilde{A}(\tilde{X}, t) \leq 0.$$

On démontre aisément que s'il existe une fonction de Liapounov alors $X = 0$ est un point d'équilibre stable. Pour avoir la stabilité asymptotique, il faut supposer des propriétés supplémentaires.

PROPOSITION 3.3.5
Soit V une fonction de Liapounov vérifiant les hypothèses suivantes :

- *il existe* $W : B(0, R) \to \mathbb{R}^-$ *continue telle que*
$$\forall t,\ \forall X, \quad \frac{\mathrm{d}V}{\mathrm{dt}}(X, t) \leq \tilde{W}(X),$$

- $\tilde{W}(X) = 0$ *si et seulement si* $X = 0$,

- $\lim\limits_{X \to 0} \sup\limits_{t \geq 0} V(X, t) = 0$,

alors $X = 0$ est asymptotiquement stable.

Si enfin le choix de W et $\tilde{W}$ permet d'assurer que W est une norme et que
$$\frac{\mathrm{d}V}{\mathrm{dt}}(X, t) \leq -\alpha V(X, t),$$
alors, d'après le lemme de Gronwall, le système est exponentiellement stable.

Comme premier exemple, reprenons le système linéaire autonome dont les pôles sont tous de partie réelle négative. On peut aussi démontrer sa stabilité par une fonction de Liapounov. En effet, soit $Q = \int_0^\infty e^{t\tilde{A}^T} N e^{t\tilde{A}}\, dt$ définie à partir de N une norme adaptée à l'état. Les pôles étant tous de partie réelle négatives, cette expression a un sens et
$$\tilde{A}^T Q + Q\tilde{A} = \int_0^\infty \frac{\mathrm{d}}{\mathrm{dt}}\left(e^{t\tilde{A}^T} N e^{t\tilde{A}}\right) dt = -N$$
Posons alors $V(\tilde{X}, t) = \tilde{X}^T Q \tilde{X}$. C'est une fonction de Liapounov car
$$\frac{\mathrm{d}V}{\mathrm{dt}} = \tilde{X}^T(\tilde{A}Q + Q\tilde{A}^T)\tilde{X} = -\tilde{X}^T N \tilde{X} = -\|\tilde{X}\|_N^2$$
ce qui donne une nouvelle démonstration de l'exponentielle stabilité de ce système.

Cette fonction assure aussi la stabilité d'un système non-linéaire dont le linéarisé vérifie les mêmes propriétés. Il permet alors de démontrer le théorème 3.3.3 énoncé plus haut. En effet dans ce cas, le système se met sous la forme $\dot{\tilde{X}} = \tilde{A}\tilde{X} + q(\tilde{X}, t)$ avec $q(\tilde{X}, t) = o(\|\tilde{X}\|_N)$. Donc il existe un voisinage de l'origine tel que $\frac{\mathrm{d}V}{\mathrm{dt}} \leq -\frac{1}{2}\|\tilde{X}\|_N^2$.

Dans notre étude, deux exemples de fonctionnelle de Liapounov apparaissent naturellement et nous allons voir que c'est autour de ces deux fonctionnelles que se dessineront nos futurs estimateurs.

□ ÉNERGIE MÉCANIQUE – En effet, nous avons vu au chapitre 2 que l'énergie mécanique du système total vérifiait une propriété de conservativité y compris en mécanique non-linéaire, que ce soit pour le système de référence continu, ou ses discrétisations au moins en espace, et même en temps avec un choix judicieux de schéma. Sans anticiper sur la suite, on voit déjà que sans filtre cette propriété se transmet directement au système de l'erreur en élastodynamique **linéaire**. Par exemple pour le système discret, on a pour

$$N = \begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix}, \tag{3.19}$$

tel que l'énergie globale (somme des énergies cinétiques et de déformation) $\mathcal{E} = \mathcal{E}_c + \mathcal{E}_m$ du système soit donnée par

$$\|\tilde{X}\|^2_{\mathcal{E}} = \frac{1}{2}\int_{\Omega_0} \rho \left|\dot{\underline{\tilde{y}}}\right|^2 d\Omega + \frac{1}{2}\int_{\Omega_0} \underline{\underline{\varepsilon}}(\underline{\tilde{y}}) : \underline{\underline{\underline{\underline{A}}}} : \underline{\underline{\varepsilon}}(\underline{\tilde{y}})\, d\Omega = \frac{1}{2}\tilde{X}^T N \tilde{X},$$

une fonction de Liapounov vérifiant

$$\frac{\mathrm{d}}{\mathrm{dt}}\|\tilde{X}\|^2_{\mathcal{E}} = -\dot{\tilde{Y}}^T C \dot{\tilde{Y}}.$$

En revanche, pour la mécanique non-linéaire, la propriété de dissipativité du système de référence ne s'étend pas en tant que telle à un quelconque système d'erreur. Seuls les linéarisés autour de 0 retrouveront cette propriété naturellement.

□ RICCATI DES ESTIMATEURS LQ – En fait, nous ne l'avions pas évoqué au chapitre 1 mais les estimateurs moindres carrés linéaires vérifient pour le système de l'erreur une propriété de stabilité via une fonction de Liapounov. En effet, reconsidérons le filtre de Kalman continu et définissons

$$V(t) = \tilde{X}^T P^{-1} \tilde{X} = \|\tilde{X}\|^2_{P^{-1}}. \tag{3.20}$$

On vérifie alors que

$$\dot{V} = 2\tilde{X}^T P^{-1} \dot{\tilde{X}} + \tilde{X}^T \frac{\mathrm{d}}{\mathrm{dt}}(P^{-1}))\tilde{X}.$$

Or l'inverse de la covariance vérifie dans le cas de l'estimation de condition initiale LQ sans erreur de modèle (cf. équation (1.61))

$$\frac{\mathrm{d}}{\mathrm{dt}}(P^{-1})) = -P^{-1}\dot{P}P^{-1} = -P^{-1}A - A^T P^{-1} + H^T W^{-1} H. \tag{3.21}$$

Donc

$$\dot{V} = 2\tilde{X}^T P^{-1} A \tilde{X} - \tilde{X}(A^T P^{-1} + P^{-1} A + H^T W^{-1} H)\tilde{X} = -\tilde{X}^T H^T W^{-1} H \tilde{X} \leq 0.$$

En fonction des propriétés d'observabilité, qui reprennent ici tout leur sens, on peut alors assurer la convergence de l'estimateur. Par exemple, la condition d'observabilité (forte) suivante permet de conclure

$$\exists \alpha, T, \quad \forall t, \quad \int_t^T \tilde{X}^T H^T W^{-1} H \tilde{X} \geq \alpha \mathbb{1}. \tag{3.22}$$

A noter enfin que, dans ce cadre linéaire, la fonction de Liapounov est exactement la fonction définie dans le cadre de la programmation dynamique Section 1.4.1 (vérifiant HJB) appliquée à l'erreur.

b. Cadre discret et schéma en temps

Nous avons déjà parlé de l'intérêt vis-à-vis de l'énergie mécanique de vérifier aussi une propriété de stabilité au niveau discret car elle fournit ensuite une condition suffisante, qui, couplée à la consistance, assure la convergence du schéma suivi. Nous avons vu aussi que, dans le cas des estimateurs LQ, l'estimateur LQ discret offrait une piste pertinente de discrétisation car il vérifiait une propriété de contrôle optimal au niveau discret. En fait, on peut aller beaucoup plus loin dans la justification de ces schémas en prouvant qu'ils vérifient *in fine* une condition de stabilité au sens de Liapounov au niveau discret. Une fois encore c'est le lemme d'inversion matricielle qui nous permet de retrouver cette propriété technique dont nous donnons ici une démonstration simplifiée.

Pour cela, on reprend les équations « one-step » de l'estimateur discret (pour H indépendant du temps),

$$\hat{X}_{n+1}^- = A_{n+1|n}X_n^- + A_{n+1|n}K_n(Z_n - H\hat{X}_n^-),$$

dont le système de l'erreur vérifie la dynamique d'équation homogène

$$\tilde{X}_{n+1}^- = A_{n+1|n}(\mathbb{1} - K_nH)\tilde{X}_n^-.$$

Si on pose alors

$$V_n = (\tilde{X}_n^-)^T(P_n^-)^{-1}\tilde{X}_n^-, \tag{3.23}$$

on vérifie

$$\begin{aligned} V_{n+1} &= (\tilde{X}_{n+1}^-)^T(P_{n+1}^-)^{-1}\tilde{X}_{n+1}^- \\ &= (\tilde{X}_n^-)^T(\mathbb{1} - K_nH)^TA_{n+1|n}^T(P_{n+1}^-)^{-1}A_{n+1|n}(\mathbb{1} - K_nH)\tilde{X}_n^-. \end{aligned}$$

Or

$$P_{n+1}^- = A_{n+1|n}P_n^+A_{n+1|n}^T \Rightarrow P_n^+ = A_{n+1|n}^{-T}(P_{n+1}^-)^{-1}A_{n+1|n}^{-1},$$

donc

$$V_{n+1} = (\tilde{X}_n^-)^T(\mathbb{1} - K_nH)^T(P_n^+)^{-1}(\mathbb{1} - K_nH)\tilde{X}_n^-.$$

En supposant $(\mathbb{1}-K_nH)$ inversible et en utilisant la dernière identité de (1.52) sur la covariance d'erreur a posteriori :

$$V_{n+1} = (\tilde{X}_n^-)^T(P_n^- + (\mathbb{1} - K_nH)^{-1}K_nW_nK_n(\mathbb{1} - K_nH)^{-T})^{-1}\tilde{X}_n^-,$$

le lemme d'inversion matricielle fournit alors

$$\begin{aligned} V_{n+1} &= (\tilde{X}_n^-)^T(P_n^-)^{-1}\tilde{X}_n^- - (\tilde{X}_n^-)^TQ_n\tilde{X}_n^- \\ &= V_n - (\tilde{X}_n^-)^TQ_n\tilde{X}_n^-, \end{aligned}$$

où Q_n est symétrique positive ce qui conduit aux mêmes résultats que pour la stabilité en temps continu.

3.3.3 Stabilisation et contrôle

La notion de stabilisation de système est une notion fondamentale du contrôle en boucle fermée qui apparaît dans le cadre de l'observation sur le système de l'erreur. Elle apparaît donc très naturellement dans toute la bibliographie du contrôle et nous incite à utiliser des résultats de contrôle des systèmes mécaniques ou des EDP en général afin de formuler des nouveaux filtres. D'ailleurs, la stabilisation n'est finalement rien d'autre qu'une forme de contrôlabilité en temps long du système. Comme illustration théorique de ce que nous avions évoqué dès l'introduction de ce document, le théorème suivant permet de mieux cerner l'équivalence entre observation de système (donc identification), contrôle et stabilisation. Pour mémoire, les opérateurs antisymétriques que ce théorème utilise sont évidemment ceux qui nous intéressent. Nous avons en effet beaucoup écrit

$$A = \begin{pmatrix} 0 & 1 \\ -M^{-1}K & -M^{-1}C \end{pmatrix},$$

mais nous avons vu aussi apparaître (par exemple avec l'algorithme variationnel) une forme antisymétrique en multipliant la deuxième équation par M et la première par K,

$$A_{\text{skew}} = \begin{pmatrix} 0 & K \\ -K & -C \end{pmatrix}.$$

En l'absence de viscosité ($C = 0$) ce système considéré par rapport à sa norme énergétique $N = \left(\begin{smallmatrix} K & 0 \\ 0 & M \end{smallmatrix}\right)$ est bien antisymétrique. Plus généralement les systèmes hyperboliques d'ordre 2 (équation des ondes, mécaniques des milieux continus) vérifient ce type de formulation.

THÉORÈME 3.3.6
Soit A un opérateur anti-symétrique et (A, B) une paire contrôlable alors $-B'$ stabilise le système, i.e la solution de

$$\begin{cases} \dot{X} = AX - BB'X \\ X(0) \end{cases}$$

est exponentiellement stable.

■ ***Principe de la démonstration :*** *On donne ici les quelques grandes idées de la preuve donnée par Micu et Zuazua (2004). Ce type de dynamiques permet d'écrire un loi d'évolution sur la norme de la solution*

$$\frac{1}{2}\frac{\mathrm{d}\|X\|^2}{\mathrm{dt}} = -\|B'X\|^2 \leq 0,$$

donc

$$\|X(T)\|^2 - \|X(0)\|^2 = -2\int_0^T X^T BB'X\, dt.$$

L'exponentielle stabilité est alors facile à démontrer si on peut assurer qu'il existe c et T tel que

$$\|X(0)\|^2 \leq c\int_0^T X^T BB'X\, dt, \tag{3.24}$$

conduisant à

$$\|X(T)\|^2 \leq (1 - \frac{2}{c})\|X(0)\|^2,$$

puis l'exponentielle stabilité[4]*. On voit donc apparaître une condition de contrôlabilité sur le système. Pour la prouver sur le système qui nous intéresse, on décompose ce dernier en* $X = \varphi + \eta$ *avec*

$$\begin{cases} \dot{\varphi} = A\varphi \\ \varphi(0) = X(0) \end{cases}$$

et

$$\begin{cases} \dot{\eta} = A\eta - BB'X \\ \eta(0) = 0 \end{cases}$$

Le premier système est contrôlable donc vérifie une condition de contrôlabilité du type

$$\|X(0)\|^2 \leq C \int_0^T \varphi^T BB'\varphi \, dt,$$

permettant d'obtenir

$$\|X(0)\|^2 \leq 2C \int_0^T \|B'X\|^2 \, dt + 2C \int_0^T \|B'\eta\|^2 \, dt.$$

Pour le deuxième système l'utilisation du lemme de Gronwall permet de majorer le deuxième terme en fonction de X *sous la forme*

$$\int_0^T \|B'\eta\|^2 \, dt \leq T\|B'\|^4 e^T \int_0^T \|B'X\|^2 \, dt,$$

assurant la condition (3.24). ■

On comprend désormais le lien profond entre stabilisation et contrôle et de même le lien entre le contrôle de ces systèmes et la stabilisation. Mais au travers du lien entre observation et stabilisation on imagine maintenant une recherche potentielle de stabilisateur pour un opérateur d'observation H en choisissant « simplement » H'.

En dimension infinie, le théorème 3.3.6 s'étend et les principaux résultats sur les systèmes mécaniques sont synthétisés dans Lasiecka et Triggiani (2003). La condition de contrôlabilité est certes plus abstraite, mais peut être testée numériquement (voir Bourquin (1995)). Pour assurer la stabilisation exponentielle, il faut lui adjoindre une autre condition comme exprimé dans le théorème énoncé dans Lasiecka et Triggiani (2003) (voir aussi Bourquin (1998)).

THÉORÈME 3.3.7
Soit A *un opérateur anti-adjoint et* $(\mathcal{A}, \mathcal{B})$ *vérifiant la propriété de contrôlabilité exacte*

$$\forall x, \quad \int_0^T \|\mathcal{B}' e^{\mathcal{A}t} x\|_{\mathcal{U}}^2 \, dt \geq C_T \|x\|_{\mathcal{E}}^2,$$

où $\|\cdot\|_{\mathcal{U}}$ *est la norme associée à l'espace de contrôle et* $e^{\mathcal{A}t}$ *est le semi-groupe engendré par* $\mathcal{A}$*. Si on suppose de plus que l'opérateur* $\mathcal{L}$ *défini par*

$$\mathcal{L}u(t) = \int_0^t e^{\mathcal{A}(t-\tau)} \mathcal{B}u \, ds,$$

[4]voir aussi dans ce manuscrit Section 4.2.4.c. pour des manipulations conduisant à l'exponentielle stabilité à partir de ce type de majorations

est tel que $\mathcal{B}'\mathcal{L}$ est continue de $L^2([0,T];\mathcal{U})$ dans lui-mêmes alors le système

$$\begin{cases} \dot{x} = \mathcal{A}x - \mathcal{B}\mathcal{B}'x \\ x(0) \end{cases}$$

est exponentiellement stable.

Par rapport à la dimension finie, la condition fondamentale supplémentaire à la démonstration de ce type de résultat est une condition de régularité sur l'opérateur $\mathcal{B}$. Lorsque ce dernier est borné, le théorème ci-dessus s'étend sans difficulté. Cependant, dans d'autres cas, des opérateurs définis uniquement sur la frontière nécessitent un résultat supplémentaire de régularité, en général difficile. En effet, le contrôle de l'énergie de la solution ne permet pas **a priori** de contrôler la norme de termes de vitesse sur le bord du domaine. Cependant, le simple fait de pouvoir considérer le système stabilisé peut apporter un terme de régularité cachée permettant l'extension des théorèmes précédents (voir les résultats et contre-exemples présentés par Lasiecka et Triggiani (2003)). On retiendra donc que, pour des opérateurs de contrôle (d'observation) frontière, le cadre théorique est plus délicat mais que les résultats perdurent si on peut démontrer qu'on dispose de la régularité suffisante pour contrôler solution et termes frontières. Ces notions apparaîtront en pratique dans notre exposé en considérant successivement des opérateurs d'observation volumiques « standards » puis des opérateurs frontières imaginés à partir de la description (cf. chapitre 2) des mesures types attendues en cardiologie.

3.4 Estimation d'état en élastodynamique à partir de vitesses

3.4.1 Rappels sur la position du problème

Bien que l'objectif soit de réaliser une estimation pour le système cardiaque que nous avons vu être un système hautement non-linéaire, nous avons désormais conscience qu'un pré-requis indispensable est la définition d'un estimateur en élastodynamique linéaire. De plus, les premiers éléments de compréhension autour de la stabilisation de systèmes indiquent que la viscosité a un rôle fondamental dans l'estimation. Ainsi il est à la fois crucial et plus difficile, de pouvoir stabiliser les systèmes faiblement visqueux. Pour autant, trop de viscosité tue la stabilisation en ramenant des pôles sur l'axe imaginaire. Il faut donc réaliser une étude en fonction du taux de viscosité. *A fortiori*, l'estimation que nous proposons s'inscrit, nous l'espérons, dans un réflexion plus globale que simplement l'assimilation de données cardiaques pour toucher d'autres cadres applicatifs en élastodynamique voire pour tout système hyperbolique du second ordre (propagation d'onde en général par exemple). L'ensemble de ces considérations justifient le choix du système « témoin » proposé ci-dessous. Nous proposerons alors un premier estimateur pour des mesures de vitesses, malgré la difficulté liée à leur obtention en pratique. Le flot optique en échographie est un objectif raisonnable même si nous commencerons par regarder des vitesses observées directement par rapport à la configuration Lagrangienne ce qui, en pratique, est peu évident sans poser une hypothèse (irréaliste en cardiaque) de petits déplacements. En revanche, la différence entre deux tags successifs dans le cas (non applicable aujourd'hui) d'une résolution temporelle assez forte s'inscrit dans cette motivation. Peu importe, ce sont les linéarisations futures et le fait que les vitesses vont fournir un estimateur

beaucoup plus direct et naturel que les déplacements qui incitent à les considérer en premier. L'ensemble des résultats présentés ici sont compilés des premières parties des articles Moireau et al (2008b) et Chapelle et al (2009).

□ MODÈLE DE CONTRACTION – On considère donc un système d'élasticité linéaire écrit sous forme variationnelle

$$\int_{\Omega_0} \rho \frac{d\underline{y}}{dt} \cdot \underline{v}\, d\Omega = \int_{\Omega_0} \rho \dot{\underline{y}} \cdot \underline{v}\, d\Omega, \quad \forall \underline{v} \tag{3.25}$$

$$\int_{\Omega_0} \rho \frac{d\dot{\underline{y}}}{dt} \cdot \underline{v}\, d\Omega = -\int_{\Omega_0} \underline{\underline{\Sigma}}(\underline{y}, \dot{\underline{y}}) : \mathrm{d}_{\underline{y}} \underline{\underline{e}} \cdot \underline{v}\, d\Omega_0 + \mathcal{P}^{\mathrm{ext}}(\underline{v}), \quad \forall \underline{v} \tag{3.26}$$

où on identifie dans le cadre de petits déplacements la variation infinitésimale du tenseur de Green-Lagrange $\mathrm{d}_{\underline{y}} \underline{\underline{e}} = \underline{\underline{\nabla y}} + (\underline{\underline{\nabla y}})^T = \mathrm{d}_{\underline{y}} \underline{\underline{\varepsilon}}$ et le second tenseur de Piola-Kirchhoff $\underline{\underline{\Sigma}}$ au tenseur des contraintes de Cauchy. De plus, on suppose que la contrainte suit uniquement une loi passive en petits déplacements alors que le phénomène de contraction cardiaque est supposé être la conséquence d'une précontrainte active (fonction uniquement de l'activation électrique dans le tissu). De plus on choisit de négliger les autres forces extérieures, qui n'apportent pas de difficultés supplémentaires côté estimation mais alourdissent la structure, on choisit donc

$$\mathcal{P}^{\mathrm{ext}} = \int_{\Omega_0} = \sigma_0 w(\underline{\mathrm{x}}, t)\, \mathrm{tr}(\, \mathrm{d}_{\underline{y}} \underline{\underline{\varepsilon}} \cdot \underline{v})\, d\Omega,$$

et w est une fonction représentant la propagation de la contrainte dans le tissu sous l'effet de l'onde électrique. Durant la thèse, nous avons essayé différents exemples de propagation :

- une onde plane suivant $O\mathrm{z}$
$$w(\underline{\mathrm{x}}, t) = w(\mathrm{z} - ct),$$
pour un coeur dont le grand axe est recalé sur cet axe ;
- une onde sphérique initiée au point C de la Figure 3.3
$$w(\underline{\mathrm{x}}, t) = w(|\mathrm{x} - \mathrm{x}_C| - ct);$$

Physiologiquement, ce choix revient à négliger la dépendance de la contrainte active par rapport à la déformation et on peut donc intégrer le modèle Bestel-Clément-Sorine directement. La fonction w est donc issue de l'intégration de

$$\sigma_c = \tau_c \text{ et } \dot{\tau}_c = \sigma_0 \left|u\right|_+ - \left|u\right| \tau_c.$$

Au final on a donc linéarisé le modèle cardiaque par rapport au déplacement (puisqu'on considère des petits déplacements) mais aussi par rapport à σ_0. Les raisons de ce choix apparaîtront clairement au chapitre suivant avec la question de l'estimation paramétrique mais pour l'estimation d'état l'intérêt est avant tout de simplifier la dynamique dans la présentation[5]. La raideur est donc uniquement passive et nous la choisissons de type Saint-Venant Kirchhoff avec un amortissement proportionnel, soit

$$\underline{\underline{\Sigma}} = \lambda_i\, \mathrm{tr}(\underline{\underline{\varepsilon}} + \eta_i \dot{\underline{\underline{\varepsilon}}}) \underline{\underline{1}} + 2\mu_i (\underline{\underline{\varepsilon}} + \eta_i \dot{\underline{\underline{\varepsilon}}}), \tag{3.27}$$

[5]Si ce cas-test ne servait pas au chapitre 4 nous aurions tout autant pu choisir une véritable raideur active dépendant du temps pour modéliser la contraction

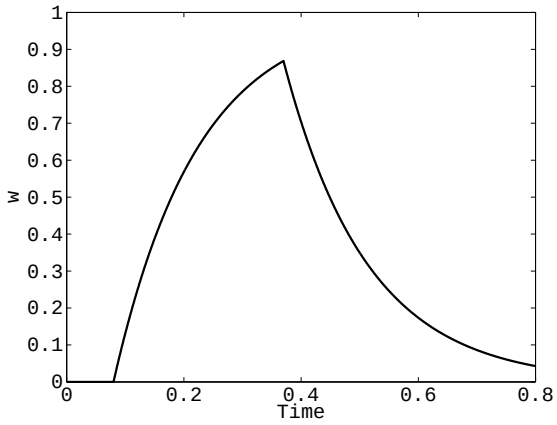


FIGURE 3.1: Fonction $w(t)$ pour la précontrainte active

avec ainsi une viscosité choisie de type Rayleigh.

Sous forme d'opérateur, la dynamique étant linéaire, on peut démontrer (cf. chap. 6 de Marsden et Hughes (1983)) que $\mathcal{A}$ définit un semi-groupe de contraction sur $H^1 \times L^2$ donnant un sens à l'équation formelle

$$\dot{x} = \mathcal{A}x + \mathcal{R}, \tag{3.28}$$

avec

$$\mathcal{A} = \begin{pmatrix} 0 & \mathbb{1} \\ \rho^{-1}\,\mathrm{div}(\underline{\underline{\underline{A}}} \cdot \underline{\underline{\nabla}}) & \rho^{-1}\,\mathrm{div}(\underline{\underline{\underline{A}}}^v \cdot \underline{\underline{\nabla}}) \end{pmatrix}.$$

□ OBSERVATIONS – L'opérateur d'observation sur les champs de déplacements et vitesses x consiste en une série de mesures de vitesse uniquement sur q cellules à support disjoints.

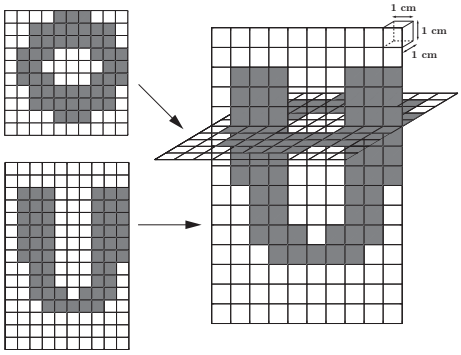


FIGURE 3.2: Rappel : Opérateur d'observation par cellules

Cet opérateur est donc l'analogue de celui présenté Section 2.5.1.c. mais pour les vitesses,

$$Z = \mathcal{H}x + \chi, \text{ avec } \mathcal{H} = \begin{pmatrix} 0 & \mathcal{H}^v \end{pmatrix} \begin{pmatrix} \underline{y} & \underline{\dot{y}} \end{pmatrix},$$

avec Z_i les q composantes de Z données par

$$Z_i = \int_{\Omega_m} s_i \underline{\dot{y}}\, d\Omega,$$

où $\|s_i\|_{L^2(\omega_m)} = 1$. On choisit, de plus, les s_i constantes pour simplifier les expériences numériques de ce cas-test. Ces mesures sont perturbées par un bruit blanc en espace de covariance $\alpha_i^3 I_3$ sur chacune des composantes. Soit donc W la covariance du bruit en espace, alors

$$W = \sum_{i=1}^{q} (\alpha_i)^2 \big[V_i^1 (V_i^1)^T + V_i^2 (V_i^2)^T + V_i^3 (V_i^3)^T\big],$$

où V_i^j représente le vecteur colonne avec 1 sur la j-ème coordonnée ($j = 1, 2, 3$) de la i-ème cellule et 0 ailleurs. Nous supposerons de plus le bruit blanc en temps $W\delta(t-s)$ afin de faciliter le formalisme dans les estimations finales en temps continu. Insistons sur le fait que malgré les simplifications formelles que le bruit blanc apporte, il ne facilite pas pour autant les estimations en particulier pour ce qui est de sa discrétisation.

□ Discrétisation et exemple numérique – La discrétisation en espace conduit directement à un opérateur A linéaire, indépendant du temps et à la dynamique (3.5). Les tests numériques seront réalisés sur la géométrie elliptique de la Section 2.4.1 rappelée Figure 3.3. Les valeurs numériques utilisées sont en pratique

$$E_i = 12.6\,10^3\,\mathrm{Pa}, \quad \nu_\mathrm{i} = 0.3, \quad \eta_\mathrm{i} = 1.6\,10^{-4}\,\mathrm{s} \quad \forall \mathrm{i} \in \{1, \ldots, 17\}. \tag{3.29}$$

On remarquera que la viscosité est nettement moins importante que dans le cas cardiaque afin de souligner les performances de l'estimateur. En effet pour cette viscosité, le système a un comportement vibratoire permettant de suivre aisément un écart à une trajectoire de référence. On étudiera néanmoins sous forme de remarques l'efficacité de l'estimateur en fonction de l'amortissement structurel initial.

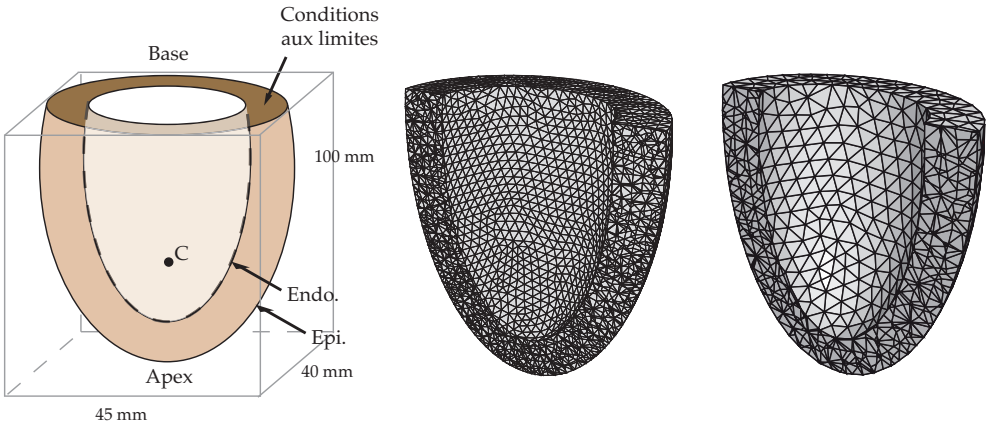


Figure 3.3: Le modèle elliptique avec des conditions aux limites simplifiées (gauche). Maillage de référence pour le problème direct cible (centre). Maillage désiré pour l'estimation (droite).

Les mesures, synthétiques, correspondent à une image tridimensionnelle de résolution 1 cm dans toutes les directions. En pratique, elles conduiront à 50 cellules. Ces observations sont crées à partir d'un maillage de référence plus fin que celui utilisé pour l'estimation. En effet, le système réel est remplacé par un système lui-même discret mais de raffinement très

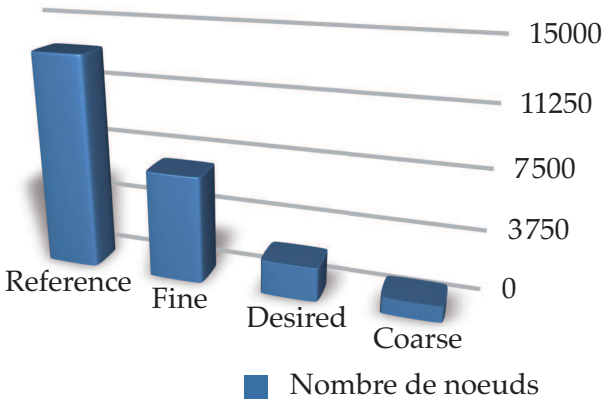


FIGURE 3.4: Nombres de noeuds pour les différentes discrétisations envisagées. Multiplier par 3 pour avoir approximativement le nombre de DOFs N_d en déplacement puis $N = 2N_d$

supérieur aux maillages désirés pendant la phase d'estimation. On analysera d'ailleurs l'efficacité de l'estimateur en fonction de la discrétisation. Pour fixer les idées nous utiliserons les différentes dénominations de la Figure 3.4 pour les maillages.

Nous utiliserons deux types de grandeurs scalaires pour suivre l'évolution du système : son énergie et le volume de sa cavité. Le premier est un indicateur mécanique évident que nous comparerons aux énergies des erreurs d'estimation. Le deuxième est plus « physiologique » et nous aide à suivre la convergence de la trajectoire de l'estimateur sur le système réel. Les courbes de références Figure 3.5 correspondent au maillage le plus raffiné, i.e. la cible. La « contractilité » $\sigma_0 = 3.10^4$Pa a d'ailleurs été choisie en cohérence avec les ordres de grandeurs des volumes éjectés. Pour le volume les courbes correspondent à l'onde sphérique ou à l'onde planaire. On remarque que, bien que la viscosité soit faible, peu de modes sont excités par l'activation électrique ce qui donne une allure régulière à la courbe de volume (notamment pour l'onde plane). Cette régularité disparaîtra dès que la condition initiale de notre système ne sera plus nulle et, quand la viscosité est faible, les fréquences excitées par la condition initiale se maintiendront sur toute la fenêtre de simulation, voir Figure 3.16.

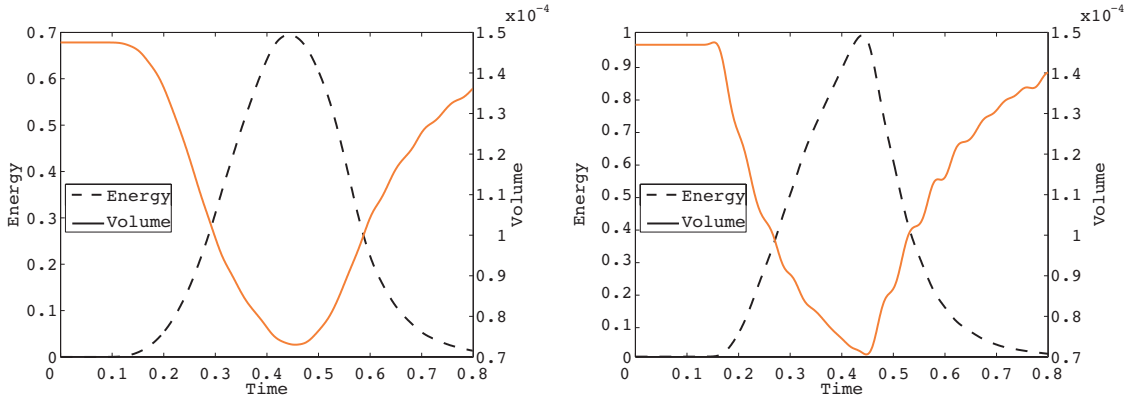


FIGURE 3.5: Energie et volume du système cible : Onde d'activation planaire (gauche), sphérique (droite)

3.4.2 Définition de l'estimateur DVF

En contrôle des structures, un contrôleur classique consiste à mesurer la vitesse en un certain nombre de points de la structure et appliquer, via un actionneur situé au même endroit, une force directement opposée à la vitesse mesurée, soit

$$\underline{f}_{\text{DVF}}(\underline{\mathrm{x}}) = -\gamma \underline{\dot{y}}_{\mathfrak{m}}(\underline{\mathrm{x}}).$$

Le fait que le contrôle s'applique en chaque point de la mesure est appelé collocalisation. On parle de contrôle collocalisé. Ce type de contrôle s'appliquant sur des vitesses est alors appelé *Direct Velocity Feedback* qui, en pratique, peut être réalisé de deux façons différentes :

- **de manière active** : on utilise des éléments piézo-électriques[6] comme capteurs (qui mesurent la vitesse) et comme actionneurs (pour appliquer la force). On trouve désormais beaucoup de structures vibrantes qui utilisent ce type de matériaux pour stabiliser les premiers modes de vibration : raquettes de tennis, skis, vitres insonorisées, ...
- **de manière passive** via tout simplement des vérins. C'est notamment le principe de stabilisation des vibrations d'une voiture par ses amortisseurs.

L'effet de ce contrôle sur les vibrations d'une structure, au moins à travers le vérin passif, est très intuitif. En effet, sur une masse isolée, la dynamique

$$m\ddot{\mathrm{y}} + k\mathrm{y} = f^{\text{ext}} + f_{\text{DVF}} \Rightarrow m\ddot{\mathrm{y}} + \gamma\dot{y} + k\mathrm{y} = f$$

est directement amortie. Pour les structures distribuées, les propriétés du DVF, notamment son efficacité en tant qu'amortisseur, sont cependant plus complexes et très largement étudiées à la fois du point de vue ingénierie mais aussi théorique, nous y reviendrons. Pour les systèmes linéaires, nous avons vu l'équivalence entre contrôle, stabilisation et observations (d'ailleurs, le DVF est un contrôle pour stabiliser une structure et pas un contrôle exact). Notre idée est d'appliquer le même type de feedback pour imaginer un filtre d'observation tel que l'erreur soit stabilisée vers 0. En effet, pour toute vitesse mesurée en un point, nous décidons d'appliquer un filtre générant une force opposée à la différence de vitesses entre la vitesse courante de notre estimateur et la vitesse du système cible :

$$\underline{f}_{\text{DVF}}(\underline{\mathrm{x}}) = \gamma(\underline{\dot{y}}_{\mathfrak{m}}(\underline{\mathrm{x}}) - \underline{\dot{\hat{y}}}(\underline{\mathrm{x}})).$$

Dans le cas de capteurs et d'actionneurs réels ayant une extension spatiale comme nos cellules de tags, la force devient donc

$$\underline{f}^{(i)}_{\text{DVF}} = \gamma s_i \int_{\Omega_{\mathfrak{m}}} s_i(\underline{\dot{y}}_{\mathfrak{m}} - \underline{\dot{\hat{y}}})\, d\Omega,$$

qui s'exprime de manière variationnelle par

$$\mathcal{P}^{\text{ext}}_{\text{DVF}} = \int_{\Omega_0} \underline{f}_{\text{DVF}} \cdot \underline{v}\, d\Omega = \sum_i \gamma \int_{\Omega_{\mathfrak{m}}} s_i(\underline{\dot{y}}_{\mathfrak{m}} - \underline{\dot{\hat{y}}})\, d\Omega \cdot \int_{\Omega_{\mathfrak{m}}} s_i \underline{v}\, d\Omega.$$

[6] L'effet piézo-électrique a été découvert en 1880 par les frères Curie et lie la déformation mécanique entre deux faces d'un cristal en fonction du champ électrique qui lui est appliqué.

Du point de vue éléments finis, le calcul de ces forces de feedback est extrêmement simple puisqu'il se limite à définir

$$K_X = \begin{pmatrix} 0 \\ \gamma M^{-1}(H^v)' \end{pmatrix}, \tag{3.30}$$

où H^v est l'opérateur $X \to (Z_i = \int_{\Omega_m} s_i \underline{\dot{y}}\, d\Omega)$. L'adjoint de H^v n'est pas en général le transposé de H^v au sens des matrices puisqu'il est pris au sens de la formulation variationnelle comme expliqué plus haut. L'adjoint est donc pris au travers de la norme L^2 adaptée à l'espace des observations que l'on peut noter W^{-1}

$$(H^v)' = (H^v)^T W^{-1}.$$

Ainsi, ce feedback est directement implémentable dans n'importe quel code élément fini, disons y compris des codes industriels où l'introduction de nouveaux opérateurs n'est possible que par manipulation de formulations classiques existantes. La dynamique de l'estimateur est donc

$$M\ddot{\hat{Y}} + \big(C + \gamma(H^v)'H^v\big)\dot{\hat{Y}} + K\hat{Y} = F + \gamma(H^v)'Z, \tag{3.31}$$

et celle de l'erreur

$$M\ddot{\tilde{Y}} + \big(C + \gamma(H^v)'H^v\big)\dot{\tilde{Y}} + K\tilde{Y} = \gamma(H^v)'(\epsilon_h + \chi). \tag{3.32}$$

Ainsi, l'énergie mécanique (rappel $N = \left(\begin{smallmatrix} K & 0 \\ 0 & M \end{smallmatrix}\right)$) vérifie

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}\tilde{X}^T N \tilde{X} = -(\dot{\tilde{Y}})^T (C + \gamma(H^v)'H^v)\dot{\tilde{Y}} - \gamma(\dot{\tilde{Y}})^T (H^v)'(\epsilon_h + \chi), \tag{3.33}$$

et elle est une fonction de Liapounov pour la partie homogène $(\epsilon_h + \chi) = 0$ de la dynamique de l'erreur.

On remarque déjà qu'on a ajouté un opérateur stabilisant le système de l'erreur en plus de son seul amortissement interne (et donc on peut même stabiliser un système initialement conservatif $C = 0$). La question est de déterminer la performance de cette stabilisation. En effet pour l'instant, l'équation (3.33) fournit une propriété évidente de stabilité $\frac{\mathrm{d}}{\mathrm{dt}}\tilde{X}^T N \tilde{X} < 0$ mais pas nécessairement d'exponentielle stabilité.

3.4.3 Estimations en dimension infinie issues du contrôle

Comme nous l'avons dit, stabilisation et contrôle sont intimement liés et le DVF a donné lieu à des études approfondies dans le cadre du contrôle des structures (réelles) c'est pourquoi nous commençons par donner un certain nombre d'éléments de justification sur les systèmes d'EDP. Afin de présenter certains résultats théoriques de manière compacte nous nous limitons au cas a priori le plus défavorable où initialement il n'y a pas d'amortissement dans le système. Le seul amortissement à considérer est donc celui apporté par le filtre DVF. De plus, puisque nous considérons la stabilisation nous regardons dans une première partie les systèmes homogènes. Soit donc une fonction $\nu(\underline{x})$ positive ou nulle sur Ω représentant l'amortissement introduit par le filtre et $\underline{\underline{\underline{\underline{A}}}}$ le tenseur d'elasticité linéaire symétrisé. La formulation forte du système de

l'élasticité linéaire s'écrit

$$\begin{cases} \dfrac{\partial^2 \underline{y}}{\partial t^2} + \nu(\underline{x})\dfrac{\partial \underline{y}}{\partial t} - \rho^{-1}\underline{\mathrm{div}}(\underline{\underline{\underline{A}}} : \underline{\underline{\nabla}}\,\underline{y}) = \underline{f} \\ \underline{y}|_{\Gamma_D} = 0, \quad \dfrac{\partial \underline{y}}{\partial \underline{n}}|_{\Gamma\setminus\Gamma_D} = 0 \\ \underline{y}(\underline{x},0) = y_0, \quad \dfrac{\partial}{\partial t}\underline{y}(\underline{x},0) = y_1 \end{cases} \tag{3.34}$$

définissant un semi-groupe d'opérateur $\mathcal{A} : D(\mathcal{A}) = D(\underline{\mathrm{div}}(\underline{\underline{\underline{A}}} : \underline{\underline{\nabla}}.)) \times H_0^1(\Omega) \subset E = H_0^1(\Omega) \times L^2(\Omega) \to E$. Depuis le début, Ω est normalement défini comme un ouvert borné de $\mathbb{R}^3$. A partir de ce système, nous évoquerons aussi la version 1D en espace, déjà présenté au chapitre 2 pour modéliser une fibre cardiaque. Sous l'hypothèse que σ_c est donné par une précontrainte active, le système de l'erreur revient simplement à considérer l'équation d'onde 1D :

$$\begin{cases} \left(\dfrac{\partial^2}{\partial t^2} + \nu(\mathrm{x})\dfrac{\partial}{\partial t} - c^2\dfrac{\partial^2}{\partial x^2}\right) y(\mathrm{x},t) = f \\ y(0,t) = 0, \quad \dfrac{\partial y}{\partial n}(L,t) = 0 \\ y(\mathrm{x},0) = y_0, \quad \dfrac{\partial y}{\partial t}(\mathrm{x},0) = y_1 \end{cases} \tag{3.35}$$

Plus généralement, il faut garder à l'esprit que y compris en 3D, l'ensemble des résultats que nous obtenons pour l'élasticité pourrait aussi être envisagé pour d'autres systèmes hyperboliques du second ordre, donc notamment l'équation des ondes. D'ailleurs les questions de stabilisation que nous recherchons sont le plus souvent étudiées en premier pour les ondes

$$\begin{cases} \left(\dfrac{\partial^2}{\partial t^2} + \nu(\mathrm{x})\dfrac{\partial}{\partial t} - \rho^{-1}(\mathrm{x})\,\mathrm{div}(A(\mathrm{x})\cdot\nabla\cdot)\right) y(\underline{\mathrm{x}},t) = f \\ y|_{\Gamma_D} = 0, \quad \dfrac{\partial y}{\partial n}|_{\Gamma\setminus\Gamma_D} = 0 \\ y(\underline{\mathrm{x}},0) = y_0, \quad \dfrac{\partial}{\partial t}y(\underline{\mathrm{x}},0) = y_1 \end{cases} \tag{3.36}$$

Pour résumer, le fait d'avoir un semi-groupe d'opérateur noté par la suite $\mathcal{T}$ est classiquement obtenu en démontrant que $A + \|\nu\|_\infty$ est maximal accrétif et en utilisant le théorème de Hille-Yosida (voir Yosida (1995) et par exemple Burq et Gérard (2002) pour une démonstration simple pour l'équation des ondes). Par extension avec la dimension finie on note $\mathcal{T}(t) = e^{\mathcal{A}t}$ et on rappelle donc que $\mathcal{T}$ commute avec son générateur infinitésimal $\mathcal{A}$. Pour ces différentes EDP, le bilan énergétique, équivalent continu de (3.33), peut s'écrire

$$\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{E}(t) = -\int_\Omega \nu(\mathrm{x})\left|\frac{\partial y}{\partial t}\right|^2 d\mathrm{x}, \tag{3.37}$$

où $\mathcal{E}$ est la somme de l'énergie mécanique et de l'énergie cinétique. Pour rappel, pour l'équation d'ondes $\mathcal{E}$ correspond à la somme des normes H^1 de y et L^2 de $\partial_t y$:

$$\mathcal{E}(y,t) = \frac{1}{2}\int_\Omega (|\partial_t y|^2 + c^2\,|\nabla y|^2)\,d\mathrm{x}.$$

□ Principe du contrôle – les sections précédentes conduisent à deux approches pour évaluer la performance du filtre au regard de l'équation (3.37). La première est une approche

strictement orientée autour de la stabilisation. La question est de savoir comment l'amortissement sur une partie du domaine entraîne d'une manière où d'une autre la stabilisation peut-être exponentielle de toute l'énergie du système. Or ce type de résultat n'a a priori rien d'évident et on pourrait croire que l'énergie du système stabilisé (qui pour nous est notre erreur d'estimation) ne diminue qu'au lieu des points où l'on observe et applique le filtre (dans le cas de la collocalisation ceux où la mesure peut s'effectuer). Or il n'en est rien pour le comportement des modes du système amortis vis-à-vis de la zone d'amortissement nous donne déjà un premier aperçu de la raison profonde de cette stabilité.

THÉORÈME 3.4.1 (STABILISATION FAIBLE)
Il y a équivalence entre :

1. *toute solution vérifie* $\lim_{t\to\infty} \mathcal{E}(y,t) = 0$ *;*
2. *tout vecteur propre* ψ *de* $\mathcal{A}$ *vérifie* $\lim_{t\to\infty} \mathcal{E}(y,t) = 0$ *pour* $y(0) = \psi$ *;*
3. *tout vecteur propre* $\psi = \begin{pmatrix} \psi^d \\ \psi^v \end{pmatrix}$ *de* $\mathcal{A}$ *vérifie* $\nu(\mathrm{x})\,|\psi^v| = 0 \Rightarrow \psi = 0$ *;*
4. *pour tout vecteur propre* ϕ *de l'opérateur elliptique* $\underline{\mathrm{div}}(\underline{\underline{\underline{A}}} : \underline{\underline{\nabla}}.)$ *(resp.* Δ*) sur* $L^2(\Omega)$ *pour l'équation des ondes vérifiant les conditions aux limites, on a* $\nu(\mathrm{x})\phi = 0 \Rightarrow \phi = 0$.

On trouvera une démonstration directe dans Burq et Gérard (2002) pour l'équation d'ondes à coefficients constants qui se généralise pour les coefficients variables et l'élasticité linéaire. C'est avant tout l'implication $(iv) \Rightarrow (i)$ qui nous est utile pour comprendre le mécanisme de stabilisation. En effet les théorèmes de prolongement unique type Holmgren (voir Ciarlet (1988) ; Lions (1988) pour l'élasticité) assurent que si un mode d'un opérateur elliptique s'annule sur un ouvert alors il est identiquement nul. Donc d'après le théorème ci-dessus toute solution de l'équation stabilisée tend bien vers 0 au cours du temps. Cependant, cette approche directe ne fournit pas de résultat sur l'exponentielle stabilité du système, ce pourquoi nous ne parlons ici que de stabilisation faible.

La deuxième approche pour évaluer (3.37) est de se reporter à la Section 3.3.3 où ce type de relation sur l'énergie issue d'un problème de stabilisation faisait référence à un problème de contrôle sous-jacent et à une condition de contrôlabilité (observabilité). On se ramène alors à un problème de contrôlabilité exacte. Ces méthodes ont largement été étudiées par Lions avec la méthode HUM et on trouvera dans Lions (1988) une présentation complète. Ces méthodes ont aussi bénéficié des travaux de Komornik (1994) sur la méthode des multiplicateurs avec applications à l'élasticité comme notamment présenté dans Guesmia (2000). Plus tard, le développement de l'analyse microlocale (voir appendice 2 de Lions (1988)) a permis de préciser les conditions sous lesquelles la contrôlabilité exacte ou la stabilisation étaient possibles. Ces conditions appelées condition de contrôle géométrique ont été initialement introduites par Bardos et al (1992). L'analyse microlocale permettant la formalisation des règles de l'optique géométrique (il existe un certain nombre de limitations dans le cas des domaines Ω non réguliers), ces règles indiquent que la stabilisation exponentielle a lieu si tout rayon de la solution rencontre dans le futur le support de ν. Ces approches conduisent alors à un résultat d'exponentielle stabilité que l'on peut écrire ici sur le semi-groupe sous la forme

$$\exists C, \delta, \quad \|\mathcal{T}(t)\| \leq Ce^{\delta t}. \tag{3.38}$$

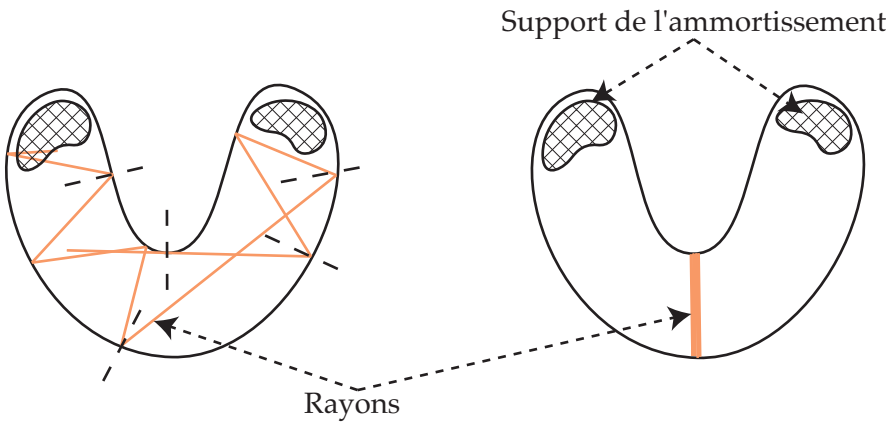


FIGURE 3.6: Condition de contrôle géométrique. A gauche, un rayon atteignant la zone d'amortissement. A droite, un rayon cloisonné dans une partie du domaine

Reste que ces méthodes ne fournissent pas de résultats *quantitatifs* sur l'exponentielle stabilité, i.e les constantes C et surtout δ. Pour obtenir des estimations plus précises sur le taux de décroissance à l'image de ce qu'on peut espérer avoir en dimension finie à partir de la jordanisation de l'opérateur, on en revient, y compris en dimension infinie, à des tentatives de décomposition spectrale. On cherche alors à prouver que l'opérateur $\mathcal{A}$ peut se décomposer sur ses vecteurs propres généralisés formant une base de Riez. On rappelle que les vecteurs propres généralisés sont exactement les extensions en dimension quelconque des vecteurs qui, en dimension finie, forment une base des sous-espaces caractéristiques[7] :

$$\mathcal{L}_i = \mathrm{Ker}(\mathcal{A} - \lambda_i \mathbb{1})^{m_i} = \{V \mid (\mathcal{A} - \lambda_i \mathbb{1})^{m_i} V = 0\}.$$

De plus, une famille (ψ_n) de vecteurs est une base de Riez d'un espace H si elle est complète dans H, qu'on peut lui associer une famille biorthogonale $(\bar{\psi}_n)$ (i.e. $< \bar{\psi}_n, \psi_m >= \delta_{m,n}$) complète dans H, et que pour tout f de H, elles vérifient (cf. rappels de Cox et Zuazua (1994))

$$\sum_n | < \psi_n, f > |^2 \leq \infty, \quad \sum_n | < \bar{\psi}_n, f > |^2 \leq \infty.$$

On pourra se reporter à Curtain et Zwart (1995) pour une présentation complète de la notion d'opérateur Riez-spectral même si la définition ne semble pas suffisamment générale pour prendre en compte les cas où la dimension de l'espace $\mathcal{L}_i$ est supérieure à 1 ($m_i = 1$ aussi). Dans ce cas contraire, les résultats sont pourtant valides si on peut toujours définir une chaîne de Jordan comme base de $\mathcal{L}_i$ comme rappelé dans Shubov (1997), c'est-à-dire des $(\psi_{i,k})_{1 \leq k \leq m}$

$$\mathcal{A}\psi_{i,k} = \lambda_i \psi_{i,k} + \psi_{i,k-1}.$$

Nous insistons sur le fait que ces résultats ne sont pas évident car A n'est pas un opérateur auto-adjoint donc l'amortissement « ne préserve pas » la base de l'opérateur elliptique sous-jacent. Cependant si on peut démontrer que les vecteurs propres généralisés forment une base de Riez, alors toute condition initiale se décomposera sous la forme

$$\begin{pmatrix} \underline{y}_0 \\ \underline{\dot{y}}_0 \end{pmatrix} = \sum_{i=\pm 1}^{\pm\infty} \sum_{k=0}^{m_i - 1} \gamma_{i,k} \psi_{i,k},$$

[7] qui, en dimension finie, génèrent toujours tout l'espace d'après Cayley-Hamilton

conduisant à la solution à tout temps (cf. preuve du théorème 3.3.2 en dimension finie)

$$\begin{pmatrix} \underline{y} \\ \underline{\dot{y}} \end{pmatrix}(t) = \sum_{\pm 1}^{i=\pm\infty} e^{\lambda_n t} \sum_{k=0}^{m_i-1} \gamma_{i,k} \sum_{l=0}^{k} \frac{t^{(k-l)}}{(k-l)!} \psi_{i,k}.$$

On en déduit, en supposant une multiplicité maximale d des valeurs propres (voir un exemple dans Cox et Zuazua (1995)),

$$\mathcal{E}(t) \leq C\mathcal{E}(0)(1+t^d)e^{2\lambda_0 t} \tag{3.39}$$

ce qui assure que

$$\delta = \inf\left(\frac{1}{t}\log\|\mathcal{T}\|\right) = \sup_{n\geq 1}(\Re(\lambda_n)). \tag{3.40}$$

En pratique, il existe extrêmement peu de résultats de cette nature. La difficulté n'est d'ailleurs pas la complétude des vecteurs propres généralisés puisque, comme rappelé dans Cox et Zuazua (1994), le théorème 10.1 du chapitre 5 de Gohberg et Kreĭn (1969) assure que toute perturbation bornée d'un opérateur antisymétrique (le cas sans viscosité) vérifie cette propriété. L'enjeu est d'obtenir les estimations de la base de Riez. Pour le problème qui nous occupe, en 1D, le résultat est apporté par Cox et Zuazua (1994) qui traite le système hyperbolique de l'équation d'onde avec amortissement interne. Donc pour (3.35), on peut démontrer que la famille des vecteurs propres généralisés de $\mathcal{A}$ est une base de Riez assurant les résultats (3.39), (3.40) où C correspond à l'équivalence de normes entre la base de Riez et $H^1([0,L]) \times L^2([0,L])$. Il existe quelques résultats en dimension 2 d'espace pour certaines géométries par Shubov mais à notre connaissance aucun résultat à ce jour en dimension 3 pour l'équation d'onde et donc *a fortiori* en élastodynamique linéaire. En d'autres termes, ce que la méthode spectrale apporte en précision, elle le perd en généralité. Or nous verrons que ce type de résultat est extrêmement puissant pour assurer la convergence de notre estimateur vis-à-vis de l'approximation.

3.4.4 Problème aux valeurs propres quadratiques

Le comportement de l'EDP stabilisée justifie, en particulier lorsqu'on bénéficie d'une propriété de type opérateur Riez-spectral, de se concentrer sur le problème aux valeurs propres discret avec l'espoir de conserver les propriétés sur le spectre à la limite $h \to 0$ sous certaines conditions, et donc de pouvoir assurer la performance uniforme par rapport à la discrétisation du filtre. En dimension finie, la propriété $\delta = \inf\left(\frac{1}{t}\log\|e^{A}t\|\right) = \sup_{n\geq 1}(\Re(\lambda_n))$ est directe ; reste maintenant à étudier l'effet du filtre sur le lieu des vecteurs propres de l'opérateur non stabilisé. Par rapport aux résultats que nous venons d'énoncer en continu, nous traiterons en dimension finie le cas plus général avec amortissement proportionnel.

Commençons par un premier calcul de spectre sur la dynamique de l'erreur sans filtre $K_X = 0$ (i.e. $\tilde{A} = A$). Déterminer les valeurs propres de A consiste à résoudre en (λ, V),

$$\begin{pmatrix} 0 & K \\ -K & -C \end{pmatrix} V = \lambda \begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix} V. \tag{3.41}$$

Cette résolution est équivalente à résoudre le problème aux valeurs propres quadratique *QEP* suivant (voir Tisseur et Meerbergen (2001) pour un aperçu des problèmes de type QEP)

$$\lambda^2 MV^d + \lambda CV^d + KV^d = 0. \tag{3.42}$$

Il n'existe pas de solution analytique à ce problème sans hypothèse sur C. En effet, pour des questions de coréduction d'opérateurs symétriques, on ne peut se ramener à un problème de $N/2$ équations séparables que pour un amortissement C proportionnel $C = \alpha M + \beta K$ ou toute combinaison polynomiale de M et K. Dans le cas simplement proportionnel, les valeurs propres se déduisent, en revanche, directement à partir de celles du système non amorti. Soit en effet les valeurs propres positives (μ_i) associées aux vecteurs propres $(\Phi_i^d)_i$ de R^{N_d} de $K \in S_n^+(\mathbb{R})$ par rapport à la norme définie par $M \in S_n^+(\mathbb{R})$. On pose $\mu_i = \omega_i^2$

$$K\Phi_i^d = \omega_i^2 M \Phi_i^d,$$

alors les valeurs propres de 3.42 vérifient

$$\lambda^2 + (\alpha + \omega_i^2 \beta)\lambda + \omega_i^2 = 0.$$

L'étude des racines de ce polynôme comprend plusieurs cas

- Si le discriminant $\Delta = (\alpha + \omega_i^2\beta)^2 - 4\omega_i^2$ est négatif, alors les valeurs propres sont complexes. Elles sont alors disposées sur un cercle de centre $c = -\frac{1}{\beta}$, $R = \frac{1}{\beta}\sqrt{1-\alpha\beta}$. En effet on vérifie alors que
$$\left(c + \frac{\alpha + \omega_i^2\beta}{2}\right)^2 + \frac{\Delta}{4} = R^2.$$
Cette région n'existe donc que si $\alpha < \frac{1}{\beta}$ (ce qui sera le cas pour nous en pratique).
- Pour ω_i plus petit, les valeurs propres sont sur la droite réelle.
- Pour ω_i plus grand, elle reviennent sur la droite réelle de part et d'autre de $-R - \frac{1}{\beta}$; notamment pour $\omega_i \to \infty$ celles situées à droite s'accumulent vers $-\frac{1}{\beta}$.

On rappelle que d'après la décomposition de la preuve du théorème 3.3.2, lorsque les valeurs propres ont une partie imaginaire non nulle, elles sont dites *sous-critiques* et *a contrario* elles sont *sur-critiques* si $\Im(\lambda_i) = 0$. En effet dans le premier cas, les modes sous-critiques associés sont vibratoires alors que les sur-critiques suivent une dynamique de premier ordre dans le second cas (voir Le Tallec (2000)).

En résumé, les valeurs propres sont disposées comme schématisé Figure 3.7, et les flèches indiquent comment elles se déplacent quand la viscosité augmente. Ainsi le système de l'erreur est déjà stable voire asymptotiquement stable, même sans filtre, mais les constantes de temps d'amortissement sont à comparer au temps T sur lequel s'effectue l'estimation. Or notre besoin (et il se renforcera lorsque nous examinerons l'estimation paramétrique) est bien de réaliser l'estimation sur des temps de l'ordre du dixième du battement soit 100 ms. Si, pour les valeurs numériques des constantes choisies au chapitre 2, ce type de constante de temps était vérifié avec le seul amortissement interne, cela signifierait que le système est entièrement sur-critique. On aurait alors un problème quasiment parabolique et non plus hyperbolique du second ordre.

Afin d'illustrer cette répartition du spectre, nous présentons Figure 3.8 un exemple pour un maillage relativement grossier (500 degrés de libertés) permettant de calculer l'ensemble du spectre. Pour obtenir à la fois des modes **numériques** sous-critiques et sur-critiques, la viscosité est augmentée par rapport à notre problème-test : $\eta = 1.6\,10^{-3}$.

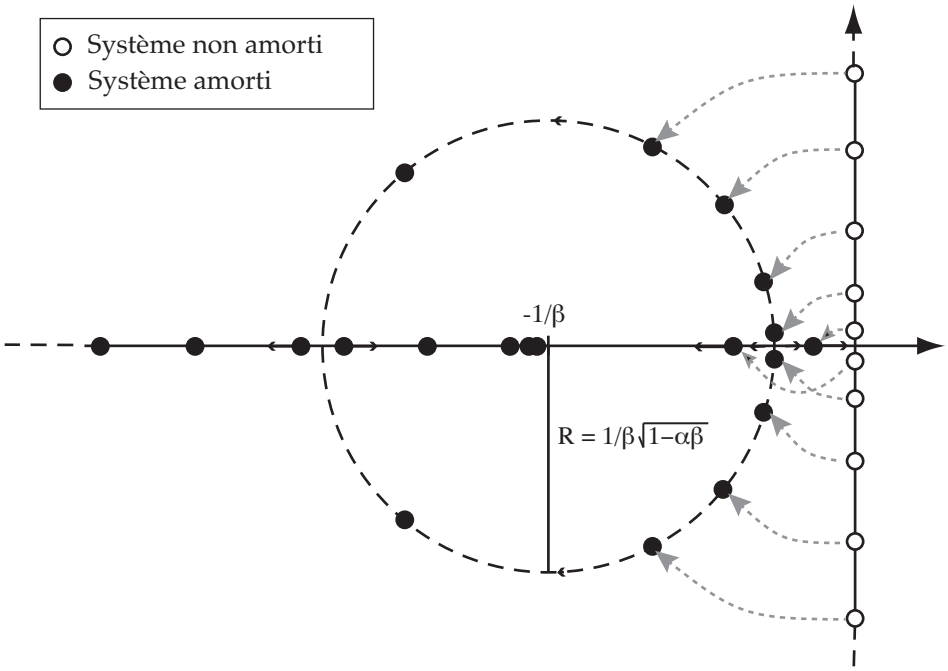


FIGURE 3.7: Pôles du système mécanique linéaire classique, sans amortissement ou avec amortissement de Rayleigh.

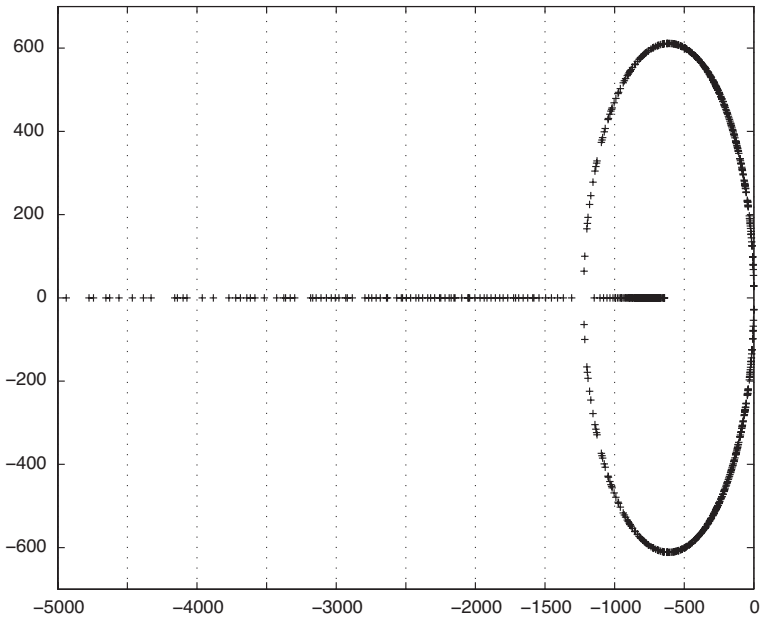


FIGURE 3.8: Exemple numérique de pôles pour un système avec amortissement de Rayleigh proportionnel à K.

a. Comportement du spectre du DVF

Une fois ajouté le terme de stabilisation, le problème aux valeurs propres à étudier devient

$$\begin{pmatrix} 0 & K \\ -K & -C-\gamma(H^v)'H^v \end{pmatrix} V = \lambda \begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix} V, \tag{3.43}$$

et nous avons vu que dans ce cas les valeurs propres sont solutions d'un QEP. Cependant, puisque l'amortissement du système de l'erreur n'est plus proportionnel, nous ne pouvons plus obtenir les valeurs propres analytiquement. Pour le DVF, nous renvoyons donc aux nombreuses études qualitatives sur le spectre Preumont (2002) ; Collet et al (2003) dont nous allons résumer les principales caractéristiques.

□ EXEMPLES NUMÉRIQUES – Mais commençons par nous fixer les idées autour de quelques exemples numériques que nous interprétons successivement dans les paragraphes suivants. Tout d'abord sur la Figure 3.9 sont représentés, à gauche, les pôles de la structure non stabilisée avec celle d'une structure stabilisée. Le choix du paramètre $\gamma = 0.9$ correspond dans le premier cas au décalage maximal des modes de basse fréquence. A droite dans le deuxième cas, avec $\gamma = 2$ ou $\gamma = 5$ au contraire, on cherche plutôt un décalage plus important sur l'ensemble du spectre quitte à ce que les toutes premières fréquences soient relativement moins stabilisées.

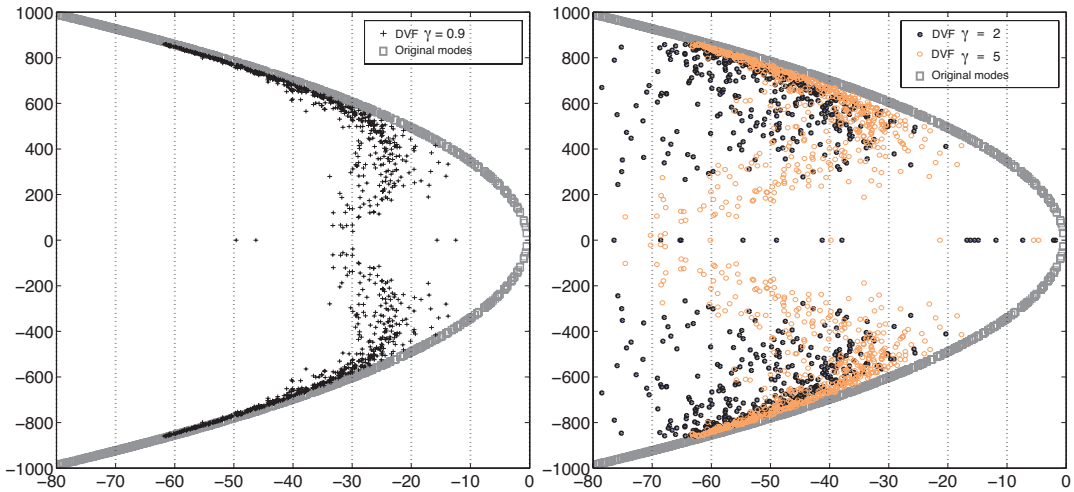


FIGURE 3.9: QEPs en fonction du paramètre γ réalisés sur le maillage « desired mesh ». 1000 modes sont représentés

Dans la Figure 3.10, nous observons le comportement du spectre pour deux viscosités différentes. La première correspond à la viscosité utilisée dans le problème-test alors que la seconde correspond à un système initial conservatif. Le maillage est grossier (500 degrés de liberté) afin de calculer l'ensemble des valeurs propres. On constate que les hautes fréquences représentent une véritable difficulté dans le cas conservatif car elles restent « plaquées » sur l'axe imaginaire. Inversement l'amortissement interne de notre problème-test permet un contrôle des hautes fréquences justifiant de se concentrer uniquement sur les basses fréquences contenues dans le rectangle en pointillés.

La dernière série de figures concerne l'évolution du spectre en fonction de la discrétisation.

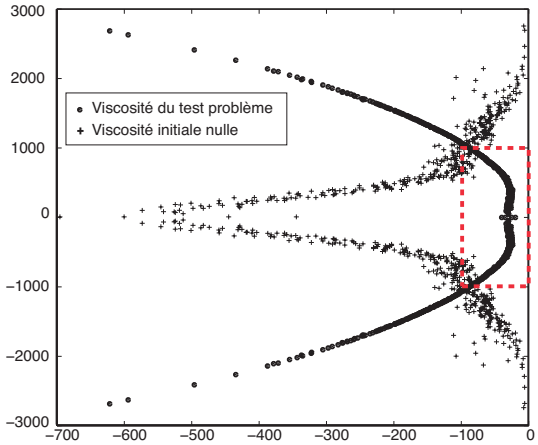


FIGURE 3.10: Spectre pour deux situations de viscosité interne, celle du problème-test et une viscosité initiale nulle. ($\gamma = 0.9$)

La Figure 3.11 présente les différentes évolutions du spectre dans le cas $\gamma = 2$ pour les discrétisations présentées tableau 3.4. Nous nous limitons à un nombre restreint de modes (100) pour limiter le coût de calcul sur les maillages raffinés. On constate que sur les premiers modes en tout cas, la convergence est déjà largement acquise pour les discrétisations envisagées.

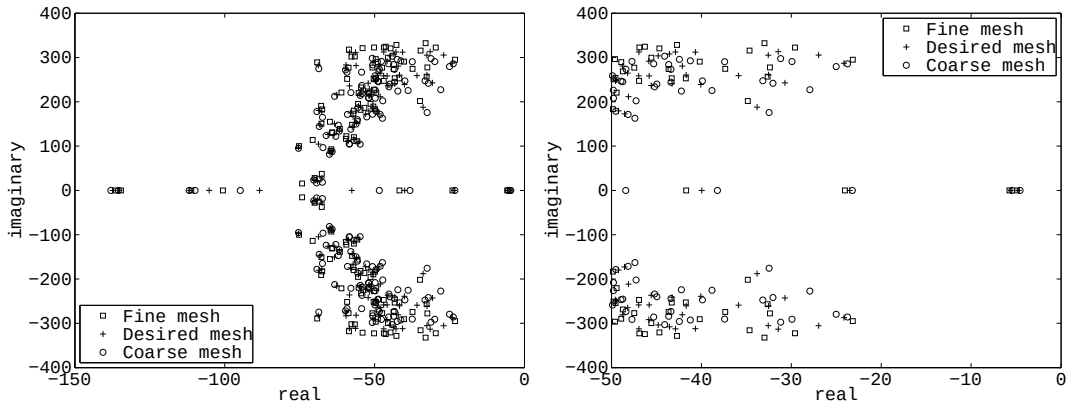


FIGURE 3.11: Exemple d'évolution du spectre en fonction de la discrétisation (zoom à droite)

□ HOLMGREN DISCRET – La première question légitime sur les modes discrets est de savoir si les théorèmes d'unicité de type Holmgren permettent comme en dimension infinie d'assurer que tous les modes discrets sont tels que leur partie réelle est strictement négative à partir des modes de la structure non contrôlée (cf. théorème 3.4.1). Le résultat dans ce cas n'est pas du tout évident en dehors d'arguments de consistance par rapport au problème continu sachant que les modes discrets de hautes fréquences « n'ont pas de sens physique »... Par exemple sur une structure $1D$ élastique, pour que la nullité d'un mode sur une partie du domaine discret implique que tout le mode est nul il suffit que 2 points successifs soient nuls. En effet, l'information se propage alors de proche en proche à partir de la relation « elliptique » discrète

$$-\Phi^d_{i,\mathrm{x}+\Delta\mathrm{x}} + \Phi^d_{i,\mathrm{x}} - \Phi^d_{i,\mathrm{x}-\Delta\mathrm{x}} = 0.$$

Il suffit donc que le raffinement soit suffisant pour assurer que la zone d'observation réelle comporte au moins deux points de discrétisation. En dimension supérieure à 1 cependant, on imagine aussi immédiatement la difficulté de ce type de raisonnement de proche en proche pour une discrétisation spatiale quelconque.

□ ANALYSE DE SENSIBILITÉ – Une deuxième analyse du comportement du spectre de l'opérateur stabilisé consiste à réaliser une étude de sensibilité en fonction de γ. Ce calcul peut être mené pour la sensibilité autour de $\gamma = 0$. Pour $\Psi_i(\gamma)$ un mode de la structure stabilisée on a

$$(\lambda^2(\gamma)M + \lambda(\gamma)\gamma H'H + K)\Psi_i^d = 0.$$

donc autour de $\gamma = 0$,

$$\Big(2\lambda(0)\frac{\mathrm{d}\lambda}{\mathrm{d}\gamma}(0)M + \lambda(0)H'H\Big)\Psi_i^d(0) + \Big(\lambda^2(0)M + K\Big)\frac{\mathrm{d}\Psi_i}{\mathrm{d}\gamma}(0) = 0.$$

Soit $\Phi_i^d = \Psi_i^d(0)$ le mode de la structure non stabilisée et $\lambda(0) = -\omega_i^2$ (tel que $K\Phi = \omega_i^2 M$) sa valeur propre associée. En testant cette identité sur Φ_i^d, la deuxième partie s'annule et on obtient alors

$$2\omega_i^2\frac{\mathrm{d}\lambda}{\mathrm{d}\gamma}(0)\Phi_i^T M\Phi_i^d + \omega_i^2\Phi_i^T H'H\Phi_i^d = 0.$$

En choisissant des modes tel que le déplacement soit normalisé par M, on obtient alors

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\gamma}(0) = -\frac{(\Phi_i^d)^T H^T H\Phi_i^d}{2}. \tag{3.44}$$

On remarque que seule la partie réelle des valeurs propres est affectée au premier ordre par l'ajout de la stabilisation et que celles-ci sont bien toutes décalées vers les valeurs négatives. Cependant le résultat précédent ne permet pas d'aller beaucoup plus loin car c'est simplement une sensibilité et met en jeu les modes de la structure stabilisée testés à travers $H'H$.

□ RETOUR VERS L'AXE – La sensibilité est une chose, mais les ingénieurs savent aussi que lorsque le gain est grand les valeurs propres de la structure stabilisée ont tendance à revenir vers l'axe imaginaire. Ainsi, alors qu'on espérait stabiliser la structure, on modifie le spectre de telle façon que les partie réelles des modes stabilisés deviennent moins négatives donc le système « moins » exponentiellement stable. Cette propriété est classiquement justifiée par Preumont (2002) à partir d'une formalisation autour de la fonction de transfert du système vibratoire représenté Figure 3.12.

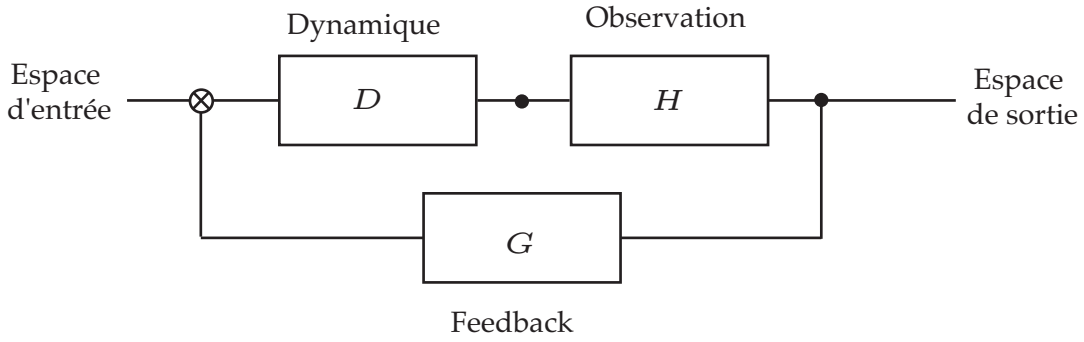


FIGURE 3.12: Schéma du système entrée-sortie définissant la fonction de transfert.

Afin de simplifier le formalisme pour obtenir des résultats heuristiques, nous nous limitons à une structure 1D contrôlée en un point (possible uniquement en dimension 1 d'ailleurs). Cette simplification entraîne que la fonction de transfert considérée est celle d'un système SISO (single input - single output). On rappelle que pour la fonction de transfert considérée Figure 3.12, l'entrée est la force appliquée au point d'observation. On note $T(s)$ cette fonction de transfert. Elle s'écrit de façon très générale

$$T(s) = \frac{H(s)D(s)}{1 - G(s)H(s)D(s)},$$

où H, D et G sont les fonctions de transfert de chacun des blocs. Puisque l'observation est une vitesse on a

$$H = s(0, \ldots, 0, 1, 0, \ldots, 0),$$

et $G = H^T$. Pour la dynamique, la fonction de transfert HD est classique (voir chap. 2 de Preumont (2002)) est reproduite Figure 3.13. Elle est notamment croissante entre deux pôles successifs (ceux qui annulent $Ms^2 + K$) et les zéros s'intercalent avec les pôles. De plus, dans le cas considéré, les zéros sont des vibrations qui s'annulent au lieu du capteur. C'est donc par définition un pôle de la structure où on a bloqué le point d'emplacement de l'actionneur.

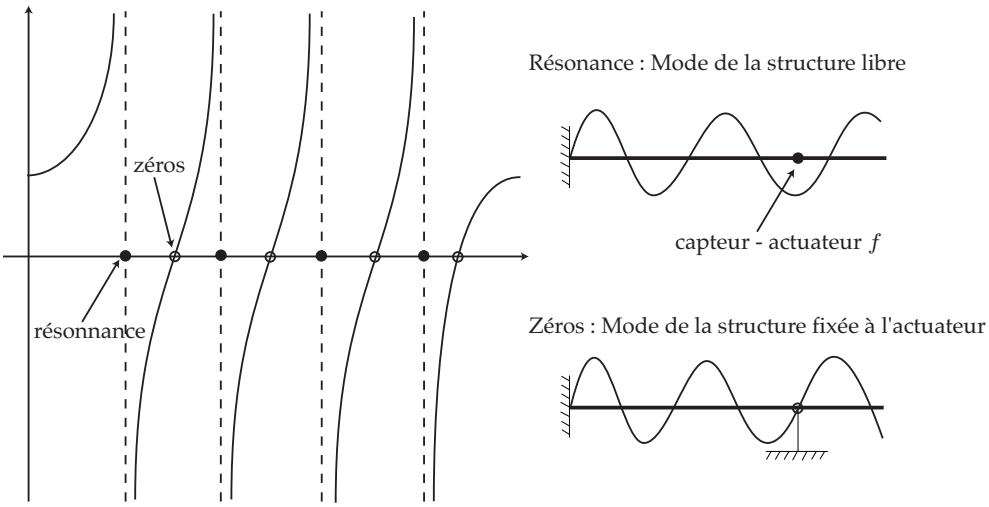


FIGURE 3.13: A gauche, fonction de transfert d'un structure 1D avec un actionneur, capteur collocalisé. A droite des modes associés aux résonances et anti-résonances de la fonction de transfert.

Ces remarques nous permettent d'analyser les pôles de la fonction de transfert totale. Pour cela, il suffit de localiser les zéros de $1 - \gamma GHD$ car les pôles de HD sont aussi pôles de $1 - \gamma GHD$ et donc apparaisent au numérateur et dénominateur de la fonction de transfert. Or les zéros de $1 - \gamma GHD$ en fonction de γ parcourent le chemin entre les deux cas limites :

- $\gamma \to 0$ alors $GHD = \frac{1}{\gamma} \to \infty$ donc les zéros de $1 - \gamma GHD$ sont pôles de GHD. Or G et H n'admettent pas de pôles donc ce sont les pôles de la dynamique.
- $\gamma \to \infty$ alors $GHD = \frac{1}{\gamma} \to 0$ donc les zéros de $1 - \gamma GHD$ sont zéros de GHD. Ils sont donc soit zéros de la dynamique soit correspondent à s=0.

Ces cas limites nous permettent de caractériser approximativement le lieu des pôles de la fonction de transfert en fonction de γ. Ceux-ci passent donc des pôles aux zéros de la dynamique observée sans filtre comme illustré Figure 3.14. Si on considère dans le plan complexe les arguments α_i de chaque contribution du numérateur $(z-z_i)$ et β_j de celle du dénominateur $(z-p_j)$ pour la fonction de transfert GHD,

$$GHD = \frac{\Pi_i(z-z_i)}{\Pi_j(z-p_j)}.$$

On peut alors préciser la courbe cherchée car

$$GHD = \frac{1}{\gamma} \Rightarrow \arg(\gamma GHD) = 0 \Rightarrow \sum \alpha_i - \sum \beta_j = 0.$$

En négligeant l'influence (donc qualitativement) des autres zéros et pôles, on a en tout point de la courbe décrite par les valeurs propres entre un pôle et un zéro successif

$$\delta\Psi + \delta\Phi = 0,$$

ce qui signifie que l'angle $\alpha_i - \beta_i$ est constant. Cette propriété caractérise alors un cercle.

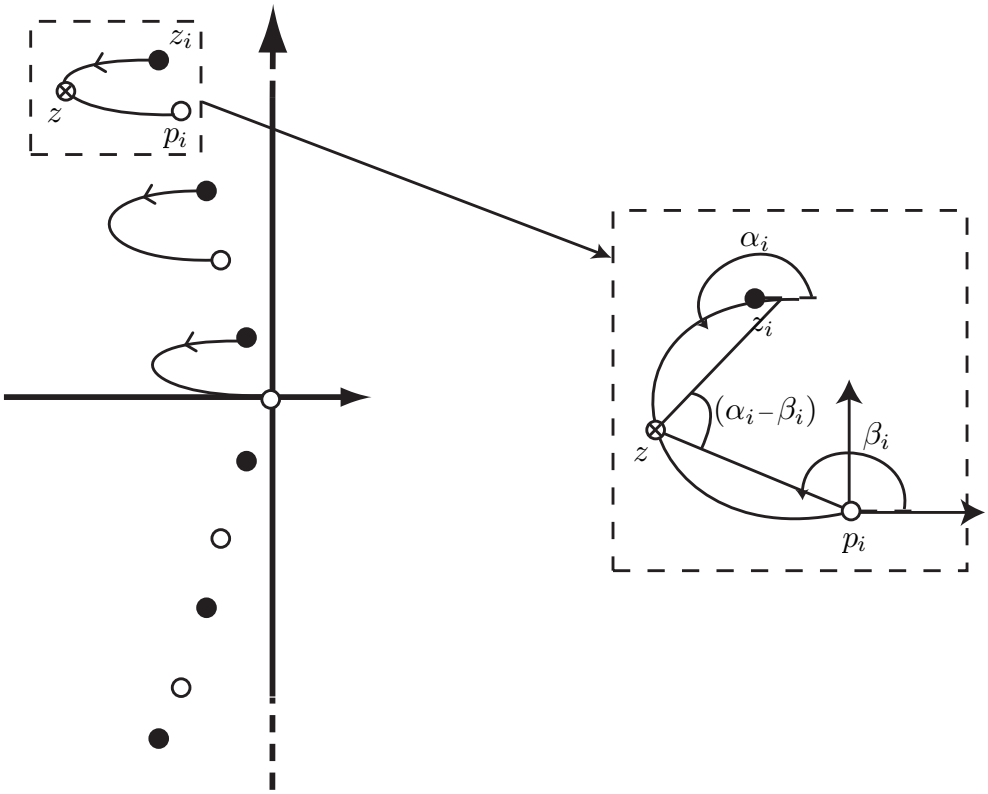


FIGURE 3.14: Evolution des zéros de la dynamique contrôlées en fonction de γ

□ GAIN OPTIMAL – Le gain optimal est donc déterminé pour obtenir l'éloignement maximum des modes entre leur position de pôles de la dynamique non stabilisée et leur retour sur le zéro. Cependant cette condition est nécessairement soumise à la répartition globale des zéros et pôles et à une dépendance homogène en fonction de γ. Ainsi il est préférable de se

concentrer sur l'atténuation maximum des premiers modes de la structure sous condition que la viscosité interne du système (toujours présente) stabilise durablement les hautes fréquences. Notamment le retour sur l'axe des modes de basse fréquence est clairement limitant dans la détermination du gain optimal car ce sont potentiellement les modes les plus représentés dans la dynamique. Ainsi le gain optimal est à notre avis à choisir tel que les premiers modes soient à la limite de leur régime surcritique c'est-à-dire rejoignent l'axe réel.

□ LIMITE EN DIMENSION INFINIE – Reste la question de l'évolution du spectre avec la discrétisation. Dans le cas où le spectre de l'opérateur continu est discret (ce qui est le cas pour nous puisque le système est à résolvante compacte) alors on peut assurer la convergence du spectre discret approché vers celui-ci par la théorie des perturbations (voir Kato (1995)). On pourrait alors supposer que le gain optimal est stable par rapport à la discrétisation. Cependant les limites de convergence du spectre discret ne sont pas uniformes par rapport à la discrétisation et les hautes fréquences discrètes sont souvent parasites comme le rappellent notamment les travaux nombreux comme ceux de Banks et Fabiano (1998) ; Roder Tcheugoué Tébou et Zuazua (2007). On voit, notamment, un repliement sur l'axe imaginaire des hautes fréquences caractéristiques, intuitivement, de l'absence de propriété de type Holmgren discret. Dans le cas d'une viscosité suffisante, les hautes fréquences sont toutes contrôlées par l'amortissement interne et dans ce cas, le gain optimal est stable par rapport à la discrétisation car il est imposé par le comportement des basses fréquences. Dans le cas où la viscosité interne est nulle Roder Tcheugoué Tébou et Zuazua (2007) démontre qu'il est nécessaire pour l'équation d'ajouter à la discrétisation un terme de viscosité numérique en h^2K afin d'éviter le repliement des hautes fréquences sur l'axe imaginaire. Dans les différents cas pratiques étudiés par la suite, nous considérerons donc toujours un système numérique qui comporte au moins ce niveau de viscosité interne.

b. Cas de la discrétisation en temps

Du point de vue discrétisation en temps, le feedback correspond à un opérateur de viscosité défini variationellement comme n'importe quel autre opérateur mécanique. Le schéma en temps est donc immédiat puisqu'on lui applique directement le schéma de Newmark pour faire apparaître le terme

$$\gamma\Big((H^v)^T H^v\Big)\Big(\frac{\dot{Y}_{n+1}+\dot{Y}_n}{2}\Big).$$

Le système récursif à résoudre s'écrit alors matriciellement

$$\begin{cases} A_1 X_{n+1} = A_0 X_n + \tilde{R} \\ X_{n=0} = X_0 + \zeta^X \end{cases}$$

avec

$$A_1 = \begin{pmatrix} \frac{1}{\Delta t} & -\frac{1}{2} \\ \frac{1}{2}K & \frac{1}{2}C + \frac{\gamma}{2}(H^v)^T H^v + \frac{1}{\Delta t}M \end{pmatrix}, \quad A_0 = \begin{pmatrix} \frac{1}{\Delta t} & \frac{1}{2} \\ -\frac{1}{2}K & -\frac{1}{2}C - \frac{\gamma}{2}(H^v)^T H^v + \frac{1}{\Delta t}M \end{pmatrix}.$$

Une question importante se pose cependant : celle de l'efficacité de la stabilisation y compris au niveau discrétisé en temps. Cette question est légitime comme le montrent de nombreux

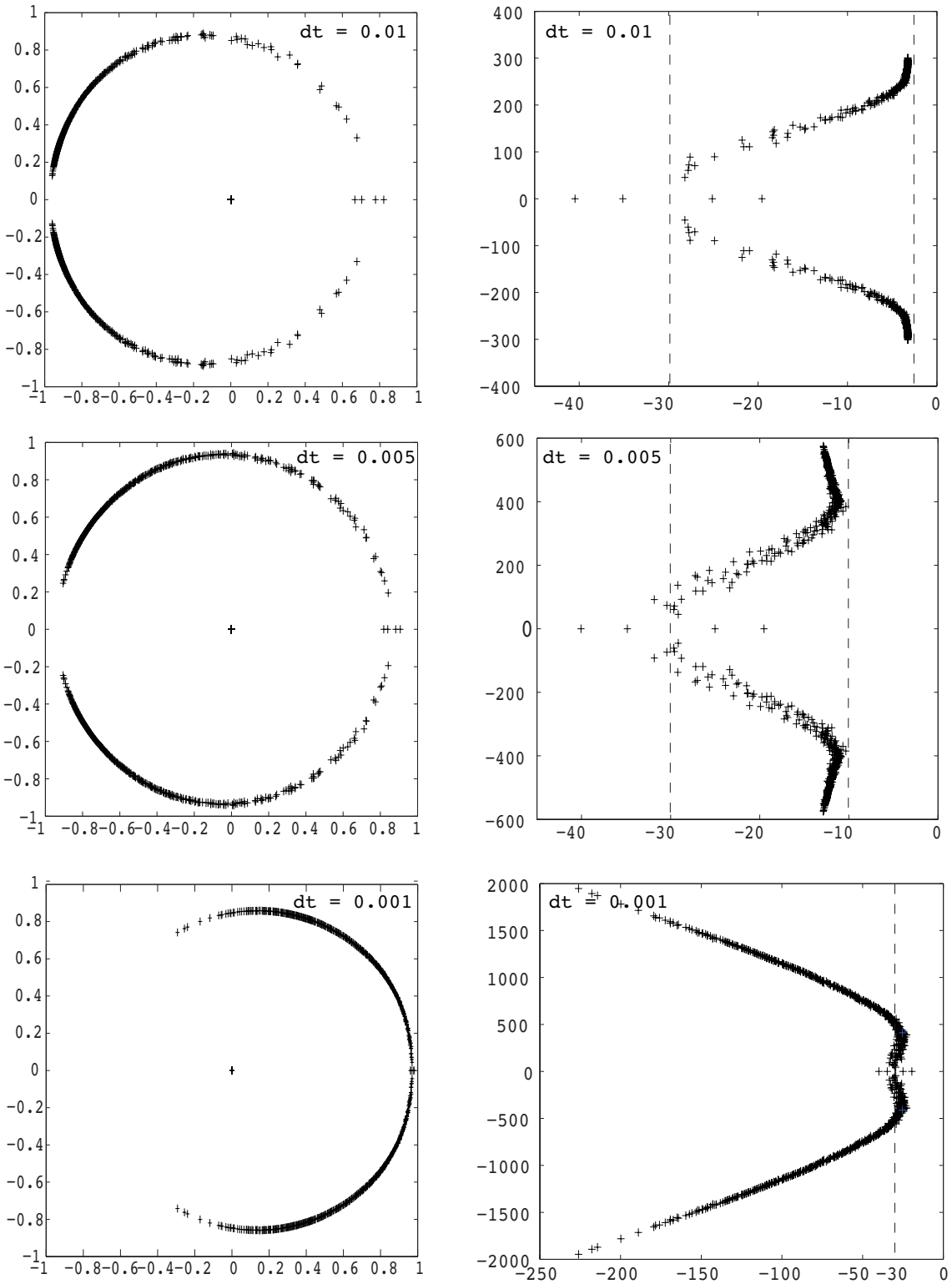


FIGURE 3.15: Pôles du système discrétisé en temps. A gauche pour l'opérateur de transition et à droite leur correspondance en temps continu.

travaux tels que ceux de Ervedoza et Zuazua (2008). Le système récursif obéit au même type de propriété de stabilité fonction de la position des pôles de l'opérateur de transition. Cependant, dans ce cas, la condition de placement se transforme en « toutes les valeurs propres sont de module strictement inférieur à 1 ». En effet, soit $V_{\Delta t}$ un pôle discret

$$A_1^{-1} A_0 V_{\Delta t} = \lambda_{\Delta t} V_{\Delta t},$$

tel que pour tout n, le système récursif initialisé à $V_{\Delta t}$ vérifie

$$V_n = (\lambda_{\Delta t})^n V_{\Delta t}.$$

A ces valeurs propres, on peut associer des valeurs en temps continu en considérant que V_n approche un mode de pôle λ_{eq} à $n\Delta t$ soit

$$(\lambda_{\Delta t})^n = e^{\lambda_{\text{eq}} n \Delta t} \Rightarrow \lambda_{\text{eq}} = \frac{\ln(\lambda_{\Delta t})}{\Delta t}.$$

Figure 3.15, nous présentons plusieurs résultats numériques de calcul de pôles pour différents pas de temps avec un maillage grossier de 500 degrés de libertés. Tous les pôles de cette structure sont ainsi représentés. On constate que, bien que le système reste stable comme garanti par la conservativité du schéma, le choix de Δt est déterminant par rapport au placement global des pôles et que, ici, $\Delta t = 10^{-3}$ est nécessaire pour retrouver le profil des pôles attendu en continu (visible pour cette discrétisation Figure 3.10). Ce résultat est cohérent avec le fait que les pôles numériques obtenus après discrétisation spatiale sont plafonnés en fréquence en fonction du pas de maillage. Pour que le système discret préserve l'exponentielle stabilité vis-à-vis des perturbations, il faut donc que le pas de temps soit adapté au contenu fréquentiel de ces perturbations, au moins jusqu'au niveau des fréquences numériques maximales des modes du système discret en espace. Dans notre cas, la propagation de l'activation électrique nous avait contraint à prendre un pas de temps $\Delta t = 10^{-3}$ qui assure la stabilisation attendue des 1000 premiers modes de la structure (*overkill*). Pour d'autres problèmes mécaniques où l'on souhaiterait relaxer cette forme de condition CFL pour la stabilisation des modes, le schéma en temps devrait être revu en conséquence, comme exprimé dans Ervedoza et Zuazua (2008), et peut constituer un enjeu.

3.4.5 Estimation d'erreur en fonction des bruits

Afin d'obtenir une estimation d'erreur globale pour notre observateur nous le décomposons suivant les différentes incertitudes intervenant dans la dynamique. Pour illustrer les résultats théoriques présentés, nous nous appuyons sur un certain nombre d'expériences numériques à partir du modèle simplifié présenté Section 3.4.1. Nous considérons notamment deux types de conditions initiales différentes.

- **Pressure I.C.** : On impose une surpression supplémentaire au niveau de l'endocarde de l'ordre de $10^3\ Pa$ soit l'ordre de grandeur de la pression auriculaire. Dans ce cas l'erreur de condition initiale correspond à une erreur de l'ordre de 1% de l'énergie maximale (en contraction complète) et à une augmentation de volume de l'ordre 30%. C'est typiquement une erreur de condition initiale cohérente du point de vue physiologique si on suppose un erreur dans la définition de la géométrie de référence ou sur la pression auriculaire.

- **Modal I.C.** : Ce type de condition est plus théorique et consiste à définir un déplacement statique initial formé par une combinaison linéaire de quelques modes. Ces modes pourront être, soit des modes de la structure initiale, soit de la structure stabilisée. Dans ce cas, on considérera une erreur en énergie de l'ordre de 10% afin d'évaluer pleinement l'effet de l'estimation.

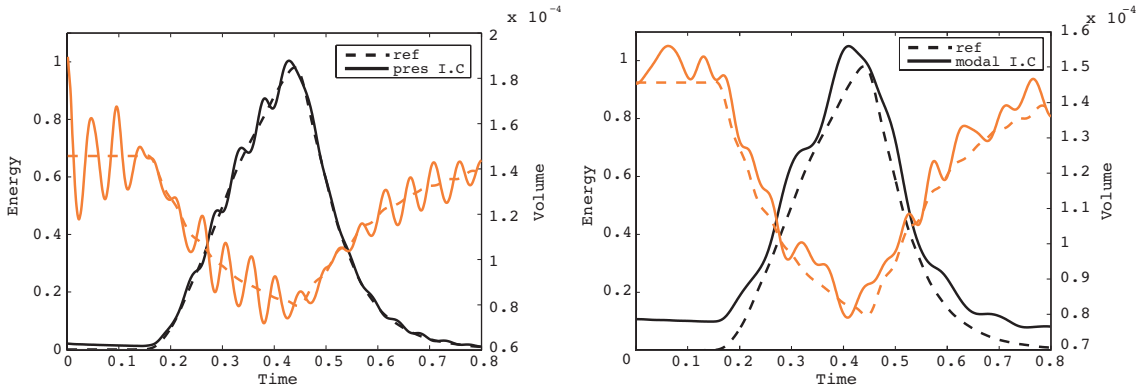


FIGURE 3.16: Courbe d'énergie et volume pour les conditions initiales « Pressure I.C. » (gauche), « Modal I.C. » (droite).

Une remarque est nécessaire à ce stade sur ces choix d'exemples numériques. Nous avons choisi comme condition initiale pour le système de référence la configuration $X(0) = 0$. Les conditions initiales ci-dessus sont alors celles des estimateurs testés, i.e. $\hat{X}(0)$ est soit la « Pressure I.C. » soit « Modal I.C. ». Nos notations impliquent donc que X_0 n'est pas nulle et change en fonction de la condition initiale choisie puisque

$$\hat{X}(0) = X_0,$$

et

$$X(0) = X_0 + \zeta^X = 0 \Rightarrow \zeta^X = -X_0 = -\hat{X}(0).$$

Ceci peut sembler surprenant sur le point de vue théorique par rapport aux notations fixées dans tout le document d'autant que, logiquement, X_0 représente une sorte de moyenne statistique sur les configurations de référence. On aurait mieux imaginé 0 comme valeur par défaut. Ce choix est motivé uniquement par des considérations pratique sans incidence sur la qualité des exemples. En effet, dans ce cas, une seule simulation de référence (faite sur un maillage beaucoup plus fin) permet ensuite de tester plusieurs types d'incertitudes dans l'estimation d'état.

a. Erreur de condition initiale

Pour l'erreur de condition initiale, examinons le système homogène défini sur $\mathbb{R}^N$

$$\begin{cases} \dot{\tilde{X}}_{hg} = (A - K_X H)\tilde{X}_{hg} \\ \tilde{X}_{hg}(0) = \zeta^X \end{cases} \tag{3.45}$$

Nous avons déjà fait un certain nombre de remarques sur le stabilité de ces systèmes, et dit qu'en 1D on disposait même d'estimations précises du taux décroissance car on pouvait démontrer d'après Cox et Zuazua (1994) que l'opérateur $A - KH$ était Riez-spectral. Les discrétisations spatiales de Roder Tcheugoué Tébou et Zuazua (2007) et temporelles de Ervedoza et Zuazua (2008) fournissent alors un taux de décroissance exponentielle uniforme par rapport à la discrétisation. On obtient donc

$$\exists C,\ \forall h,\ \|\mathcal{T}^h(t)\| \leq Ce^{-\delta t}, \tag{3.46}$$

où δ peut être calculé numériquement à partir d'un modèle de dimension finie raisonnable. On obtient alors une majoration du type

$$\|\tilde{X}_{hg}(t)\|_{\mathcal{E}} \leq C'e^{-\delta t}\|\tilde{X}_{hg}(0)\|_{\mathcal{E}}.$$

Nous complétons l'ensemble des remarques ci-dessus par une approche plus spécifique au cas considéré[8]. En effet, on peut supposer que ζ^X n'est pas une indétermination quelconque car elle représente la condition initiale du système réel. Ainsi, on peut supposer que cette dernière est suffisamment régulière pour que $X(0)$ et, donc *a fortiori* ζ^X, soit bien approché par quelques modes de la structure initiale (i.e. non stabilisée). Notons toujours (Φ_i, μ_i) ces modes où, dans le cas d'un amortissement proportionnel, les Φ_i sont les mêmes que pour une structure classique non amortie. Les modes de la structure stabilisée seront notés (Ψ_i, λ_i). Nous proposons dans ces conditions le résultat suivant

PROPOSITION 3.4.2
Supposons que pour le système (3.45) *la condition initiale peut être décomposée sur la base modale* $\{\Phi_i\}$ *du système non stabilisé*

$$\tilde{X}_{hg}(0) = \sum_{i=1}^{q} \alpha_i \Phi_i + r_q.$$

Alors pour tout temps t *on a l'estimation*

$$\begin{aligned}\|\tilde{X}_{hg}(t)\|_{\mathcal{E}} \leq C(q')e^{-\delta(q')t}\big(1 + d(q,q')\big)\left(\|\tilde{X}_{hg}(0)\|_{\mathcal{E}} + \|r_q\|_{\mathcal{E}}\right) + \|r_q\|_{\mathcal{E}} \\ + d(q,q')\left(\|\tilde{X}_{hg}(0)\|_{\mathcal{E}} + \|r_q\|_{\mathcal{E}}\right),\end{aligned} \tag{3.47}$$

où les constantes sont définies comme suit

$$\delta(q') = \inf_{i \leq q'} \left(-\Re(\lambda_i)\right),$$

$$d(q,q') = \sup_{V \in \mathrm{span}(\Phi_i)_{i=1}^{q}} \ \inf_{W \in \mathrm{span}(\Psi_i)_{i=1}^{q'}} \frac{\|V - W\|_{\mathcal{E}}}{\|V\|_{\mathcal{E}}},$$

$$C(q') = \frac{\gamma_2(q')}{\gamma_1(q')} \textit{ tel que } \forall (\alpha_i)_i,\ \gamma_1(q')\sum_{i=1}^{q'}|\alpha_i| \leq \Big\|\sum_{i=1}^{q'}\alpha_i\Psi_i\Big\|_{\mathcal{E}} \leq \gamma_2(q')\sum_{i=1}^{q'}|\alpha_i|.$$

[8]Je remercie vivement Frédéric Bourquin, LCPC, pour la discussion avec lui très fructueuse à ce sujet

■ ***Démonstration :*** *Si on décompose chaque* Φ_i, $1 \leq i \leq q$ *sur les* Ψ_i, $1 \leq i \leq q'$,

$$\sum_{i=1}^{q} \alpha_i \Phi_i = \sum_{j=1}^{q'} \beta_j \Psi_j + r_d.$$

Puis la condition inf-sup entre les sous-espaces permet de contrôler r_d

$$\begin{aligned} \|r_d\|_{\mathcal{E}} &\leq d(q,q')\|\sum_{i=1}^{q} \alpha_i \Phi_i\|_{\mathcal{E}} \\ &\leq d(q,q')(\|\tilde{X}_{hg}(0)\|_{\mathcal{E}} + \|r_q\|_{\mathcal{E}}). \end{aligned}$$

En utilisant la résolvante du système notée ici $\mathcal{T}^h(t)$ *pour des extensions possibles au semi-groupe en dimension infinie*

$$\tilde{X}_{hg}(t) \leq \sum_{i=1}^{q'} \beta_i e^{\lambda_i t} \Psi_i + \mathcal{T}^h(t)(r_d + r_q),$$

avec

$$\|\mathcal{T}^h(t)(r_d + r_q)\|_{\mathcal{E}} \leq \|r_d + r_q\|_{\mathcal{E}}.$$

Ainsi

$$\begin{aligned} \|\tilde{X}_{hg}(t)\|_{\mathcal{E}} &\leq \gamma_2(q') \sum_{i=1}^{q'} |\beta_i| \, e^{\lambda_i t} + \|r_d + r_q\|_{\mathcal{E}} \\ &\leq \gamma_2(q') e^{-\delta(q')t} \sum_{i=1}^{q'} |\beta_i| + \|r_d + r_q\|_{\mathcal{E}} \\ &\leq \frac{\gamma_2(q')}{\gamma_1(q')} e^{-\delta(q')t} \|\sum_{i=1}^{q'} \beta_i \Psi_i\|_{\mathcal{E}} + \|r_d + r_q\|_{\mathcal{E}}, \end{aligned}$$

qui en lui réinjectant

$$\begin{aligned} \|\sum_{i=1}^{q'} \beta_i \Psi_i\|_{\mathcal{E}} &\leq \|\sum_{i=1}^{q} \alpha_i \Phi_i\|_{\mathcal{E}} + \|r_d\|_{\mathcal{E}} \\ &\leq \|\tilde{X}_{hg}(0)\|_{\mathcal{E}} + \|r_q\|_{\mathcal{E}} + \|r_d\|_{\mathcal{E}}, \end{aligned}$$

donne le résultat (3.47). ■

L'interprétation du théorème est claire. Une fois la condition initiale décomposée sur les premiers pôles de la structure à un reste près, on peut alors projeter chacun de ces modes sur les pôles de la structure stabilisée. La constante $d(q,q')$ représente la distance entre les sous-espaces de vecteurs propres de la structure initiale $\{\Phi_i\}$ et celle de la structure stabilisée $\{\Psi_i\}$ et $C(q')$ la constante d'équivalence des normes entre la norme d'énergie et celle associée aux composantes modales. On peut vérifier numériquement en calculant les quotients de Rayleigh (voir section suivante) que $d(q,q')$ est petite dès que $q' \geq q$, et que $C(q')$ reste finie, et ce indépendamment du maillage considéré. C'est en quelque sorte une vérification numérique de la propriété d'opérateur Riez-spectral concentré sur les basses fréquences qui nous intéressent

plus spécifiquement ici, puisque nous poursuivons un système réel supposé régulier. Ainsi, nous sommes conduits à une estimation de la forme

$$\|\tilde{X}_{hg}(t)\|_{\mathcal{E}} \leq C_1 e^{-\delta_1 t} \|\tilde{X}_{hg}(0)\|_{\mathcal{E}} + \epsilon,$$

où δ_1 est donné par l'étude du spectre de l'opérateur pour les basses fréquences. On évite ainsi les éventuels comportements indésirables des haute fréquence lors de la convergence de la discrétisation spatiale. On note au passage qu'un raisonnement similaire pourrait s'effectuer sur les modes du système discrétisé en temps. Reste tout de même que ces constantes doivent être quantifiées numériquement (voir les résultats ci-dessous). Enfin, nous avons démontré ce résultat en décomposant sur les vecteurs propres, mais on pourrait tout aussi bien le faire sur les vecteurs propres généralisés avec des termes de stabilisation comportant des polynômes en facteur des exponentielles.

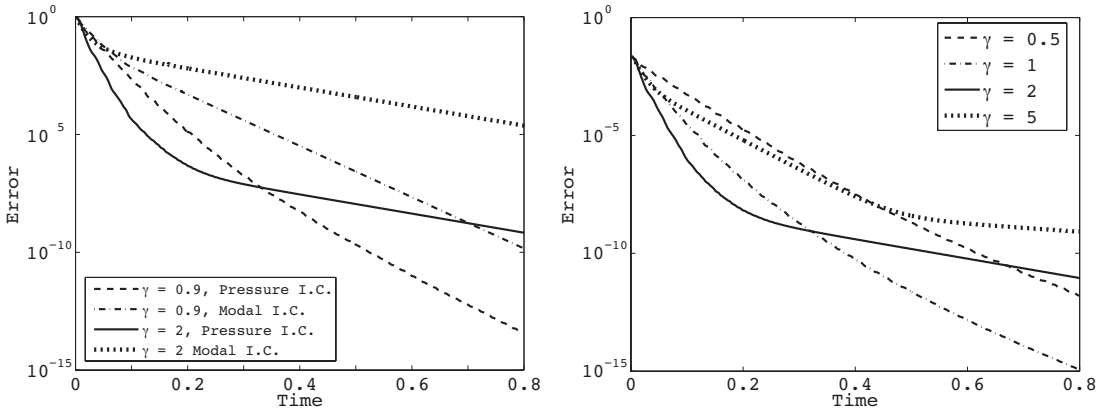


FIGURE 3.17: Énergie relative de l'erreur de condition initiale pour deux choix de condition initiale et de paramètre γ (gauche). Erreur absolue spécifique de la condition initiale en surpression dont le contenu fréquentiel implique de choisir $\gamma = 2$ pour un taux de décroissance maximum (droite).

Passons à des résultats numériques d'estimation pour le système d'erreur $\tilde{X}_{hg}$ précédent où l'incertitude se limite à la condition initiale. Sur la condition initiale « Pressure I.C. », nous montrons Figure 3.17, à droite, la décroissance de l'énergie en fonction du choix de γ. Nous voyons que dans ce cas c'est $\gamma = 2$ qui est optimal car il permet une meilleur taux de décroissance sur l'ensemble des composantes spectrales contenues dans la condition initiale. Cependant si la condition initiale contient les premiers modes de la structure stabilisée (« Modal I.C. ») alors on constate que $\gamma = 0.9$ permet un meilleur taux de décroissance (courbes de gauche de la Figure 3.17). Ceci illustre parfaitement ce que nous avions déjà anticipé par le spectre.

Quantifions maintenant numériquement la proposition 3.4.2. Pour cela nous imaginons que la condition initiale cible est suffisamment régulière pour se décomposer sur $q = 20$ modes de la structure non amortie. Nous observons alors l'évolution de la distance de sous-espaces avec jusqu'à $q' = 50$ modes de la structure stabilisée pour différentes valeurs de γ. Pour réaliser ce calcul numérique, nous utilisons les matrices grammiennes des modes

$$S_{\Phi\Phi} = (< \Phi_i, \Phi_j >_{\mathcal{E}})_{i,j}, \quad S_{\Psi\Psi} = (< \Psi_i, \Psi_j >_{\mathcal{E}})_{i,j}, \quad S_{\Phi\Psi} = (< \Phi_i, \Psi_j >_{\mathcal{E}})_{i,j} = S_{\Psi\Phi}^T.$$

On rappelle que ces produits scalaires se calculent par exemple comme $\Psi_i^T N \Phi_j$ pour les vecteurs de degrés de liberté. Le carré de la distance est alors égal à :

$$d^2(q,q') = \sup_{V \in \mathrm{span}(\Phi_i)_{i=1}^q} \inf_{W \in \mathrm{span}(\Psi_i)_{i=1}^{q'}} \frac{\|V-W\|_{\mathcal{E}}^2}{\|V\|_{\mathcal{E}}^2},$$

$$= \sup_{V_\alpha} \inf_{W_\alpha} \frac{V_\alpha S_{\Phi\Phi}\, V_\alpha^T + W_\alpha S_{\Psi\Psi} W_\alpha^T - 2V_\alpha S_{\Phi\Psi} W_\alpha^T}{V_\alpha S_{\Phi\Phi} V_\alpha^T}.$$

Le terme au numérateur est minimisé sur W_α par la projection (il suffit de dériver la forme quadratique)

$$W_\alpha = (S_{\Psi\Psi})^{-1} S_{\Psi\Phi} V_\alpha,$$

donnant

$$d^2(q,q') = \sup_{V_\alpha} \frac{V_\alpha (S_{\Phi\Phi} - S_{\Phi\Psi} S_{\Psi\Psi}^{-1} S_{\Psi\Phi}) V_\alpha^T}{V_\alpha S_{\Phi\Phi} V_\alpha^T}.$$

On reconnaît ici un quotient de Rayleigh, donc la distance relative au carré est simplement la plus grande valeur propre du problème aux valeurs propres généralisées

$$(S_{\Phi\Phi} - S_{\Phi\Psi} S_{\Psi\Psi}^{-1} S_{\Psi\Phi}) V = \lambda S_{\Phi\Phi} V.$$

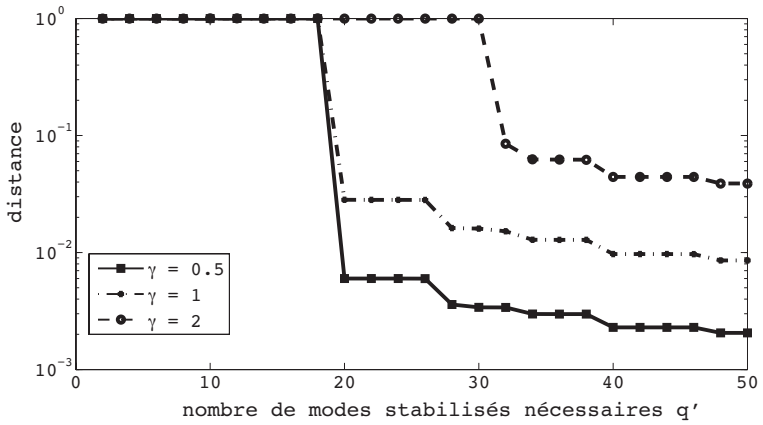


FIGURE 3.18: Carré de la distance de sous-espace entre 20 modes initiaux et jusqu'à 50 modes stabilisés pour différentes valeurs de γ.

La Figure 3.18 montre les résultats numériques du calcul de distance pour le maillage « desired mesh ». La distance décroît avec un net décrochage après les 20 premiers modes justifiant l'appariement des modes entre eux. La valeur propre du mode se déplace donc d'un pôle à un zéro de la fonction de transfert mais le mode est structurellement toujours le même. Le résultat est légèrement détérioré quand γ augmente mais il n'est pas sûr que le problème ne soit pas uniquement algorithmique. En effet, il est toujours délicat d'ordonner du point de vue numérique des modes. De plus, les méthodes de calcul de valeurs propres classiques se limitent à celles de plus grand ou plus petit module. Or quand la stabilisation est importante les valeurs propres de plus petit module ne sont plus nécessairement celles de plus petite partie imaginaire ce qui complique encore la classification. Le résultat global reste cependant tout à fait suffisant dans notre cas, car numériquement on contrôle largement le spectre des 100 premiers modes stabilisés pour les γ sélectionnés.

b. Erreur de mesures

L'aspect asymptotiquement stable (sur l'ensemble du spectre) reste cependant fondamental dans l'analyse de l'influence des erreurs de mesure sur le système. En effet le membre de droite dans la dynamique contient le terme $K\chi$, or le contenu fréquentiel de χ est très large notamment haute fréquence. Donc il est nécessaire que la dynamique stabilise ces hautes fréquences. Cependant, en fonction de l'ordre de grandeur du bruit de mesures, on peut considérer avoir moins besoin d'une estimation quantitative exacte du taux de décroissance pour la stabilisation de ces quantités. On se place donc dans le cas où la résolvante (extension possible au semi-goupe) $\mathcal{T}^h(t)$ générée par l'opérateur $A - K_X H$ vérifie

$$\|\mathcal{T}^h(t)\| \leq Ce^{-\delta t},$$

avec plusieurs solutions envisageables

$$\begin{cases} \delta \,.\, T >> 1 & \text{si exponentielle stabilité} \\ \delta = 0 & \text{si stabilité} \end{cases} \tag{3.48}$$

Et on se concentre sur le système à la dynamique suivante

$$\begin{cases} \dot{\tilde{X}}_\chi = (A - K_X H)\tilde{X}_\chi - K_X.\chi \\ \tilde{X}_\chi(0) = 0 \end{cases} \tag{3.49}$$

De plus, on suppose le bruit χ blanc de covariance le Dirac $\delta(t-t')W$. Alors $\tilde{X}_\chi$ est gaussien de moyenne nulle $\mathbb{E}(\tilde{X}_\chi) = 0$ et d'énergie moyenne

$$\mathbb{E}(\|\tilde{X}_\chi\|^2_{\mathcal{E}}) = \int_0^t \operatorname{tr}\big(W(\mathcal{T}^h(t-s))^T N \mathcal{T}^h(t-s)\big)\, ds.$$

D'après le choix d'opérateur de mesures et de bruit que nous avons fait Section 3.4.1 on obtient directement

$$\mathbb{E}(\|\tilde{X}_\chi\|^2_{\mathcal{E}}) \leq C_2 T_2 \gamma^2 \sum_{i=1}^{q} \alpha_i^2, \tag{3.50}$$

avec

$$T_2 = \begin{cases} 1/\delta \ll T & \text{dans le cas uniforme exponentielle stabilité} \\ T & \text{dans le cas de simple stabilité} \end{cases} \tag{3.51}$$

Comme résultat numérique nous choisissons d'insister Figure 3.19 sur l'influence du bruit de mesure en fonction de gain. Le bruit ajouté est un bruit blanc correspondant à 10% de la vitesse maximale sur le cycle pour un taux d'échantillonage de $50\ ms$. Ce bruit remis à l'échelle pour le pas de temps de la simulation ($\Delta t = 0.001s$) comme indiqué en introduction de ce chapitre est alors de 70% de la valeur de référence. On constate nettement que le gain amplifie comme prévu le bruit présent dans les mesures. Cependant, nous verrons que cette erreur reste limitée pour les gains sélectionnés par rapport aux autres contributions d'erreur, notamment celle liée à la discrétisation.

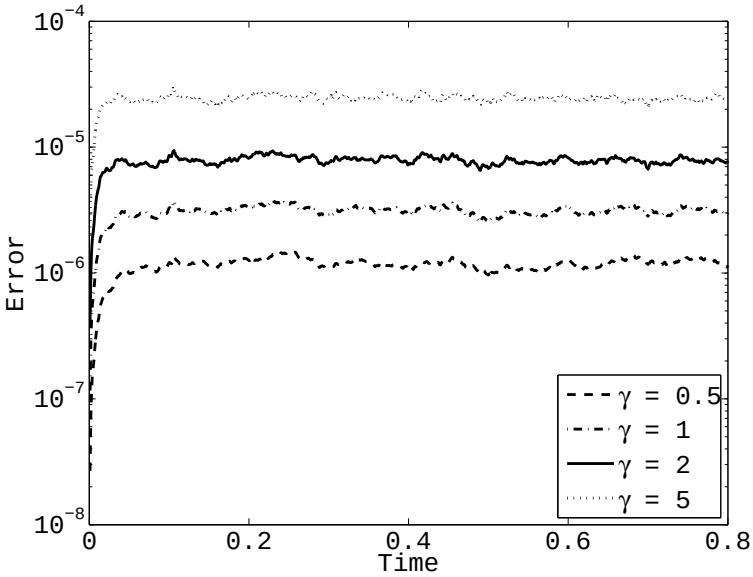


FIGURE 3.19: Erreur induite par le bruit de mesure en fonction de γ

c. Erreur de discrétisation

Il en est de même pour l'influence de la discrétisation ϵ_h dont le contenu fréquentiel peut lui aussi être très large. Là aussi, on introduit un système à la dynamique partielle

$$\begin{cases} \dot{\tilde{X}}_d = (A - K_X H)\tilde{X}_d - K_X \epsilon_h \\ \tilde{X}_0 = 0 \end{cases} \tag{3.52}$$

et on utilise la formule de variation de la constante (ou de Duhamel)

$$\tilde{X}_d(t) = \int_0^t \mathcal{T}^h(t-s) K_X \epsilon_h(s)\, ds,$$

dont on déduit une estimation

$$\|\tilde{X}_d(t)\|_{\mathcal{E}} \leq C\gamma\sqrt{T_2}\|x - x_h\|_{L^2([0,T];\mathcal{E})}.$$

Au niveau numérique, l'évolution de l'erreur de filtrage en fonction de la discrétisation est plus caractéristique pour le système complet. Nous renvoyons donc à la Figure 3.21 pour l'évolution de l'erreur totale en fonction du choix du maillage.

REMARQUE 3.4.1 (ECHANTILLONNAGE DE LA MESURE)
Un autre terme exactement de même nature est aussi à considérer en toute rigueur dans cette estimation et concerne la gestion de l'échantillonnage temporel des mesures. Ainsi ϵ_h ne comporte pas simplement l'erreur de discrétisation spatiale vue à travers l'observation mais aussi une erreur $\epsilon_{\Delta T}$ de l'ordre de ΔT car Z est le produit d'une interpolation à tout temps t de données initialement échantillonnées tous les ΔT.

d. Estimation finale

L'ensemble des contributions partielles que nous venons de considérer :

- erreur de condition initiale
- seuillage haute fréquence
- bruit de mesures
- erreur de discrétisation

décomposent la dynamique de l'erreur globale. Une fois ajoutées, on obtient l'estimation d'erreur globale pour notre estimateur d'état (en énergie moyenne)

$$\mathbb{E}(\|\tilde{X}\|_{\mathcal{E}}^2) \leq C\Big[e^{-2\delta_1 t}\mathbb{E}(\|\zeta^X\|_{\mathcal{E}}^2) + T_2\gamma^2\|x - x_h\|_{L^2([0,T];\mathcal{E})}^2 + T_2\gamma^2\sum_{i=1}^{q}\alpha_i^2 + \epsilon^2\Big]. \tag{3.53}$$

Figure 3.21, nous présentons les très bons résultats de notre estimation. En effet, nous lui associons l'erreur d'interpolation et de discrétisation pour le problème sans erreur de condition initiale, c'est-à-dire les situations optimales classiques en analyse numérique, sans incertitudes mais aussi sans mesure. On rappelle que l'erreur de discrétisation correspond à l'erreur entre la solution du problème de référence et la solution simulée avec le maillage désiré. Cette erreur est calculée en interpolant la solution grossière sur le système de référence. L'erreur d'interpolation est, elle, la solution du problème de référence directement projetée sur le maillage désiré puis réinterpolée sur le maillage de référence. C'est en quelque sorte l'étalon des erreurs numériques les plus petites que l'on puisse espérer.

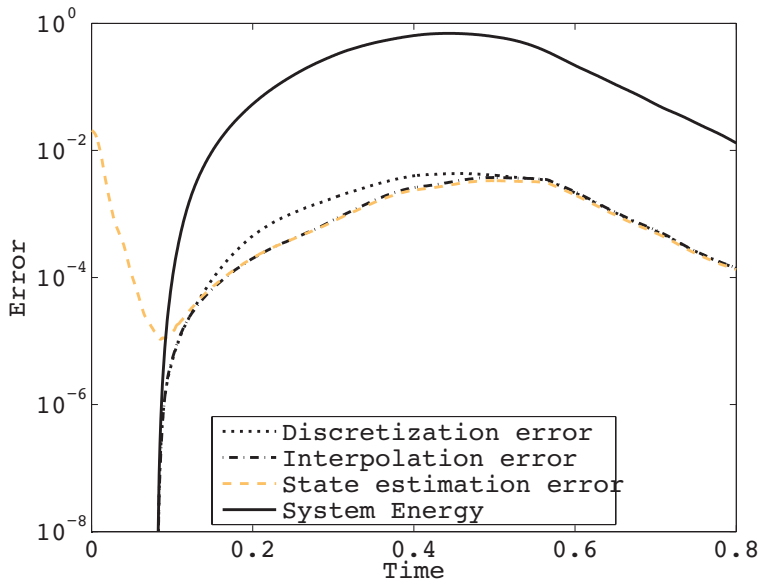


FIGURE 3.20: Erreur globale de l'estimateur et comparaisons avec l'erreur de discrétisation ou d'interpolation pour le problème d'analyse numérique considéré

Or on observe que cette erreur d'interpolation est rattrapée en fin de période d'estimation par l'erreur globale de notre estimateur. Cette dernière est donc meilleure que l'erreur de discrétisation dans le cas d'une connaissance parfaite de la condition initiale. Ce résultat indique qu'il est préférable, même avec une bonne condition initiale, d'utiliser les mesures disponibles

si le bruit est raisonnable car celles-ci améliorent la qualité de l'approximation. Nous interprétons ce résultat par le fait que le système est ici nourri par le contenu fréquentiel du système réel et ne se limite pas à sa propre approximation des modes réels. On rattrape ainsi les classiques erreurs de déphasage de l'approximation numérique classique. Cette remarque laisse enfin penser que notre inégalité triangulaire

$$\|x - \hat{x}_h\| \leq \|x - x_h\| + \|x_h - \hat{x}_h\|.$$

n'est pas optimale et nous renvoyons aux modifications apportées section suivante, dans le cas de l'opérateur frontière, qui, nous imposera par ailleurs de passer par un autre système que l'approximation x_h ordinaire.

Enfin, nous voyons Figure 3.21 que cette estimation est très largement robuste au choix de γ ou de la discrétisation, ce qui était dans les propriétés fondamentales de notre cahier des charges.

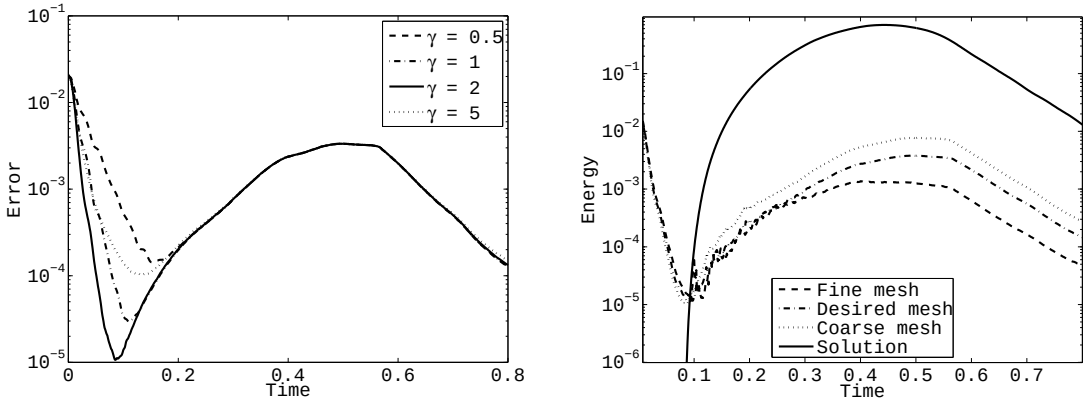


FIGURE 3.21: Erreur d'estimation en fonction du choix de γ (gauche) ou de la discrétisation (droite) pour la condition initiale « Pressure I.C. »

3.4.6 Observateur frontière

Nous venons de présenter l'ensemble des résultats d'analyse de l'estimateur pour des mesures de types cellules qui représentent des données comparables à des tags (même s'il est vrai qu'elles sont pour l'instant un peu artificiellement définies pour donner des vitesses). Tout en conservant des vitesses, on pourrait imaginer disposer uniquement d'observations frontières si par exemple nous ne faisions confiance qu'à ce type de mesures en sortie du flot optique. Dans ce cas, l'opérateur d'observation peut être modélisé sous la forme

$$\mathcal{H}\underline{y} = \underline{y}|_{\partial\Gamma_{\mathfrak{m}}}.$$

où $\Gamma_{\mathfrak{m}}$ est la partie de la surface où l'on dispose des mesures. Cette situation est idéalisée car, en toute rigueur, les mesures disponibles, y compris sur la surface, font intervenir des cellules d'une certaine extension. Ici on suppose disposer de tout le champ sur $\Gamma_{\mathfrak{m}}$. Ce choix est motivé par des questions théoriques que nous appréhenderons plus particulièrement au chapitre suivant et correspond finalement à une sorte de cas d'observation limite. Voici quelques remarques qui motivent son introduction.

a. Espaces de définition de H

Pour l'instant, on a toujours supposé que la dimension de l'espace d'observation était inférieure à celle de l'état. En pratique c'est effectivement souvent le cas mais pas dans ce cas limite. En effet

$$H = \mathcal{H}\mathcal{I}_h : \mathbb{R}^N \to L^2(\Gamma_{\mathfrak{m}}),$$

et donc notamment H n'est plus surjectif.

En plus de cet inconvénient pratique, un opérateur frontière présente les difficultés théoriques déjà mentionnées à la Section 3.3.3 et synthétisées dans Lasiecka et Triggiani (2003). Plus largement, le contrôle frontière a toujours donné lieu à des résultats spécifiques car les opérateurs frontières agissent sur les EDP via leurs conditions aux limites (il n'y a plus d'opérateur de contrôle B pour le contrôle direct sur l'EDP) et font donc intervenir des espaces de solutions spécifiques.

b. Bruit blanc et opérateur frontière

Dans le cas d'un opérateur frontière, il est en fait impossible de considérer un bruit blanc comme source d'erreur de mesures. Afin d'illustrer cette remarque, reprenons le système le plus défavorable, c'est-à-dire celui non amorti

$$M\ddot{Y} + KY = \dot{F},$$

où F représente un champ régulier en espace, soit volumique soit surfacique, mais correspondant à un processus de Wiener en temps de telle sorte que $\dot{F}$ est un bruit blanc. A noter que nous restons dans le cadre dimension finie en espace mais le même type d'argument fonctionne formellement en dimension infinie avec les variables cylindriques introduits par Bensoussan (1971). Le système se réécrit sous forme du premier ordre

$$\dot{X} = AX + \dot{R},$$

avec toujours

$$A = \begin{pmatrix} 0 & I \\ -M^{-1}K & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ M^{-1}F \end{pmatrix}.$$

On note comme précédemment le semi-groupe $\mathcal{T}^h$, avec la formule de Duhamel

$$X(t) = \mathcal{T}_h X(0) + \int_0^t \mathcal{T}^h(t-s)\dot{R}(s)\,ds,$$

et on calcule l'énergie de la solution au cours du temps en moyenne

$$\begin{aligned}\mathbb{E}(\|X\|_{\mathcal{E}}^2) &= \mathbb{E}(X^T N X) \\ &= X(0)^T \mathcal{T}^h(t)^T N \mathcal{T}^h(t) X(0) + \int_0^t \int_0^t \mathbb{E}\big(\dot{R}(\tau)^T \mathcal{T}^h(t-\tau)^T N \mathcal{T}^h(t-s)\dot{R}(s)\big)\,d\tau ds,\end{aligned}$$

le système étant conservatif, le premier terme donne

$$X(0)^T \mathcal{T}^h(t)^T N \mathcal{T}^h(t) X(0) = \|\mathcal{T}^h(t)X(0)\|_{\mathcal{E}}^2 = \|X(0)\|_{\mathcal{E}}^2.$$

Soit alors W la matrice de covariance du processus de Wiener R, on sait que $\mathbb{E}\big(\dot{R}(s)\dot{R}(\tau)^T\big) = W\,\delta(t-s)$ qui combiné avec la relation

$$\dot{R}(\tau)^T\mathcal{T}^h(t-\tau)^T N\mathcal{T}^h(t-s)\dot{R}(s) = \operatorname{tr}\big(\dot{R}(s)\dot{R}(\tau)^T\mathcal{T}^h(t-\tau)^T N\mathcal{T}^h(t-s)\big),$$

donne

$$\begin{aligned}
\mathbb{E}(\|X\|_{\mathcal{E}}^2) &= \|X(0)\|_{\mathcal{E}}^2 + \int_0^t \operatorname{tr}\big(W\mathcal{T}^h(t-\tau)^T N\mathcal{T}^h(t-\tau)\big)\,d\tau,\\
&= \|X(0)\|_{\mathcal{E}}^2 + \int_0^t \mathbb{E}\big(R(\tau)^T\mathcal{T}^h(t-\tau)^T N\mathcal{T}^h(t-\tau)R(\tau)\big)\,d\tau,\\
&= \|X(0)\|_{\mathcal{E}}^2 + \int_0^t \mathbb{E}\big(R(\tau)^T N R(\tau)\big)\,d\tau,\\
&= \|X(0)\|_{\mathcal{E}}^2 + \int_0^t \tau\operatorname{tr}(WN)\,d\tau,\\
&= \|X(0)\|_{\mathcal{E}}^2 + \frac{t^2}{2}\operatorname{tr}(WN).
\end{aligned}$$

Or

$$t\operatorname{tr}(WN) = \mathbb{E}\big(F^T M^{-1} F\big), \tag{3.54}$$

donc l'énergie est bornée en dimension infinie (ou uniformément bornée par rapport à la discrétisation en dimension finie) si et seulement si F est un champ volumique appartenant au dual de $L^2(\Omega)$. Ainsi un bruit blanc concentré sur la surface associée à un opérateur d'observation surfacique est théoriquement inadmissible du point de vue énergétique. Tous les avantages du bruit blanc du point de vue formel pour la formulation probabiliste du filtre et les estimations sont donc caducs. Nous verrons au chapitre suivant, Section 4.2.7, comment on peut contourner cette difficulté théorique par l'approche déterministe et le *contrôle robuste* où on dispose de plus de liberté sur les espaces dans lesquels sont définis les bruits.

c. Opérateur de feedback frontière

La définition de l'opérateur frontière est identique à ce que nous avons vu pour les observations sous forme de cellule. L'avantage, en revanche, est que sa structure est plus immédiate et démontre pourquoi ces filtres sont très avantageux en pratique dans un code élément-finis classique à la différence des opérateurs de Kalman ou autres méthodes variationnelles. Le DVF consiste à appliquer une force opposée à la vitesse mesurée au point de la mesure, donc le terme $K(Z-HX)$ s'écrit variationnellement

$$\gamma\int_{\Gamma_{\mathrm{m}}} s(\underline{\xi})(\underline{\dot{y}} - \underline{\dot{\hat{y}}})\cdot\underline{v}\,dS,$$

avec γ un gain et s une éventuelle fonction poids (que nous prendrons égale à 1) et où $\underline{\dot{y}}$ est la vitesse mesurée sur le système réel. Ainsi le feedback est

$$\gamma\int_{\Gamma_{\mathrm{m}}} s(\underline{\xi})\underline{\dot{\hat{y}}}\cdot\underline{v}\,dS,$$

c'est-à-dire un terme identique à celui introduit pour les conditions aux limites Section 2.3.5. Par conséquent, la mise en place de ce type de contrôle dans un code éléments finis ne requiert

pas de développement important et suit la même logique de développement que d'autres termes variationnels utilisés dans la simulation numérique. En plus de sa complexité, la mise en oeuvre de ces filtres est donc un atout majeur par rapport aux filtres de Kalman ou même à des formulations à partir de l'adjoint qui nécessitent, en plus de l'adjoint, des méthodes de descente de gradient performantes.

Pour conserver notre structure habituelle nous définissons l'adjoint $(H^v)'$ par rapport à la norme L^2 pondérée sur $\Gamma_{\mathfrak{m}}$:

$$\|\underline{\dot{y}}\|_{\mathfrak{m}}^2 = \int_{\Gamma_{\mathfrak{m}}} s(\xi)\, |\underline{\dot{y}}|^2 \, dS,$$

on peut alors écrire sous forme matricielle

$$K_X = \begin{pmatrix} 0 \\ -\gamma M^{-1}(H^v)' \end{pmatrix},$$

et l'opérateur de viscosité supplémentaire est alors

$$\gamma (H^v)' H^v = \gamma (T_{\mathfrak{m}})^T \mathfrak{M} T_{\mathfrak{m}},$$

où $\mathfrak{M}$ est la matrice associée à la norme $\|\cdot\|_{\mathfrak{m}}$ et $T_{\mathfrak{m}}$ l'interpolateur sur la frontière. Nous utilisons la notation $\mathfrak{M}$ pour rappeler que cette matrice est un forme de matrice de masse sur la frontière correspondant aux mesures. Par rapport aux autres présentations faisant apparaître une matrice de covariance sur des mesures discrètes en espace, $\mathfrak{M}$ est homogène à W^{-1}.

d. Analyse de l'erreur d'estimation et espace d'erreur

L'estimation d'erreur repose sur les mêmes idées maîtresses que précédemment. Cependant la structure frontière de l'opérateur nous a conduit à modifier le système de référence afin de bénéficier de la régularité suffisante pour manipuler les termes frontières. On introduit donc le système de référence modifié

$$\begin{cases} \dot{\bar{X}} = A\bar{X} + R + K_X(\bar{Z} - H\bar{X}) \\ X(0) = X_0 + \zeta^X \end{cases} \tag{3.55}$$

où $\bar{Z} = \mathcal{H}x$ est la mesure parfaite. Ce système n'est pas calculable en pratique mais il nous sert de système intermédiaire dans le calcul de l'erreur. En effet, ce dernier nous permet de contrôler l'erreur de discrétisation vue au travers de l'erreur d'observation $\epsilon_h = \mathcal{H}(x - \bar{x}_h)$. Avec un tel système, l'estimation d'état peut alors être contrôlée pour une norme plus riche

$$\|\tilde{X}\|_{\mathcal{E}'} = \frac{1}{2}\tilde{X} N' \tilde{X},$$

avec

$$N' = N + 2H'H = \begin{pmatrix} K & 0 \\ 0 & M + 2(T_{\mathfrak{m}})^T \mathfrak{M} T_{\mathfrak{m}} \end{pmatrix}.$$

□ Contrôle de l'erreur de discrétisation – Nous présentons l'analyse numérique nécessaire pour considérer le système ci-dessus. Cette présentation se limite au cas où le système réel n'a pas de viscosité interne afin de simplifier la présentation, mais la considérer ne

modifierait pas le principe des estimations obtenues. Nous rappelons que les formulations variationnelles satisfaites par le système réel et son approximation élément finis vérifient

$$\int_\Omega \rho \ddot{\underline{y}} \cdot \delta \underline{y} \, d\Omega + \int_\Omega \underline{\underline{\epsilon}}(\underline{y}) : \underline{\underline{\underline{\underline{A}}}} : \delta \underline{\underline{\epsilon}} \, d\Omega = \int_\Omega \underline{f} \cdot \delta \underline{y} \, d\Omega, \quad \forall \delta \underline{y}, \tag{3.56}$$

$$\int_\Omega \rho \ddot{\underline{y}}_h \cdot \delta \underline{y}_h \, d\Omega + \int_\Omega \underline{\underline{\epsilon}}(\underline{y}_h) : \underline{\underline{\underline{\underline{A}}}} : \delta \underline{\underline{\epsilon}} \, d\Omega = \int_\Omega \underline{f} \cdot \delta \underline{y}_h \, d\Omega + \gamma \big(\mathcal{H}^v(\dot{\underline{y}} - \dot{\underline{y}}_h), \mathcal{H}^v(\delta \underline{y}_h) \big)_{\mathfrak{m}}, \quad \forall \delta \underline{y}_h. \tag{3.57}$$

Où nous notons $(\cdot,\cdot)_{\mathfrak{m}}$ le produit scalaire associé à la norme d'observation.

On définit l'opérateur de projection Π_h entre les espaces de déplacements continus et discrets pour le produit scalaire associé à la raideur. En particulier

$$\int_\Omega \underline{\underline{\epsilon}}\big(\Pi_h \underline{y}\big) : \underline{\underline{\underline{\underline{A}}}} : \delta \underline{\underline{\epsilon}} \, d\Omega = \int_\Omega \underline{\underline{\epsilon}}(\underline{y}) : \underline{\underline{\underline{\underline{A}}}} : \delta \underline{\underline{\epsilon}} \, d\Omega, \quad \forall \delta \underline{y}_h. \tag{3.58}$$

Soustrayant (3.57) de (3.56) et en choisissant $(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)$ comme fonction test, on obtient alors

$$\begin{aligned} \frac{d}{dt} \|\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\|_{\mathrm{In}}^2 + \frac{d}{dt} \|\Pi_h \underline{y} - \underline{y}_h\|_{\mathrm{St}}^2 + \gamma \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)\|_{\mathfrak{m}}^2 \\ = 2\big(\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}, \Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\big)_{\mathrm{In}} + \gamma \big(\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}), \mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h) \big)_{\mathfrak{m}}, \end{aligned} \tag{3.59}$$

où $(\cdot,\cdot)_{\mathrm{In}}$ et $(\cdot,\cdot)_{\mathrm{St}}$ sont les produits scalaires pour les formes bilinéaires d'inertie et de raideur, avec $\|\cdot\|_{\mathrm{In}}$ et $\|\cdot\|_{\mathrm{St}}$ les normes associées. D'après l'inégalité classique

$$|\big(\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}), \mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)\big)_{\mathfrak{m}}| \leq \frac{1}{2}\big(\|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_{\mathfrak{m}}^2 + \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)\|_{\mathfrak{m}}^2\big),$$

on a

$$\begin{aligned} \frac{d}{dt} \|\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\|_{\mathrm{In}}^2 + \frac{d}{dt} \|\Pi_h \underline{y} - \underline{y}_h\|_{\mathrm{St}}^2 + \frac{\gamma}{2} \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)\|_{\mathfrak{m}}^2 \\ \leq 2\big(\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}, \Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\big)_{\mathrm{In}} + \frac{\gamma}{2} \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_{\mathfrak{m}}^2, \end{aligned} \tag{3.60}$$

qui implique

$$\frac{d\mathcal{E}_h}{dt} \leq 2\big(\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}, \Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\big)_{\mathrm{In}} + \frac{\gamma}{2} \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_O^2,$$

avec

$$\mathcal{E}_h(t) = \|\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\|_{\mathrm{In}}^2 + \|\Pi_h \underline{y} - \underline{y}_h\|_{\mathrm{St}}^2,$$

et nous pouvons utiliser le lemme de Gronwall pour obtenir

$$\mathcal{E}_h(t)^{\frac{1}{2}} \leq \left\{ \mathcal{E}_h(0) + \int_0^t \frac{\gamma}{2} \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_{\mathfrak{m}}^2 \, d\tau \right\}^{\frac{1}{2}} + \int_0^t \|\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}\|_{\mathrm{In}} \, d\tau. \tag{3.61}$$

Ainsi,

$$\mathcal{E}_h(t) \leq 2\mathcal{E}_h(0) + \int_0^t \gamma \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_{\mathfrak{m}}^2 \, d\tau + 2t \int_0^t \|\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}\|_{\mathrm{In}}^2 \, d\tau, \tag{3.62}$$

et en utilisant (3.60) nous avons aussi l'erreur d'observation

$$\gamma \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h)\|_{\mathfrak{m}}^2 \leq C \left\{ \gamma \|\mathcal{H}^v(\Pi_h \dot{\underline{y}} - \dot{\underline{y}})\|_{\mathfrak{m}}^2 + \|\Pi_h \ddot{\underline{y}} - \ddot{\underline{y}}\|_{\mathrm{In}}^2 + \|\Pi_h \dot{\underline{y}} - \dot{\underline{y}}_h\|_{\mathrm{In}}^2 \right\}. \tag{3.63}$$

Bien sûr, afin d'obtenir des estimations similaires pour $(\underline{y}-\underline{y}_h)$, nous devons borner les erreurs de projection $(\underline{y}-\Pi_h\underline{y})$ dans les mêmes normes. Puisque les projections sont définies comme la solution discrète d'un problème standard d'élasticité (cf. (3.58)), ce type de borne est classique et dépend uniquement de la régularité du domaine et du chargement. Donc finalement ϵ_h peut être borné en utilisant (3.63) et nous avons

$$\|\epsilon_h\|_{W^{-1}} \leq C\|\epsilon_h\|_{\mathfrak{m}} \leq C\Big(\|\mathcal{H}^v(\Pi_h\underline{\dot{y}} - \underline{\dot{y}}_h)\|_{\mathfrak{m}} + \|\mathcal{H}^v(\Pi_h\underline{\dot{y}} - \underline{\dot{y}})\|_{\mathfrak{m}}\Big). \tag{3.64}$$

□ RÉSULTATS DE CONTRÔLE FRONTIÈRE – Même si le système cible change, le système de l'erreur suit toujours une dynamique stabilisée qui oriente son analyse vers les résultats de contrôle frontière. Ceux-ci, comme évoqués, sont largement distincts de ceux pour un contrôle réparti mais existent dans la configuration qui nous intéresse. En particulier Cox et Zuazua (1995) réalisent l'extension pour un contrôle frontière en 1D du résultat de Cox et Zuazua (1994) (leur résultat) déjà cité pour nos observations sous formes de cellules volumiques. Ainsi, l'analyse d'erreur est la même que pour les mesures de tags et passe par le QEP.

Nous présentons ainsi Figure 3.22 le comportement du spectre dans le cas d'un observateur frontière sur tout l'endocarde et l'épicarde, ou simplement sur tout l'endocarde. Cette deuxième situation est motivée par les situations où l'imagerie et la segmentation ne permettent pas le suivi de l'épicarde alors qu'elles sont de bonne qualité pour l'endocarde en raison du fort contraste sur les images entre les tissus et le sang contenu dans les cavités.

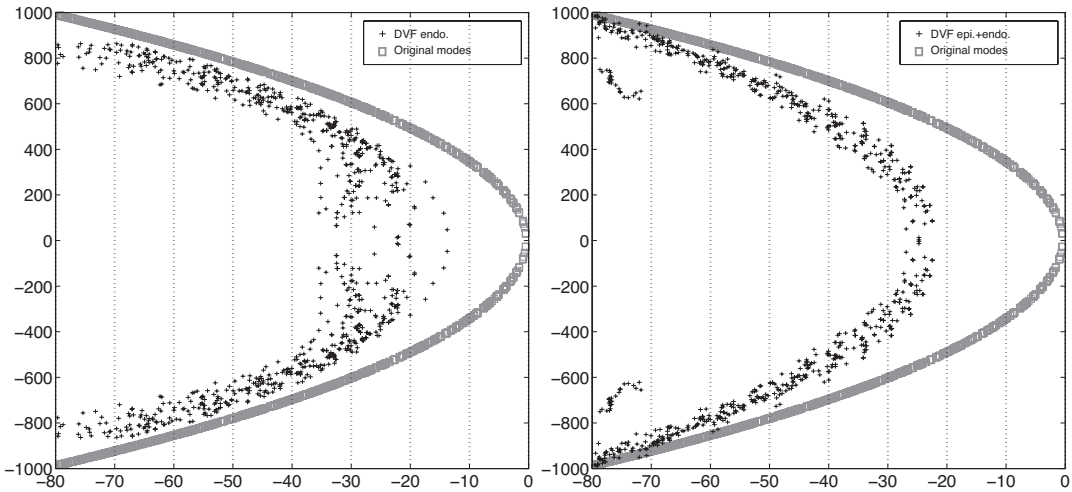


FIGURE 3.22: QEP pour le filtre DVF avec des mesures limitées à l'endocarde (gauche) ou sur les deux frontières (droite). Le gain a été optimisé dans les deux cas.

Nous passons alors directement à l'estimation finale où nous comparons Figure 3.23 l'erreur d'estimation à l'erreur d'interpolation et de discrétisation. Les résultats sont très positifs et le paragraphe précédent nous fait prendre conscience de l'intérêt de choisir le système cible dans l'analyse numérique. En effet l'erreur d'interpolation est toujours l'étalon optimal pour analyser l'erreur, mais on peut se demander ce qu'il en est de l'erreur de discrétisation, notamment en fonction du système de référence. En effet, nous avions choisi jusque là le système numérique correspondant simplement à une discrétisation classique du problème de référence,

sans erreur mais sans prendre en compte non plus les mesures. Or, tout indique ici qu'il est nécessaire d'introduire les mesures dans le système discrétisé afin que l'erreur de discrétisation soit dans les bons espaces. Mais ce système a d'autres avantages. En effet, puisqu'il est alimenté par les mesures, celui-ci est forcé par les observations contenant exactement le contenu fréquentiel du système de référence. Or, en analyse numérique standard, on sait que les fréquences sont sur-évaluées, la structure discrète étant plus raide. Ce résultat se retrouve en effet très simplement à partir des quotients de Rayleigh où le minimum est pris sur un espace de plus en plus grand avec la diminution du pas de discrétisation en espace. On obtient alors, pour des problèmes d'évolution en temps, des trajectoires déphasées entre le système réel et sa discrétisation classique. Même si la période ne permet pas de l'observer sur les figures, l'erreur de discrétisation peut alors être très importante. Au contraire en choisissant comme système discrétisé un système ayant intégré les mesures, le forçage permet de rattraper ce déphasage. Autrement dit notre estimateur est plus performant qu'une discrétisation classique car il intègre les mesures et donc le contenu fréquentiel réel. Le système cible intermédiaire permettant l'analyse numérique dès l'introduction, via une simple inégalité triangulaire, devrait donc, dans tous les cas, intégrer lui aussi les mesures pour que l'estimation finale soit la plus fine possible.

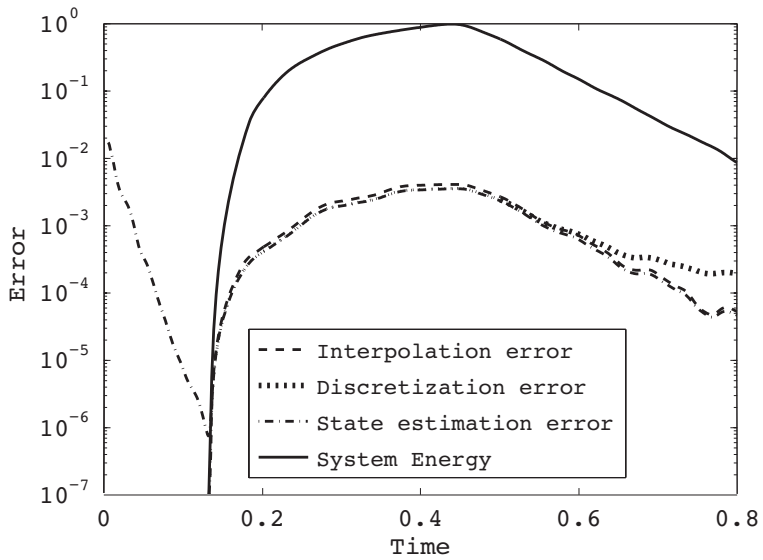


FIGURE 3.23: Erreur finale d'estimation pour notre problème-test dans le cas de mesures à l'endocarde. Comparaison avec l'erreur d'interpolation et de discrétisation

3.4.7 Extension en dynamique non-linéaire

a. Formulation du problème et de l'estimateur

Finalement nous proposons de tester nos estimateurs dans le cadre de la visco-élasticité non-linéaire. Afin d'évaluer l'estimateur, nous commençons par un modèle mécanique qui est simplement une extension en grands déplacements du précedent. Nous généralisons donc le système présenté Section 3.4.1 à

- un matériau hyperélastique de type Ciarlet-Geymonat

$$W^e(\underline{\underline{e}}) = \kappa_1(J_1 - 3) + \kappa_2(J_2 - 3) + \kappa(J - 1) - \kappa \ln J,$$

- un terme visqueux de type pseudopotentiel

$$W^v(\underline{\underline{e}}) = \eta Tr(\dot{\underline{\underline{e}}}^2),$$

- un terme de précontrainte active

$$\delta\mathcal{W}^{PS} = \int_\Omega \sigma_0 w(\underline{x}, t)\ \mathrm{tr}(\ \mathrm{d}_{\underline{y}}\underline{\underline{e}} \cdot \underline{v})\, d\Omega.$$

On s'approche donc progressivement du modèle cardiaque, avec pour les exemples numériques les mêmes valeurs de κ_1, κ_2, κ mais ici on regarde $0.68Pa.s \leq \eta \leq 6.8Pa.s$. Pour la contractilité, celle-ci est augmentée à $\sigma_0 = 4.5\ 10^4$Pa en raison des non-linéarités du système afin de conserver des volumes éjectés physiologiques. Le système non-linéaire discrétisé en espace est donc

$$M\ddot{Y} + C(Y)\dot{Y} + F^i(Y) = F^{PS}(Y, t), \tag{3.65}$$

et l'estimateur proposé consiste à garder le même type de feedback linéaire

$$M\ddot{\hat{Y}} + C(\hat{Y})\dot{\hat{Y}} + F^i(\hat{Y}) = F^{PS}(\hat{Y}, t) + \gamma(H^v)'(Z - H^v\dot{\hat{Y}}). \tag{3.66}$$

L'étude de cet estimateur est beaucoup plus délicate. Tout d'abord, on sait que la dynamique de l'estimateur (y compris dans sa discrétisation) n'a pas été déstabilisée par l'ajout du feedback puisqu'il correspond toujours exactement à un opérateur de dissipation. Nous considérons cette propriété comme un gage au moins de robustesse. Cependant, comme nous l'avons déjà largement vu, c'est le système de l'erreur dont on doit assurer la convergence. Or l'erreur vérifie

$$M\ddot{\tilde{Y}} + (C(Y)\dot{Y} - C(\hat{Y})\dot{\hat{Y}}) + \gamma(H^v)'H^v\dot{\tilde{Y}} + F^i(Y) - F^i(\hat{Y}) = F^{PS}(Y, t) - F^{PS}(\hat{Y}, t) + \gamma(H^v)'Z.$$

Ce système peut être réécrit en ne faisant intervenir que la dynamique réelle (supposé régulière, bornée etc...) et celle de l'erreur. Par exemple dans le cas de l'opérateur de type cellules, on a alors un système couplé sur les variables X et $\tilde{X}$

$$\begin{cases} M\ddot{Y} + C(Y)\dot{Y} + F^i(Y) = F^{PS}(Y, t) \\ M\ddot{\tilde{Y}} + \Big(C(Y)\dot{Y} - C(Y - \tilde{Y})(\dot{Y} - \dot{\tilde{Y}})\Big) + \gamma(H^v)'H^v\tilde{Y} \\ \qquad + \Big(F^i(Y) - F^i(Y - \tilde{Y})\Big) = \Big(F^{PS}(Y, t) - F^{PS}(Y - \tilde{Y}, t)\Big) \end{cases} \tag{3.67}$$

La preuve de la stabilité de ce système n'est obtenue que localement à travers l'étude du linéarisé. Comme déjà exprimé dans un contexte très général Section 3.3.1.b., on doit choisir le système sur lequel on définit l'équilibre et le linéarisé. Si on linéarise le système (3.67) dans son ensemble, alors l'équilibre à considérer est $(X, \tilde{X}) = (0, 0)$, ce qui signifie que le système réel est analysé pour des petits déplacements. Or justement, une dynamique non-linéaire a été privilégiée ici justement pour s'affranchir de cette limitation. Notre idée est plutôt de considérer la deuxième équation seule et d'appliquer les propriétés de stabilité des systèmes non autonomes déjà citées. Le linéarisé de ce système est

$$M\ddot{\widetilde{\delta Y}} + (C(Y) + \gamma(H^v)'H^v)\dot{\widetilde{\delta Y}} + K(Y, \dot{Y})\widetilde{\delta Y} = 0, \tag{3.68}$$

où $K(Y)$ est la raideur tangente du système

$$K(Y,\dot{Y}) = \frac{\partial}{\partial Y}(F^i + C(\cdot)\dot{Y} - F^{PS})(Y).$$

Or même si K et C sont désormais dépendants du temps (donc l'erreur n'est plus autonome), la dynamique de l'erreur linéarisée est toujours stabilisée car tous les arguments développés précédemment sont valides tant que $K(Y)$ reste positif. Dans ce cas, le linéarisé est donc stable et il existe un voisinage de 0 pour l'erreur de condition initiale tel que la trajectoire de l'estimateur converge vers la trajectoire réelle.

Les situations où $K(Y)$ n'est pas définie positive restent une difficulté. Plaçons-nous dans le cas C indépendant de la déformation, alors « $K(Y) < 0$ » a lieu lorsque la précontrainte est une précontrainte de compression $\frac{\partial}{\partial Y}F^{PS}(\hat{X}^d) < 0$ et où σ_0 est de l'ordre du module d'Young

$$\left|\frac{\partial}{\partial Y}F^{PS}(\hat{X}^d)\right| > \left|\frac{\partial}{\partial Y}F^{i}(\hat{X}^d)\right|.$$

Cette situation est un cas de flambement sur la structure illustré Figure 3.24. Nous sommes

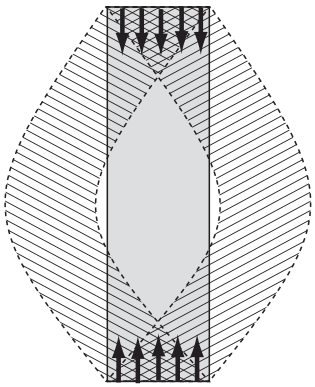


FIGURE 3.24: Exemple de flambement sur une poutre (fibre) en compression

alors dans une situation délicate de bifurcation y compris dans l'analyse numérique du système cible lui-même soumis aux même types d'efforts. Une fois le schéma en temps choisi, les opérateurs linéarisés sont en effet les opérateurs intervenant dans la résolution du Newton de la cible. Une question intéressante à traiter serait d'examiner si l'estimation favorise une sélection dans la bifurcation. En effet, si le système cible a été soumis au même type de contraintes, les mesures donnent une indication de la branche de bifurcation choisie lors du flambement. Celles-ci, injectées dans l'estimateur pourraient permettre à ce dernier de choisir la même branche que la cible.

b. Résultats numériques

Nous présentons Figure 3.25 le résultat d'estimation pour le filtre DVF obtenu sur une erreur de condition initiale en surpression. Afin de valider la robustesse de notre filtre, nous avons de plus augmenté la surpression d'un facteur 1.8 par rapport aux simulations précédentes, ce qui

se traduit par une augmentation de volume bien au delà des conditions physiologiques. Les résultats sont encore une fois très satisfaisants pour une complexité algorithmique négligeable par rapport au problème direct puisqu'il a suffi d'ajouter un opérateur classique de viscosité **linéaire**.

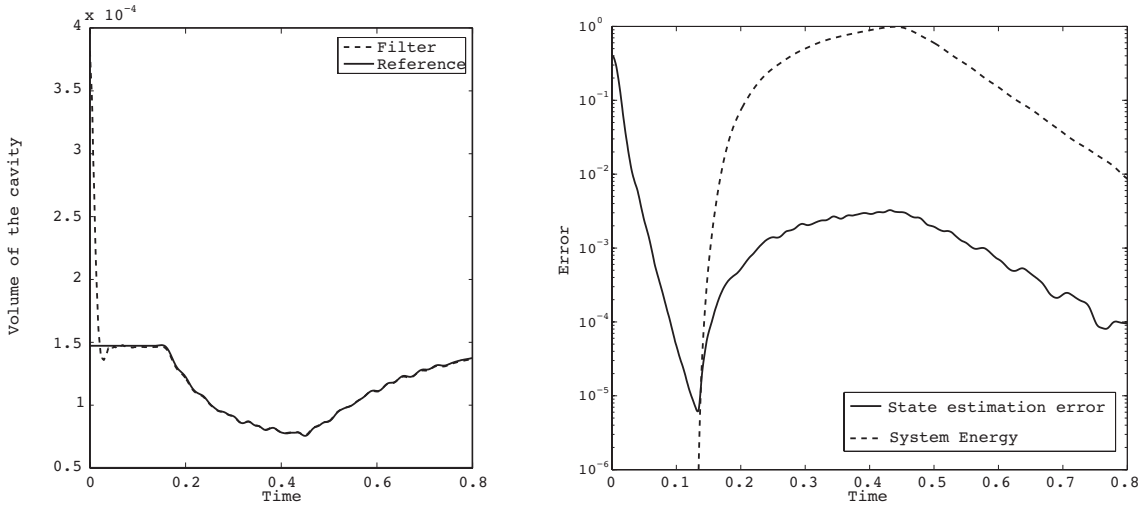


FIGURE 3.25: ***Dynamique non-linéaire :*** **Volume avec et sans filtre (gauche) et énergie de l'erreur (droite)**

3.5 Estimation d'état en élastodynamique à partir de déplacements

Dans toute cette section, nous nous concentrons désormais sur la définition d'un estimateur à partir des déplacements. Ainsi

$$Z = \mathcal{H}x + \chi, \text{ avec } \mathcal{H} = \begin{pmatrix} \mathcal{H}^d & 0 \end{pmatrix} \begin{pmatrix} \underline{y} \\ \underline{\dot{y}} \end{pmatrix}.$$

Comme dans le cas de mesures de vitesses, $\mathcal{H}^d$ peut représenter les différentes cellules de tags, ou correspondre à un opérateur frontière, mais il applique directement sur des déplacements. D'après le chapitre 2, les tags fournissent en effet plus directement des déplacements. Ainsi, les estimateurs à partir de mesures de vitesses étaient un peu artificiels, sauf à véritablement considérer des données de flot optique sur des petits déplacements, à de très fortes résolutions temporelles. Toutefois, les vitesses offraient directement à partir de contrôleurs un filtre efficace via le DVF. Il va être plus délicat de construire un opérateur stabilisant (donc d'amortissement) à partir des déplacements puisque la propriété de conjugaison énergétique entre vitesses et forces ne tient plus pour des mesures de déplacements. Des possibilités existent cependant puisque le filtre de Kalman est capable de transformer des déplacements en opérateur d'amortissement. En effet, par définition, il peut opérer sur n'importe quel type de mesures. Nous verrons comment cette idée a été une source d'inspiration pour la création de ce que nous appellerons ensuite « *Direct Displacement feedback* » (DDF) par analogie avec les DVF. Cette fin de chapitre développe ainsi les résultats contenus dans Moireau et al (2008a).

3.5.1 Définition de l'estimateur

Comme dans tout le chapitre on introduit $H = \mathcal{H} \circ \mathcal{I}_h$ tel que $Z = HX + \chi + \epsilon_h$ et on se concentre sur la définition de l'estimateur pour le système discrétisé

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + R + K_X(Z - H\hat{X}) \\ \hat{X}(0) = X_0 \end{cases}$$

L'objectif est toujours la décroissance de l'énergie de l'erreur

$$\begin{cases} \dot{\tilde{X}} = (A - K_X H)\tilde{X} - K_X(\chi + \epsilon_h) \\ \tilde{X}(0) = \zeta^X \end{cases}$$

Dans le DVF, l'aspect dissipatif du contrôle est lié, du point de vue mécanique, au fait qu'on applique, via les actionneurs, une force proportionnelle à la vitesse mesurée par les capteurs. Or les forces sont les conjugués énergétiques des vitesses en mécanique ce qui offre directement un opérateur monotone $\gamma(H^v)^T H^v$ dissipatif dans le calcul de l'énergie de l'erreur $\gamma\dot{\tilde{Y}}(H^v)^T H^v \dot{\tilde{Y}}$. Si avec des déplacements nous opérons comme en (3.30), en posant

$$K_X = \begin{pmatrix} 0 \\ \gamma M^{-1}(H^d)' \end{pmatrix},$$

on a

$$M\ddot{\hat{Y}} + C\dot{\hat{Y}} + \big(K + \gamma(H^d)'H^d\big)\hat{Y} = R + \gamma(H^d)'Z. \tag{3.69}$$

La dynamique du second ordre pour l'erreur à l'image de (3.32) devient alors

$$M\ddot{\tilde{Y}} + C\dot{\tilde{Y}} + \big(K + \gamma(H^d)'H^d\big)\tilde{Y} = 0, \tag{3.70}$$

où à K s'ajoute une nouvelle rigidité par l'opérateur monotone $(H^d)'H^d$. Dans l'absolu, cette rigidité ajoutée peut avoir pour les systèmes déjà amortis naturellement un effet sur le positionnement général des parties réelles des pôles comme le montre un simple calcul de sensibilité. En effet la rigidité ajoutée fait monter le système en fréquence, or l'amortissement naturel augmente avec la fréquence, donc le système est globalement plus amorti. Cependant ce déplacement des pôles est assez indirect (et de relative faible ampleur) et nécessite de traiter ensuite un problème aux fréquences plus élevées et donc moins physiques. Entre autres difficultés, la discrétisation en temps imposerait un pas de temps plus petit pour représenter les vibrations de la structure suivant la règle empirique de conserver une vingtaine de pas de temps pour la période du mode représenté de plus haute fréquence. Nous devons donc nous écarter de la stratégie DVF inspirée du contrôle et choisissons de bénéficier de la situation précise de notre problématique : définir **un observateur numérique**. En effet, comme notre estimateur est numérique, nous n'avons aucune restriction sur le choix et le positionnement des actionneurs. Les capteurs sont imposés par la mesure mais les actionneurs sont virtuels à tel point que nous sommes libres de modifier aussi la première équation de A qui donne initialement la vitesse du système comme la dérivée du déplacement. A l'instar de ce que la Section 3.3.3 permettait d'imaginer si on relâchait la contrainte de la faisabilité d'un tel contrôle, on définirait donc

$$K_X = \begin{pmatrix} \gamma K^{-1}(H^d)' \\ 0 \end{pmatrix}, \tag{3.71}$$

pour que, en reprenant notre notation $\hat{X} = \begin{pmatrix} \hat{X}^d \\ \hat{X}^v \end{pmatrix}$, le système devienne

$$\begin{cases} \dot{\hat{X}}^d = \hat{X}^v + \gamma K^{-1}(H^d)'(Z - H^d\hat{X}^d) \\ M\dot{\hat{X}}^v + C\hat{X}^v + K\hat{X}^d = F \end{cases} \tag{3.72}$$

Nous insistons sur cette notation déjà introduite pour le filtre de Kalman 3.2.2, puisque la première équation ne permet plus d'écrire « $\dot{\hat{X}}^d = \hat{X}^v$ ». D'après la relation (3.12), le filtre de Kalman opérait en effet lui aussi sur cette équation qui en contrôle est pourtant inviolable.

Précisons l'importance du K^{-1} au travers du calcul de l'énergie de l'erreur. En effet, la dynamique de l'erreur est

$$\begin{cases} K\dot{\tilde{X}}^d = K\tilde{X}^v - \gamma(H^d)'H^d\tilde{X}^d - \gamma(H^d)'(\epsilon_h + \chi) \\ M\dot{\tilde{X}}^v + C\tilde{X}^v + K\tilde{X}^d = 0 \end{cases} \tag{3.73}$$

donc, pour obtenir l'énergie du système, on multiplie la première équation par $(\tilde{X}^d)^T$ et la seconde par $(\tilde{X}^v)^T$ qui sont directement les conjugués énergétiques

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}}(\frac{1}{2}\tilde{X}^T N\tilde{X}) &= \frac{\mathrm{d}}{\mathrm{dt}}\big(\frac{1}{2}(\tilde{X}^v)^T M\tilde{X}^v + \frac{1}{2}(\tilde{X}^d)^T K\tilde{X}^d\big) \\ &= -\gamma(\tilde{X}^d)^T(H^d)'H^d\tilde{X}^d - (\tilde{X}^v)^T C\tilde{X}^v - \gamma(\tilde{X}^d)^T(H^d)'(\epsilon_h + \chi), \end{aligned} \tag{3.74}$$

où cette fois on retrouve un terme de dissipation à partir des déplacements et non plus des vitesses. Afin de souligner la structure de l'estimateur par rapport à l'énergie de l'erreur nous le définirons donc sous la forme

$$\begin{cases} K\dot{\hat{X}}^d = K\hat{X}^v + \gamma(H^d)'(Z - H^d\hat{X}^d) \\ M\dot{\hat{X}}^v + C\hat{X}^v + K\hat{X}^d = F \end{cases} \tag{3.75}$$

Reste à considérer, au regard des critères d'efficacité, les deux éléments de « réglage » à disposition dans cet estimateur qui sont bien sûr le gain γ mais aussi le choix de l'espace dans lequel l'adjoint $(H^d)'$ est défini. Nous pouvons continuer comme pour le DVF à considérer pour l'adjoint une norme L^2 sur l'espace d'observation notée W^{-1}. Par exemple pour un observateur frontière, nous avons vu apparaître un opérateur de masse sur le bord. Cependant utiliser la même norme pour H^d introduit une dissipation en déplacement $(\tilde{X}^d)^T(H^d)'H^d\tilde{X}^d$ à comparer avec le $(\tilde{X}^v)^T(H^v)'H^v\tilde{X}^v$ du DVF. Or Pour chaque mode de pulsation ω, le DVF fait apparaître en utilisant la vitesse un facteur d'amplification en ω^2. Dans l'optique d'une telle amplification, l'idée serait d'utiliser pour l'adjoint de H^d un opérateur de raideur au lieu de l'opérateur de masse, c'est-à-dire une norme de type H^1 au lieu de L^2. Cependant, nous ne pouvons pas utiliser la véritable raideur du système sur l'espace des observations puisque ces dernières ne sont que partielles par rapport au vecteur d'état complet. Pour contourner la dimension de l'espace d'observation, nous décidons donc d'effectuer un relèvement de cet espace sur l'espace d'état afin de pouvoir lui appliquer l'opérateur de raideur. Pour fixer les idées, on peut considérer par exemple un opérateur d'observations frontière sur $\partial\Omega$. Alors on décide que tout vecteur dans cet espace est relevé sur l'espace des déplacements sur tout Ω par

$$\begin{cases} \mathrm{div}(\underline{\underline{\underline{A}}} \cdot \underline{\underline{\nabla y}}) = 0, & \text{sur } \Omega \\ \underline{y} = \underline{y}_{\mathfrak{m}}, & \text{sur } \partial\Omega \end{cases}$$

On applique alors sur ce relèvement l'opérateur de raideur qui fournit directement un champ de forces aux bord. Ce type de relèvement (harmonique) est très classique en décomposition de domaine. L'objectif est de généraliser cette idée pour n'importe quel opérateur d'observation. Les inverses généralisés (« statiques ») introduits Section 1.2.1 vont nous aider à cet effet. En effet, pour n'importe quel opérateur H nous pouvons définir le relèvement comme l'inverse généralisé de H^d pour la norme K, ou encore par la minimisation équivalente de

$$\min_{X \mid Z=H^d X} \frac{1}{2} X^T K X. \tag{3.76}$$

Ainsi à chaque pas de temps on effectue une estimation statique de l'observation pour trouver, pour cette donnée uniquement, le meilleur état statique (sans présupposé de la dynamique donc) au sens de l'énergie mécanique. En pratique, cette minimisation peut être obtenue à partir du problème mixte

$$\begin{pmatrix} 0 & H^d \\ (H^d)^T & K \end{pmatrix} \begin{pmatrix} \Lambda_Z \\ X_Z \end{pmatrix} = \begin{pmatrix} Z \\ 0 \end{pmatrix}, \tag{3.77}$$

par complément de Schur, ce qui pour un observateur frontière rappelle encore une pratique classique en décomposition de domaine. En résolvant ce système en Λ_Z on définit donc de manière unique l'opérateur S tel que $\Lambda_Z = -SZ$. Le filtre est alors donné par

$$(H^d)' Z = K X_Z = (H^d)^T S Z.$$

Ainsi cette estimation statique est reinjectée dans la dynamique qui devient

$$\begin{cases} K \dot{\hat{X}}^d = K \hat{X}^v + \gamma (H^d)^T S (Z - H^d \hat{X}^d) \\ M \dot{\hat{X}}^v + C \hat{X}^v + K \hat{X}^d = F \end{cases} \tag{3.78}$$

conduisant à

$$\frac{\mathrm{d}}{\mathrm{dt}} (\frac{1}{2} \tilde{X}^T N \tilde{X}) = -\gamma (\tilde{X}^d)^T (H^d)^T S H^d \tilde{X}^d - (\tilde{X}^v)^T C \tilde{X}^v - \gamma (\tilde{X}^d)^T (H^d)^T S (\epsilon_h + \chi). \tag{3.79}$$

Nous appelerons finalement ce filtre « Schur Displacement Feedback » (SDF) pour ne conserver l'appellation DDF que dans le cas où l'adjoint est L^2. Ce filtre n'est plus collocalisé mais créé fondamentalement pour que $H^d S H^d$ se comporte comme K. En fait, si on note X_{corr} le relèvement de l'erreur d'observation $Z - H^d X^d$ alors le complément de Schur permet d'écrire sur la première équation

$$\dot{\hat{X}}^d = \hat{X}^v + \gamma X_{\text{corr}}.$$

Ainsi si on avait eu le déplacement complet comme observation (ce qui limite dans ce cas l'intérêt d'un estimateur !) on aurait pris exactement $K = (H^d)' H^d$ et on aurait simplement corrigé la première équation par l'erreur d'observation

$$\dot{\hat{X}}^d = \hat{X}^v + \gamma \tilde{Z}.$$

Pour mettre en oeuvre ce filtre, on peut commencer par écrire le système complet de l'estimateur

$$\begin{pmatrix} 0 & & & 0 \\ & 0 & & \\ & & K & \\ 0 & & & M \end{pmatrix} \begin{pmatrix} \dot{\widehat{\Lambda_Z}} \\ X_Z \\ X^d \\ X^v \end{pmatrix} = \begin{pmatrix} 0 & H^d & -H^d & 0 \\ (H^d)^T & K & 0 & 0 \\ -(H^d)^T & 0 & 0 & K \\ 0 & & -K & -C \end{pmatrix} \begin{pmatrix} \Lambda_Z \\ X_Z \\ X^d \\ X^v \end{pmatrix} + \begin{pmatrix} Z \\ 0 \\ 0 \\ R \end{pmatrix}, \tag{3.80}$$

On voit alors qu'en pratique, en tout cas avec une schéma implicite et un solveur direct, on a largement augmenté la dimension du système à résoudre pour passer de N_d à $3N_d + m$, conséquence de la résolution du problème mixte initialement défini en (3.77). Là encore c'est ce que nous avons vu au premier chapitre qui va nous offrir sur le même principe une formulation largement préférable. On décide en effet de ne pas définir l'adjoint via la minimisation (3.76) mais par une forme pénalisée

$$\min_X \frac{1}{2}\left\{ X^T K X + \kappa \|Z - H^d X\|^2_{W^{-1}} \right\}. \tag{3.81}$$

Par rapport au chapitre 1, on exprime simplement le fait que l'a priori sur l'état se mesure au travers de la matrice de K et on voit là encore l'analogie entre la covariance au sens probabiliste et la norme pour laquelle on caractérise les espaces en déterministe. Nous avons aussi vu au chapitre 1 que cette formulation permettait pour tout Z de définir une solution (caractère bien posé du problème) car même un Z qui n'est pas sous la forme $H^d X$ va être projeté sur cet espace. Cette remarque assure une solution y compris lorsque H^d n'est pas surjectif (par exemple pour l'opérateur frontière continu introduit Section 3.4.6). Elle est aussi encourageante vis-à-vis du bruit en assurant à coup sûr la projection de Z sur l'espace image de H^d. Du point de vue algébrique, les calculs ont déjà été mené et le système à résoudre devient

$$\begin{pmatrix} -\frac{1}{\kappa}W & H^d \\ (H^d)^T & K \end{pmatrix} \begin{pmatrix} \Lambda_Z \\ X_Z \end{pmatrix} = \begin{pmatrix} Z \\ 0 \end{pmatrix}. \tag{3.82}$$

On peut alors éliminer le multiplicateur de Lagrange pour obtenir une expression familière (où K a remplacé P_0)

$$(H^d)^T S = \kappa K \big(K + \kappa (H^d)^T W^{-1} H^d\big)^{-1} (H^d)^T W^{-1}, \tag{3.83}$$

conduisant à l'estimateur SDF

$$\begin{cases} K\dot{\hat{X}}^d = K\hat{X}^v + \gamma\kappa K\big(K + \kappa (H^d)^T W^{-1} H^d\big)^{-1} (H^d)^T W^{-1} (Z - H^d \hat{X}^d) \\ M\dot{\hat{X}}^v + C\hat{X}^v + K\hat{X}^d = F \end{cases} \tag{3.84}$$

A la différence du DDF, ici nous pouvons désormais simplifier la première équation par K pour obtenir une structure classique

$$\begin{cases} \dot{\hat{X}}^d = \hat{X}^v + \gamma\kappa\big(K + \kappa (H^d)^T W^{-1} H^d\big)^{-1} (H^d)^T W^{-1} (Z - H^d \hat{X}^d) \\ M\dot{\hat{X}}^v + C\hat{X}^v + K\hat{X}^d = F \end{cases} \tag{3.85}$$

mais dans tous les cas, DDF ou SDF, la première équation traduit en continu une relation « point par point » du type

$$\dot{\hat{X}}^d = \hat{X}^v + \gamma X_{\text{corr}}.$$

3.5.2 Analyse de l'estimateur

a. QEP

Les pôles du système sont donnés cette fois par le problème aux valeurs propres

$$\begin{pmatrix} \gamma (H^d)' H^d & K \\ -K & -C \end{pmatrix} V = \lambda \begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix} V, \tag{3.86}$$

et nous commençons ici par quelques résultats numériques pour des observations sur les cellules de tags.

□ DDF – Nous observons numériquement que pour des valeurs de γ telles que tout le spectre est déplacé vers la gauche alors le taux de décroissance est nettement moins bon que pour le DVF. Ainsi, afin d'obtenir une décroissance suffisante pour la plupart des modes nous sommes contraints d'augmenter artificiellement le gain de telle sorte que les tout premiers modes sont ramenés vers l'axe réel (cf. Figure 3.26). Ainsi un ζ^X qui contient ces modes sera faiblement amorti. Pour illustrer cette idée nous utiliserons plus bas les deux types de conditions initiales présentées pour le DVF Section 3.4.5 que sont la condition initiale en pression et celle à partir d'une combinaison linéaire des premiers modes amortis. Pour mémoire, la première CI est donc un exemple de condition que l'on pourrait rencontrer en pratique, alors que la deuxième nous permet d'étudier le cas le moins favorable.

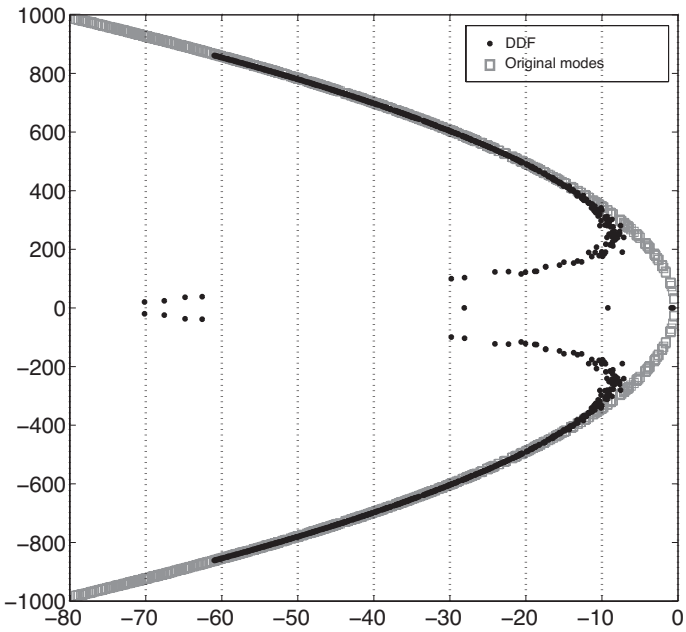


FIGURE 3.26: Pôles de l'estimateur DDF pour l'observateur de type cellules volumiques. Comparaison avec les pôles de la structure initiale.

□ SDF – Pour ce deuxième estimateur, les performances, comme attendu, sont nettement plus intéressantes. Pour comprendre la Figure 3.27 remarquons tout d'abord qu'on déduit de (3.83) l'identité

$$(H^d)^T S H^d + K\big(K + \kappa (H^d)^T W^{-1} H^d\big)^{-1} K = K, \tag{3.87}$$

donc[9]

$$0 \leq (H^d)^T S H^d.$$

Dans le cas idéal $(H^d)^T S H^d = K$ et dans les autres cas, sur le principe, H^d a beau être de beaucoup plus petite dimension que K, l'opérateur $(H^d)^T S H^d$ est proche de K au moins pour les basses fréquences. Donc il est pertinent de comparer le spectre à celui justement où on choisit K. Dans ce cas, nous pouvons alors mener les calculs de spectre analytiquement conduisant

[9]On note au passage que (3.87) permet de retrouver immédiatement que $(H^d)^T S H^d$ est symétrique.

au QEP

$$\lambda^2 MV + \lambda(\gamma M + \eta K)V + (\gamma\eta + 1)KV = 0. \tag{3.88}$$

Ces pôles se déduisent en effet du problème avec amortissement proportionnel

$$\tilde{\lambda}^2 MV + \tilde{\lambda}(\alpha M + \beta K)V + KV = 0, \tag{3.89}$$

avec $\tilde{\lambda} = \frac{\lambda}{\sqrt{\gamma\eta+1}}$, $\alpha = \frac{\gamma}{\sqrt{\gamma\eta+1}}$, $\beta = \frac{\eta}{\sqrt{\gamma\eta+1}}$. Alors les résultats précédemment obtenus donnent donc un gain optimal qui au premier ordre est

$$\gamma_{\text{opt}} = \sqrt{2}\,\omega_0 + \text{h.o.t}, \tag{3.90}$$

où les termes d'ordre supérieur sont en $\mathcal{O}(\omega_0\eta)^2$ et ω_0 est la première fréquence propre du système (dans l'exemple numérique présenté $\omega_0\eta \sim 10^{-2}$). On constate que cette estimation du gain optimal pour le SDF est indépendante de la discrétisation à partir du moment où le premier mode est bien représenté, ce qui est un gage de robustesse.

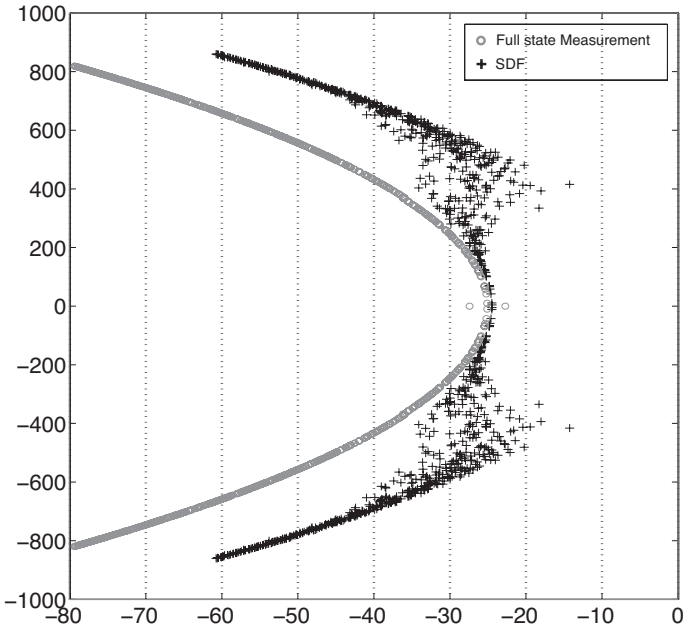


FIGURE 3.27: Pôles de l'estimateur SDF pour l'observateur de type cellules volumiques. Comparaison avec les pôles du SDF à observation complète.

b. Remarques sur les systèmes fortement visqueux

Le fait de modifier la première équation avec le filtre peut poser problème dans le cas des systèmes fortement visqueux qui sont plus proches d'un système parabolique que d'un système hyperbolique. Le système écrit sous forme du premier ordre en déplacements vitesses devient alors singulier. Dans ce cas limite on vérifie que la correction doit donc être introduite comme une force directement dans la principe de la dynamique. L'estimateur correspond à (3.70) où le terme d'inertie est négligé

$$C\dot{\hat{Y}} + \big(K + \gamma(H^d)'H^d\big)\hat{Y} = R + \gamma(H^d)'Z,$$

soit pour l'erreur

$$C\dot{\tilde{Y}} + \big(K + \gamma(H^d)'H^d\big)\tilde{Y} = 0.$$

On contrôle ainsi la norme du système en testant cette équation sur les déplacements (et non plus les vitesses)

$$\frac{\mathrm{d}}{\mathrm{dt}}(\frac{1}{2}\tilde{Y}^T C\tilde{Y}) = \tilde{Y}^T K\tilde{Y} + \gamma\tilde{Y}(H^d)'H^d\tilde{Y}.$$

Ce type de relation donne clairement le principe du filtre envisageable pour les systèmes paraboliques généraux.

Pour ce qui nous concerne, la question reste de savoir si une viscosité trop importante peut détériorer le filtre SDF ou DDF. Pour des viscosités jusqu'à la limite du sur-critique (donc une classe à laquelle appartient le système cardiaque) on peut se convaincre aisément que ce n'est pas le cas en réécrivant formellement sous forme du second ordre notre estimateur SDF (3.72). Tous calculs faits, on a

$$M\ddot{X}^d + (C + \gamma MK^{-1}H'H)\dot{\hat{X}}^d + (K + \gamma CK^{-1}H'H)\hat{X}^d = \\ F + \gamma MK^{-1}H'\dot{Z} + \gamma CK^{-1}H'Z, \quad (3.91)$$

Cette expression permet de comprendre comment agit le filtre sur les différents opérateurs mécaniques. Il ajoute en effet indirectement à la fois une contribution à la viscosité mais aussi à la raideur. L'interprétation détaillée de (3.91) est délicate dans le cas général puisque tous les opérateurs ne sont plus monotones. Cependant, dans le cas d'observations parfaites, on obtient

$$M\ddot{X}^d + (C + \gamma M)\dot{\hat{X}}^d + (K + \gamma C)\hat{X}^d = F + \gamma M\dot{Z} + \gamma CZ$$

qui, pour $C = \eta K$, conduit au calcul de pôles (3.88). Dans ce cas, on peut complètement justifier que le filtre agit efficacement en remontant le seuil critique des modes (voir Figure 3.7 où on remonte le sens des flèches). De façon générale, cette contribution sur la raideur agit donc dans le sens des premières remarques de ce paragraphe pour les systèmes très visqueux. Formellement et numériquement nos estimateurs ont donc un bon comportement quelle que soit la viscosité du système.

L'expression (3.91) permet aussi de constater l'action de notre estimateur sur les observations. Nous savons désormais que les filtres sont simples à concevoir dès qu'on dispose de vitesses. Dans le cas présent, nous n'avons que des déplacements mais formellement le DDF ou SDF dérivent en temps cette donnée de déplacement (terme en $\gamma MK^{-1}H'\dot{Z}$) pour concevoir de nouveau une forme de DVF.

c. Remarques théoriques

Tout d'abord l'ensemble des résultats présentés pour le DVF pourrait s'appliquer dans le cadre DDF-SDF sous condition que le semi-groupe engendré par la dynamique (3.72) soit exponentiellement stable. En effet, l'observateur formulé correspond à un système stabilisé tel qu'analysé dans le théorème 3.3.6 (avec $B' = (H^d\ 0)$) liant stabilisation et contrôlabilité exacte. À la différence cependant du DVF, il n'existe pas de résultat de contrôlabilité exacte dans ce cas puisque, pour un système réel, aucun actionneur n'est capable de corriger le fait que la vitesse est la dérivée de la position ! Notre observateur numérique, en revanche, permet de s'affranchir

de telles limites. Les résultats particuliers pour ce système sont donc ouverts, soit autour de l'exponentielle stabilité classique, soit autour de la formulation d'un opérateur Riez-spectral. Sous condition de cette stabilité, on obtiendrait alors, comme pour le DVF, une estimation *a priori* du type

$$\|\tilde{X}\|_{\mathcal{E}} \leq C\big(e^{-T/\tau}\|\zeta^X\|_{\mathcal{E}} + \gamma\sqrt{\tau}\|\epsilon_h + \chi\|_{L^2([0,T],\mathfrak{m})}\big). \tag{3.92}$$

Une autre piste intéressante pour assurer le résultat en fonction de la discrétisation est de s'appuyer non pas sur le système non stabilisé et donc sur le spectre de l'opérateur elliptique associé, mais de regarder le système comme une perturbation au moins pour les basses fréquences du système stabilisé par K. En effet, en dimension infinie aussi, on pourrait utiliser le fait que les modes de ce système non auto-ajoint restent ceux du problème elliptique donc forment une base de Riez. Puis pour tout γ, y compris γ fini et non pas simplement en étude de sensibilité, on pourrait espérer en décomposant hautes et basses fréquences contrôler les premières par la viscosité interne (comme pour le DVF finalement) et les basses par cette perturbation.

d. Autres opérateurs d'observation

□ OBSERVATION FRONTIÈRE – On reprend l'opérateur frontière introduit avec le DVF mais cette fois on l'applique sur des déplacements. Notamment, on considère les deux situations où les mesures sont disponibles sur endocarde et épicarde ou seulement sur l'endocarde. Les performances du filtre sont bonnes dans les deux cas, assurant le fait que les observations sur la frontière interne sont suffisantes du point de vue de l'observabilité. Lorsque les mesures sont disponibles sur l'ensemble des frontières, on remarque que, dans ce cas, les pôles du filtre SDF sont très proches des pôles du filtre à observation complète.

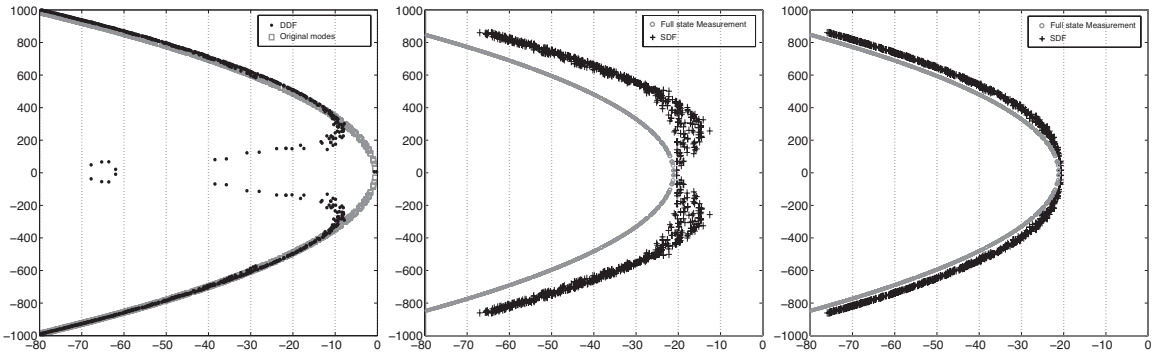


FIGURE 3.28: Pôles pour le DDF and SDF en cas de mesures à l'endocarde (gauche et centre), pour le SDF dans le cas de mesures sur toute la frontière (droite)

□ OBSERVATION NORMALE SUR LA FRONTIÈRE – Dernier type d'opérateur d'observation : on considère seulement une composante du déplacement sur la frontière, en particulier le déplacement normal. Cette idée est motivée par les mesures issues par exemple du flot optique où on peut raisonnablement supposer que la seule grandeur accessible est le déplacement des points de la surface dans la direction orthogonale à cette surface. Le feedback issu de ce type

de mesure s'écrit alors

$$\int_{\Gamma_{\mathrm{m}}} (\underline{\tilde{y}} \cdot \underline{\nu})(\underline{\nu} \cdot \underline{\delta y})\, dS. \tag{3.93}$$

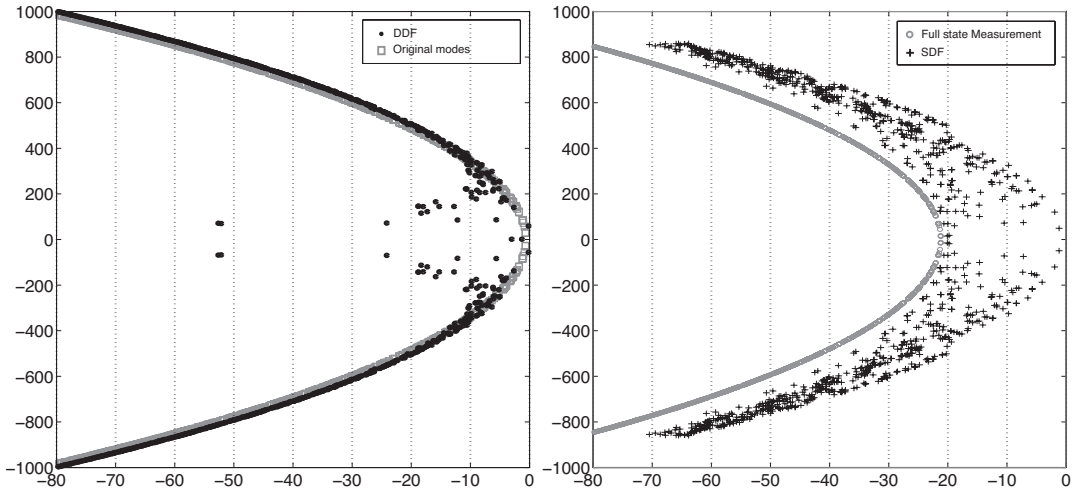


FIGURE 3.29: Pôles pour le DDF (gauche) and SDF (droite) en cas de mesures des déplacements normaux à la surface

Figure 3.29, on constate que les performances du filtre sont nettement moins bonnes. De plus des essais avec d'autres directions, tangentes à la surface ou une des directions $\underline{e}_{\mathrm{x}}$, $\underline{e}_{\mathrm{y}}$, $\underline{e}_{\mathrm{z}}$, montrent le même type de comportement spectral. Ce n'est seulement qu'en ajoutant trois directions orthogonales qu'on retrouve les performances du paragraphe précédent. Cette performance est la conséquence directe d'un manque d'observabilité. Supposons, par exemple, que notre ellipsoïde soit symétrique (il l'est presque) et en rotation autour de son axe d'inertie alors les déplacements normaux seraient nuls et ne permettraient pas le contrôle de la structure.

3.5.3 Schéma en temps

Comme pour le DVF, nous profitons du fait que l'énergie est une fonction de Liapounov pour le filtre pour dériver le schéma en temps de l'estimateur des règles de différences finies trapézoïdales du schéma de Newmark. Le schéma est donc donné par

$$\begin{cases} K\dfrac{\hat{X}^d_{n+1} - \hat{X}^d_n}{\Delta t} = K\dfrac{\hat{X}^v_{n+1} + \hat{X}^v_n}{2} + \gamma (H^d)'\left(Z_{n+\frac{1}{2}} - H^d \dfrac{\hat{X}^d_{n+1} + \hat{X}^d_n}{2}\right) \\ M\dfrac{\hat{X}^v_{n+1} - \hat{X}^v_n}{\Delta t} + C\dfrac{\hat{X}^v_{n+1} + \hat{X}^v_n}{2} + K\dfrac{\hat{X}^d_{n+1} + \hat{X}^d_n}{2} = F_{n+\frac{1}{2}} \\ \hat{X}^d_{n=0} = Y_0, \quad \hat{X}^v_{n=0} = \dot{Y}_0 \end{cases} \tag{3.94}$$

qui fournit une relation exacte de décroissance de l'énergie discrète

$$\begin{aligned} \frac{1}{2\Delta t}(\tilde{X}^T_{n+1} N \tilde{X}_{n+1} - X^T_n N X_n) = -\gamma (\tilde{X}^d_{n+\frac{1}{2}})^T (H^d)' H^d \tilde{X}^d_{n+\frac{1}{2}} + (\tilde{X}^v_{n+\frac{1}{2}})^T C \tilde{X}^v_{n+\frac{1}{2}} \\ - \gamma (\tilde{X}^d_{n+\frac{1}{2}})^T (H^d)' (\chi + \epsilon_h)_{n+\frac{1}{2}}. \end{aligned} \tag{3.95}$$

Cependant, à la différence du DVF, ce schéma est implicite en temps sur les deux équations avec des opérateurs non diagonaux. Autrement dit, on ne doit pas résoudre à la volée la première équation pour la remplacer dans la seconde mais inverser ce système à $N = 2N_d$ degrés de libertés. Comme évoqué précédemment, la complexité de l'inversion des systèmes linéaires est en N fois la largeur de bande au carré donc par rapport à une simulation mécanique classique où on inverse l'opérateur de Newmark $\frac{1}{2}K + \frac{1}{\Delta t}C + \frac{2}{\Delta t^2}M$ (cf Section 2.4.2b.) l'inversion directe du système **coûte approximativement 8 fois plus cher**.

Pour contourner cette difficulté, dans le cas du SDF, nous pouvons choisir de résoudre le système (3.94) par un algorithme itératif si nous trouvons un préconditionneur efficace. On rappelle que dans ce cas le principe de la résolution du système nécessite de pouvoir calculer

$$D_{\text{prec}}^{-1} \cdot D_{\text{SDF}} V,$$

pour tout vecteur V où D_{SDF} est l'opérateur à inverser initialement dans le schéma en temps, soit pour nous

$$D_{\text{SDF}} = \begin{pmatrix} K + \frac{\gamma\Delta t}{2}(H^d)^T S H^d & -\frac{\Delta t}{2}K \\ \frac{K}{2} & \frac{M}{\Delta t} + \frac{C}{2} \end{pmatrix} \begin{pmatrix} \hat{X}_{n+1}^d \\ \hat{X}_{n+1}^v \end{pmatrix} = \text{RHS}.$$

D'après les analyses précédentes, nous choisissons naturellement le préconditionneur

$$D_{\text{pre}} = \begin{pmatrix} K + \frac{\gamma\Delta t}{2}K & -\frac{\Delta t}{2}K \\ \frac{K}{2} & \frac{M}{\Delta t} + \frac{C}{2} \end{pmatrix} \begin{pmatrix} \hat{X}_{n+1}^d \\ \hat{X}_{n+1}^v \end{pmatrix} = \text{RHS},$$

qui est plus simple à inverser car on peut simplifier toute la première équation par K (voir ci-dessous). Ce système correspond à notre estimateur où on a remplacé $(H^d)'H^d$ par K, c'est-à-dire l'estimateur en cas d'observations complètes sur l'état. Or nous avons vu que les deux spectres de D_{SDF} et D_{pre} étaient comparables. Pour quantifier la performance du préconditionneur, on élimine $\hat{X}_{n+1}^v$ dans la deuxième équation pour obtenir une formulation de type Newmark

$$K_{\text{SDF}}\hat{X}_{n+1}^d = \text{RHS}, \quad \text{avec } K_{\text{SDF}} = K + \frac{\gamma\Delta t}{2}(H^d)^T S H^d + \frac{\Delta t}{4}K\left(\frac{M}{\Delta t} + \frac{C}{2}\right)^{-1} K. \tag{3.96}$$

Pour le préconditionneur on doit donc considérer

$$K_{\text{pre}}\hat{X}_{n+1}^d = \text{RHS}, \quad \text{avec } K_{\text{pre}} = K + \frac{\gamma\Delta t}{2}K + \frac{\Delta t}{4}K\left(\frac{M}{\Delta t} + \frac{C}{2}\right)^{-1} K,$$

et puisque (3.87) donne

$$0 \leq (H^d)^T S H^d \leq K,$$

on a pour ces matrices symétriques (cf. (3.87)) positives

$$K_{\text{SDF}} \leq K_{\text{pre}}.$$

De l'autre côté, on obtient aussi

$$\frac{K_{\text{pre}}}{1 + \frac{\gamma\Delta t}{2}} = K + \frac{\frac{\Delta t}{4}}{1 + \frac{\gamma\Delta t}{2}}K\left(\frac{M}{\Delta t} + \frac{C}{2}\right)^{-1} K \leq K_{\text{SDF}}.$$

donc,

$$K_{\text{SDF}} \leq K_{\text{pre}} \leq \big(1 + \frac{\gamma \Delta t}{2}\big) K_{\text{SDF}},$$

donc le conditionnement est proche de 1 quand $\frac{\gamma \Delta t}{2}$ est petit.

Maintenant du point de vue algorithmique, nous avons dit que le préconditionneur était plus simple à inverser que le SDF car on pouvait simplifier par K la première équation. Plus précisément, d'après la forme (3.84) du filtre SDF, la première multiplication $D_{\text{SDF}} V$ avant l'inversion de D_{pre} nécessite l'inversion $(K + \kappa (H^d)^T W^{-1} H^d)^{-1}$ et introduit l'opérateur K sur la première ligne. Cet opérateur peut alors être simplifié lors de la résolution de

$$D_{\text{pre}} \tilde{V} = D_{\text{SDF}} V = \begin{pmatrix} K* \\ * \end{pmatrix}.$$

On n'a donc pour inverser D_{pre} qu'à résoudre un système proche d'une dynamique classique

$$\begin{cases} \dfrac{\hat{X}^d_{n+1} - \hat{X}^d_n}{\Delta t} = \dfrac{\hat{X}^v_{n+1} + \hat{X}^v_n}{2} - \gamma \dfrac{\hat{X}^d_{n+1} + \hat{X}^d_n}{2} + R_{n+\frac{1}{2}} \\ M \dfrac{\hat{X}^v_{n+1} - \hat{X}^v_n}{\Delta t} + C \dfrac{\hat{X}^v_{n+1} + \hat{X}^v_n}{2} + K \dfrac{\hat{X}^d_{n+1} + \hat{X}^d_n}{2} = F_{n+\frac{1}{2}} \end{cases} \tag{3.97}$$

se ramenant uniquement à l'inversion de K_{pre}. On a donc au final à chaque pas de l'algorithme itératif d'inversion, deux inversions de matrices ($(K + \kappa (H^d)^T W^{-1} H^d)$ puis K_{pre}) de taille N_d.

Nous avons mise en oeuvre cette idée avec un algorithme de GMRES préconditionné. Nous avons fixé la tolérance à 10^{-8} et obtenu convergence avec au maximum 4 itérations. Le nombre d'itérations décroît de plus au cours du temps pour atteindre rapidement 2 itérations lorsque l'estimateur converge vers la véritable trajectoire. Nous présentons Figure 3.30, la différence en norme de l'énergie entre les deux solutions données par les méthodes directes et itératives.

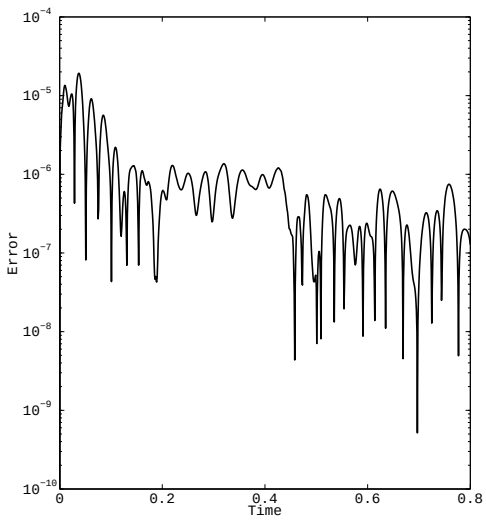


FIGURE 3.30: Différence entre les deux types d'inversion directe et itératif

3.5.4 Résultats d'estimation

Passons maintenant à des exemples concrets d'estimation dans le cas de mesures issues des cellules de tags. Comme pour le DVF, nous commençons Figure 3.31 par l'analyse de l'erreur de condition initiale, c'est-à-dire la contribution du système homogène de condition initiale ζ^x.

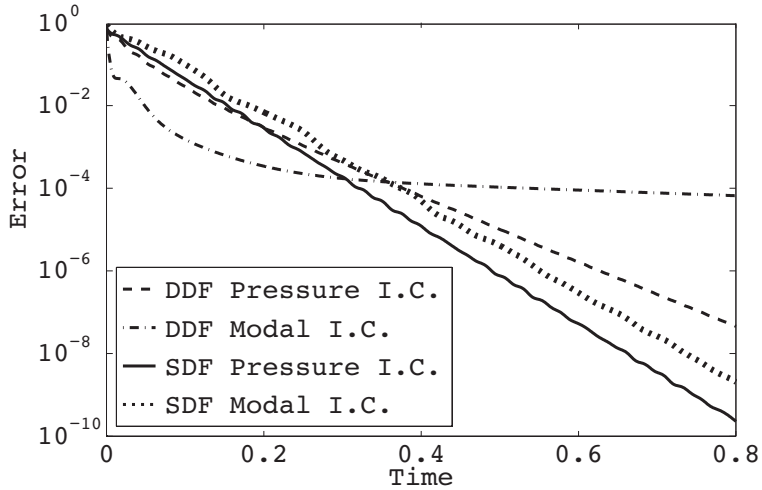


FIGURE 3.31: Énergie relative de l'erreur pour le système homogène SDF and DDF (remis à l'échelle par l'énergie de la condition initiale)

Ensuite, nous passons sur la contribution liée au bruit et à la discrétisation qui sont du même type que pour le DVF. Nous nous intéressons donc directement à l'estimation finale, où nous présentons Figure 3.32 la poursuite de la trajectoire cible au travers des courbes de volume puis au travers des estimations d'analyse numérique classiques Figure 3.33. On constate, comme attendu, que les deux filtres, DDF et SDF, sont très efficaces. Cependant, Figure 3.31, le filtre SDF permet de stabiliser efficacement tous les types de conditions initiales. Ainsi, sur la Figure 3.33, le filtre SDF permet de se rapprocher en fin de période d'estimation de l'erreur d'interpolation alors que le filtre DDF, lui, reste au niveau de l'erreur de discrétisation. Autrement dit, à l'image du DVF pour les mesures de vitesses, le SDF pour les déplacements encourage à assimiler les données dès qu'elles sont disponibles, y compris dans une situation de simulation directe classique « sans incertitude ».

3.5.5 Cas d'observations eulériennes

a. Définition de l'estimateur

Pour l'instant, nous avons toujours supposé les observateurs linéaires afin de définir, mais surtout d'analyser, les performances de nos estimateurs. Pour le DVF, nous avons aussi été jusqu'à regarder le comportement de l'estimateur dans le cas d'une dynamique non-linéaire,

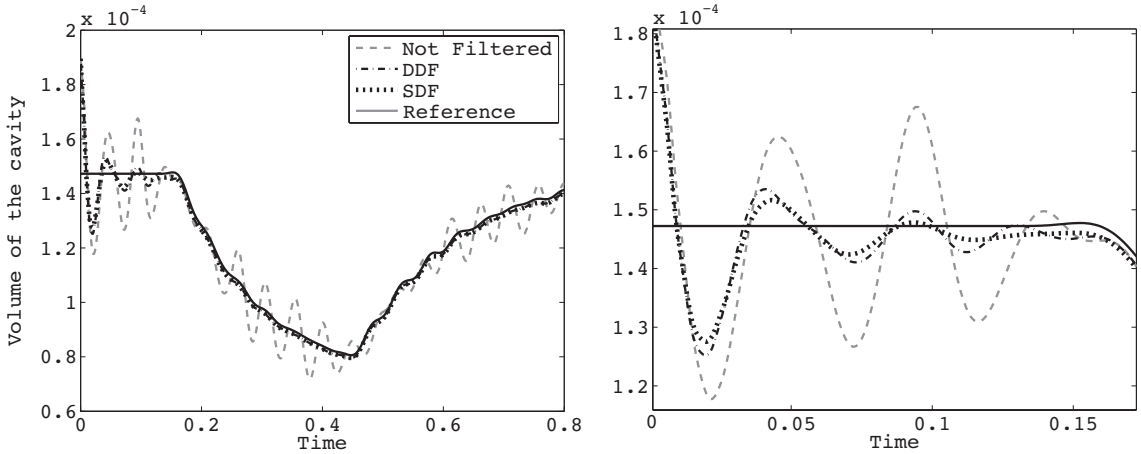


FIGURE 3.32: Volume avec et sans filtre (gauche) et zoom autour de $t = 0$ (droite)

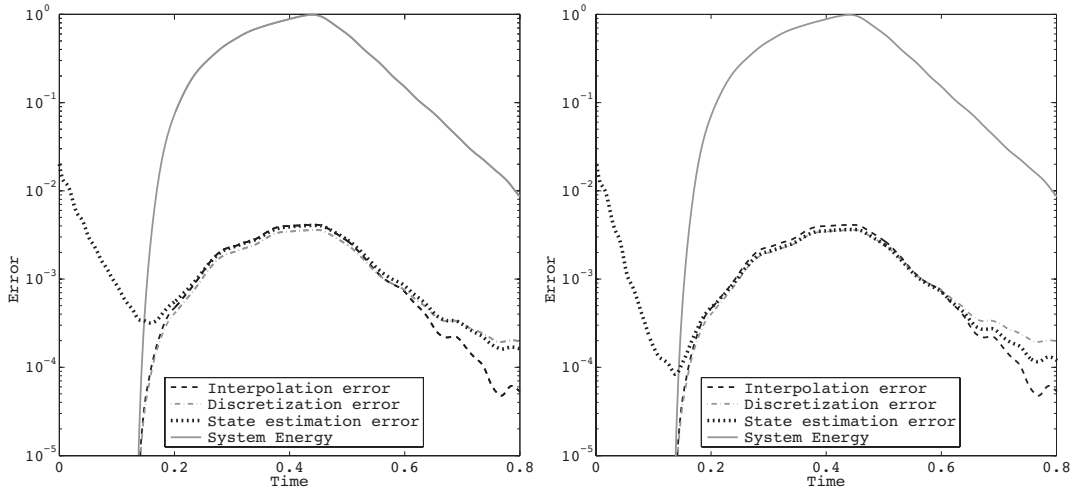


FIGURE 3.33: Comparaison des erreurs d'estimation, d'interpolation et de discrétisation : DDF (gauche) et SDF (droite)

mais toujours avec un opérateur d'observation linéaire. Or, nous avons vu au chapitre 2 que les données d'observation étaient parfois extrêmement complexes et non-linéaires. C'est notamment le cas quand on choisit de conserver la structure profondément Eulerienne de la majorité des données d'imagerie, par exemple les segmentations de ciné-IRM. En fait les tags font un peu figure d'exception puisque même les données de flot optique devraient en réalité être considérées sous une angle Eulerien. Ici nous nous concentrons sur le cas qui nous semble le plus représentatif et d'enjeu le plus important, c'est-à-dire l'utilisation directe de données d'IRM classique. En effet, de nombreuses équipes de traitement d'image utilisent des modèles biomécaniques pour construire des méthodes de segmentation leur images. On pourra consulter parmi d'autres Cohen et al (1991) ; Montillo et al (2002) ; Papademetris et al (2004). Pour cela, elles utilisent en général ce qu'elles appellent des forces images. Ces forces sont en fait proportionnelles à la distance entre le contour de l'image et le modèle biomécanique afin d'attirer ce dernier vers le contour réel. Bien évidemment, puisque le contour de l'image est recherché (donc inconnu !), ces méthodes tentent de le détecter au travers du gradient de l'image mais, indépendamment des techniques de plus en plus abouties pour caractériser ce

contour, le principe est bien de faire converger la géométrie d'un modèle mécanique (appelé modèle déformable dans la communauté de l'imagerie) sur ce contour. Cette méthode peut être effectuée de manière quasi-statique ou au cours du temps, voir notamment Sermesant (2003) ; Rouchdy et al (2007) et leurs références. Le cadre dynamique en imagerie médicale est largement circonscrit au système cardiaque mais on peut aussi trouver des besoins pour le système respiratoire comme décrit dans von Berg et al (2007). Schématiquement ces méthodes sont donc des estimations de la forme

$$\hat{Y} = K^{-1} F_{\text{Image}}(Y, t),$$

où

$$M\ddot{\hat{Y}} + C\dot{\hat{Y}} + K\hat{\hat{Y}} = K^{-1} F_{\text{Image}}(\hat{Y}, t),$$

avec $F_{\text{Image}}(Y, t)$ cette force image proportionnelle à la distance entre le contour estimé et le contour perçu dans l'image. Remarquons qu'il n'existe pas à notre connaissance de version en grands déplacements de cette méthode tenant compte de l'inertie. Néanmoins, Rouchdy et al (2007) proposent de résoudre, par une méthode de type Fourier, une dynamique sur-critique (parabolique) avec énergie hyperélastique de type Saint-Venant.

Plaçons-nous du point de vue méthodologique et examinons alors la ressemblance de ces méthodes avec nos procédés d'estimation. Pour cela on considère que le travail a déjà été fait, c'est-à-dire que la segmentation a déjà eu lieu. Ceci évite de gérer une méthode pertinente de définition implicite du contour, ce qui est un sujet de recherche en soi. Ensuite, cela se justifie aussi dans un processus à deux étapes où on commence par des segmentations classiques avant d'y plonger un modèle mécanique qui va apporter une cohérence cinématique et même mécanique si le modèle est pertinent. Les segmentations étant données au cours du temps, on veut réaliser une estimation à partir uniquement de ces surfaces extraites comme nous l'avons fait pour d'autres données d'observation. Ces surfaces, notées (S_k), correspondant à la segmentation sont fournies avec un échantillonnage en temps, tous les $k\Delta T,\ 1 \leq k \leq N_{\Delta T}$. Cet échantillonnage est plus grossier que le pas de temps de simulation. Nous prenons soin de le considérer à nouveau car la non-linéarité des opérateurs d'observation mis en jeu complexifie la notion d'interpolation des mesures en temps vue en introduction de ce chapitre.

Comme déjà exprimé au chapitre 2, ces surfaces ne peuvent s'exprimer simplement sous la forme $Z = \mathcal{H}(x) + \chi$. En revanche, elles peuvent être utilisées pour déterminer si un point du modèle est placé sur cette surface ou non en calculant la distance de ce point à S_k. Ainsi, un estimateur suit le système réel au sens de l'observation si

$$\text{dist}(\underline{\text{x}}, S_k) = 0, \quad \forall \underline{\text{x}} \in \partial\hat{\Omega}_t, \tag{3.98}$$

où $\text{x}(t, \underline{\xi}) = \underline{\xi} + \underline{y}(t, \underline{\xi})$ est la position en configuration déformée pour l'estimateur (i.e sur $\hat{\Omega}_t$). Cette relation réécrite en configuration Lagrangienne et sur tous les points des surfaces du modèle éléments finis s'écrit donc sous la forme $D_k(Z, X) = 0$ à comparer en linéaire à $Z_k - HX = 0$. En l'occurrence, la présence de bruit transforme cette relation en

$$D_k(Z, X) = \chi,$$

comme déjà introduit au chapitre 1. Afin de prendre en compte l'échantillonnage en temps, et de formuler un opérateur (généralisé) d'observations à tous les pas de temps, on introduit

alors l'interpolation sur ces distances

$$D(\hat{X}^d, t) = \alpha_k(t) D(\hat{X}^d, S_k) + (1 - \alpha_k(t)) D(\hat{X}^d, S_{k+1}), \quad \forall t \in [k\Delta T, (k+1)\Delta T]. \tag{3.99}$$

Cette interpolation a du sens si la distance est signée, c'est-à-dire si la surface est orientée de telle sorte qu'on puisse déterminer si on est à l'intérieur ou à l'extérieur du domaine Ω_t poursuivi. Dans ce cas, cette interpolation est légitime même si la surface $\partial\hat{\Omega}_t$ est comprise entre deux surfaces successives S_k et S_{k+1} comme schématisé Figure 3.34. Insistons sur le fait

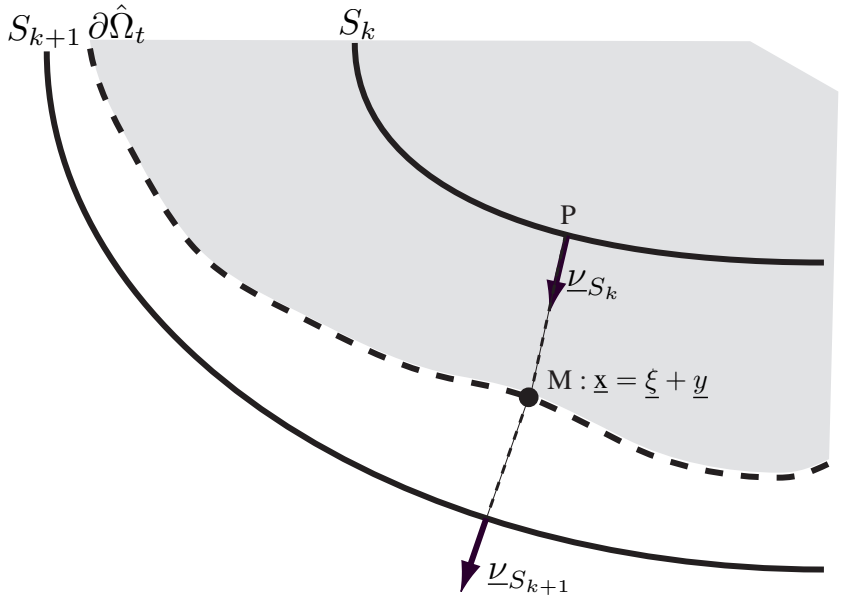


FIGURE 3.34: Projection du contour estimé entre deux surfaces d'images successives

que ce choix d'interpolation en temps est directement le pendant non-linéaire de celle que nous avions effectuée en linéaire en introduisant

$$(\alpha_k(t) Z_k + (1 - \alpha_k(t)) Z_{k+1}) - H(\hat{X}) = \alpha_k(t)(Z_k - H(\hat{X})) + (1 - \alpha_k(t))(Z_{k+1} - H(\hat{X})).$$

Elle aura donc les mêmes implications en terme d'erreur d'estimation avec l'ajout d'un terme $\epsilon_{\Delta T}$ en $\mathcal{O}(\Delta T)$.

Au chapitre 1, nous avions vu que la définition d'un filtre dans le cas d'un opérateur d'observation non-linéaire faisait apparaître la dérivée de D par rapport à l'état. Ainsi si $-H^T W^{-1}(Z - HX)$ correspond à la dérivée du terme $\frac{1}{2}\|Z - HX\|^2_{W^{-1}}$, alors $\frac{\partial D}{\partial Y}^T W^{-1} D(Z, X)$ est l'analogue pour un critère comprenant $\frac{1}{2}\|D(Z, X)\|^2_{W^{-1}}$.

Nous dérivons donc notre distance en utilisant les notations introduites Figure 3.34. On a, d'après l'orthogonalité de $\underline{MP}$) et $\underline{\nu}_{S_k}$

$$\begin{aligned}
\frac{\partial \mathrm{dist}(\underline{\xi} + \underline{y}, S_k)}{\partial \underline{y}} \cdot \underline{dy} &= -\frac{\partial((\underline{MP}) \cdot \underline{\nu}_{S_k})}{\partial \underline{y}} \cdot \underline{dy} \\
&= \left(\frac{\partial \underline{M}}{\partial \underline{y}} \cdot \underline{dy}\right) \cdot \underline{\nu}_{S_k} - \left(\frac{\partial \underline{P}}{\partial \underline{y}} \cdot \underline{dy}\right) \cdot \underline{\nu}_{S_k} - \underline{MP} \cdot \left(\frac{\partial \underline{\nu}_{S_k}}{\partial \underline{y}} \cdot \underline{dy}\right) \\
&= \left(\frac{\partial \underline{M}}{\partial \underline{y}} \cdot \underline{dy}\right) \cdot \underline{\nu}_{S_k}.
\end{aligned}$$

En d'autres termes le gradient de la distance est la normale au contour au point de projection

$$\nabla_{\underline{y}}(\text{dist}(\underline{\xi}+\underline{y}, S_k)) = \underline{\nu}_{S_k}. \tag{3.100}$$

Par conséquent dans le cas où l'on considère la norme $L^2(\Gamma_{\mathrm{m}})$, le terme $\frac{\partial D}{\partial Y} W^{-1} D(Z,Y)$ s'écrit variationnellement

$$\alpha_k(t)\int_{\Gamma_{\mathrm{m}}} \text{dist}(\underline{\xi}+\underline{y}, S_k)(\underline{\nu}_{S_k}\cdot\underline{v})\,dS, +(1-\alpha_k(t))\int_{\Gamma_{\mathrm{m}}} \text{dist}(\underline{\xi}+\underline{y}, S_{k+1})(\underline{\nu}_{S_{k+1}}\cdot\underline{v})\,dS. \tag{3.101}$$

On reconnaît ici la structure de force image puisqu'on tire effectivement le contour de l'estimateur dans la direction du contour réel et proportionnellement à la distance à celui-ci. La différence est que notre compréhension des mécanismes de stabilisation d'erreur avec des données en déplacements nous conduit à définir un estimateur DDF ou SDF et non pas à appliquer ces forces directement au second membre de la dynamique. L'estimateur que nous proposons est en effet

$$\begin{cases} K\dot{\hat{X}}^d = K\hat{X}^v - \gamma\left(\dfrac{\partial D}{\partial Y}\right)' D(Z,\hat{X}^d) \\ M\dot{\hat{X}}^v + C\hat{X}^v + K\hat{X}^d = F \end{cases} \tag{3.102}$$

et nous pouvons espérer une convergence au moins pour le problème linéarisé alors que les termes de force image standards ne font qu'ajouter de la raideur dans le linéarisé de l'erreur (ce qui, en toute rigueur, ne signifie d'ailleurs pas qu'il ne converge pas puisqu'un système non-linéaire peut-être stable alors que son tangent ne l'est pas). Avant de passer justement au linéarisé, mentionnons aussi que comme précédemment la version SDF du filtre que nous proposons n'est plus collocalisée puisque moralement il introduit une correction sur la première équation de la forme

$$\dot{\hat{X}}^d = \hat{X}^v + \gamma X_{corr},$$

où X_{corr} est le relèvement sur tout le volume de l'écart d'observation mesuré à la frontière.

Le système de l'erreur, comme d'habitude en non-linéaire, n'est plus autonome mais peut être linéarisé autour de la trajectoire cible qui par définition vérifie $D(Z,X)=0$

$$\begin{cases} K\dot{\tilde{X}}^d = K\tilde{X}^v - \gamma\left(\dfrac{\partial D}{\partial Y}(X^d,t)\right)'\left(\dfrac{\partial D}{\partial Y}(X^d,t)\right)\tilde{X}^d \\ M\dot{\tilde{X}}^v + C\tilde{X}^v + K\tilde{X}^d = 0 \end{cases} \tag{3.103}$$

Cette dynamique est exactement dissipative puisqu'elle correspond à l'équation (3.75) où on a simplement remplacé H par $\frac{\partial D}{\partial Y}$. L'opérateur de dissipation ajouté a pour forme variationnelle pour le cas DDF

$$\int_{\Gamma_{\mathrm{m}}} (\underline{\tilde{y}}\cdot\underline{\nu}_{S_k})(\underline{\nu}_{S_k}\cdot\underline{\delta y})\,dS,$$

et on constate qu'il est donc très proche du cas d'observation suivant la normale au contour. La différence vient cependant du fait qu'ici le terme fait intervenir le vecteur normal $\underline{\nu}_{S_k}$ qui varie au cours du temps.

En pratique, nous avons mis en oeuvre cet algorithme en définissant S_k comme des surfaces maillées en triangles. Ainsi, il nous a été possible d'utiliser l'algorithme de Baerentzen et Aanaes (2005). Cet algorithme avait déjà été mis à profit pour les différentes opérations géométriques nécessaires à la mise en oeuvre des modèles directs (fibres, régions AHA etc.). Son

avantage sur d'autres est qu'il permet effectivement un calcul de distance signée exacte pour maillages surfaciques (orientables) triangulaires. Or on rappelle que le signe revêt ici toute son importance afin d'interpoler correctement les contributions des surfaces successives.

b. Discrétisation en temps

Puisque l'opérateur $\Big(\frac{\partial D}{\partial Y}(Y,t)\Big)D(Y,t)$ est non-linéaire, une discrétisation point milieu

$$\left(\frac{\partial D}{\partial Y}\Big(\frac{\hat{X}^d_{n+1}+\hat{X}^d_n}{2},\frac{t_{n+1}+t_n}{2}\Big)\right)' D\Big(\frac{\hat{X}^d_{n+1}+\hat{X}^d_n}{2},\frac{t_{n+1}+t_n}{2}\Big).$$

n'est plus équivalente à une règle des trapèzes

$$\frac{1}{4}\left(\frac{\partial D}{\partial Y}(\hat{X}^d_{n+1},t_{n+1})+\frac{\partial D}{\partial Y}(\hat{X}^d_n,t_n)\right)'\Big(D(\hat{X}^d_{n+1},t_{n+1})+D(\hat{X}^d_n,t_n)\Big).$$

En utilisant la règle point milieu, l'avantage est que le système de l'erreur linéarisée va être du même type que (3.94) donc exactement dissipatif avec, à la place de $(H_d)'H_d$, un terme en

$$(\frac{\partial D}{\partial Y}(\hat{X}^d_{n+\frac{1}{2}},t_{n+\frac{1}{2}}))'(\frac{\partial D}{\partial Y}(\hat{X}^d_{n+\frac{1}{2}},t_{n+\frac{1}{2}})).$$

Cependant ce schéma implique qu'on résout un système non-linéaire (donc un schéma de Newton) même lorsque la dynamique initiale est linéaire. Ceci contredit le cahier des charges initial où nous souhaitions laissé de côté le critère optimal les estimateurs variationnels (pour qui d'ailleurs le même problème de discrétisation se pose au niveau des équations adjointes) au profit d'un gain de complexité algorithmique. Afin de pallier cette difficulté, nous proposons de linéariser ce terme avec une extrapolation du second ordre afin de conserver la propriété de consistance quadratique du schéma de Newmark. On introduit donc

$$X_e=\frac{3X_n-X_{n-1}}{2},$$

tel que $X_e-X_{n+\frac{1}{2}}=O(\Delta t^2)$ et remplaçons

$$\left(\frac{\partial D}{\partial Y}(\hat{X}^d_{n+\frac{1}{2}},t_{n+\frac{1}{2}})\right)' D(\hat{X}^d_{n+\frac{1}{2}},t_{n+\frac{1}{2}})\rightarrow$$
$$\left(\frac{\partial D}{\partial Y}(\hat{X}^d_e,t_{n+\frac{1}{2}})\right)'\left(D(\hat{X}^d_e,t_{n+\frac{1}{2}})+\frac{1}{2}\Big(\frac{\partial D}{\partial Y}(\hat{X}^d_e,t_{n+\frac{1}{2}})\Big)(\hat{X}_{n+\frac{1}{2}}-\hat{X}_e)\right).$$

Pour le système réel discrétisé en temps nous avons

$$\begin{aligned}0&=D(Y_{n+\frac{1}{2}},t_{n+\frac{1}{2}})\\&=D(Y_e,t_{n+\frac{1}{2}})+\Big(\frac{\partial D}{\partial Y}(Y_e,t_{n+\frac{1}{2}})\Big)(Y_{n+\frac{1}{2}}-Y_e)\\&\quad+\Big(\frac{\partial^2 D}{\partial Y^2}(Y_e,t_{n+\frac{1}{2}})\Big):(Y_{n+\frac{1}{2}}-Y_e)\otimes(Y_{n+\frac{1}{2}}-Y_e)+o(\Delta t^4).\end{aligned}$$

Donc cette fois pour le système de l'erreur linéarisé on obtient

$$\begin{aligned}-\gamma\left(\frac{\partial D}{\partial Y}(Y_e,t_{n+\frac{1}{2}})\right)'\Big(&\Big(\frac{\partial D}{\partial Y}(Y_e,t_{n+\frac{1}{2}})\Big)(\tilde{X}_{n+\frac{1}{2}})\\&+\Big(\frac{\partial^2 D}{\partial Y^2}(Y_e,t_{n+\frac{1}{2}})\Big):\tilde{X}_e^d\otimes(Y_{n+\frac{1}{2}}-Y_e)\Big)+o(\Delta t^4),\end{aligned}\tag{3.104}$$

qui est dissipatif à l'ordre 2 en temps à $O(\Delta t^2)$.

REMARQUE 3.5.1
Nous avons aussi utilisé une forme explicite classique $X_e = X_n$ qui donne cette fois une consistance en $\mathcal{O}(\Delta t)$ et en pratique pour les pas de temps que nous avons utilisés nous n'avons pas vu de véritables différences numériques pour cette consistance.

3.5.6 Résultats

Comme pour le DVF, la Figure 3.35 présente, via les courbes de volume, un premier aperçu de la capacité du filtre à poursuivre la trajectoire du système cible. Plus précisément, nous présentons Figure 3.36, l'erreur d'estimation obtenue avec les filtres DDF et SDF pour la condition initiale de surpression intra-cavitaire, et pour la condition initiale à base de modes de la structure non amortie Figure 3.37. Dans les deux cas, toute la frontière est considérée, c'est-à-dire qu'on utilise moralement une segmentation complète de l'organe. Pour des raisons pratiques, les surfaces utilisées sont directement issues de la simulation directe avec le maillage de référence sans ajout de bruit de mesures. Nous traiterons plus loin, Section 3.5.8, une configuration où nous introduisons un bruit réaliste pour ce type de données d'observation. Les performances des filtres sont excellentes dans les deux cas, même si, comme en linéaire, le DDF est plus sensible à la configuration modale. En effet, son gain artificiellement élevé permet un très bon comportement dans le cas de la surpression, mais se paye sur certains autres modes de basse fréquence. Malgré le fait que le linéarisé de cet estimateur corresponde à un observateur du déplacement normal à la frontière, le comportement en non-linéaire est bien meilleur que celui en linéaire. Notre interprétation de cette performance est que les vecteurs $\underline{\nu}_{S_k}$ sont en fait mobiles au cours de l'estimation, donc balaient l'espace des directions et augmentent d'autant l'observabilité du système par rapport au cadre linéaire présenté Section 3.5.2.d. Cette remarque justifie aussi que les systèmes linéaires sont souvent réputés pour justifier la stabilisation mais n'apportent que des critères suffisants mais non nécessaires (voir Coron (2007) pour de nombreux exemples).

3.5.7 Extension en dynamique non-linéaire

a. Formulation

Nous reprenons l'extension du problème ci-dessus en grands déplacements déjà vu Section 3.4.7. Nous réécrivons cependant la formulation introduite à partir des notations non ambiguës X^d,

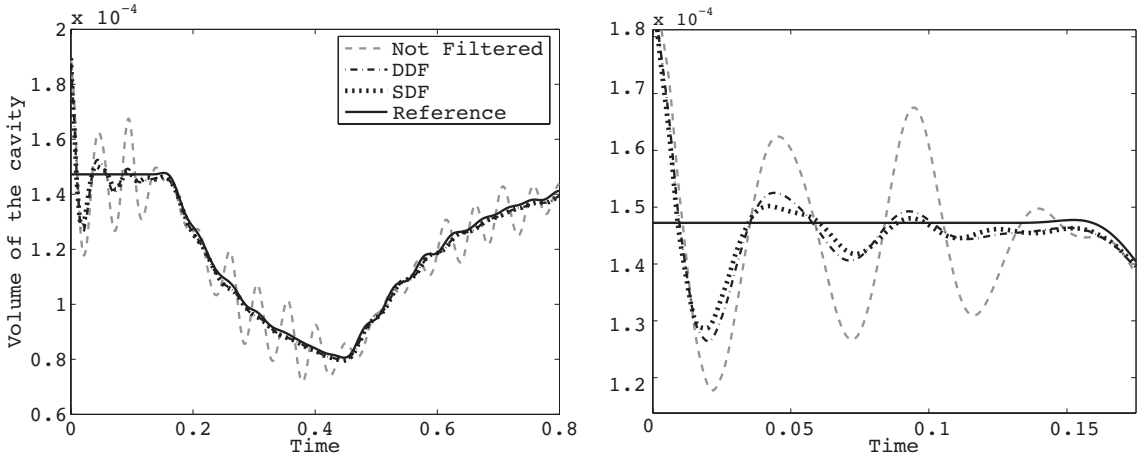


FIGURE 3.35: Volume avec et sans filtre (gauche) et zoom autour de $t = 0$ (droite)

X^v :

$$\begin{cases} \dot{X}^d = X^v \\ M\dot{X}^v + C(X^d)X^v + F^i(X^d) = F^{PS}(X^d, t) \end{cases} \tag{3.105}$$

Introduisons alors $K(X^d, X^v)$ le linéaire tangent de l'opérateur de raideur **par rapport à la position courante**

$$K(X^d, X^v) = \frac{\partial}{\partial Y}(F^i + C(X^d)X^v - F^{PS})(X^d).$$

Nous proposons alors l'estimateur suivant en non-linéaire

$$\begin{cases} K(\hat{X}^d, \hat{X}^v)\dot{\hat{X}}^d = K(\hat{X}^d, \hat{X}^v)\hat{X}^v + \gamma(H^d)'(Z - H^d\hat{X}^d) \\ M\dot{\hat{X}}^v + C(\hat{X}^d)\hat{X}^v + F^i(\hat{X}^d) = F^{PS}(\hat{X}^d, t) \end{cases} \tag{3.106}$$

Commençons par considérer le linéarisé autour de la trajectoire réelle, c'est-à-dire uniquement pour la dynamique non autonome de la seule variable d'erreur. Celle-ci admet comme dynamique

$$\begin{cases} K(Y, \dot{Y})\dot{\tilde{X}}^d = K(Y, \dot{Y})\tilde{X}^v - \gamma(H^d)'H^d\tilde{X}^d - \gamma(H^d)'(\epsilon_h + \chi) \\ M\dot{\tilde{X}}^v + C(Y)\tilde{X}^v + K(Y, \dot{Y})\tilde{X}^d = 0 \end{cases} \tag{3.107}$$

Or, comme Section 3.4.7, même si K et C sont désormais dépendants du temps, la dynamique de l'erreur linéarisée est toujours stabilisée tant que $K(X^d)$ reste positif. Dans ces cas, les théorèmes de stabilisation pour le linéarisé (cf. Section 3.3.1.b.) s'appliquent et assurent l'existence d'un voisinage de $\tilde{X}(0) = \zeta^X$ pour lequel l'estimateur converge vers le système. Nous insistons de plus sur le fait que dans le cas du SDF, $K(X^d, X^v)$ est utilisé dans la définition du filtre.

Dans le cas où $K(X^d, X^v)$ n'est pas définie positive, c'est-à-dire typiquement des cas de forte compression et de flambement que nous avons déjà mentionnés, nous proposons de remplacer $K(X^d, X^v)$ par $K(X^d(0), 0)$. Dans ce cas, la linéarisation de l'erreur se fait pour le système complet de l'erreur $(X, \tilde{X})$ autour de $(0, 0)$. Bien que cette justification soit peu convaincante sur le fond puisque le système était justement supposé travailler en grands déplacements, l'estimateur garde le mérite de rester stable et il est forcé par les déplacements du système réel. Donc notre estimateur a l'avantage d'être robuste et on peut même espérer

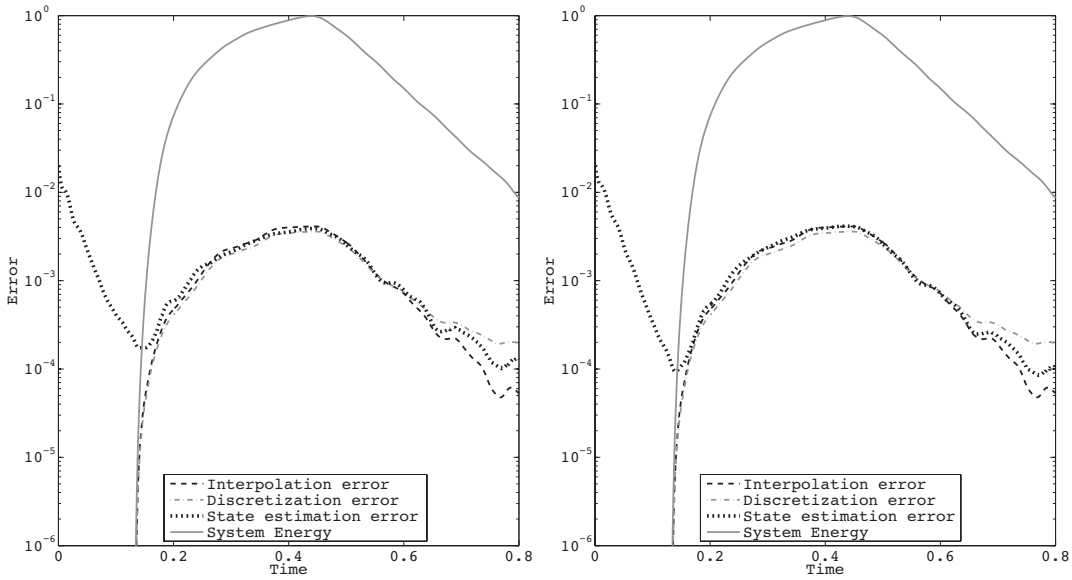


FIGURE 3.36: Erreur d'estimation pour des images segmentées à partir d'une erreur de condition initiale en surpression : DDF (gauche) et SDF (droite)

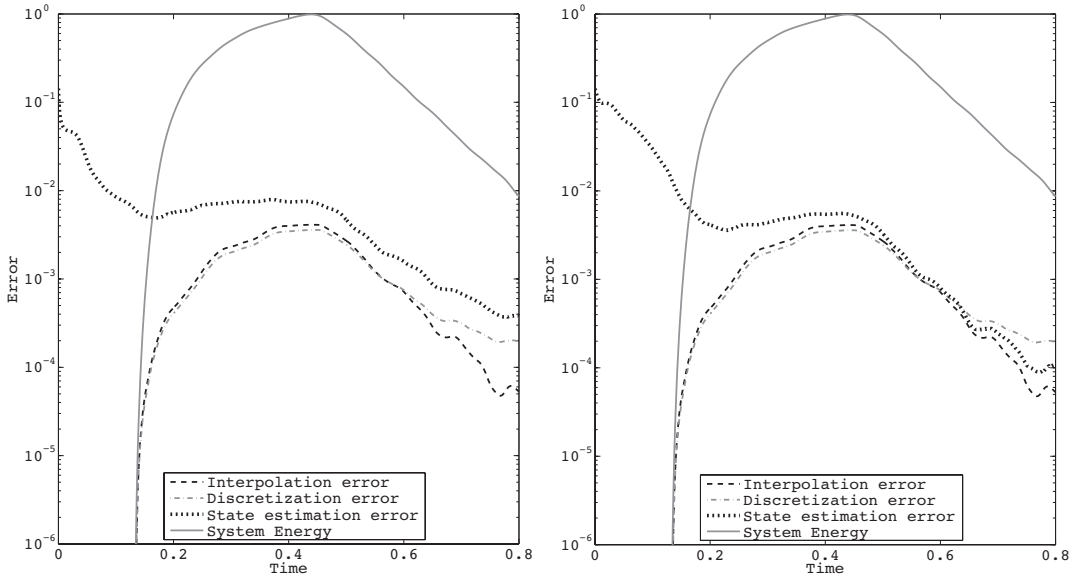


FIGURE 3.37: Erreur d'estimation pour des images segmentées à partir d'une erreur de condition initiale modale : DDF (gauche) et SDF (droite)

qu'il suive la bonne bifurcation imposée par le flambement du système réel si les observations le permettent alors qu'une simulation numérique simple sans observation pourrait diverger. En pratique nous avons d'ailleurs testé les deux estimateurs avec $K(\hat{X}^d, \hat{X}^v)$ et $K(\hat{X}^d(0), 0)$ et nous n'avons observé que peu de différence dans les résultats y compris dans des configurations où nous avons forcé l'erreur de condition initiale.

b. Modifications du schéma en temps standard

L'une des qualités des filtres proposés est qu'ils sont implémentables à peu de frais dans un code éléments finis car ils sont imaginés à partir de contrôleurs physiques donc de termes variationnels classiques vis-à-vis du problème simulé. Pour mémoire, les opérateurs de stabilisation frontière ont la même implémentation que nos opérateurs de conditions aux limites élastiques. Le cas des filtres en déplacement est cependant un peu particulier par rapport à la philosophie des codes d'élastodynamique car ils corrigent l'équation en vitesse. Or celle-ci discrétisée suivant le schéma de Newmark est habituellement résolue « à la volée » dans un code de mécanique. Autrement dit la vitesse $\dot{Y}_{n+1}$ est remplacée dans tous les termes de viscosité et d'accélération par $\dot{Y}_{n+1} = \frac{2}{\Delta t}(Y_{n+1} - Y_n) + \dot{Y}_n$. Ici cette relation n'est plus vérifiée et nous avons donc dit qu'il fallait conserver la variable de vitesse comme une variable à part entière dans le code. La conséquence la plus grave pour l'implémentation concerne alors les opérateurs de viscosité en non-linéaire. Classiquement ils se simplifient en utilisant le schéma de Newmark comme indiqué en (2.61), et permettent le calcul de l'opérateur élément-finis uniquement à partir des déplacements. Pour l'implémentation du DDF ou SDF, ce paradigme n'est plus valable et le code doit pouvoir gérer indépendamment les variables de déplacements X^d et de vitesses X^v, et renvoyer des opérateurs tangents par rapport à chacune d'elle. En particulier l'opérateur de viscosité discrétisé s'écrit dans notre cas

$$\Big(C(X^d_{n+\frac{1}{2}})X^v\Big)^T V = \int_\Omega \eta \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}(\underline{x}^d_{n+\frac{1}{2}}) \cdot \underline{\dot{x}}^v_{n+\frac{1}{2}} \, : \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}(\underline{x}^d_{n+\frac{1}{2}}) \cdot \underline{v} \, d\Omega,$$

dont les opérateurs tangents sont par rapport au déplacement au point milieu

$$\left(\Big(\frac{\partial C_{n+\frac{1}{2}}}{\partial Y} X^v_{n+\frac{1}{2}}\Big) U^d\right)^T V = \int_\Omega \eta(\, \mathrm{d}^2_{\underline{y}}\underline{\underline{e}} : \underline{u}^d \otimes \underline{\dot{x}}^v_{n+\frac{1}{2}}) : \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}(\underline{x}^d_{n+\frac{1}{2}}) \cdot \underline{v} \, d\Omega +$$
$$\int_\Omega \eta \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}_{n+\frac{1}{2}} \cdot \underline{\dot{x}}^v_{n+\frac{1}{2}} : (\, \mathrm{d}^2_{\underline{y}}\underline{\underline{e}} : \underline{u}^d \otimes \underline{v}) \, d\Omega,$$

non symétrique et, par rapport à la vitesse au point milieu

$$\Big(C_{n+\frac{1}{2}} U^v\Big)^T V = \int_\Omega \eta \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}(\underline{x}^d_{n+\frac{1}{2}}) \cdot \underline{u}^v \, : \, \mathrm{d}_{\underline{y}}\underline{\underline{e}}(\underline{x}^d_{n+\frac{1}{2}}) \cdot \underline{v} \, d\Omega,$$

qui lui reste symétrique. On rappelle que la différentielle d'ordre 2 du tenseur de déformation ne dépend pas du déplacement et s'exprime

$$\mathrm{d}^2_{\underline{y}}\underline{\underline{e}} : \underline{u} \otimes \underline{v} = \underline{\underline{\nabla}}\, \underline{u} : \underline{\underline{\nabla}}\, \underline{v}.$$

c. Résultats

Nous présentons Figure 3.38 le résultat d'estimation pour le filtre SDF obtenu sur une erreur de condition initiale en surpression. Afin de valider la robustesse de notre filtre, nous avons de nouveau augmenté la suppression d'un facteur 1.8. Nous avons aussi tracé les résultats d'estimation dans le cas où nous utilisons la matrice de raideur $K(X^d(0), 0)$ au lieu de $K(\hat{X}^d, \hat{X}^v)$ dans la définition du filtre. Dans notre exemple, les deux situations sont en fait très proches car nous commençons l'estimation pour une trajectoire cible nulle pendant le premier dixième de seconde de la simulation.

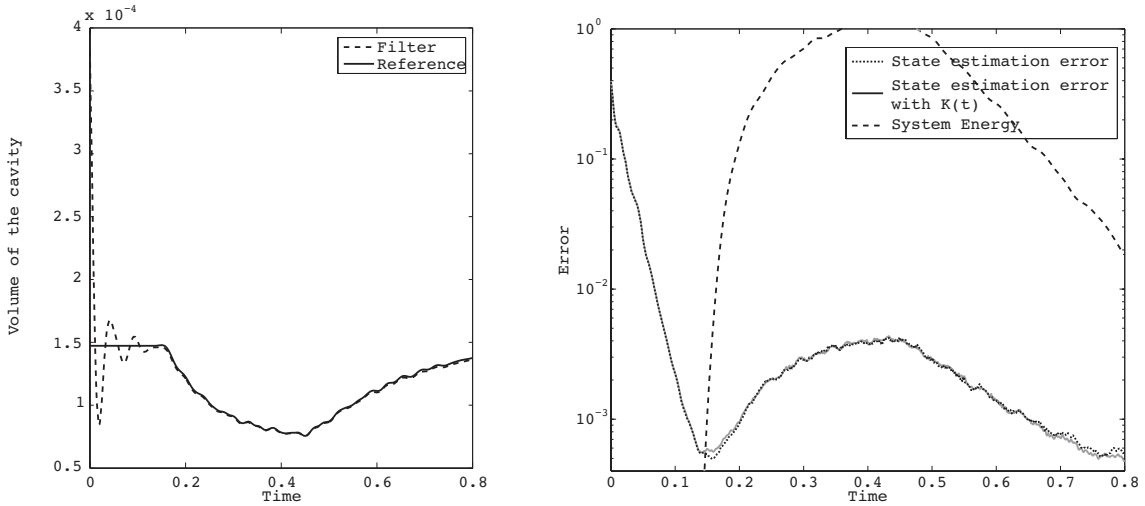


FIGURE 3.38: ***Dynamique non-linéaire :*** **Volume avec et sans filtre (gauche) et énergie de l'erreur (droite)**

3.5.8 Résultats d'estimation d'état pour le cœur

Pour clore ce chapitre nous proposons un exemple d'estimation dans une configuration que nous espérons la plus proche possible de l'application. Cet exemple incorpore alors les éléments les plus techniques des méthodes proposées entre modélisation, mesure, et filtrage. Mais, avant cela, intéressons-nous au cas de l'estimation des variables d'état de pression qui au chapitre 2 étaient présentées comme des variables du système cardiaque à part entière.

a. Estimation des variables de pression

Dans un cas 0D (le piston de la Section 2.6 avec un seul degré de liberté), nous avions en effet constaté que sauf à mesurer la pression ventriculaire avec un haut degré de confiance, laisser le système estimer la trajectoire en lui permettant de changer de phase engendrait des instabilités sur la pression qui « sautait » d'une phase à l'autre. Nous illustrons cette idée Figure 3.39, avec un résultat d'estimation d'état par filtre de Kalman sur ce système 0D. Dans le premier cas, on ne dispose pas de mesures de pression et la pression ventriculaire oscille alors à chaque pas de temps entre les différentes phases. Dans le deuxième cas, en revanche, mesurer la pression ventriculaire avec une bonne confiance permet de la stabiliser et de contrôler les changements

de phase. Cet exemple tend à démontrer qu'il est donc indispensable de disposer d'une mesure de pression pour le système cardiaque complet.

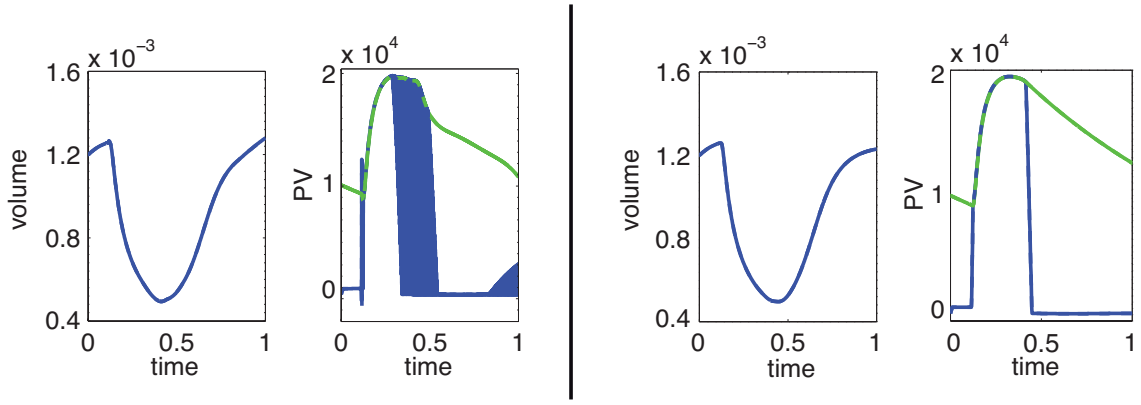


FIGURE 3.39: Cas d'estimation de déplacement et pression par filtre de Kalman pour un système cardiaque 0D**. A droite on dispose de mesures de pression mais pas à gauche ou la pression estimée est instable**

Si cette donnée de pression disponible est, grâce à un cathéter, la pression ventriculaire $P^v_{\mathfrak{m}}$, il faut imaginer un filtre valide y compris en grande dimension. Dans la configuration où les valves artérielles sont ouvertes et en supposant le modèle de valve parfait, c'est-à-dire sans fuite, le système cardiaque complet se résume alors à l'équation de la dynamique et au Windkessel déjà présenté lors de la mise en place des schéma en temps du problème direct

$$\begin{cases} M\ddot{Y} + C(Y)\dot{Y} + F^i(Y) = -P^{ar}S(Y) \\ \dfrac{L_wC_w}{R_w}\ddot{P}^{ar} + C_w\dot{P}^{ar} + \dfrac{1}{R_w}(P^{ar} - P_{sv}) = S(Y)^T\dot{Y} \end{cases} \tag{3.108}$$

Dans ce cas, le filtre de Luenberger s'écrit en plus du filtre sur les déplacements ou vitesses

$$\begin{cases} \dot{\hat{X}} = A(\hat{X},t) + B(\hat{X})\hat{P}^{ar} + \gamma H'(Z - H\hat{X}) \\ \dfrac{L_wC_w}{R_w}\ddot{\hat{P}}^{ar} + C_w\dot{\hat{P}}^{ar} + \dfrac{1}{R_w}(\hat{P}^{ar} - P_{sv}) = -B(\hat{X})^T\hat{X} + \gamma_P(P_{\mathfrak{m}} - \hat{P}^{ar}) \end{cases} \tag{3.109}$$

avec $B(X) = \begin{pmatrix} 0 \\ -M^{-1}S(X^d) \end{pmatrix}$. L'erreur linéarisée devient alors

$$\begin{cases} \dot{\tilde{X}} = A\tilde{X} + B\tilde{P}^{ar} - \gamma H'H\tilde{X} \\ \dfrac{L_wC_w}{R_w}\ddot{\tilde{P}}^{ar} + C_w\dot{\tilde{P}}^{ar} + \dfrac{1}{R_w}\tilde{P}^{ar} = -B^T\tilde{X} - \gamma_P\tilde{P}^{ar} \end{cases} \tag{3.110}$$

avec le bilan énergétique suivant :

$$\frac{\mathrm{d}}{\mathrm{dt}}(\mathcal{E}_c + \mathcal{E}_m + \tfrac{1}{2}(\tilde{P}^{ar})^2) = -\gamma\tilde{X}H'H\tilde{X} - \gamma_P(\tilde{P}^{ar})^2. \tag{3.111}$$

Ce bilan assurera une exponentielle stabilité du système cardiaque complet. Le même type de raisonnement permet de formuler un estimateur quand c'est la valve auriculaire qui est ouverte que ce soit lorsqu'on dispose d'un modèle de pression auriculaire ou qu'elle est donnée comme une fonction du temps (i.e mesure parfaite dans ce cas !). Cependant, lors des phases

isovolumiques, la pression ventriculaire est un multiplicateur de Lagrange ce qui change fondamentalement la nature du couplage pressions - déformations, donc il est plus difficile d'imaginer un estimateur.

Dans l'hypothèse où la donnée de pression n'est pas disponible, on peut essayer de se contenter du filtre mécanique $\gamma H'(Z - H\hat{X})$ seul. Le système de l'erreur linéarisée assure toujours la stabilité (faire $\gamma_P = 0$ dans (3.111)) mais l'exponentielle stabilité reste à démontrer.

REMARQUE 3.5.2 (PRESSION AU DOIGT)
Reste enfin la configuration où il n'y a pas eu pose de cathéter dans l'examen médical et où, nous en parlions au chapitre 2, il est possible d'utiliser une mesure de pression au doigt P^f_{m}. Elle est intégrée au processus d'assimilation de données

- soit sous forme d'opérateur d'observation distant, mais dans ce cas l'observateur de Luenberger est à imaginer ;
- soit en couplant au modèle existant le modèle de propagation dans l'arbre artériel de Crépeau et Sorine (2007) ; Laleg et al (2007), afin de conserver un estimateur simple de la forme $\gamma_P(P^f - P^f_{\mathrm{m}})$.

b. Cas pratique à valve aortique ouverte

□ POSITION DU PROBLÈME – Après ces premières remarques sur le traitement des variables de pression, nous choisissons comme exemple final un cas d'estimation à valves artérielles ouvertes. Nous supposons que la pression aortique est mesurée sans erreur et donc nous la négligeons en tant que variable d'état du système. Le reste du modèle cardiaque est conservé, y compris la loi de contraction *Clément-Bestel-Sorine* avec les valeurs numériques du chapitre 2 notamment une viscosité relativement plus importante que dans nos problèmes tests. Les indicateurs apparaissent Figure 3.40 sur un intervalle de simulation de 0.6 secondes correspondant approximativement à la systole. Enfin pour l'erreur de condition initiale dans l'estimation, nous avons choisi une erreur en surpression de 900 Pa sur les deux cavités.

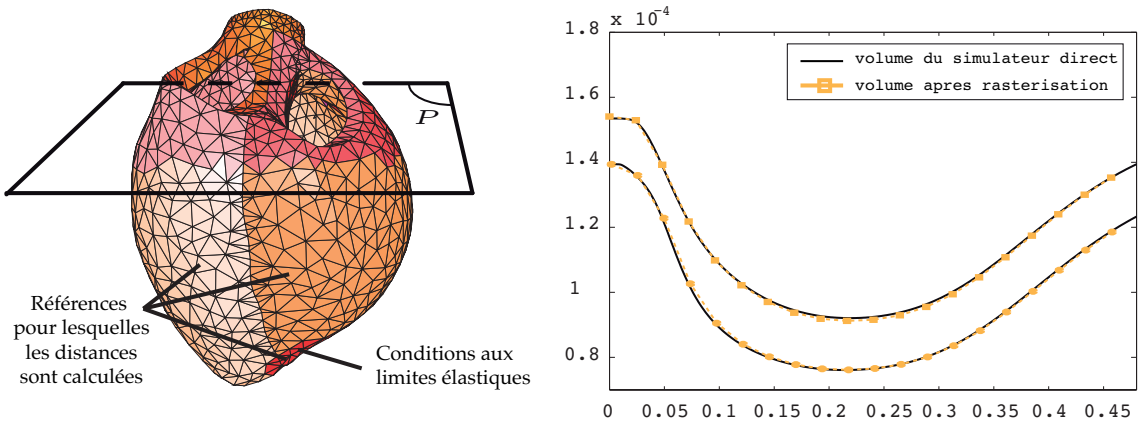


FIGURE 3.40: Maillage (référencé) utilisé pour les simulations directes et courbes de volumes associées.

Côté observations nous souhaitons de même nous placer dans des conditions très proches de l'application. Nous considérons donc des mesures sous forme de surfaces issues de segmentations. Mais, plus que dans notre problème-test, nous avons imaginé un bruit de mesure le

plus réaliste possible. Pour ce faire, nous effectuons des rastérisations de la simulation directe toutes les $24ms$ pour une résolution spatiale d'un demi-millimètre. Ces rastérisations correspondent donc à des résultats idéaux de segmentation mais introduisent une erreur liée à la pixélisation de l'image (cf. Figure 2.53). Figure 3.40, nous avons d'ailleurs superposé aux calculs de volume en sortie du simulateur direct des calculs de volume à partir des rastérisations comme ils auraient pu être effectués après une segmentation en milieu clinique.
Comme notre algorithme réalisant l'estimation en configuration Eulerienne nécessite des données de surface sous forme de maillage triangulaires, nous avons alors reconstitué des maillages à partir des images rastérisées. A cette fin, nous n'avons pas utilisé d'algorithme de marching-cube mais les outils développés par Pons et al (2007) et leur code `CGAL`. Nous remercions M. Yvinnec et toute l'équipe-projet Géometrica de l'INRIA dirigée par J.D. Boissonnat pour avoir réalisé les maillages à partir des images de la simulation directe. Figure 3.41, nous donnons un exemple de maillage issu de `CGAL` comparé au maillage initial en sortie de la simulation directe. La qualité du maillage (voir un exemple Figure 3.41) n'a pas besoin d'être optimisée côté mesure puisque le maillage n'est pas utilisé directement comme maillage de calcul du problème inverse mais uniquement afin de calculer des distances. Les maillages `CGAL` échantillonnés toutes les $24ms$ représentent pour nous des mesures parfaitement admissibles à l'issue de la segmentation d'un cycle cardiaque. La seule différence avec une configuration réelle est donc le fait que le modèle est parfait (i.e. nous ne présenterons d'ailleurs pas de résultats en fonction de la discrétisation puisque nous avons déjà assuré son contrôle dans les autres cas-tests). A titre de comparaison nous utiliserons aussi les mesures de surfaces directement issues du problème direct d'où nous avons extrait soit l'endocarde et l'épicarde soit uniquement l'épicarde.

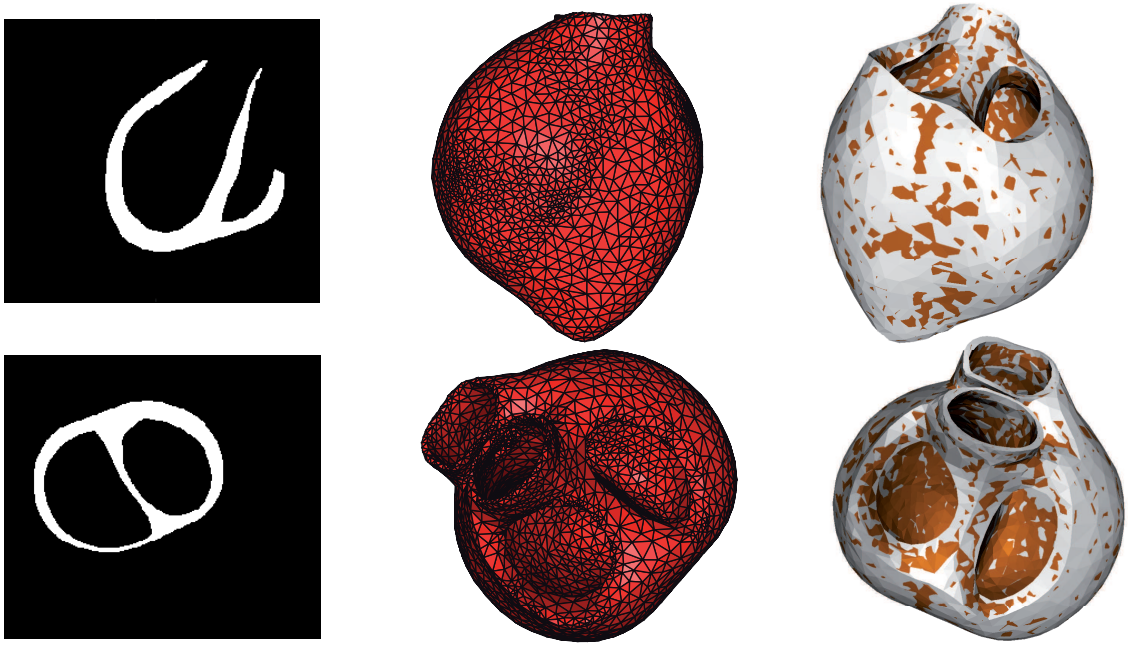


FIGURE 3.41: Données d'observations utilisées : rastérisation puis maillage direct de l'image avec `CGAL`. A droite, comparaison avec les surfaces idéales issues de la simulation directe

□ DIFFICULTÉS PRATIQUES – Nous avons été confrontés à une difficulté très bien connue de ceux qui essaient de faire de la segmentation de séquences d'images cardiaques à partir de modèle qui consiste à se tromper de surface sur laquelle l'estimateur essaie de tirer (cf. Sermesant (2003)). En effet, quand l'erreur d'état est importante[10] et que la donnée se limite à la surface globale du cœur (endocarde et épicarde) sans référencement des différentes parties intérieures et extérieures il peut arriver qu'un point de l'endocarde (resp. l'épicarde) de l'observateur apparie un point de l'épicarde (resp. l'endocarde) de l'observation. Afin d'éviter au maximum cette situation, nous avons déjà limité nos recherches de distances aux points de l'épicarde situés sous le plan de coupe géométrique introduit Section 2.4.1 (plan représenté Figure 3.40). Les maillages `CCGAL` sont conservés en l'état mais sur les surfaces idéales nous avons pu, soit les conserver entièrement, soit en extraire l'épicarde grâce aux références déjà présentes dans le maillage du problème direct. Pour une application réelle ce type de difficulté et les résultats observés Figure 3.44 démontrent la nécessité du référencement des surfaces après la segmentation. Ensuite il faut cumuler deux estimateurs, l'un d'endocarde à endocarde et l'autre d'épicarde à épicarde.

□ RÉSULTATS – Nous illustrons les performances de l'algorithme en montrant Figure 3.42 trois instants de la simulation. Le contour estimé est représenté en orange et il « colle » effectivement très rapidement à la donnée observée (contour bleu). Cette comparaison intervient dans l'espace d'observation et se retrouve dans l'estimation des courbes de volume de la Figure 3.43.

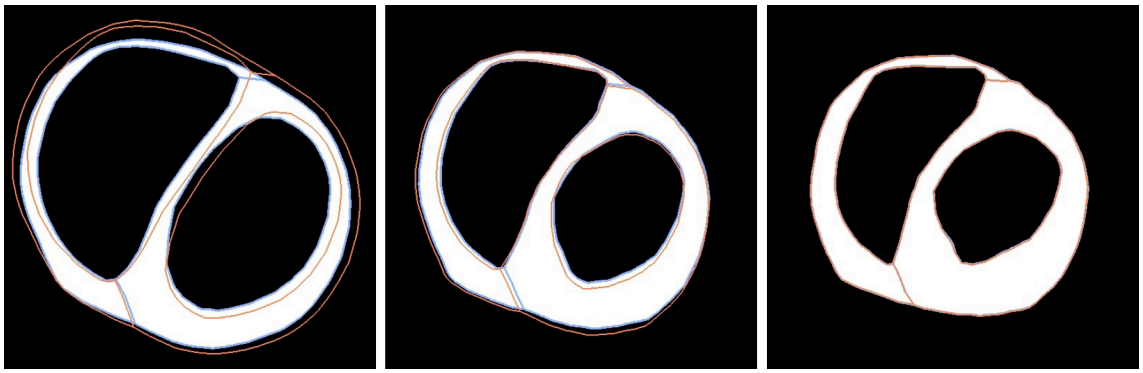


FIGURE 3.42: Contours estimés (orange) comparés avec la segmentation à $t = 0,\ 100,\ 200$ **ms**

Cependant, les erreurs doivent être comparées dans l'espace d'état et non l'espace d'observation ce qui, on le rappelle, n'est possible qu'en estimation sur données synthétiques. On obtient alors les courbes de la Figure 3.44 présentant les caractéristiques suivantes :

- l'erreur décroît exponentiellement avec une constante de temps très courte (voir aussi le zoom Figure 3.44). Le choix de γ a été effectué à partir du problème linéarisé pour lequel nous avons obtenu jusqu'à $\sup Re\lambda = -70s^{-1}$;
- l'erreur décroît jusqu'au palier de l'erreur de mesure en comparaison des autres estimations à surfaces idéales ;

[10] ou que l'échantillonnage des mesures est important

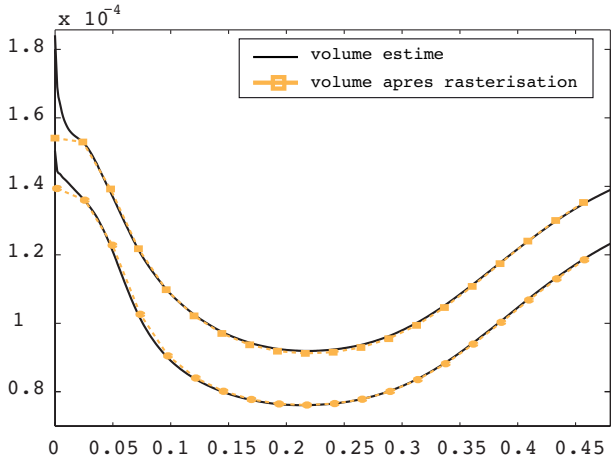


FIGURE 3.43: Estimation du volume à partir de l'observateur et comparaison avec les volumes obtenus directement « en comptant les pixels » dans l'image

- on constate des oscillations dans l'erreur nettement visibles ici car l'échantillonnage entre les images est très important (2.5 fois plus grand que dans le cas-test précédent) ; Ainsi l'interpolation linéaire des distances aux images successives introduit une erreur visible au travers de ces oscillations ;
- les quelques pics sur la courbe d'erreur sont caractéristiques de la difficulté évoquée plus haut où certains points de l'épicarde vont être appariés à des points de l'endocarde.

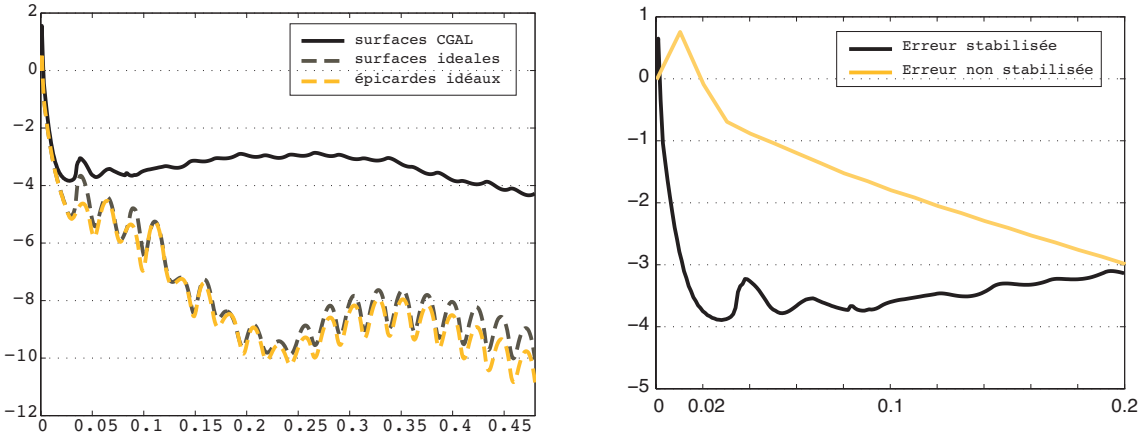


FIGURE 3.44: A gauche, différentes courbes d'erreur d'estimation pour les surfaces CGAL issues de la segmentation ou les surfaces idéales. A droite, taux de décroissance de l'erreur. Le palier rapidement atteint par l'estimateur correspond à l'erreur générée par le bruit de mesure.

Nous profitons aussi de ce dernier cas « réaliste » pour montrer Figure 3.45 sur le premier échantillon d'image les corrections d'images surfaciques. Les corrections visualisées sont les termes avant le relèvement. On rappelle de plus qu'à la différence de forces images classiques elles apparaissent dans la dynamique au niveau de la première équation donnant initialement

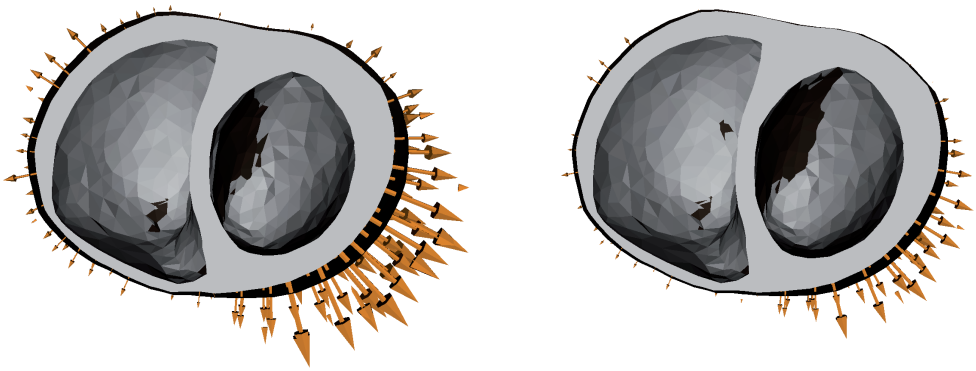


FIGURE 3.45: Deux exemples de corrections surfaciques introduites par les distances aux surfaces cibles (en noires)

la vitesse comme dérivée du déplacement. L'épicarde cible apparaît en noir et les corrections sont effectivement d'autant plus importantes que la distance est grande.

3.6 Conclusion et perspectives

Dans ce chapitre nous avons donc développé un ensemble de filtres efficaces en mécanique que ce soit dans la configuration où les mesures disponibles sont des vitesses ou des déplacements, volumiques ou sur les frontières. Ces méthodes consistent à développer un observateur de Luenberger naturel (également relié au méthode de « Forward Nudging » en assimilation de données, cf. Blum et al (2008)) à partir de l'opérateur d'observation, grâce notamment au fait que l'estimateur est un objet numérique sur lequel nous avons toute latitude pour positionner les actionneurs. Les performances de ces filtres ont été démontrées dans le cas de mesures de vitesses à partir des résultats existant nombreux en théorie du contrôle sur de tels systèmes. Ils ont aussi été justifiés par des résultats numériques de calcul de spectre dans le cas de mesures de déplacements avec des pistes pour assurer le même type de résultat théoriques que pour les vitesses.

Ces filtres développés au départ sur des systèmes linéaires s'étendent à la mécanique non-linéaire et à des observateurs eux aussi non-linéaires. Ainsi, nous avons été capables de formuler des méthodes originales d'estimation à partir d'une séquence de surfaces. Ces dernières peuvent représenter différentes segmentations d'une séquence IRM ou échographique et permettent alors en contexte clinique d'assurer une cohérence cinématique et mécanique à ces extractions de surfaces. L'originalité par rapport à ce que certains utilisent déjà en segmentation à base de modèle est, encore une fois, d'avoir placé la correction image non pas comme une force mais comme une correction de la dérivée en temps du déplacement. C'est pour nous une différence majeure qui permet de justifier la convergence de notre estimateur et qui le positionne plus dans la problématique de l'assimilation de données que du traitement d'images.

Nous insistons enfin sur le fait que les résultats présentés ont une portée qui peut dépasser le cadre de l'estimation en cardiologie par deux aspects :

- il existe d'autres problématiques d'estimation, y compris à base d'images, en mécanique

des structures comme par exemple ce que propose Hild et Roux (2006) ;

- les résultats présentés sont valides dans le cadre très général des systèmes hyperboliques du second ordre donc pour des modèles d'acoustique, ou vibratoires en général.

De plus cette notion de filtre numérique est un outil qui peut globalement être mis en place dans n'importe quelle situation car il favorise la diminution de l'erreur numérique en recalant le système sur le contenu fréquentiel du système réel. Autrement dit c'est un outil d'analyse numérique à part entière qui ne génère pas de développements algorithmiques hors de portée si on peut imaginer des filtres « physiques », donc développables à partir des briques du code direct, aussi efficaces et plus robustes que ceux, plus lourds dans leur mise en oeuvre, du contrôle optimal.

CHAPITRE

4

Estimation conjointe état-paramètres pour les systèmes distribués

On mesure l'intelligence d'un individu à la quantité d'incertitudes qu'il est capable de supporter

— Emmanuel Kant

Sommaire du chapitre

4.1 Introduction **246**

4.2 Cas linéaire **248**

4.2.1 Approche variationnelle 249

4.2.2 Estimation état-paramètre par filtrage réduit 250

4.2.3 Application à l'estimation paramétrique 252

4.2.4 Analyse de l'estimateur 255

4.2.5 Schéma numérique associé et analyse de stabilité 263

4.2.6 Résultats 265

4.2.7 Opérateur frontière et Formulation H infini 267

4.3 Cas bilinéaire et extensions du filtre SEEK **276**

4.3.1 Expression à partir de la sensibilité 276

4.3.2 Analyse d'erreur 277

4.3.3 Premiers résultats 277

4.3.4 Amélioration des performances de l'estimateur conjoint en non-linéaire 280

4.3.5 Equivalence avec EKF et le filtrage non-linéaire 282

4.4 Cas non-linéaire et filtre SEIK **283**

4.4.1 Filtre SEIK 284

4.4.2 Application à l'estimation paramétrique 293

4.5 Perspectives autour du système cardiaque **299**

4.5.1 Estimation pour *des paramètres dynamiques* 299

4.5.2 Implications pour l'estimation cardiaque 301

4.1 Introduction

L'Estimation d'état est une notion fondamentale en météorologie puisqu'au fond la préoccupation principale est l'état du système, c'est-à-dire le temps qu'il fait (diagnostic) et donc qu'il va faire (pronostic). En mécanique, cependant, même si nous pouvons donner quelques exemples de besoins en estimation d'état, il apparaît très vite que le diagnostic est plus centré sur l'identification des paramètres physiques de la loi de comportement ou des chargements. On pourra consulter Bui (1994) pour une introduction complète des problématiques de problèmes inverses en mécanique. Pour schématiser, on peut dire que même si en météorologie on prend en compte des erreurs de modèles, les paramètres physiques associés à un modèle de comportement de l'atmosphère sont souvent bien connus. *A contrario* en mécanique, il est difficile de faire de l'identification sur une structure réelle loin des expériences en éprouvettes.

L'identification paramétrique en mécanique cardiaque est donc l'objectif principal de l'estimation, que ce soit pour identifier quantitativement les lois de comportements dans des configurations saines ou pathologiques. Bien que moins développés par rapport à la modélisation directe, on peut citer les travaux de Augenstein et al (2005) à partir de méthodes variationnelles sur des fantômes imagés par IRM taggés. Ils identifient les paramètres d'une loi de comportement hyperélastique (passive) pour un tissu cardiaque en fonctionnement nominal. Les méthodes séquentielles ont aussi été mises en application par Tong et Shi (2006), avec un filtre H^∞ sur données synthétiques ou à partir de mesures de déplacements extraits d'images IRM de cœurs de chien. Les résultats se limitent à des coupes 2D de tissu puisqu'ils reconnaissent la complexité trop importante d'un filtre H^∞ optimal en 3D.

□ Dimension de l'espace paramétrique – la localisation spatiale des paramètres est clairement moins distribuée que celle de l'état. Autrement dit, même si les équations permettent dans l'absolu une description paramétrique distribuée, on considère que les paramètres sont plus homogènes que les champs de déplacements et vitesses de l'état. De plus, le nombre de paramètres n'a pas vocation à tendre vers l'infini comme les degrés de libertés de l'état car il est directement assujetti à l'identifiabilité du problème. Or une représentation paramétrique spatiale trop dense, notamment par rapport à la densité des mesures, pourrait « noyer » l'information disponible. Reste que fixer des paramètres à des valeurs erronées peut, d'un autre côté, entraîner une erreur de modèle irréversible. Ainsi le choix du nombre de paramètres dans la modélisation, c'est-à-dire le dimensionnement du problème est absolument critique. En estimation cardiaque, nous envisageons pouvoir estimer au maximum quelques paramètres (<5 maximum) pour quelques régions (<100 maximum). On emploie donc un vecteur $\theta \in \mathbb{R}^p$ de dimension finie dans lequel on agrège les paramètres des différentes zones d'intérêt, avec p de l'ordre de la centaine. Pour le coeur nous avons choisi, pour des raisons de commodité qui n'ont aucune incidence sur la méthodologie, de définir la distribution paramétrique constante par morceaux ($\mathbb{P}_0$) en fonction des régions AHA définies Section 2.2.1. Nous pourrions tout autant définir les paramètres via des fonctions linéaires par morceaux ($\mathbb{P}_1$) sur un maillage grossier de quelques éléments.

□ Nature des paramètres à identifier – Puisque dans θ nous ne choisissons que des paramètres incertains pertinents dans l'ensemble des paramètres du modèle, le choix est donc contraint par l'application souhaitée. Pour la mécanique cardiaque, l'identification doit notam-

ment servir le diagnostic de pathologies mécaniques relatives à une ischémie, c'est-à-dire à la diminution de l'apport en sang dans l'organe. Les ischémies ont principalement un effet sur

- certaines concentrations en $[ATP]$, $[K+]$, contenues pour nous dans la variable $u(t)$,
- la capacité à se contracter représentée dans le modèle par σ_0,
- la partie passive, c'est-à-dire le module d'Young ou κ_1 dans la loi hyperélastique, car de nombreux tissus passifs viennent « coloniser » les zones musculaires endommagées.

Du point de vue de l'identifiabilité, nous pouvons déjà noter que $u(t)$ et σ_0 apparaissant toujours en produits et il ne peuvent donc pas être identifiés séparément. Ceci traduit l'ambiguïté qui peut exister entre les infarctus de nature profondément électrique à ceux de nature mécanique. Nous n'envisagerons que dans les perspectives de cette thèse l'estimation de $u(t)$ en fonction du temps (notamment car elle nécessite la prise en compte du modèle électrique donc d'un état couplé électro-mécanique hors de portée de notre sujet) et nous nous concentrons sur l'identification de paramètres au sens classique. Ainsi, les seules incertitudes seront pour nous d'ordre strictement mécanique et s'appliqueront à la contractilité σ_0 ou aux raideurs passives.

L'objectif de ce chapitre est alors de proposer des méthodes de filtrages pour l'identification paramétrique d'un vecteur de paramètres qui, vu sa dimension, rend envisageable n'importe quelle méthode de filtrage, y compris Kalmanienne. Cependant, même si l'estimation d'état n'est pas l'objectif principal devant l'identification paramétrique, nous allons montrer qu'elle est en fait indispensable. On peut déjà rappeler deux avantages de l'estimation d'état évoqués au chapitre précédent :

- elle permet de rattraper les éventuelles erreurs de conditions initiales pour des systèmes mécaniques où l'amortissement naturel n'offrirait pas de garanties suffisantes de stabilité pour l'estimation ;
- elle permet de gérer éventuellement les erreurs de modèles ou le bruit de mesures. En effet, à tout instant, celles-ci agissent de façon à produire sur l'état une erreur de condition initiale pour le système évoluant à partir de cet instant.

Cette notion d'**estimation conjointe état-paramètres** est donc fondamentale et nous verrons, dès les premiers paragraphes dans la formulation moindres carrés, qu'elle traduit le caractère bien posé du problème d'estimation. Ainsi en pratique, faire de l'identification paramétrique sans stabiliser l'état, c'est avant tout miser sur l'absence de bruit sur le système et donc imaginer des algorithmes qui ne fonctionnent que sur données synthétiques parfaites. Nous verrons notamment Section 4.3 un exemple frappant dans une situation par ailleurs très simple en terme de modélisation.

Cette approche conjointe état-paramètres nous conduit ainsi à une formulation dite d'**état étendu** introduite en filtrage dans le classique *filtrage de Kalman étendu*. Elle consiste à considérer le modèle dynamique d'état $x(\zeta^x, \theta)$

$$\begin{cases} \dot{x} = \mathcal{A}(x, \theta) \\ x(0) = x_0 + \zeta^x \end{cases}$$

où le vecteur de paramètres θ peut être appréhendé de manière probabiliste ou déterministe. Fidèles à notre formalisme nous distinguons dans θ notre a priori, ici sa moyenne statistique

$$\theta = \mathbb{E}(\theta) + (\theta - \mathbb{E}(\theta)) = \theta_0 + \zeta^\theta.$$

Nous pouvons alors réécrire les variables du système sous la forme $x^e(\zeta^x, \zeta^\theta) = \left(\begin{smallmatrix} x \\ \theta \end{smallmatrix}\right)$ vérifiant la dynamique étendue

$$\begin{cases} \dot{x} = \mathcal{A}(x, \theta) \\ \dot{\theta} = 0 \\ x(0) = x_0 + \zeta^x \\ \theta = \theta_0 + \zeta^\theta \end{cases}$$

avec x un champ et $\theta \in \mathbb{R}^p$ un vecteur paramétrique de dimension finie. Cette formulation sur la variable x^e permet « de se ramener au problème précédent » de l'estimation d'état. Ainsi il est possible d'obtenir directement les stratégies variationnelles ou Kalmanienne pour ce problème, ce qui conduit d'ailleurs à dire que, formellement, l'identification paramétrique n'est qu'une sous classe de problème d'observation. Cependant il est plus difficile de développer une stratégie autour d'observateurs de Luenberger pour ce nouvel état puisque la dynamique des paramètres est un peu artificielle et on perd la propriété d'antisymétrie de l'opérateur (conservatif) global. Autrement dit, ce que nous avons gagné au chapitre 3 en terme de mise en oeuvre pour les filtres d'état sur les systèmes hyperboliques du second ordre, nous l'avons perdu en terme d'extension directe à l'identification. La problématique de ce chapitre est donc de savoir **comment nous pouvons cumuler les bénéfices des observateurs de Luenberger sur les variables d'état de grande dimension avec une identification paramétrique conjointe**.

4.2 Cas linéaire

Comme toujours, nous commençons par étudier le cas linéaire conjointement pour l'état et les paramètres. Autrement dit, la dynamique du système mécanique sera donnée dans un premier temps par

$$\dot{x} = \mathcal{A}x + \mathcal{B}\theta + \mathcal{R}, \tag{4.1}$$

dont la discrétisation en espace est notée

$$\dot{X} = AX + B\theta + R. \tag{4.2}$$

Par rapport aux différents modèles mécaniques que nous avons déjà manipulés, nous précisons ici la dépendance paramétrique du second membre $R \to B\theta + R$ décrivant, pour les opérateurs étudiés, une dépendance paramétrique du chargement mécanique $F^{\text{ext}} \to B\theta + F^{\text{ext}}$. Par exemple, pour notre modèle[1] linéaire de contraction vu au chapitre 3, nous pouvons envisager estimer σ_0 dans le terme de précontrainte. En effet, on peut préciser la dépendance de σ_0 sur les différentes régions AHA sous la forme

$$\mathcal{P}^{\text{ext}} = \sum_{1 \leq i \leq 17} \int_{\Omega_i^{AHA}} \theta_i \sigma_0 w(\underline{x}, t) \ \text{tr}(\nabla \underline{v}) \, d\Omega = \delta Y^T \cdot B\theta. \tag{4.3}$$

[1] nous avions précisé que c'était un problème-test plutôt qu'un véritable modèle cardiaque

On rappelle aussi que, dans ce problème-test, nous n'avions pas pris en compte d'autres forces extérieures donc $\mathcal{R} = 0$ dans (4.1).

C'est l'occasion de préciser pourquoi ce problème nous intéresse dans la perspective de l'estimation cardiaque. Nous avons vu qu'une condition suffisante à la stabilisation (pour nous au sens de l'erreur d'observation) de problèmes non-linéaires est de prouver la stabilité du linéarisé. Ce n'est seulement qu'en cas d'échec de cette première stratégie qu'il faut regarder plus finement les propriétés intrinsèques du non-linéaire comme l'explique en détails Coron (2007). Or, dans le cas d'une estimation de la contractilité σ_0 pour le modèle cardiaque complet *Bestel-Clément-Sorine*, la linéarisation par rapport à l'état étendu déplacement, vitesses et paramètres fait apparaître un opérateur de sensibilité paramétrique ne dépendant pas (par définition de la linéarisation) du déplacement et de la vitesse. Ce terme est donc bien une précontrainte au sens mécanique.

En termes de notations, nous utiliserons pour l'état étendu du problème linéaire les opérateurs

$$\begin{cases} \dot{X}^e = A^e X^e + R^e \\ X^e(0) = X_0^e + \zeta^e \end{cases} \tag{4.4}$$

où

$$X^e = \begin{pmatrix} X \\ \theta \end{pmatrix}, \quad A^e = \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix}, \quad R^e = \begin{pmatrix} R \\ 0 \end{pmatrix}, \tag{4.5}$$

et

$$X_0^e = \begin{pmatrix} X_0 \\ \theta_0 \end{pmatrix}, \quad \zeta^e = \begin{pmatrix} \zeta^X \\ \zeta^\theta \end{pmatrix}. \tag{4.6}$$

Suivant ces notations l'objectif d'estimation conjointe état-paramètres est donc ramené à un objectif d'estimation d'état et plus particulièrement de condition initiale ζ^e.

4.2.1 Approche variationnelle

a. Critère

Nous reprenons l'approche contrôle optimal classique comme référence pour nos estimateurs futurs. Le critère variationnel devient pour l'état étendu (d'ailleurs y-compris en non-linéaire)

$$J(\xi^X, \xi^\theta) = \frac{1}{2}\|\xi^X\|^2_{(P^X)^{-1}} + \frac{1}{2}\|\xi^\theta\|^2_{(P^\theta)^{-1}} + \frac{1}{2}\int_0^T \|Z - HX\|^2_{W^{-1}}\, dt, \tag{4.7}$$

qui est bien la reformulation de

$$J(\xi^e) = \frac{1}{2}\|\xi^e\|^2_{(P^e)^{-1}} + \frac{1}{2}\int_0^T \|Z - H^e X^e\|^2_{W^{-1}}\, dt,$$

dans le cas où la norme sur la condition initiale est tensorisée. Les bruits de condition initiale sur l'état et les paramètres sont donc jugés indépendants. Le gradient du critère est obtenu en introduisant la variable adjointe de la dimension de l'état étendu $p^e = \left(\begin{smallmatrix} p^X \\ p^\theta \end{smallmatrix}\right)$ dont les équations

(adjointes) sont

$$\begin{cases} \dot{p}^X + A^T p^X = -H^T W^{-1}(Z - HX) \\ p^X(T) = 0 \\ \dot{p}^\theta + B^T p^X = 0 \\ p^\theta(T) = 0 \end{cases}$$

On obtient alors

$$\mathrm{d}_{\xi^\theta} J.\delta\xi^\theta = {\xi^\theta}^T (P^\theta)^{-1} \delta\xi^\theta - \int_0^T {p^X}^T B \delta\xi^\theta \, dt = {\xi^\theta}^T (P^\theta)^{-1} \delta\xi^\theta - p^\theta(0)^T \delta\xi^\theta, \tag{4.8}$$

et

$$\mathrm{d}_{\xi^X} J.\delta\xi^X = {\xi^X}^T (P^X)^{-1} \delta\xi^X - p^X(0)^T \delta\xi^X. \tag{4.9}$$

On remarque dans ces expressions la similitude formelle entre l'adjoint sur les paramètres et ce que nous avions obtenu dans le cas d'un bruit de modèle. En effet, permettre la variation paramétrique est une façon de modéliser l'incertitude de modèle par rapport à l'état classique.

b. Kalman étendu

On peut effectuer le même type de formulation par filtrage à partir de ce critère et obtenir ce qui est classiquement appelé *Filtre de Kalman étendu*

$$\begin{cases} \dot{X} = AX + B\theta + P^{XX} H^T W^{-1}(Z - HX) \\ \dot{\theta} = P^{\theta X} H^T W^{-1}(Z - HX) \end{cases} \tag{4.10}$$

La matrice de covariance

$$P^e = \begin{pmatrix} P^X & P^{X\theta} \\ P^{\theta X} & P^\theta \end{pmatrix},$$

suit toujours l'équation de Riccati, par exemple en l'absence de bruit de modèle,

$$\dot{P}^e = A^e P^e + P^e A^{eT} - P^e H^{eT} (W^e)^{-1} H^e P^e.$$

On vérifie alors directement que les blocs de P vérifient les équations dont la structure réapparaîtra plus tard dans l'exposé

$$\begin{cases} \dot{P}^\theta = -P^{\theta X} H^T W^{-1} H P^{X\theta} \\ \dot{P}^{X\theta} = AP^{X\theta} + BP^\theta - P^X H^T W^{-1} H P^{X\theta} \\ \dot{P}^X = AP^X + BP^{\theta X} + P^X A^T - P^{X\theta} B^T + P^X H^T W^{-1} H P^X \end{cases}$$

4.2.2 Estimation état-paramètre par filtrage réduit

Le filtrage de Kalman offre cet avantage de fournir pour n'importe quel système un filtre efficace qui s'étend naturellement au cas de l'estimation conjointe état-paramètre. Cependant, sa mise en oeuvre numérique est hors de portée en raison des champs de déplacements et vitesses. Pour l'état seul, nous avions choisi au chapitre 3 d'orienter notre recherche de filtre autour des propriétés mécaniques intrinsèques du système mais la question se pose désormais

de savoir si ce fut au détriment d'extensions potentielles aux paramètres. La réponse est non et elle est fournie par les *filtres d'états réduits*, en particulier ceux proposés par Pham et al (1997). Le principe des filtres réduits est de supposer qu'une partie de l'état est stable et donc ne nécessite pas de filtrage supplémentaire. Nous allons nous inspirer de cette idée pour construire un filtre à deux étages sur l'état augmenté *état-paramètres*. L'idée est d'utiliser l'observateur de Luenberger développé précédemment pour restreindre la covariance d'erreur aux paramètres et appliquer alors un filtre réduit. Dans la communauté automaticienne, cette démarche de construction de filtres paramétriques comme supplément du filtre d'état est souvent appelée *filtrage adaptatif*. La présentation qui suit reprend largement l'article Moireau et al (2008b) en lui ajoutant quelques compléments numériques, notamment des précisions sur le schéma en temps, et des extensions.

a. Filtrage de rang réduit

□ CAS CONTINU EN TEMPS – Supposons que P est symétrique de rang p, alors il existe une décomposition de type *Singular Value Decomposition* SVD de la forme

$$P^e = L^e U^{-1} L^{eT},$$

avec $L^e \in \mathbb{M}_{n,p}(\mathbb{R})$ et $U \in \mathbb{M}_p(\mathbb{R})$ inversible. Essayons de déterminer une équation différentielle remarquable sur chacune de ces matrices à travers le Riccati vérifié par P. On a

$$\begin{aligned}\dot{L}^e U^{-1} L^{eT} + L^e U^{-1} (\dot{L}^e)^T - L^e U^{-1} \dot{U} U^{-1} L^{eT} = \\ A^e L^e U^{-1} L^{eT} + L^e U^{-1} L^{eT} A^{eT} - L^e U^{-1} L^{eT} H^T W^{-1} H L^e U^{-1} L^{eT}. \quad (4.11)\end{aligned}$$

On peut chercher L^e satisfaisant la dynamique

$$\dot{L}^e = A^e L^e,$$

permettant la mise en facteur de L^e de part et d'autre de l'identité. On obtient alors une équation admissible sur $U \in \mathbb{M}_p(\mathbb{R})$

$$U^{-1} \dot{U} U^{-1} = U^{-1} L^{eT} H^T W^{-1} H L^e U^{-1} \Rightarrow \dot{U} = L^{eT} H^T W^{-1} H L^e.$$

Ces dynamiques sont compatibles avec la condition initiale P_0^e de rang p qui, elle-même, peut donc s'écrire sous la forme $L_0 U_0^{-1} L_0^T$. On a donc les conditions initiales respectives

$$L^e(0) = L_0^e, \quad U(0) = U_0.$$

Cette décomposition n'est pas unique, donc L^e et U ne sont pas uniques non plus au cours du temps. En revanche, l'unicité de la solution de l'équation de Riccati impose que $P = L^e U^{-1} L^{eT}$ est unique. Elle reste donc de rang réduit sur toute la fenêtre en temps. Cependant on constate dès maintenant que cette propriété n'est possible qu'en l'absence de bruit de modèle qui ajouterait à l'équation de Riccati une matrice, éventuellement pleine, S^e telle que

$$\dot{P}^e = P A^{eT} + A^e P^e - P^e H^{eT} W^{-1} H^e P^e + S^e.$$

Dans ce cas la décomposition n'est clairement plus possible sauf à projeter la matrice S^e sur le sous-espace réduit.

□ FILTRE SEEK – Historiquement, la démonstration de cette décomposition dans un contexte d'assimilation de données a initialement été effectuée en discret par Pham et al (1997). Ils en déduisent un filtre dit de rang réduit qui s'étend d'ailleurs au cas de l'EKF d'où sa dénomination de *filtre SEEK*. Le cas discret en temps offre donc la même décomposition de la matrice de covariance de prédiction[2]

$$P_n^{e-} = L_n^e U_{n-1}^{-1} L_n^{e\,T}, \tag{4.12}$$

avec

$$\begin{cases} L_{n+1}^e = A_{n+1,n}^e L_n^e \\ U_{n+1} = U_n + L_{n+1}^{e\,T} (H_{n+1}^e)^T W^{-1} H_{n+1}^e L_{n+1}^e. \end{cases} \tag{4.13}$$

La démonstration est calquée sur le continu à partir de l'équation récursive « one-step » (1.44) sur P_{n+1}^- rappelée ici avec les opérateurs augmentés

$$P_{n+1}^{e-} = A_{n+1,n}^e P_n^{e-} (A_{n+1,n}^e)^T - A_{n+1,n}^e P_n^{e-} H_n^{e\,T} (H_n^e P_n^{e-} H_n^{e\,T} + W_n)^{-1} H_n^e P_n^{e-} (A_{n+1,n}^e)^T \tag{4.14}$$

En effet, si on suppose que $P_k^{e-} = L_k^e U_{k-1}^{-1} L_k^{e\,T}$ pour $k = n, n+1$, alors on obtient directement comme conditions suffisantes de (4.14),

$$L_{n+1}^e = A_{n+1,n}^e L_n^e,$$

et

$$U_n^{-1} = U_{n-1}^{-1} - U_{n-1}^{-1} L_n^{e\,T} H_n^{e\,T} (H_n^e P_n^{e-} H_n^{e\,T} + W_n)^{-1} H_n^e L_n^e U_{n-1}^{-1}.$$

Or le lemme d'inversion matricielle 1.1.6 introduit au chapitre 1 simplifie directement cette équation en

$$U_n = U_{n-1} + L_n^{e\,T} (H_n^e)^T W_n^{-1} H_n^e L_n^e.$$

Enfin, notons que la covariance a posteriori s'exprime immédiatement en fonction des opérateurs U_n et L_n^e

$$P_n^{e+} = L_n^e U_n^{-1} L_n^{e\,T}. \tag{4.15}$$

REMARQUE 4.2.1 (INITIALISATION)
Comme au premier chapitre, en cas d'a priori et d'observations au temps initial, on a normalement

$$U_{n=0} = U_0 + L_0^{e\,T} (H_0^e)^T W_0^{-1} H_0^e L_0^e.$$

Cependant, le plus souvent les algorithmes en temps discret sont présentés avec un a priori qui prend déjà en compte l'observation initiale et nous avons respecté dans la suite cette convention. Ainsi la prédiction sera donnée en premier.

4.2.3 Application à l'estimation paramétrique

Reprenons notre système linéaire et définissons le système noté désormais $\bar{X}$ bénéficiant d'un filtre d'état développé au chapitre 3 en fonction des mesures Z de vitesses ou déplacements

[2]Nous avons utilisé une matrice U qui est l'inverse de celle de Pham et al (1997), afin d'insister sur la propriété de rang de P

fournies. Cet état convergerait vers le système cible si ses paramètres étaient corrects. Définissons alors le problème de l'adaptation paramétrique de ce système à travers l'écriture de la dynamique de l'état augmenté.

$$\begin{cases} \dot{\bar{X}}^e = A^e \bar{X}^e + R^e + K_X^e (Z - H^e \bar{X}^e) \\ \bar{X}^e(0) = X_0^e + \zeta_\theta^e, \end{cases} \tag{4.16}$$

où

$$\zeta_\theta^e = \begin{pmatrix} 0 \\ \zeta^\theta \end{pmatrix}, \quad K_X^e = \begin{pmatrix} K_X \\ 0 \end{pmatrix}, \quad H^e = \begin{pmatrix} H & 0 \end{pmatrix}.$$

L'idée serait d'appliquer le filtre de Kalman à ce système mais deux éléments nous empêchent une application immédiate

- Z n'est pas l'observation de ce système mais du système réel.
- Z apparaît au second membre de l'équation d'état donc introduit une forme de bruit de modèle sur l'état.

Nous choisissons donc de fixer à notre estimateur un nouveau système cible « virtuel » donné par

$$\begin{cases} \dot{\bar{X}}_a^e = A^e \bar{X}_a^e + R^e + K_X^e (\bar{Z} - H^e \bar{X}_a^e) \\ \bar{X}_a^e(0) = X_0^e + \zeta_\theta^e. \end{cases} \tag{4.17}$$

où $\bar{Z}$ est l'observation parfaite définie par $\bar{Z} = HX$. Associé à ce système, on considère les observation « virtuelles »

$$Z_a = H\bar{X}_a^e + \chi, \tag{4.18}$$

On définit $\hat{X}_a^e$, l'estimateur de Kalman correspondant à ces incertitudes et observations. La covariance associée à cet observateur

$$P_a^e \stackrel{\text{def}}{=} \mathbb{E}((\bar{X}_a^e - \hat{X}_a^e)(\bar{X}_a^e - \hat{X}_a^e)^T | Z_a) \tag{4.19}$$

vérifie alors

$$P_a^e(0) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{E}(\zeta^\theta \zeta^{\theta^T}) \end{pmatrix}. \tag{4.20}$$

Sur ce système, il est donc possible d'appliquer le filtre de Kalman réduit présenté au paragraphe précédent. En l'absence de bruit de modèle nous avons vu que la matrice de covariance se décomposait suivant une SVD. Ici on choisit $L^e(0) = \left(\begin{smallmatrix} 0 \\ \mathbb{1}_p \end{smallmatrix}\right)$ et $U(0) = \left(\mathbb{E}(\zeta^\theta \zeta^{\theta^T})\right)^{-1}$. La covariance reste de rang réduit p, dimension du vecteur paramétrique, et se décompose à partir de :

$$\begin{cases} P_a^e = L^e U^{-1} L^{eT} \\ \dot{L}^e = (A^e - K_X^e H) L^e \\ L^e(0) = (0 \quad \mathbb{1}_p)^T \\ \dot{U} = L^{eT} H^{eT} W^{-1} H^e L^e \\ U(0) = \left(\mathbb{E}(\zeta^\theta \zeta^{\theta^T})\right)^{-1} \end{cases} \tag{4.21}$$

Si nous décomposons maintenant $L = (\, L^X L^\theta \,)$ on obtient

$$\begin{cases} \dot{L}^X = (A - K_X H)L^X + BL^\theta \\ \dot{L}^\theta = 0 \\ L^X(0) = 0 \\ L^\theta(0) = \mathbb{1}_p \end{cases} \tag{4.22}$$

Ainsi $\forall t,\ L^\theta = \mathbb{1}_p$, donc seule la dynamique de L^X

$$\dot{L}^X = (A - K_X H)L^X + B, \tag{4.23}$$

et de U

$$\dot{U} = {L^X}^T H^T W^{-1} H L^X \tag{4.24}$$

sont nécessaires à la définition de l'observateur

$$\begin{cases} \dot{\hat{X}}_a = A\hat{X}_a + B\hat{\theta}_a + R + K_X(\bar{Z} - H\hat{X}_a) + L^X\dot{\hat{\theta}}_a \\ \dot{\hat{\theta}}_a = U^{-1}{L^X}^T H^T W^{-1}(Z_a - H\hat{X}_a) \\ \hat{X}(0) = X_0 \\ \hat{\theta}(0) = \theta_0 \end{cases} \tag{4.25}$$

Malheureusement, cet estimateur $\hat{X}_a$ ne peut pas être utilisé en l'état puisque ni $\bar{Z}$ ni Z_a ne sont disponibles. Cependant, une idée naturelle est de remplacer ces derniers par Z pour définir notre propre estimateur sous la forme

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B\hat{\theta} + R + K_X(Z - H\hat{X}) + L^X\dot{\hat{\theta}} \\ \dot{\hat{\theta}} = U^{-1}{L^X}^T H^T W^{-1}(Z - H\hat{X}) \\ \dot{L}^X = (A - K_X H)L^X + B \\ \dot{U} = {L^X}^T H^T W^{-1} H L^X \\ \hat{X}(0) = X_0 \\ \hat{\theta}(0) = \theta_0 \\ L^X(0) = 0 \\ U(0) = \left(\mathbb{E}(\zeta^\theta {\zeta^\theta}^T)\right)^{-1} \end{cases} \tag{4.26}$$

Cette substitution est légitime car Z et $\bar{Z}$ ne diffèrent que par le bruit χ, et Z_a converge vers Z avec la même constante de temps que l'estimateur d'état retenu. Cependant, il est possible d'aller plus loin que cette définition Kalmanienne un peu *ad-hoc*. Afin d'offrir un cadre théorique rigoureux à (4.26), nous pouvons fournir à la fois

- un nouveau critère variationnel de minimisation moindres carrés,
- une analyse de l'erreur de convergence de l'estimateur.

Cet estimateur existe déjà dans la littérature même si son lien avec le filtre SEEK n'avait pas été présenté avant Moireau et al (2008b). Il a été formulé dans Zhang et Clavel (2001) et

Zhang (2002) dans le cadre de l'identification adaptative pour des systèmes de faible dimension. L'intuition derrière la définition de Zhang (2002) se trouve dans la dynamique de L^X. En effet, L^X vérifie le même problème de Cauchy (même dynamique et même condition initiale) que la sensibilité de $\bar{X}$ par rapport aux paramètres soit

$$L^X = \frac{\partial \bar{X}}{\partial \theta}. \tag{4.27}$$

Cette identité justifie la structure $L^X\dot{\hat{\theta}}$ dans la dynamique de l'estimateur conjoint correspondant à $\frac{\partial \bar{X}}{\partial \theta}\dot{\hat{\theta}}$ afin d'exprimer le fait qu'un observateur adaptatif est souvent fondé sur la sensibilité du système par rapport aux paramètres supposés désormais variables. L'approche SEEK conduirait plutôt à écrire un terme Kalmanien du type $K^X_{\text{al}}(Z - H\hat{X})$, avec

$$K^X_{\text{al}} = L^X U^{-1} {L^X}^T H^T W^{-1}. \tag{4.28}$$

Il est remarquable que ces deux approches assez radicalement différentes, approches Kalmanienne pour le SEEK de Pham et al (1997), et approche adaptative pour Zhang (2002) conduisent, aux observations virtuelles Z_a près, au même estimateur. Ce sont ces deux approches qui vont nous fournir chacune de leur côté une justification de son efficacité.

Insistons pour achever cette formalisation sur le fait que l'ensemble des manipulations effectuées avaient pour but de définir un observateur où la covariance d'erreur se réduit à l'espace paramétrique. Cette question est sensiblement différente de l'objectif principal du filtre SEEK qui cherche à réduire la rang de la matrice aux premières valeurs singulières de la matrice de covariance. Autrement dit la base du filtre SEEK est initialement imaginée évolutive (sauf extensions proposées dans Hoteit (2001)). Nous voyons dans cette différence une des causes probables du non rapprochement entre le filtre SEEK et le filtre adaptatif de Zhang (2002).

REMARQUE 4.2.2 (OPÉRATEURS D'OBSERVATIONS MULTIPLES)
On remarque qu'il n'est pas obligatoire d'utiliser le même opérateur d'observation dans l'estimateur d'état et dans le deuxième étage d'estimation paramétrique. Notant H_1 et H_2 deux opérateurs d'observations, on peut définir un estimateur conjoint de la forme

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B\hat{\theta} + R + K_X(Z - H_1\hat{X}) + L^X\dot{\hat{\theta}} \\ \dot{\hat{\theta}} = U^{-1}{L^X}^T H_2^T W^{-1}(Z - H_2\hat{X}) \end{cases}$$

tel que

$$L^X\dot{\hat{\theta}} = K^X_{\text{al}}(Z - H_2\hat{X}) = L^X U^{-1}{L^X}^T H_2^T W^{-1}(Z - H_2\hat{X}).$$

L'avantage potentiel d'une telle définition s'envisage dans le cas où l'on dispose par exemple à la fois de mesures de vitesses et de déplacement. Or on a vu au chapitre 3 que la stabilisation de l'erreur d'état était plus directe et de plus faible complexité algorithmique. On pourrait donc s'en satisfaire pour le filtre d'état mais conserver les deux types d'observations sur les paramètres afin améliorer l'identifiabilité paramétrique du système.

4.2.4 Analyse de l'estimateur

a. Critère moindres carrés

Après la présentation du chapitre 1, nous pouvons énoncer directement ce critère sous forme de théorème.

THÉORÈME 4.2.1
Soit le critère variationnel

$$J_T(\xi) = \frac{1}{2}\xi^T U(0)\xi + \frac{1}{2}\int_0^T (Z - H^e\bar{X}^e(\xi))^T W^{-1}(Z - H^e\bar{X}^e(\xi))\,dt, \tag{4.29}$$

alors il existe un unique minimiseur $\bar{\xi}$ et le système (4.26) *vérifie*

$$\hat{X}(T) = \bar{X}^e_{\bar{\xi}}(T), \quad \forall T \geq 0. \tag{4.30}$$

■ ***Démonstration :*** *La démonstration est calquée sur celles présentées au chapitre 1. Le critère admet un minimum global en*

$$\bar{\xi}^e = P^e_a(0)p^e(0),$$

et le problème au deux bouts est

$$\begin{cases} \dot{\bar{X}}^e_{\bar{\xi}} = A^e\bar{X}^e_{\bar{\xi}} + R^e + K_X(Z - H^e\bar{X}^e_{\bar{\xi}}) \\ \dot{p^e} + (A^e - K^e_X H^e)^T p^e = -H^{eT}W^{-1}(Z - H^e\bar{X}^e_{\bar{\xi}}) \\ \bar{X}^e_{\bar{\xi}}(0) = X^e_0 + P^e_a(0)p^e(0) \\ p^e(T) = 0 \end{cases} \tag{4.31}$$

On le reformule en problème de Cauchy en utilisant le fait que $\bar{X}^e_{\bar{\xi}}$ est une fonction affine de p^e. Soit (r^e, P^e) tels que

$$\bar{X}^e_{\bar{\xi}}(t) = r^e(t) + P^e(t)p^e(t).$$

La substitution dans (4.31) *donne alors*

$$\begin{aligned} &\dot{r^e}(t) + P^e H^{eT}W^{-1}H^e r^e - (A^e - K^e_X H^e)r^e - R^e - K_X\,Z \\ &\quad +(\dot{P^e} - P^e(A^e - K^e_X H^e)^T - (A^e - K^e_X H^e)P^e + P^e H^{eT}W^{-1}H^e P^e)p^e = P^e H^e W^{-1}Z, \end{aligned}$$

donnant les deux équations pour P_e et r_e

$$\begin{cases} \dot{P^e} - P^e(A^e - K^e_X H^e)^T - (A^e - K^e_X H^e)P^e + P^e H^{eT}W^{-1}H^e P^e = 0 \\ \dot{r^e} + P^e H^{eT}W^{-1}H^e r^e - (A^e - K^e_X H^e)r^e = P^e H^{eT}W^{-1}Z + R^e + K_X\,Z \\ P^e(0) = P^e_a(0) \\ r^e(0) = X^e_0 \end{cases} \tag{4.32}$$

Tous ces éléments ont déjà été vus précédemment. La condition initiale sur P^e est de rang réduit p, donc il existe une décomposition sur toute la fenêtre en temps

$$\forall t \geq 0, \quad P^e = L^e U^{-1} L^{eT} = P^e_a. \tag{4.33}$$

Ainsi r^e vérifie le même problème de Cauchy que $\hat{X}^e$, i.e. (4.26)*, donc*

$$\forall t \geq 0, \quad r^e = \hat{X}^e, \tag{4.34}$$

et à la fin de l'intervalle de minimisation

$$\hat{X}^e(T) = \bar{X}^e_{\bar{\xi}}(T). \tag{4.35}$$

■

L'estimateur (4.26) vérifie donc un critère moindres carrés. Cependant, rien n'assure que ce critère est pertinent, car $Z \neq H^e \bar{X}^e(\bar{\xi})$ sur toute la fenêtre en temps. En fait, quand bien même les paramètres sont optimaux, les observations Z ne correspondent à $\bar{X}^e(\bar{\xi})$ qu'asymptotiquement. Cette différence est la conséquence de notre substitution $Z_a \to Z$ à partir du filtre SEEK. Cependant, ce critère reste raisonnable car la qualité de la stabilisation exponentielle de notre estimateur d'état limite la perturbation introduite. D'ailleurs, le paragraphe suivant valide définitivement cet argument à partir de l'analyse d'erreur.

b. Erreur d'estimation

Oublions l'approche contrôle optimal pour nous concentrer nous sur les propriétés de stabilisation de l'erreur d'estimation définie par $(\tilde{X}, \tilde{\theta}) = (X - \hat{X}, \theta - \hat{\theta})$. Nous nous appuyons dans ce cas sur l'analyse menée dans Zhang et Clavel (2001) ; Zhang (2002), et introduisons la grandeur auxiliaire

$$\eta = \tilde{X} - L^X \tilde{\theta}. \tag{4.36}$$

On a alors

$$\begin{aligned}\dot{\eta} &= \dot{\tilde{X}} - \dot{L^X}\tilde{\theta} - L^X \dot{\tilde{\theta}} \\ &= (A - K_X H)\tilde{X} + B\tilde{\theta} - K_X(\epsilon_h + \chi) + L^X \dot{\tilde{\theta}} - (A - K_X H)L^X\tilde{\theta} - B\tilde{\theta} - L^X\dot{\tilde{\theta}},\end{aligned}$$

donc

$$\begin{cases} \dot{\eta} = (A - K_X H)\eta - K_X(\epsilon_h + \chi) \\ \dot{\tilde{\theta}} = -U^{-1}{L^X}^T H^T W^{-1} H L^X \tilde{\theta} - U^{-1}{L^X}^T H^T W^{-1} H\eta - U^{-1}{L^X}^T H^T W^{-1}(\epsilon_h + \chi) \\ \eta(0) = \zeta^X \\ \tilde{\theta}(0) = \zeta^\theta \end{cases} \tag{4.37}$$

Ce changement de variable permet ainsi la **trigonalisation de la dynamique de l'erreur**. De plus, η, dont la dynamique est indépendante de l'erreur paramétrique, suit exactement la même dynamique stabilisée que l'erreur d'estimation d'état à paramètres exacts analysée au chapitre 3. Donc η est exponentiellement stable conduisant à l'estimation (3.53). On peut alors se concentrer sur la dynamique de l'erreur paramétrique[3] réécrite sous la forme

$$\dot{\tilde{\theta}} = -U^{-1}{L^X}^T H^T W^{-1} H L^X \tilde{\theta} + U^{-1}{L^X}^T H^T \varrho,$$

avec

$$\varrho = -W^{-1}(H\eta + \epsilon_h + \chi).$$

Alors, on a

$$\frac{d}{dt}\big(U\tilde{\theta}\big) = \dot{U}\tilde{\theta} + U\big(-U^{-1}{L^X}^T H^T W^{-1} H L^X \tilde{\theta} + U^{-1}{L^X}^T H^T \varrho\big) = {L^X}^T H^T \varrho,$$

d'où il vient

$$\tilde{\theta} = U^{-1}\left(U(0)\tilde{\theta}(0) + \int_0^t {L^X}^T H^T \varrho \, d\tau\right). \tag{4.38}$$

[3] pour laquelle notre analyse diffère de Zhang et Clavel (2001) et dont nous commenterons la stratégie section suivante

Or la dynamique de U s'intègre simplement

$$U(t) = U(0) + \Upsilon(t),$$

en faisant apparaître une forme de grammienne d'observabilité du système (on rappelle que L^X suit la dynamique stabilisée de l'état)

$$\Upsilon(t) = \int_0^t {L^X}^T H^T W^{-1} H L^X \, d\tau. \tag{4.39}$$

Définissons $\lambda_{\inf}(t)$ la plus petite solution du problème aux valeurs propres généralisé

$$\Upsilon(t)\xi = \lambda U(0)\xi,$$

où $U(0)$ est adaptée à l'évaluation de l'erreur paramétrique. Alors on peut minorer $U(t)$ par $(1 + \lambda_{\inf}(t))U_0$ donc

$$\|\tilde{\theta}\|_{U(0)} \leq \frac{1}{1 + \lambda_{\inf}(t)} \left(\|\tilde{\theta}(0)\|_{U(0)} + \left\| U(0)^{-1} \int_0^t {L^X}^T H^T \varrho \, d\tau \right\|_{U(0)} \right). \tag{4.40}$$

Ainsi dans (4.40), la convergence est gouvernée par $\lambda_{\inf}(t)$ dont l'évolution, croissante au cours du temps puisque c'est la première valeur propre de la grammienne, est directement liée aux conditions d'observabilité du système (cf Section 4.2.6).

Pour démontrer cette convergence il nous faut analyser la dépendance de l'erreur paramétrique par rapport à $L^X H^T \varrho$. Tout d'abord, la sensibilité paramétrique L^X vérifie la dynamique stabilisée, elle est donc contrôlée par son second membre

$$\|L^X(t)\theta\|_{\mathcal{E}}^2 \leq C \int_0^t \|B(\tau)\theta\|_{RHS}^2 \, d\tau, \tag{4.41}$$

où $\|\cdot\|_{RHS}$ correspond au choix adéquat de norme sur le second membre de l'équation de la dynamique. Définissant les normes naturelles sur L^X et B

$$\|L^X\|_L = \sup_{\|\theta\|_{U(0)}=1} \|L^X\theta\|_{\mathcal{E}}, \quad \|B\|_B = \sup_{\|\theta\|_{U(0)}=1} \|B\theta\|_{RHS}, \tag{4.42}$$

on obtient donc

$$\|L^X(t)\|_L^2 \leq C \int_0^t \|B(\tau)\|_B^2 \, d\tau. \tag{4.43}$$

Maintenant pour ϱ, commençons par regarder la contribution de l'erreur de discrétisation ϵ_h. On a

$$\left\| U(0)^{-1} \int_0^t {L^X}^T H^T W^{-1} \epsilon_h \, d\tau \right\|_{U(0)} \leq \int_0^t \left\| U(0)^{-1} {L^X}^T H^T W^{-1} \epsilon_h \right\|_{U(0)} d\tau, \tag{4.44}$$

et

$$\left\| U(0)^{-1} {L^X}^T H^T W^{-1} \epsilon_h \right\|_{U(0)}^2 = \epsilon_h^T W^{-1} H L^X U(0)^{-1} {L^X}^T H^T W^{-1} \epsilon_h. \tag{4.45}$$

La norme sur la sensibilité peut être caractérisée à partir des quotients de Rayleigh

$$\|L^X\|_L^2 = \sup_{\theta \neq 0} \frac{\theta^T {L^X}^T N L^X \theta}{\theta^T U(0) \theta} = \sup_{X \neq 0} \frac{X^T N L^X U(0)^{-1} {L^X}^T N X}{X^T N X}, \tag{4.46}$$

d'où on déduit

$$\begin{aligned}\left\|U(0)^{-1}L^{X^T}H^TW^{-1}\epsilon_h\right\|_{U(0)}^2 &\leq \|L^X\|_L^2\,(\epsilon_h^TW^{-1}HN^{-1}H^TW^{-1}\epsilon_h)\\ &\leq \|L^X\|_L^2\,(\epsilon_h^TW^{-1}\epsilon_h)\sup_{Y\neq 0}\frac{Y^TW^{-1}HN^{-1}H^TW^{-1}Y}{Y^TW^{-1}Y}\\ &= \|L^X\|_L^2\,\|\epsilon_h\|_{W^{-1}}^2\sup_{X\neq 0}\frac{X^TH^TW^{-1}HX}{X^TNX}.\end{aligned}$$

De même via les quotients de Rayleigh on a

$$\|HX\|_{W^{-1}} \leq C_{\text{obs}}\|X\|_{\mathcal{E}}, \tag{4.47}$$

conduisant à

$$\left\|U(0)^{-1}\int_0^t L^{X^T}H^TW^{-1}\epsilon_h\,d\tau\right\|_{U(0)} \leq C_{\text{obs}}\int_0^t \|L^X(\tau)\|_L\,\|\epsilon_h(\tau)\|_{W^{-1}}\,d\tau. \tag{4.48}$$

Or par définition de ϵ_h,

$$\|\epsilon_h(\tau)\|_{W^{-1}} \leq C_{\text{obs}}\|x-x_h\|_{\mathcal{E}},$$

donc

$$\left\|U(0)^{-1}\int_0^t L^{X^T}H^TW^{-1}\epsilon_h\,d\tau\right\|_{U(0)} \leq (C_{\text{obs}})^2\int_0^t \|L^X(\tau)\|_L\,\|(x-x_h)(\tau)\|_{\mathcal{E}}\,d\tau. \tag{4.49}$$

Considérons maintenant la contribution de η dans ϱ. On obtient

$$\begin{aligned}\left\|U(0)^{-1}\int_0^t L^{X^T}H^TW^{-1}H\eta\,d\tau\right\|_{U(0)} &\leq C_{\text{obs}}\int_0^t \|L^X(\tau)\|_L\,\|H\eta(\tau)\|_{W^{-1}}\,d\tau\\ &\leq (C_{\text{obs}})^2\int_0^t \|L^X(\tau)\|_L\,\|\eta(\tau)\|_{\mathcal{E}}\,d\tau.\end{aligned} \tag{4.50}$$

Cette erreur est valide à la fois pour les parties déterministes et éventuellement probabilistes de η et on la majore par l'analyse d'erreur de l'erreur d'estimation d'état du chapitre 3 (3.53).

Enfin, pour la partie probabiliste de ϱ, nous particularisons l'analyse dans le cas des cellules taggées (pour les opérateurs frontières voir la Section 4.2.7). On a alors

$$\begin{aligned}\mathbb{E}\left(\left\|U(0)^{-1}\int_0^t L^{X^T}H^TW^{-1}\chi\,d\tau\right\|_{U(0)}^2\right) &= \mathbb{E}\left(\int_0^t \chi^TW^{-1}HL^XU(0)^{-1}L^{X^T}H^TW^{-1}\chi\,d\tau\right)\\ &= \mathbb{E}\left(\int_0^t \operatorname{tr}\big(U(0)^{-1}L^{X^T}H^TW^{-1}\chi\chi^TW^{-1}HL^X\big)\,d\tau\right)\\ &= \int_0^t \operatorname{tr}\big(U(0)^{-1}L^{X^T}H^TW^{-1}Q_\chi W^{-1}HL^X\big)\,d\tau\\ &= \sum_{i=1}^q(\alpha_i)^2\sum_{j=1}^3\int_0^t (V_i^j)^TW^{-1}HL^XU(0)^{-1}L^{X^T}H^TW^{-1}V_i^j\,d\tau,\end{aligned}$$

où nous utilisons les mêmes identités qu'au chapitre 3 sur le bruit blanc χ. Donc les mêmes arguments que pour l'erreur de discrétisation donnent

$$\mathbb{E}\left(\left\|U(0)^{-1}\int_0^t L^{X^T}H^TW^{-1}\chi\,d\tau\right\|_{U(0)}^2\right) \leq C_{\text{obs}}^2\Big(\sum_{i=1}^q(\alpha_i)^2\sum_{j=1}^3\|V_i^j\|_{W^{-1}}^2\Big)\int_0^t \|L^X(\tau)\|_L^2\,d\tau. \tag{4.51}$$

Finalement, l'erreur paramétrique se résume à l'ajout de chacune des contributions analysées

$$\mathbb{E}(\|\tilde{\theta}\|^2_{U(0)}) \leq \frac{C}{\big(1+\lambda_{\inf}(t)\big)^2}\Bigg\{1 + \|L^X\|^2_{L^2([0,T];L)}\Bigg[C^2_{\text{obs}}\Big(\sum_{i=1}^{q}(\alpha_i)^2\sum_{j=1}^{3}\|V_i^j\|^2_{W^{-1}}\Big)$$
$$+(C_{\text{obs}})^4\Big(T_1\mathbb{E}(\|\zeta^X\|^2_{\mathcal{E}}) + (1+t\,T_2\gamma^2)\|x-x_h\|^2_{L^2([0,T];\mathcal{E})} + t\,T_2\gamma^2\sum_{i=1}^{q}\alpha_i^2 + t\,\epsilon^2\Big)\Bigg]\Bigg\}. \tag{4.52}$$

Enfin à cette estimation on ajoute l'estimation d'erreur des variables d'état d'après la simple inégalité triangulaire

$$\|\tilde{X}\|_{\mathcal{E}} \leq \|\eta\|_{\mathcal{E}} + \|L^X\|_L\|\tilde{\theta}\|_{U(0)}. \tag{4.53}$$

c. Analyse historique par fonction de Liapounov

Afin d'éclairer cette démonstration technique, nous présentons ici les arguments utilisés par Zhang (2002) et les analysons à la lumière de notre estimation. Celui-ci commence par deux hypothèses sur le système :

HYPOTHÈSE 1 (STABILITÉ DE L'ERREUR D'ESTIMATION D'ÉTAT)
Il existe K associé au couple (A, H) tel que le système η de dynamique controlée par K

$$\dot{\eta} = (A - KH)\eta$$

est exponentiellement stable.

L'objectif du chapitre 3 a été de vérifier cette hypothèse pour notre problème avec $K = K_X$. La deuxième hypothèse est plus subtile.

HYPOTHÈSE 2 (EXCITATION PERSISTANTE)
Soit $L^X \in R^p \times R^N$ la matrice vérifiant la dynamique

$$\dot{L}^X = (A - K_X H)L^X + B(t).$$

On suppose que B est une excitation persistante *c'est-à-dire qu'il existe α, et T tels que pour tout t,*

$$U(t+T) - U(t) = \int_t^{t+T} {L^X}^T H^T W^{-1} H L^X \geq \alpha\mathbb{1}.$$

Cette hypothèse est clairement une condition d'observabilité, et plus précisément d'identifiabilité car la grammienne est définie à partir de L^X. En effet, l'hypothèse 2 est une condition sur $B(t)$, la dépendance du système linéaire par rapport aux paramètres ($\dot{X} = AX + B\theta$). Ainsi on suppose que $B(t)$ permet d'observer au travers de l'état suffisamment d'informations sur les paramètres de sorte que la sensibilité paramétrique L^X vérifie une condition d'observabilité. A partir de ces deux hypothèses, Zhang (2002) assure alors l'exponentielle stabilité du système. Rappelons rapidement ses arguments qui sont très instructifs par rapport à notre démonstration.

□ SYSTÈME PARAMÉTRIQUE AUTONOME – Tout d'abord, on considère le système de l'erreur paramétrique autonome et homogène, c'est-à-dire le système de dynamique

$$\dot{\check{\theta}} = -U^{-1}L^X H^T W^{-1} H L^X \check{\theta}. \tag{4.54}$$

On démontre alors son exponentielle stabilité via la fonction de Liapounov suivante

$$V(t) = \frac{1}{2}\check{\theta}^T U \check{\theta}. \tag{4.55}$$

En effet

$$\begin{aligned}\frac{\mathrm{d}V}{\mathrm{dt}} &= \check{\theta}^T U \dot{\check{\theta}} + \frac{1}{2}\check{\theta}^T \dot{U} \check{\theta} \\ &= -\frac{1}{2}\check{\theta}^T {L^X}^T H^T W^{-1} H L^X \check{\theta}.\end{aligned}$$

Donc on a déjà $\dot{V} \leq 0$, ce qui assure la stabilité de l'erreur. L'exponentielle stabilité nécessite plus d'efforts et requiert quelques lemmes. On intègre tout d'abord la fonction de Liapounov sur l'intervalle nécessaire $[t, t+T]$ à l'observabilité :

$$\begin{aligned}V(t+T) - V(t) &= \int_t^{t+T} \dot{V}(\tau)\, d\tau \\ &= -\frac{1}{2}\int_t^{t+T} \check{\theta}(\tau)^T (L^X(\tau))^T H^T W^{-1} H L^X(\tau) \check{\theta}(\tau)\, d\tau \\ &= -\check{\theta}(t+T)^T \left(\int_t^{t+T} \mathcal{R}_{\check{\theta}}^T(\tau, t+T)(L^X(\tau))^T H^T W^{-1} H L^X(\tau) \mathcal{R}_{\check{\theta}}(\tau, t+T)\, d\tau \right) \check{\theta}(t+T).\end{aligned}$$

où $\mathcal{R}_{\check{\theta}}(\tau, t+T)$ est la résolvante du système $\check{\theta}$. Or l'hypothèse d'excitation persistante permet de démontrer qu'il existe une constante κ telle que

$$\int_t^{t+T} \mathcal{R}_{\check{\theta}}^T(\tau, t+T)(L^X(\tau))^T H^T W^{-1} H L^X(\tau) \mathcal{R}_{\check{\theta}}(\tau, t+T)\, d\tau \geq \kappa \mathbb{1}$$

puis, U étant une norme,

$$V(t+T) - V(t) \leq C\check{\theta}(t+T) U \check{\theta}(t+T) \rightarrow V(t+T) \leq \frac{1}{1+C} V(t).$$

Cette expression est le pendant semi-discret des conditions classiques de stabilité pour les fonctionnelles de Liapounov. C'est en fait la relation la plus courante sur un système dynamique observé puisque V ne peut décroître que lorsque l'intervalle de temps est assez long pour permettre une observabilité complète du système. Ainsi une fois la période d'observabilité passée, on a pour tout $\tau > T$

$$V(t+\tau) \leq \left(\frac{1}{1+C}\right)^{\lfloor \frac{\tau}{T} \rfloor} \Rightarrow V(t+\tau) \leq \left(\frac{1}{1+C}\right)^{\frac{\tau}{T}-1} V(t),$$

ce qui assure l'exponentielle stabilité de $\check{\theta}$.

Cette première démonstration, bien que plus technique, est donc une version particulière des théorèmes d'exponentielle stabilité utilisant une fonction de Liapounov issue des théories

LQ. La condition d'excitation persistante remplace alors naturellement la condition d'observabilité.

□ SYSTÈME COUPLÉ $(\eta, \tilde{\theta})$ – Cette stabilité de $\check{\theta}$ complète de manière indépendante celle que nous avions sur η (cf. hypothèse 1). On peut alors conclure si on assure l'exponentielle stabilité du système couplé. Il existe deux arguments possibles pour cela. Le premier, utilisé directement dans Zhang (2002), est à l'image de notre analyse d'erreur (sans la gestion des différents bruits). Le système de l'erreur paramétrique complet est donné par

$$\dot{\tilde{\theta}} = -U^{-1}{L^X}^T H^T W^{-1} H L^X \tilde{\theta} - U^{-1}{L^X}^T H^T W^{-1} H \eta,$$

donc en utilisant la résolvante du système autonome on a

$$\tilde{\theta} = \mathcal{R}_{\check{\theta}}^T(t,0)\zeta^\theta + \int_0^t \mathcal{R}_{\check{\theta}}^T(t,\tau) U^{-1} L^X H^T W^{-1} H \eta \, d\tau.$$

Puisque η est exponentiellement stable et L^X bornée, $U^{-1}L^X H^T W^{-1} H\eta \leq C_1 e^{-\lambda_1 t}$ est exponentiellement stable. De même la résolvante correspond à un système exponentiellement stable $\|\mathcal{R}_{\check{\theta}}^T(\tau,t)\| \leq C_2 e^{-\lambda_2(t-\tau)}$ donc

$$\|\tilde{\theta}\| \leq C_2 e^{-\lambda_2 t}\|\zeta^\theta\| + C_1 C_2 e^{-\lambda_2 t} \int_0^T e^{(\lambda_2-\lambda_1)\tau} \, d\tau.$$

Cette relation assure l'exponentielle stabilité de $\tilde{\theta}$ dans les deux cas $\lambda_2 = \lambda_1$ ou $\lambda_2 \neq \lambda_1$ (en pratique pour nous $\lambda_2 < \lambda_1$ car on attend une stabilisation beaucoup plus rapide de l'état).

REMARQUE 4.2.3
On peut se demander s'il existe une démonstration entièrement fondée sur l'utilisation de fonctions de Liapounov. Nous savons que η est exponentiellement stable donc comme expliqué Section 3.3.2, il existe Q définie positive tel que $V^\eta = \frac{1}{2}\eta^T Q \eta$ soit une fonction de Liapounov pour η vérifiant

$$\frac{\mathrm{d}V^\eta}{\mathrm{dt}} = -\frac{1}{2}\eta^T \eta.$$

De même pour $\check{\theta}$ avec $V^\theta = \frac{1}{2}\check{\theta} U \check{\theta}$. Dans ce cas, considérons pour le système couplé la fonction de Liapounov

$$V(t) = \frac{1}{2}\eta^T Q \eta + \frac{1}{2}\tilde{\theta}^T U \tilde{\theta},$$

alors on obtient immédiatement

$$\frac{\mathrm{d}V}{\mathrm{dt}} = \frac{1}{2}\begin{pmatrix}\eta^T & \theta^T\end{pmatrix}\begin{pmatrix}\mathbb{1} & H^T W^{-1} H L^X \\ {L^X}^T H^T W^{-1} H & {L^X}^T H^T W^{-1} H L^X\end{pmatrix}\begin{pmatrix}\eta \\ \theta^T\end{pmatrix}.$$

La question est de savoir si cette matrice bloc est définie positive et minorée. Le lemme suivant, rappelé dans Zhang et Xu (2001), fournit une piste de recherche de démonstration.

LEMME 4.2.2
Soit une matrice

$$M = \begin{pmatrix} P & L^T \\ L & Q \end{pmatrix}$$

avec P et Q définies positives. Alors M est définie positive si et seulement si $P - L^T Q^{-1} L$ est définie positive.

■ ***Démonstration*** : *Soit la forme quadratique*

$$V(X,Y) = X^T PX + Y^T QY + 2X^T L^T Y,$$

alors le minimum en Y *de cette forme quadratique est donné par*

$$\frac{\partial V}{\partial Y} = 0 \Rightarrow Y = Q^{-1}LX,$$

donc

$$\min_Y V(X,Y) = X^T(P - L^T Q^{-1} L)X.$$

V *est donc positive si et seulement si* $P - L^T Q^{-1} L > 0$. ■

□ BILAN – Par rapport à la preuve donnée à la section précédente, nous pouvons cependant assurer que cette exponentielle stabilité ne nous apporte pas grand chose en pratique. En effet, dans le cadre de notre estimation cardiaque, l'information disponible sur les paramètres s'accumule avec la propagation de la dépolarisation et repolarisation électrique (voir Section 4.2.6 pour des exemples numériques). Ainsi typiquement T correspond pour nous à la période du battement cardiaque. Donc l'exponentielle stabilité démontrée ici n'offre aucune amélioration par rapport à l'estimation d'erreur paramétrique obtenue précédemment car la condition d'excitation persistante est vérifiée sur une période plus grande que celle considérée par notre première preuve. En fait, ces deux démonstrations tentent toutes les deux de maîtriser la condition d'identifiabilité au cours du temps et sont finalement complémentaires.

4.2.5 Schéma numérique associé et analyse de stabilité

Une fois l'estimateur défini sur le système éléments finis, il faut encore formuler un schéma numérique adapté. Pour l'estimation d'état seule, l'utilisation de filtres physiques permettait d'utiliser des schémas numériques naturels inspirés du schéma de Newmark. Ici, une première idée serait d'utiliser des règles de type point milieu ou règle des trapèzes pour le nouveau terme $L^X U^{-1} L^{X^T} H^T W^{-1}(Z - H\hat{X})$ sur l'état. Mais les premiers tests numériques que nous avons réalisés sur ce filtre adaptatif issu de Zhang (2002) ont montré de fortes instabilités pour une discrétisation point milieu. En revanche, l'équivalence avec le filtre SEEK nous a permis de formuler naturellement un bon schéma en temps de type estimation-correction inspiré du filtre de Kalman. Là, comme au chapitre 1, on choisit de discrétiser en temps le système dynamique comme si on appliquait le filtre SEEK discret en temps (d'ailleurs il a été initialement été démontré en discret dans Pham et al (1997)) à la version discrète en temps du modèle. Notons $A^K_{n+1|n}$ l'opérateur de transition discrétisant en temps l'opérateur $A - K_X H$ et $K_X l(Z_{n+1}, Z_n)$ le choix de discrétisation de $K_X Z$. La fonction l représente formellement un combinaison linéaire de Z_{n+1} et Z_n adaptée à la discrétisation de lu filtre de Luenberger choisit. On peut donc

écrire l'estimateur discrétisé en temps sous la forme

$$\begin{cases} \text{Prédiction :} \\ \hat{X}^-_{n+1} = A^K_{n+1|n}\hat{X}^+_n + B_{n+1|n}\hat{\theta}^+_n + R_{n+1} + K_X l(Z_{n+1}, Z_n) \\ \hat{\theta}^-_{n+1} = \hat{\theta}^+_n \\ \text{Correction :} \\ L^X_{n+1} = A^K_{n+1|n}L^X_n + B_{n+1|n} \\ U_{n+1} = U_n + {L^X_{n+1}}^T H^T_{n+1}W_{n+1}H_{n+1}L^X_{n+1} \\ \hat{X}^+_{n+1} = \hat{X}^-_{n+1} + L^X_{n+1}U^{-1}_{n+1}{L^X_{n+1}}^T H^T_{n+1}W^{-1}_{n+1}\big(Z_{n+1} - H_{n+1}\hat{X}^-_{n+1}\big) \\ \hat{\theta}^+_{n+1} = \hat{\theta}^-_{n+1} + U^{-1}_{n+1}{L^X_{n+1}}^T H^T_{n+1}W^{-1}_{n+1}\big(Z_{n+1} - H_{n+1}\hat{X}^-_{n+1}\big) \end{cases} \tag{4.56}$$

La quantité $l(Z_{n+1}, Z_n)$ représente une fonction des observations Z_{n+1} et Z_n compatible avec le schéma en temps de l'estimateur d'état seul. Au chapitre 3, nous avions choisi par exemple des schémas de type Newmark donc l représentait simplement un point milieu $l(Z_{n+1}, Z_n) = \frac{1}{2}(Z_{n+1} + Z_n)$. On aura alors directement un critère optimal discret associé au schéma. Mais, plus intéressant, le fait que cette discrétisation en temps vérifie elle-même une fonction de Liapounov discrète pour l'erreur nous permet d'assurer la stabilité de l'erreur discrète $\binom{\eta_n}{\tilde{\theta}_n}$ comme pour le continu. Pour cela on reprend les équations récursives « one-step »

$$\begin{cases} \hat{X}^-_{n+1} = A^K_{n+1|n}\hat{X}^-_n + B_{n+1|n}\hat{\theta}^-_n + R_{n+1} + L^X_n U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(Z_n - H_n\hat{X}^-_n\big) + K_X l(\epsilon_n, \chi_n) \\ \hat{\theta}^-_{n+1} = \hat{\theta}^-_n + U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(Z_n - H_{n+1}\hat{X}^-_n\big) \end{cases} \tag{4.57}$$

Les erreurs vérifient donc

$$\begin{cases} \tilde{X}^-_{n+1} = A^K_{n+1,n}\tilde{X}^-_n + B_{n+1|n}\tilde{\theta}^-_n - L^X_{n+1}U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(H_n\tilde{X}^-_n + \epsilon_n + \chi_n\big) + K_X l(\epsilon_n, \chi_n) \\ \tilde{\theta}^-_{n+1} = \tilde{\theta}^-_n - U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(H_n\tilde{X}^-_n + \epsilon_n + \chi_n\big) \end{cases} \tag{4.58}$$

où ϵ_n est l'erreur de discrétisation en temps et en espace vu à travers l'opérateur d'observation. Posons alors

$$\eta_n = \tilde{X}^-_n - L^X_n\tilde{\theta}^-_n, \tag{4.59}$$

on a

$$\begin{aligned} \eta_{n+1} &= A^K_{n+1|n}\tilde{X}^-_n + B_{n+1|n}\tilde{\theta}^-_n \\ &\qquad - L^X_{n+1}U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(H_n\tilde{X}^-_n + \epsilon_n + \chi_n\big) - L^X_{n+1}\tilde{\theta}^-_{n+1} + K_X l(\epsilon_n, \chi_n) \\ &= A^K_{n+1|n}\tilde{X}^-_n + B_{n+1|n}\tilde{\theta}^-_n - L^X_{n+1}\tilde{\theta}^-_n + K_X l(\epsilon_n, \chi_n) \\ &= A^K_{n+1|n}\eta_n + K_X l(\epsilon_n, \chi_n). \end{aligned} \tag{4.60}$$

Ainsi si le schéma numérique est stable pour l'estimateur d'état seul (et on sait qu'il l'est pour les estimateurs proposés d'après le chapitre 3) alors l'extension paramétrique de ce schéma par l'algorithme prédiction-correction permet la même stabilité sur η_n. La dynamique de $\tilde{\theta}^-_n$ s'exprime alors sous la forme

$$\tilde{\theta}^-_{n+1} = \big(\mathbb{1} - U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n H_n\big)\tilde{\theta}^-_n + U^{-1}_n {L^X_n}^T H^T_n W^{-1}_n\big(\eta_n + \epsilon_n + \chi_n\big), \tag{4.61}$$

qui est exactement l'analogue discret de la deuxième relation de (4.37) analysée en continu.

4.2.6 Résultats

a. Exemples numériques

Nous reprenons notre problème-test décrit au chapitre 3. La dépendance paramétrique est ici précisé en découpant le ventricule gauche analytique suivant les régions AHA comme présenté Figure 4.1. Nous supposons alors que la $14^{\text{ème}}$ région est partiellement infarcie en fixant $\theta = 0.5$ dans cette zone. Puisqu'en théorie LQ classique, tous les bruits sont supposés blancs, nous supposons enfin l'incertitude paramétrique donnée par un bruit blanc. Comme nous n'avons aucun a priori sur la zone infarcie nous supposons donc que

$$\mathbb{E}(\zeta^\theta(\zeta^\theta)^T) = \mathbb{1}.$$

Dans ces premiers résultats les données synthétiques correspondent à notre opérateur d'observation par cellules volumiques illustré Figure 3.2, pour des mesures de vitesse.

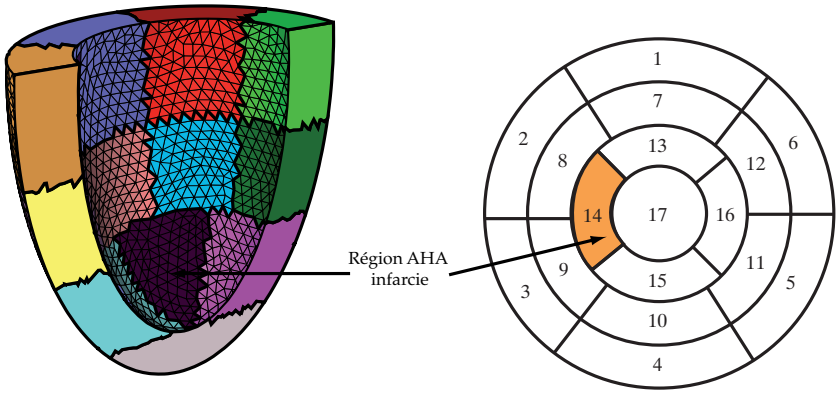


FIGURE 4.1: Maillage de référence avec ses régions AHA. La $14^{\text{ème}}$ est infarcie

□ TEST ALGORITHMIQUE – Nous commençons par un résultat de validation de l'algorithme mis en place en réalisant une estimation dite « en conditions parfaites », c'est-à-dire sans erreur de discrétisation (i.e. même maillage dans le problème direct et inverse), ni bruit de mesure, ni erreur de condition initiale. L'incertitude se limite donc strictement aux paramètres. Les résultats apparaissent Figure 4.2. La première figure montre l'ensemble des trajectoires des 17 paramètres alors que la deuxième les superpose. Plus loin nous nous limiterons à seulement quelques trajectoires différenciées en plus de la superposition pour des raisons de lisibilité mais, ici, la convergence est tellement nette que tous les paramètres restent resserrés autour de 1 sauf le $14^{\text{ème}}$ qui plonge vers 0.5 une fois que l'activation passe la région AHA concernée.

□ PROBLÈME-TEST – Une fois l'algorithme validé, nous reprenons notre problème-test en configuration de bruits réalistes. Les données synthétiques sont générées à partir du maillage de référence (donc très raffiné). La bruit est supposé blanc à 10% de la vitesse maximale pour un échantillonnage toutes les $50\ ms$. On rappelle que ramené au pas de temps Δt, ce bruit correspond alors à une incertitude de l'ordre de 70% à chaque itération en temps.

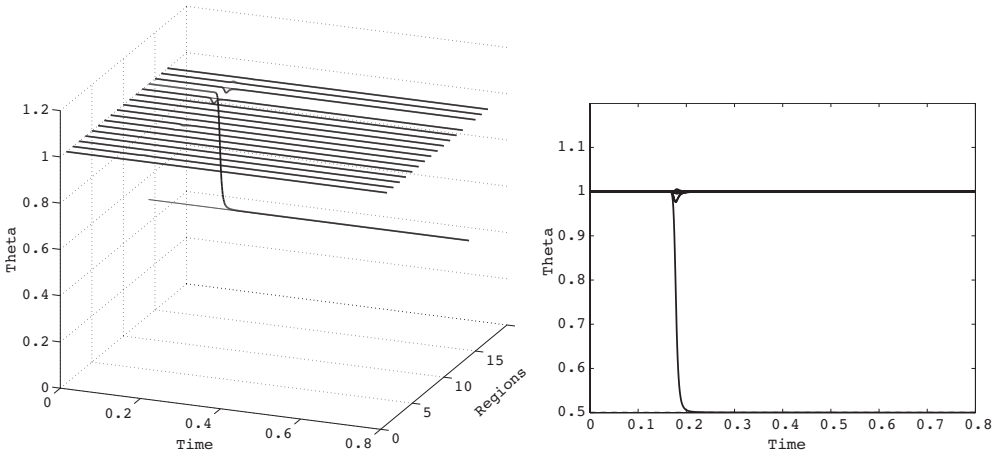


FIGURE 4.2: Test algorithmique de l'identification « en conditions idéales »

L'estimation est réalisée avec le maillage « desired mesh » (cf tableau 3.4) et elle est présentée Figure 4.3. A gauche, on montre la convergence de quelques paramètres, dont le $14^{\text{ème}}$ qui converge bien vers 0.5 alors que les autres sont restés à 1 en fin de fenêtre d'estimation. Sur la figure de droite, on superpose l'évolution des 17 paramètres afin de visualiser l'amplitude de recherche des paramètres ainsi que le décalage dans chacune des estimations.

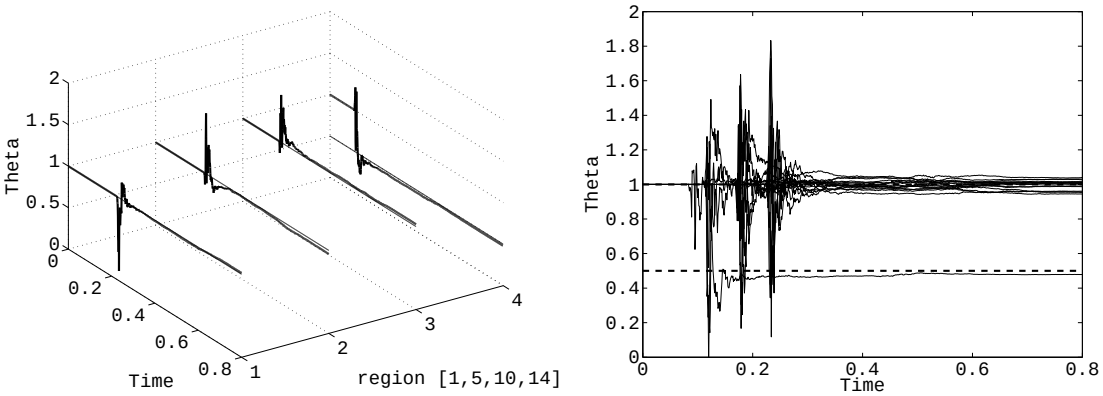


FIGURE 4.3: Identification de la zone infarcie sur le modèle linéaire pour le maillage « desired ». Sont représentées 4 trajectoires à gauche et la superposition de tous les paramètres à droite

Les résultats sont robustes avec la discrétisation, notamment Figure 4.4 on retrouve des évolutions similaires pour les deux autres maillages « coarse » and « fine ». Enfin l'évolution globale de l'erreur du système est aussi performante que celle présentée au chapitre 3, Figure 3.21.

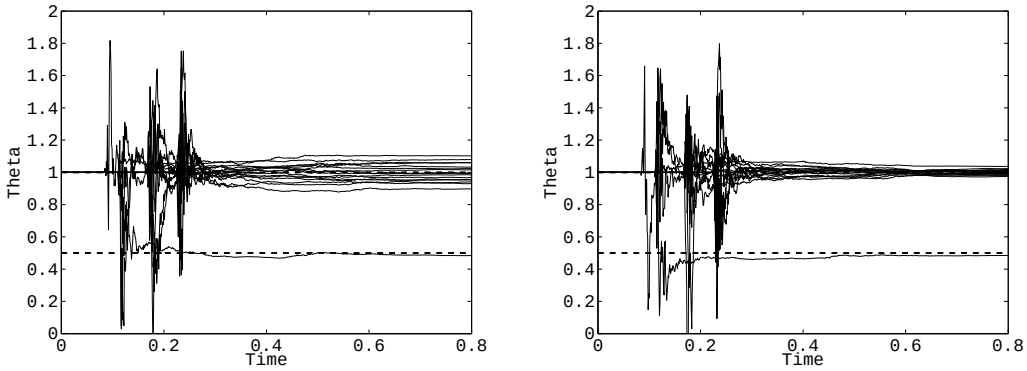


FIGURE 4.4: Identification de la zone infarcie sur le modèle linéaire pour le maillage « coarse » (gauche) et « fine » (droite)

b. Observabilité du système

Les différentes courbes que nous venons de présenter sont représentatives de l'analyse de l'estimateur que nous avions faite précédemment. Nous présentons pour cela Figure 4.5 le calcul des valeurs propres maximales et minimales de Υ (ici $U(0) = \mathbb{1}$). La valeur minimale est la plus importante pour notre estimateur et atteint après $0.3s$ une valeur de $3.5\,10^3$. Donc l'erreur paramétrique finale doit être de $1/1000$ ce que nous constatons sur la figure en condition idéale. Pour la configuration plus réaliste, on remarquera que la grammienne est toujours la même car le calcul de L_X est indépendant des conditions initiales spécifiques de l'état ainsi que des bruits. Donc les oscillations autour des valeurs attendues sont uniquement la réponse du filtre de Kalman aux différents bruits. Notamment le fait d'avoir choisi $U(0) = \mathbb{1}$ introduit une forte sensibilité paramétrique à comparer plus tard à celle utilisée dans le filtre H^∞ de la prochaine section.

Nous avons aussi décidé de tracer la plus grande valeur propre de Υ pour analyser l'évolution de l'observabilité au cours de la contraction. La propagation de l'onde électrique s'effectue en effet de l'apex vers la base en 10 ms à partir de $t = 0.2s$. Or la sensibilité paramétrique est directement dépendante de l'activation électrique puisque le paramètre de précontrainte active n'a pas d'influence sans activation. Ainsi on constate que la plus grande valeur propre devient non nulle à partir de $t = 0.2s$ alors que la plus petite l'est plus tard, une fois l'onde électrique propagée sur tout le ventricule.

4.2.7 Opérateur frontière et Formulation H infini

Nous avons déjà introduit au chapitre 3 (Section 3.4.6) l'observateur qui nous permettait d'analyser le cas « théorique » d'une observation du champ de déplacement sur une partie de la frontière. Dans ce cadre, nous avons notamment montré comment cet observateur générait un certain nombre de difficultés théoriques notamment vis-à-vis de bruits blancs de mesures. Ainsi l'analyse détaillée que nous venons d'effectuer n'est plus valide ici au sens strict. Cependant, les extensions de la théorie LQ du chapitre 1 laissent entrevoir une opportunité en adoptant le cadre plus général du contrôle robuste pour manipuler une gamme de bruit plus large

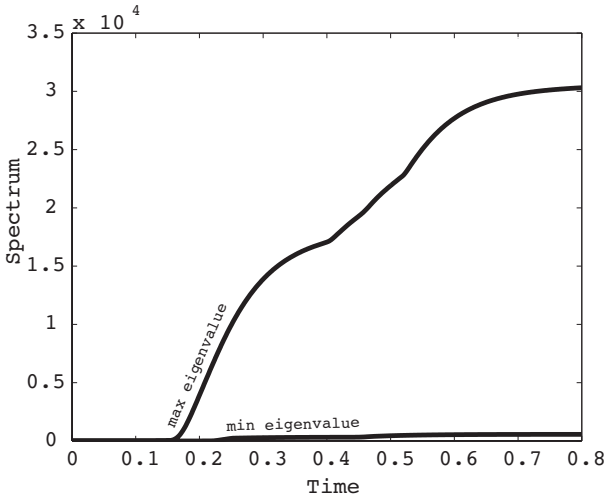


FIGURE 4.5: Evolution des valeurs propres de la grammienne d'observabilité Υ

et compatible du point de vue énergétique avec l'observateur frontière. Reste cependant la difficulté de pouvoir formuler un estimateur robuste réduit puisque ici aussi nous souhaitons bénéficier des méthodes standards de contrôle optimal uniquement sur l'espace paramétrique en conservant nos estimateurs d'état physiques développés précédemment. En effet, le filtre robuste dans le cas linéaire quadratique introduit comme le filtre de Kalman une équation de Riccati (1.91) qu'il faut pouvoir résoudre seulement sur le sous-espace paramétrique. Cette section présente donc l'extension des résultats précédent dans le formalisme du contrôle robuste et reprend les résultats publiés dans Chapelle et al (2009). Dans les exemples numériques, les observations frontières sont toujours considérées comme étant des vitesses mais peuvent de façon équivalente être des déplacements pour peu que l'on choisisse le bon estimateur d'état correspondant.

a. Formulation de l'estimateur et critère quadratique H^∞ associé

Nous considérons toujours un système étendu $X^e = \left(\begin{smallmatrix} X \\ \theta \end{smallmatrix}\right)$ auquel correspond un estimateur d'état pour les observations frontières $\bar{X}^e = \left(\begin{smallmatrix} \bar{X} \\ \theta \end{smallmatrix}\right)$ de dynamique

$$\begin{cases} \dot{\bar{X}}^e = A^e \bar{X}^e + R^e + K_X^e (Z - H^e \bar{X}^e) \\ \bar{X}^e(0) = X_0^e + \zeta_\theta^e \end{cases} \tag{4.62}$$

avec toujours

$$\zeta_\theta^e = \begin{pmatrix} 0 \\ \zeta^\theta \end{pmatrix}, \quad K_X^e = \begin{pmatrix} 0 \\ \gamma (H^v)' \end{pmatrix}, \quad H^e = (0 \quad H^v).$$

Comme précédemment, l'incertitude de condition initiale ζ^θ de ce système est réduite à sur la partie paramétrique. Nous notons alors $\bar{X}_\xi^e$ la solution de cette dynamique pour toute condition initiale $\theta(0) = \theta_0 + \xi$ et souhaitons appliquer le formalisme H^∞ présenté Section 1.4.2 sur ce système cible particulier, c'est-à-dire en l'absence d'erreur de modèle et pour une incertitude

de rang réduit. Le critère H^∞ s'écrit

$$\sup_{Z,\xi} \frac{\int_0^T \|\hat{X}^e - \bar{X}^e(\xi)\|_{S^e}^2 \, dt}{\int_0^T \|Z - H^e \bar{X}^e(\xi)\|_{\mathfrak{m}}^2 \, dt + \|\xi\|_{P_0^{-1}}^2} \leq \frac{1}{\beta}, \tag{4.63}$$

où par rapport au chapitre 1 nous appelons S^e la norme évaluant l'estimateur au numérateur (puisque cette notation est libre en l'absence de bruit de modèle). Pour notre vecteur augmenté X^e, on considérera alors une norme tensorisée

$$S^e = \begin{pmatrix} S_X & 0 \\ 0 & S_\theta \end{pmatrix}. \tag{4.64}$$

où typiquement S_X est en fait la norme de l'énergie N et S_θ une norme adéquate pour les paramètres, par exemple l'inverse de la covariance d'erreur paramétrique initiale P_0.

Dans ce cadre spécifique, nous allons démontrer au paragraphe suivant que l'estimateur SEEK précédemment présenté dans le cadre H^2 s'étend sous la forme

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B\hat{\theta} + R + K_X(Z - H\hat{X}) + L^X\dot{\hat{\theta}}, \quad \text{avec } \hat{X}(0) = X_0 \\ \dot{\hat{\theta}} = U^{-1}{L^X}^T H'(Z - H\hat{X}), \quad \text{avec } \hat{\theta}(0) = \theta_0 \\ \dot{L}_X = (A - K_X H)L^X + B, \quad \text{avec } L^X(0) = 0 \\ \dot{U} = {L^X}^T H' H L^X - \beta({L^X}^T S_X L^X + S_\theta), \quad \text{avec } U(0) = (P_0)^{-1} \end{cases} \tag{4.65}$$

La différence fondamentale entre cet estimateur et (4.26) est donc la présence des matrices S_X et S_θ dans la dynamique de U associée évidemment au terme β. Comme dans toute la théorie H^∞, ceci implique une limitation de la fenêtre en temps d'utilisation de ce filtre puisque la dynamique de U n'a de sens que tant que U reste défini positif.

PROPOSITION 4.2.3
Étant donné T_{max}, il existe une valeur maximale β^* (dépendant de T_{max}) telle que, pour tout β et T vérifiant $0 < \beta < \beta^*$ et $0 < T \leq T_{max}$, l'observateur défini en (4.65) *fournit la borne* (4.63).

■ ***Démonstration :*** *Nous définissons β^* comme la valeur maximum telle que U soit définie positive sur l'intervalle $[0, T_{max}]$. Alors, pour $0 < \beta < \beta^*$, nous définissons la fonction coût en contrôle H^∞ (cf. Section 1.4.2),*

$$J_T^\beta = \int_0^T \|\phi - \bar{X}^e(\xi)\|_{S^e}^2 \, dt - \frac{1}{\beta}\left(\int_0^T \|Z - H^e\bar{X}^e(\xi)\|_{\mathfrak{m}}^2 \, dt + \|\xi\|_{P_0^{-1}}^2\right), \tag{4.66}$$

telle que le critère robuste (4.63) *est assuré si et seulement si*

$$\sup_{\chi,\xi} J_T^\beta \leq 0. \tag{4.67}$$

On définit, comme au chapitre 1, la fonction

$$V^\beta(T) = \sup_{\xi} J_T^\beta, \tag{4.68}$$

qui correspond au supremum sur les contrôles pour une observation Z donnée, et un choix d'estimateur $\phi(Z)$ fixé (on rappelle que l'estimateur n'est qu'une fonction de Z). On introduit l'état adjoint p^e pour la trajectoire maximale $\bar{X}^e_{\bar{\xi}}$. Le problème aux deux bouts s'écrit

$$\begin{cases} \dot{\bar{X}}^e_{\bar{\xi}} = A^e \bar{X}^e_{\bar{\xi}} + R^e + K_X(Z - H^e \bar{X}^e_{\bar{\xi}}) \\ \dot{p^e} + (A^e - K^e_X H^e)^T p^e = -H^{e\prime}(Z - H^e \bar{X}^e_{\bar{\xi}}) + \beta S^e(\phi^e - \bar{X}^e_{\bar{\xi}}) \\ \bar{X}^e_{\bar{\xi}}(0) = X^e_0 + P^e_a(0) p^e(0) \\ p^e(T) = 0 \end{cases} \tag{4.69}$$

avec

$$P^e_a(0) = \begin{pmatrix} 0 & 0 \\ 0 & P_0 \end{pmatrix}.$$

Introduisant le changement de variable désormais classique

$$\bar{X}^e_{\bar{\xi}}(t) = r^e(t) + P^e(t) p^e(t), \tag{4.70}$$

on obtient le problème de Cauchy

$$\begin{cases} \dot{P^e} - P^e(A^e - K^e_X H^e)^T - (A^e - K^e_X H^e)P^e + P^e H^{e\prime} H^e P^e - \beta P^e S^e P^e = 0 \\ \dot{r^e} + P^e H^{e\prime} H^e r^e - (A^e - K^e_X H^e) r^e + \beta P^e S^e(\phi^e - r^e) = P^e H^{e\prime} Z + R^e + K_X Z \\ P^e(0) = P^e_a(0) \\ r^e(0) = X^e_0 \end{cases} \tag{4.71}$$

et on retrouve alors l'équation de Riccati H^∞ sans erreur de modèle. De plus la condition initiale sur P^e permet de vérifier, comme en filtrage de Kalman réduit, que ce dernier est décomposable sous la forme

$$P^e_a = \begin{pmatrix} L^X \\ \mathbb{1}_p \end{pmatrix} U^{-1} \begin{pmatrix} L^X & \mathbb{1}_p \end{pmatrix}.$$

Donc r^e, décomposé sur ses composantes état et paramètre, correspond au système (4.65).

Reste à démontrer que ce filtre est robuste au sens du critère (4.67). *On considère donc la fonction coût prise pour $\bar{X}^e(\bar{\xi}) = \bar{X}^e_{\bar{\xi}}$. On applique alors le principe de programmation dynamique dans le cas simplifié où il n'y a pas d'erreur de modèle. On a ici une version simplifiée de HJB qu'on retrouve par*

$$\begin{aligned} \frac{dV^\beta(T)}{dT} &= \frac{\partial J^\beta_T}{\partial \xi} \cdot \frac{d\bar{\xi}(T)}{dT} + \left(\|\phi^e - \bar{X}^e(\bar{\xi}(T))\|^2_{S^e} - \frac{1}{\beta} \|Z - H^e \bar{X}^e(\bar{\xi}(T))\|^2_{\mathfrak{m}} \right)(T) \\ &= \left(\|\phi^e - r^e\|^2_{S^e} - \frac{1}{\beta} \|Z - H^e r^e\|^2_{\mathfrak{m}} \right)(T), \end{aligned}$$

car $\bar{X}^e(\hat{\xi}(T)) = \bar{X}^e_{sup}(T) = r^e(T)$ et J^β_T est stationnaire en $\bar{\xi}$. Alors le choix particulier d'estimateur

$$\phi^e = r^e, \tag{4.72}$$

avec la condition initiale $V^\beta(0) = 0$, assure bien

$$\forall T, \quad V^\beta(T) \leq 0.$$

Ainsi nous avons un filtre H^∞ au sens de (4.63), *et il peut s'écrire sous forme réduite* (4.65). ■

b. Analyse d'erreur

Nous pouvons maintenant, comme dans le cadre H^2, effectuer l'analyse de l'erreur définie toujours par $\tilde{X} = X - \hat{X}$, $\tilde{\theta} = \theta - \hat{\theta}$. Une première remarque est que le contrôle robuste fournit immédiatement une première estimation car

$$\|Z - H^e \bar{X}^e(\zeta_\theta^e)\|_{\mathfrak{m}}^2 = \|\epsilon_h + H\check{X}\|_{\mathfrak{m}}^2,$$

et

$$\int_0^T \|\hat{X}^e - \bar{X}^e\|_{S^e}^2 \, dt \leq \frac{1}{\beta}\left(\int_0^T \|\epsilon_h + H\check{X}\|_{\mathfrak{m}}^2 \, dt + \|\zeta_\theta^e\|_{P_0^{-1}}^2\right). \tag{4.73}$$

Cependant, cette estimation d'erreur est intégrée en temps alors que nous sommes plus intéressés par une estimation au temps final qui, pour les paramètres (statiques dans la dynamique réelle), fournit un résultat plus fort. Nous reprenons donc notre changement de variable sur l'erreur $\eta = \tilde{X} - L^X\tilde{\theta}$. On a toujours la dynamique (4.37) où seule la dynamique de U est modifiée dans ce système. Ainsi η, dont la dynamique est celle de l'estimateur d'état, a la même stabilité que ce dernier. Pour $\tilde{\theta}$, nous modifions légèrement l'approche suivie dans le cas H^2 en calculant l'évolution de

$$\begin{aligned}\frac{d}{dt}(\tilde{\theta}^T U \tilde{\theta}) &= 2\tilde{\theta}^T U \dot{\tilde{\theta}} + \tilde{\theta}^T \dot{U} \tilde{\theta} \\ &= -\tilde{\theta}^T \big(L^{X^T} H' H L^X + \beta(L^{X^T} S_X L^X + S_\theta)\big)\tilde{\theta} + 2\tilde{\theta}^T L^{X^T} H' \varrho.\end{aligned}$$

En intégrant cette équation on obtient

$$\|\tilde{\theta}(t)\|_{U(t)}^2 \leq \|\zeta_\theta^e\|_{P_0^{-1}}^2 + 2\int_0^t \tilde{\theta}^T L^{X^T} H' \varrho \, d\tau. \tag{4.74}$$

Définissons alors $\lambda_U(t)$ comme la plus petite valeur propre du problème aux valeurs propres généralisées

$$U(t)\,\xi = \lambda U(0)\,\xi,$$

on obtient alors de (4.74)

$$\lambda_U(t)\|\tilde{\theta}(t)\|_{P_0^{-1}}^2 \leq \|\zeta_\theta^e\|_{P_0^{-1}}^2 + 2\int_0^t (\lambda_U(\tau))^{\frac{1}{2}} \|\tilde{\theta}(\tau)\|_{P_0^{-1}} (\lambda_U(\tau))^{-\frac{1}{2}} \|L^{X^T} H' \varrho\|_{P_0} \, d\tau,$$

qui, par application du lemme de Gronwall, donne la majoration

$$(\lambda_U(t))^{\frac{1}{2}} \|\tilde{\theta}(t)\|_{P_0^{-1}} \leq \|\zeta_\theta^e\|_{P_0^{-1}} + \int_0^t (\lambda_U(\tau))^{-\frac{1}{2}} \|L^{X^T} H' \varrho\|_{P_0} \, d\tau. \tag{4.75}$$

Nous devons alors minorer λ_U. Pour $T_{\max}$ donné, nous choisissons β_1 avec $0 < \beta_1 < \beta^\star$ et $\beta^\star$ défini par la proposition 4.2.3. On définit alors U_1 comme la solution de

$$\dot{U}_1 = L^{X^T} H' H L^X - \beta_1(L^{X^T} S_X L^X + S_\theta), \quad \text{avec } U_1(0) = (P_0)^{-1}.$$

La condition $0 < \beta_1 < \beta^\star$ implique que U_1 est définie positive sur $[0, T_{\max}]$. Donc pour tout β tel que $0 < \beta \leq \beta_1$ et U associé, nous avons

$$\frac{d}{dt}(U - U_1) = (\beta_1 - \beta)(L^{X^T} S_X L^X + S_\theta), \tag{4.76}$$

donc

$$U(t) = U_1(t) + (\beta_1 - \beta) \int_0^t (L^{X^T} S_X L^X + S_\theta)\, d\tau, \tag{4.77}$$

et puisque $U_1(t)$ est définie positive

$$\lambda_U(t) \geq (\beta_1 - \beta)\lambda_{\inf,1}(t), \tag{4.78}$$

où $\lambda_{\inf,1}(t)$ est la plus petite solution du problème aux valeurs propres généralisées

$$\left(\int_0^t (L^{X^T} S_X L^X + S_\theta)\, d\tau \right) \xi = \lambda U(0)\, \xi. \tag{4.79}$$

De même, à partir de

$$\frac{d}{dt}\left(U - \frac{\beta}{\beta_1} U_1 \right) = \frac{\beta_1 - \beta}{\beta_1} L^{X^T} H' H L^X, \tag{4.80}$$

on obtient

$$\lambda_U(t) \geq \frac{\beta_1 - \beta}{\beta_1}(1 + \lambda_{\inf,2}(t)), \tag{4.81}$$

où $\lambda_{\inf,2}(t)$ est la plus petite solution du problème aux valeurs propres généralisées

$$\left(\int_0^t L^{X^T} H' H L^X\, d\tau \right) \xi = \lambda U(0)\, \xi. \tag{4.82}$$

En assemblant (4.78) et (4.81), sachant qu'elles sont valides pour $\beta_1 < \beta^\star$, on obtient

$$\lambda_U(t) \geq \max\left((\beta^\star - \beta)\lambda_{\inf,1}(t)\ ,\ \frac{\beta^\star - \beta}{\beta^\star}(1 + \lambda_{\inf,2}(t)) \right). \tag{4.83}$$

On remarque que $\lambda_{\inf,1}(t)$ croit en $O(t)$ avec le terme S_θ dans l'intégrale, alors que l'évolution de $\lambda_{\inf,2}(t)$ correspond à la propriété d'excitation persistante de Zhang (2002).

Le reste de l'analyse d'erreur est calquée sur ce que nous avons vu dans le cadre H^2. La sensibilité paramétrique L^X est stable car elle suit la même dynamique que l'estimateur d'état et on a une borne du type

$$\|L_X(t)\theta\|^2_{\mathcal{E}'} \leq C \int_0^t \|B(\tau)\theta\|^2_{RHS}\, d\tau, \tag{4.84}$$

où $\|\cdot\|_{RHS}$ est toujours une norme adéquate sur le second membre de la dynamique. Reprenant alors les normes naturelles sur L_X and B

$$\|L_X\|_L = \sup_{\|\theta\|_{U(0)}=1} \|L_X\theta\|_{\mathcal{E}'}, \quad \|B\|_B = \sup_{\|\theta\|_{U(0)}=1} \|B\theta\|_{RHS}, \tag{4.85}$$

on obtient

$$\|L_X(t)\|^2_L \leq C \int_0^t \|B(\tau)\|^2_B\, d\tau. \tag{4.86}$$

On peut alors analyser les différentes contributions de ϱ et obtenir comme précédemment une estimation du type

$$\|L_X^T H' \varrho\|_{P_0} \leq \|L_X(\tau)\|_L \|H\eta + \epsilon_h + \chi\|_O \leq \|L_X(\tau)\|_L \left(\|\eta\|_{\mathcal{E}'} + \|\epsilon_h\|_O + \|\chi\|_O \right), \tag{4.87}$$

alors, en majorant par 1 le terme $(\lambda_U(\tau))^{-\frac{1}{2}}$ dans l'intégrale de (4.75), on déduit

$$\|\tilde{\theta}(t)\|_{P_0^{-1}} \leq (\lambda_U(t))^{-\frac{1}{2}} \Big(\|\zeta_\theta^e\|_{P_0^{-1}} + \|L_X\|_{L^2([0,t];L)} \big(\|\eta\|_{L^2([0,t];\mathcal{E}')} + \|\epsilon_h\|_{L^2([0,t];O)} + \|\chi\|_{L^2([0,t];O)} \big) \Big). \tag{4.88}$$

Cette borne associée à l'estimation sur η issue de l'analyse du chapitre 3 (voir Section 3.4.6.d.)

$$\|\eta\|_{L^2([0,t];\mathcal{E}')} \leq C \Big(\sqrt{T_1} \|\zeta^X\|_{\mathcal{E}} + \sqrt{t} \left(\gamma \sqrt{T_2} \|\chi\|_{L^2([0,t];O)} + \epsilon \right) \Big), \tag{4.89}$$

et (4.83) fournissent une estimation d'erreur complète.

Finalement, l'erreur d'estimation d'état finale se déduit de l'identité $\eta = \tilde{X} - L_X \tilde{\theta}$ par inégalité triangulaire

$$\|\tilde{X}\|_{\mathcal{E}'} \leq \|\eta\|_{\mathcal{E}'} + \|L_X\|_L \|\tilde{\theta}\|_{P_0^{-1}}. \tag{4.90}$$

c. Résultats numériques

Par rapport aux situation précédentes, nous pouvons désormais admettre d'autres types de bruit de mesure. Nous en sélectionnons trois comme suit :

1. **erreur gaussienne**, indépendante[4] par noeuds du maillage de référence et par pas de temps. Nous supposons alors une erreur de 20% de la vitesse maximum (Rappel $v_{\max} = 0.17\,\mathrm{m.s}^{-1}$) ;

2. **erreur gaussienne**, indépendante par noeuds du maillage de référence mais **identique au cours du temps**. C'est alors une forme de biais en temps pour un « bruit blanc » en espace ;

3. **erreur modale** dont le profil en espace correspond au premier mode non-amorti et constante en temps. L'amplitude de cette erreur est de 20% de l'énergie cinétique d'un champ de vitesse égal à $v_{\max}$.

Ces trois types d'erreur sont bornées dans $L^2(\Gamma_{\mathfrak{m}})$ comme indiqué dans l'analyse numérique de la Section 3.4.6. Elles sont de plus relativement importantes puisqu'elles sont définies à partir de la vitesse maximale sur la période de simulation.

Dans cette situation de contrôle robuste, le bruit n'étant plus supposé blanc, on ne modifie pas W pour respecter le rapport entre l'échantillonnage de la mesure et le pas de temps de la discrétisation. Il n'y a donc plus de motivation particulière à multiplier W par $\frac{\Delta T}{\delta t}$. De même, le covariance paramétrique n'est plus nécessairement celle d'un bruit blanc. On peut donc soit garder $P_0 = 1$ soit prendre $P_0 = \frac{1}{p}\mathbb{1}$. Dans ce deuxième cas, on limite en fait le champ de recherche de θ à une zone de rayon 1 suivant le principe que $\mathbb{E}(\|\theta\|^2) = \mathrm{tr}(P_0) = 1$. Vu notre normalisation des paramètres, on peut considérer que ce choix global pour les paramètres est plus adapté que de supposer chacune des composantes $\theta_i, 1 \leq i \leq p$ indépendantes et d'écart-type égal à 1.

[4]Pour autant on n'applique pas la règle de rééchantillonnage temporel comme pour le bruit blanc

La Figure 4.6 montre les paramètres estimés pour le premier type de bruit avec $\beta = 0$, ce qui signifie que le filtre H^∞ correspond au filtre H^2. Les résultats sont positifs sur la détection de la zone infarcie et l'estimation est rapide une fois que la contraction a lieu. Ce résultat doit être comparé à celui de la Figure 4.7 pour $\beta = 5$, une valeur proche du maximum numériquement autorisé pour assurer l'inversibilité de U sur la fenêtre de temps. Cette figure montre que la sensibilité de l'estimation paramétrique augmente de manière significative, ce qui est cohérent avec le fait que U est diminuée par la contribution H^∞ (cf. équation (4.65)). Cette sensibilité est particulièrement importante pour la région 17 qui n'est pratiquement pas observée (elle est au fond de l'endocarde et du ventricule). Cette région est donc celle où l'identifiabilité du problème est sans doute la moins bonne.

Des résultats similaires sont obtenus Figure 4.9 pour les deux autres types d'incertitude de mesure. L'estimation est moins performante dans le troisième cas qui représente véritablement un *biais* et contient une incertitude *physique*. Ce type de bruit est donc plus difficile à filtrer.

Enfin, Figure 4.9, nous présentons deux résultats d'estimation pour deux valeurs différentes du paramètre H^∞, $\beta = 2$ et $\beta = 5$ avec le biais modal dans les mesures. Ces résultats sont à comparer avec ceux de la Figure 4.8. En dehors des questions de sensibilité, nous ne constatons pas d'améliorations de l'estimation en augmentant β. Ce résultat est décevant par rapport au principe H^∞ où $\frac{1}{\beta}$ est une borne de performance comme indiqué en (4.63). C'est pourtant cohérent avec notre estimation finale (4.88), où (4.83) produit la borne minimale pour $\beta = 0$. Il n'y a en fait pas de paradoxe puisque le critère H^∞ est un critère sur toute la fenêtre en temps et il est « trop grand » pour être pertinent dans notre cas pratique. On se focalise donc sur le temps final et dans ce cas, l'algorithme H^2 suffit. Finalement, l'avantage de la démonstration H^∞ est qu'elle donne un sens à la formulation H^2, y compris pour des bruits plus généraux que les bruits blancs

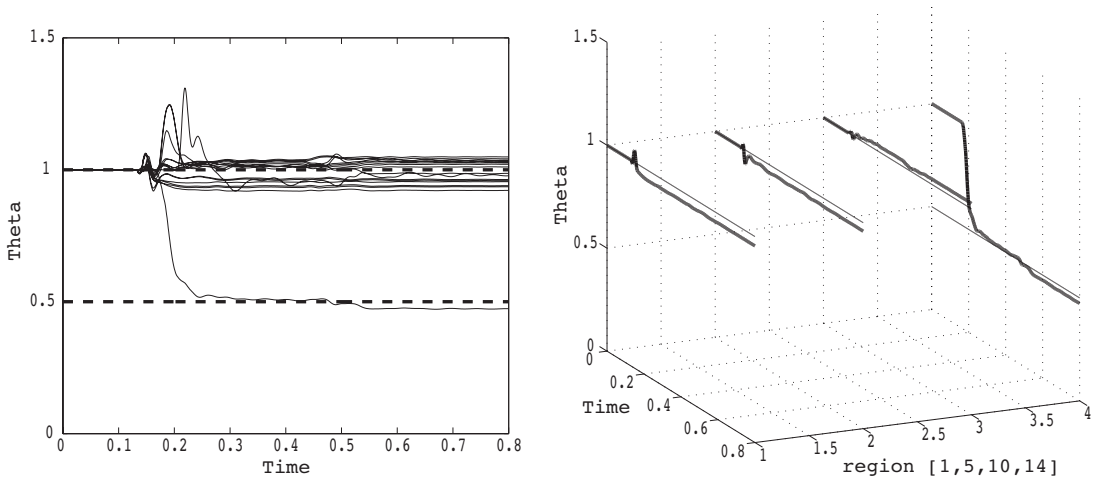


FIGURE 4.6: Incertitude du premier type avec $\beta = 0$: Paramètres estimés pour les 17 régions (gauche) et trajectoires spécifiques de 4 régions (droite)

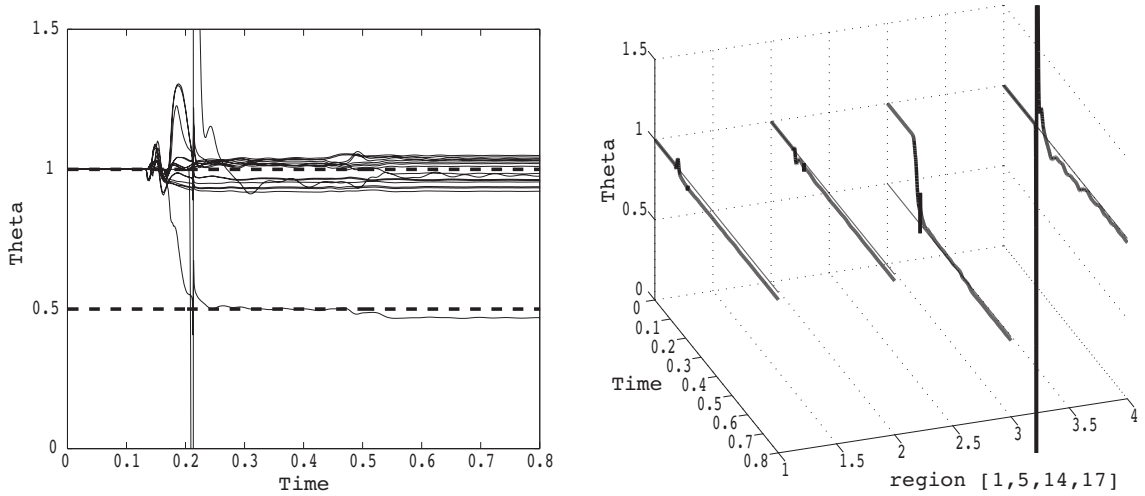


FIGURE 4.7: Incertitude du premier type avec $\beta = 5$: Paramètres estimés pour les 17 régions (gauche) et trajectoires spécifiques de 4 régions (droite)

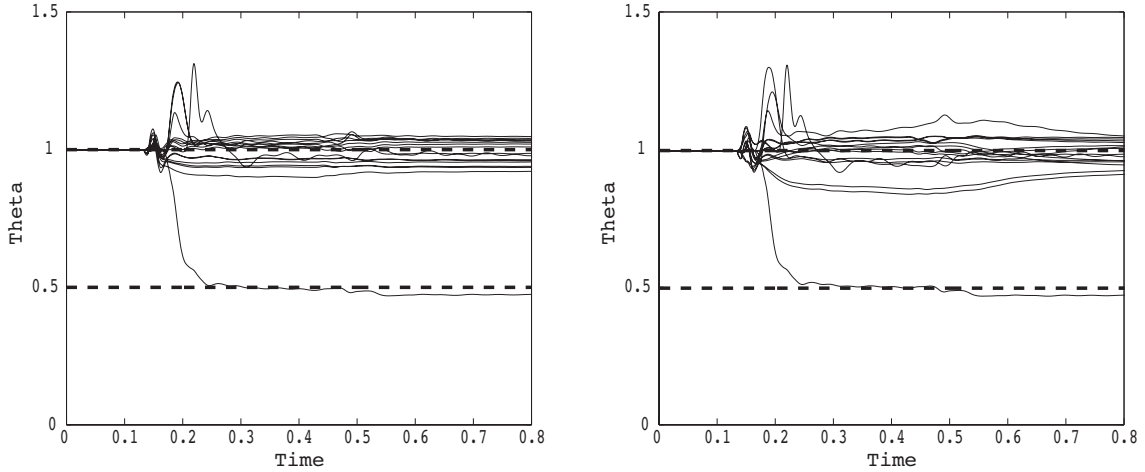


FIGURE 4.8: Incertitude du second (gauche) et troisième (droite) type avec $\beta = 0$

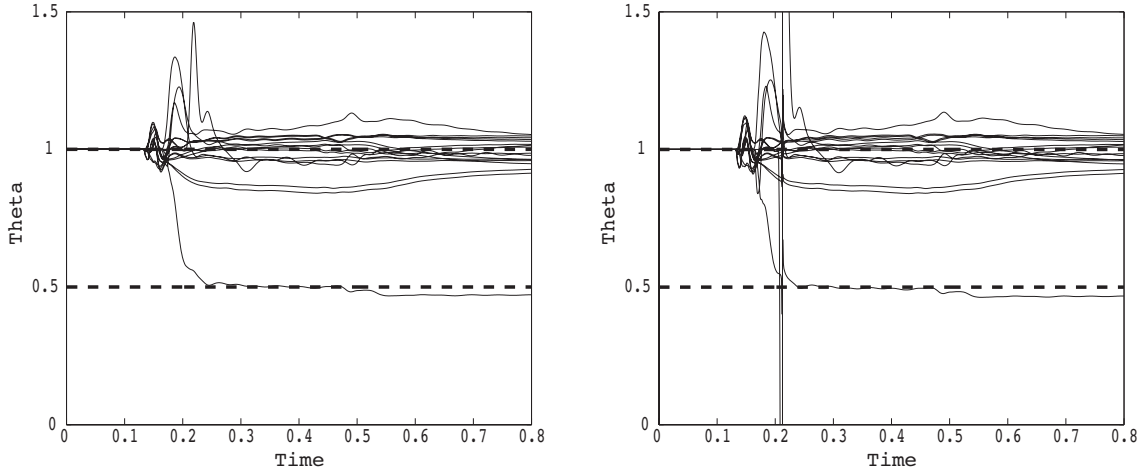


FIGURE 4.9: Incertitude du troisième type avec $\beta = 2$ ou $\beta = 5$: Superposition des 17 trajectoires

4.3 Cas bilinéaire et extensions du filtre SEEK

Nous venons d'analyser en profondeur le cas linéaire pour l'estimateur conjoint, où la linéarité concerne donc à la fois l'état (soit une hypothèse de petits déplacements) et les paramètres, et surtout leur couplage. Ceci limite donc l'identification à une dépendance paramétrique du second membre. Même si nous avons dit que finalement le linéarisé de tout système d'estimation peut se ramener à l'étude précédente, encore faut-il imaginer des filtres pour des modèles non-linéaires. A cette fin nous commençons par étudier le premier degré de non-linéarité lié au choix de paramètres identifié dans notre problématique. Nous nous intéressons en effet à l'estimation de module d'Young pour une dynamique toujours supposée linéaire. La dépendance de la raideur, donc de la dynamique, est linéaire par rapport au module d'Young mais la non-linearité vient du couplage état-paramètre puisque la puissance des efforts interne dépend linéairement du module d'Young mais aussi linéairement des déplacements, c'est-à-dire bilinéairement de l'état augmenté « état+paramètre ». Nous regardons donc un système du type

$$\begin{cases} \dot{x} = \mathcal{A}x + \mathcal{B} : \theta \otimes x + \mathcal{R} \\ x(0) = x_0 + \zeta^x \\ \theta = \theta_0 + \zeta^\theta \end{cases} \tag{4.91}$$

où θ représente le module d'Young dans chacune des régions AHA par exemple, et $\mathcal{B}$ comporte la différentielle de la puissance des efforts intérieurs par rapport au module d'Young, c'est-à-dire un terme classique de raideur où pour chaque région i, $E_i = 1$ alors que toutes les autres $E_j = 0$.

Nous pouvons explicitement écrire le système du second ordre discrétisé en espace sous la forme

$$M\ddot{Y} + C\dot{Y} + K(\theta)Y = F^{\text{ext}}(t),$$

avec ici $K(\theta)Y = \frac{\partial K}{\partial \theta} : \theta \otimes Y$ et $F^{\text{ext}}(t)$ peut contenir la contribution de type précontrainte précédemment présentée dans laquelle nous ne précisons plus la dépendance paramétrique. Nous insistons sur le fait que nous sommes encore dans cette section dans la formulation de problème simplifié afin de mettre en place les extensions non-linéaires de nos estimateurs.

4.3.1 Expression à partir de la sensibilité

La première idée pour étendre l'estimateur linéaire est de revenir à la formulation employée par Zhang et Clavel (2001) où il est remarqué que L^X est la sensibilité du système linéaire par rapport aux paramètres. Comme nous l'avions déjà noté, on a en effet pour le problème linéaire

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{dt}} \dfrac{\partial \bar{X}}{\partial \theta} = A \dfrac{\partial \bar{X}}{\partial \theta} + B \\ \dfrac{\partial X}{\partial \theta}(0) = 0 \end{cases}$$

donc $L^X = \frac{\partial \bar{X}}{\partial \theta}$. De cette interprétation nous pouvons formuler un estimateur pour le problème bilinéaire

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B : \hat{\theta} \otimes \hat{X} + R + K_X(Z - H\hat{X}) + L^X\dot{\hat{\theta}} \\ \dot{\hat{\theta}} = U^{-1}{L^X}^T H^T W^{-1}(Z - H\hat{X}) \\ \dot{L^X} = (A + B \cdot \hat{\theta} - K_X H)L^X + B \cdot \hat{X} \\ \dot{U} = (L^X)^T H^T W^{-1} H L^X \\ \hat{X}(0) = X_0 \\ \hat{\theta}(0) = \theta_0 \\ L^X(0) = 0 \\ U(0) = \left(\mathbb{E}(\zeta^\theta.{\zeta^\theta}^T)\right)^{-1} \end{cases} \tag{4.92}$$

Une première remarque, fondamentale pour l'efficacité de l'estimateur, s'impose dès maintenant. Désormais, la dynamique de L^X est dépendante des estimateurs $\hat{\theta}$ et $\hat{X}$. Autrement dit la sensibilité L^X est calculée par rapport à l'estimateur et non pas par rapport à la trajectoire cible inconnue. En linéaire, la structure particulière du système faisait qu'elles étaient égales mais ce n'est plus le cas en non-linéaire.

4.3.2 Analyse d'erreur

Nous pouvons continuer d'analyser l'erreur d'estimation via le changement de variable proposé précédemment $\begin{pmatrix} \tilde{X} \\ \tilde{\theta} \end{pmatrix} \to \begin{pmatrix} \eta \\ \tilde{\theta} \end{pmatrix}$. Dans ce cas on vérifie que pour η

$$\begin{aligned} \dot{\eta} &= \dot{\tilde{X}} - \dot{L^X}\tilde{\theta} - L^X\dot{\tilde{\theta}} \\ &= A\tilde{X} + B : \theta \otimes X - B\hat{\theta} \otimes \hat{X} - K_X H\tilde{X} + K_X(\epsilon_h + \chi) - L^X\dot{\hat{\theta}} \\ &\quad + L^X\dot{\hat{\theta}} - (A + B\hat{\theta} - K_X H)L^X\tilde{\theta} - B\tilde{\theta} \otimes \hat{X} \\ &= A\tilde{X} + B : \theta \otimes X - B\theta \otimes \hat{X} - K_X H\tilde{X} + K_X(\epsilon_h + \chi) - (A + B\hat{\theta} - K_X H)L^X\tilde{\theta} \\ &= (A + B\theta - K_X H)\eta + B : \tilde{\theta} \otimes L^X\tilde{\theta} + K_X(\epsilon_h + \chi), \end{aligned} \tag{4.93}$$

alors que la dynamique de l'erreur sur $\tilde{\theta}$ reste inchangée par rapport au linéaire

$$\dot{\tilde{\theta}} = -U^{-1}{L^X}^T H^T W^{-1} H L_X\tilde{\theta} - U^{-1}{L^X}^T H^T W^{-1} H\eta - U^{-1}{L^X}^T H^T W^{-1}(\epsilon_h + \chi). \tag{4.94}$$

Ce changement de variable ne permet plus de trigonaliser la dynamique de l'erreur. Cependant le système tangent en $\left(\begin{smallmatrix} \eta \\ \tilde{\theta} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$ reste inchangé. Donc on peut toujours appliquer le théorème 3.3.3 pour assurer l'existence d'un voisinage de 0 pour $\left(\begin{smallmatrix} \eta \\ \tilde{\theta} \end{smallmatrix}\right)$. L^X étant borné il existe donc un voisinage de 0 pour $\tilde{X}$ et $\tilde{\theta}$ tel que le système reste stable.

4.3.3 Premiers résultats

Nous nous plaçons dans les mêmes configurations numériques que précédemment mais modifions la raideur passive de la $14^{\text{ème}}$ région. Ainsi $\theta_{14} = 1.3$ alors que les autres restent à 1. Cet exemple illustre une augmentation de raideur pour un tissu nécrosé par la colonisation de

tissus conjonctifs. De plus pour des raisons d'observabilité nous ne choisissons pas P_0 « homothétique » (les écarts types dans chaque région sont égaux) mais assurons être presque sûrs des paramètres loin de la région nécrosée. L'observabilité d'une raideur est en effet toujours soumise à l'estimation d'une contrainte générale dans le tissu. Or nous n'avons pas de contrainte, ou de pression, dans les observations donc nous devons être certains de la raideur quelque part pour assurer l'identifiabilité des raideurs relativement à cet étalonnage.

□ TEST ALGORITHMIQUE – Nous commençons comme pour le cas linéaire avec un test algorithmique qui consiste à négliger tous les bruits autre que l'incertitude paramétrique. Les résultats sont présentés Figure 4.10 et préfigurent des mêmes performances qu'en linéaire. Sur les figures nous ne montrons que l'écart au cas sain $\theta_i = 1, \ \forall i$.

□ ERREUR DE CONDITION INITIALE – Nous reprenons maintenant un cas-test plus réaliste sur les incertitudes en réintroduisant un bruit blanc de mesure et une erreur de condition initiale. Nous souhaitons ici tester notre méthodologie pour plusieurs valeurs de surpression (10, 100 et 1000 Pa) pour la condition initiale « Pressure I.C. ». Les deux premières séries de figures (Figures 4.11 et 4.12) montrent les résultat pour 10 et 100 Pa. Autant pour 10 Pa l'estimation se comporte encore bien, autant dès 100 Pa, l'estimation diverge des valeurs attendues. De plus l'algorithme renvoie une erreur après avoir cherché à rendre θ négatif correspondant à une raideur non définie positive pour le système (i.e sur la figure les paramètres ont alors été bloqués à 0.5 comme expliqué ci-dessous Section 4.3.4). Ce résultat est cohérent avec l'analyse du filtre et indique simplement que le voisinage de stabilité est beaucoup plus étroit que ce que l'application impose.

Nous voyons donc surgir deux difficultés à la lumière de ces deux premiers résultats. Tout d'abord les paramètres peuvent devenir négatifs ce qui à la différence du linéaire a une certaine importance puisque dans ce cas la raideur n'est plus un opérateur positif. On pourrait imposer une plus faible tolérance sur P_0 mais dans ce cas on ne pourrait pas poursuivre suffisamment loin les grandes valeurs positives de θ. Plus fondamentalement, la sensibilité paramétrique L^X est pour le modèle bilinéaire fonction de $\hat{X}$ et $\hat{\theta}$ donc c'est la dépendance paramétrique de l'estimateur mais pas du vrai système. Quand l'erreur de condition initiale est importante, l'observateur est « très faux » au départ donc L^X est génératrice d'erreur dans la grammienne. En pratique dans notre exemple, la méthode cherche à adapter les paramètres à l'estimateur d'état en mouvement alors que la cible est statique au début de la simulation. Les paramètres s'ajustent alors en tendant vers 0 pour « stabiliser à leur manière » l'observateur par rapport aux observations fournies. Nous découvrons avec ces premiers résultats la nécessité fondamentale de coupler l'estimateur paramétrique à l'estimateur d'état. S'il n'y a pas d'erreur sur l'état initialement, l'estimateur paramétrique fonctionne très bien car dans ce cas la sensibilité de l'estimateur par rapport aux paramètres est alors très proche de celle du système réel. Puisque cette sensibilité est alors bien évaluée, l'estimateur sait comment faire évoluer positivement ces paramètres en fonction de l'identifiabilité du système réel. Dans les autres cas, l'erreur sur l'état empêche de façon irréversible la détermination d'une sensibilité cohérente avec la cible. Il faut donc la corriger et le plus rapidement possible.

De ce point de vue, le cas linéaire était un peu exceptionnel car la sensibilité paramétrique L^X suivait une dynamique indépendante de $\hat{X}$ et $\hat{\theta}$, donc indépendante de l'erreur

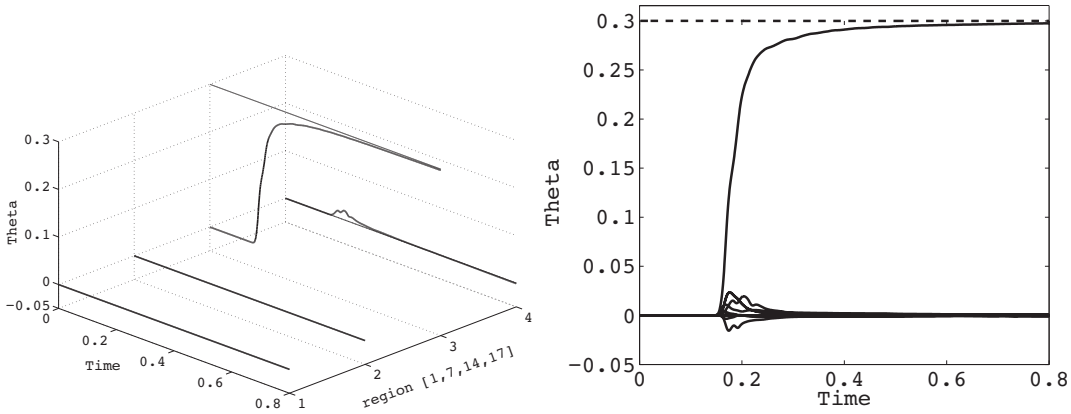


FIGURE 4.10: Test algorithmique de l'identification bilinéaire « en conditions idéales »

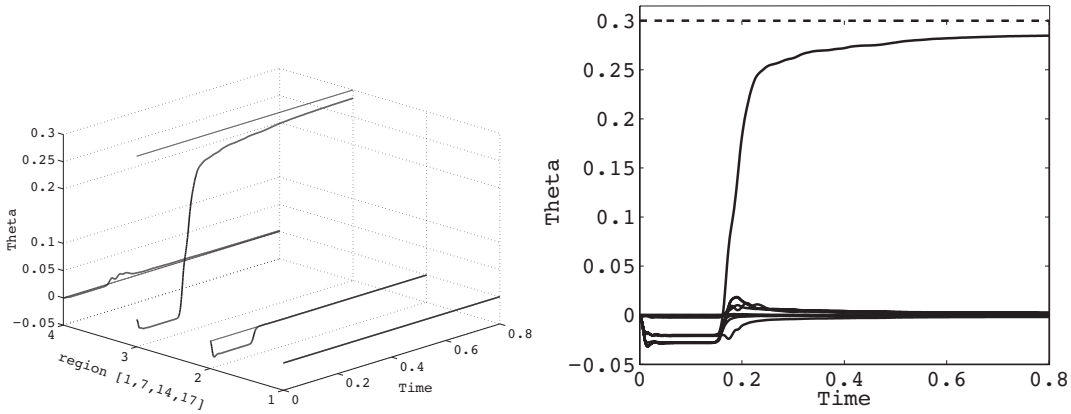


FIGURE 4.11: Estimation conjointe pour le problème bilinéaire avec une erreur de condition initiale générée par une surpression de 10 Pa

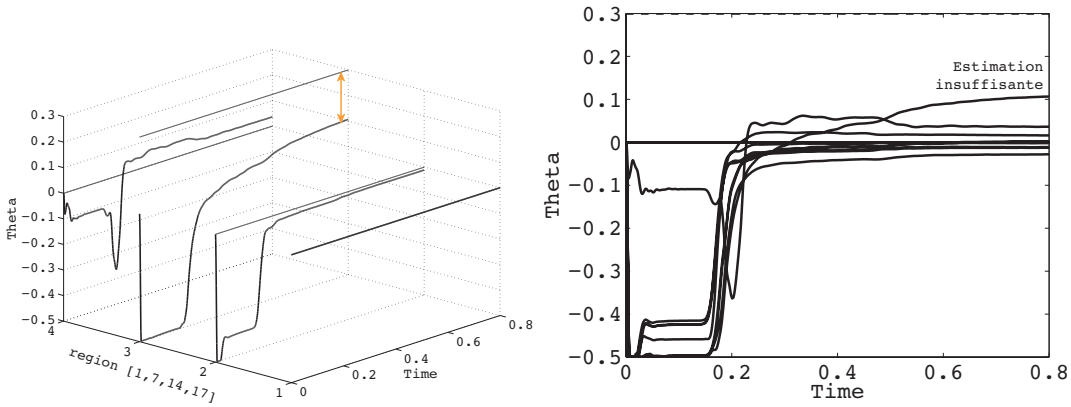


FIGURE 4.12: Estimation conjointe pour le problème bilinéaire avec une erreur de condition initiale générée par une surpression de 100 Pa

de condition initiale. Par ailleurs, il faut bien avoir conscience que ce type de sensibilité apparaît plus ou moins indirectement dans tous les algorithmes d'identification. Par exemple pour un algorithme variationnel classique c'est p_θ la variable adjointe correspondant aux paramètres qui joue ce rôle. Donc cette nécessité d'une maîtrise conjointe de l'estimation d'état et de l'estimation paramétrique est indispensable en assimilation de données quelle que soit la méthode. Voyons maintenant comment nous pouvons aménager notre méthodologie pour pouvoir prendre en compte le même type d'erreur de condition initiale que pour le problème linéaire.

4.3.4 Amélioration des performances de l'estimateur conjoint en non-linéaire

a. Estimation paramétrique différée

La première idée pour éviter cette « divergence » du filtre au début de l'estimation consiste à ne pas faire confiance à l'estimateur d'état en début d'estimation mais à attendre sa convergence rapide vers le véritable état. En effet l'estimation d'état seule est relativement robuste vis-à-vis de l'erreur paramétrique. Autrement dit ce n'est pas, toutes choses égales par ailleurs, une erreur de 30% sur les paramètres de raideur qui empêcherait l'état de se rapprocher sensiblement de la cible en début d'estimation. Après cette première phase, l'état de l'estimateur se retrouve alors dans le voisinage de stabilité pour le problème bilinéaire donc on peut réaliser l'estimation conjointe. Cette idée peut être mise en oeuvre directement dans l'algorithme proposé ci-dessus en retardant le déclenchement du calcul de la sensibilité d'un temps t_s tant que celle-ci est « particulièrement » fausse.

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B : \hat{\theta} \otimes \hat{X} + R + K_X(Z - H\hat{X}) + L^X \dot{\hat{\theta}} \\ \dot{\hat{\theta}} = U^{-1} L^{X^T} H^T W^{-1}(Z - H\hat{X}) \\ \dot{L^X} = 0, \quad \text{si } t \leq t_s \\ \dot{L^X} = (A + B \cdot \hat{\theta} - K_X H)L^X + B \cdot \hat{X}, \quad \text{si } t > t_s \\ \dot{U} = (L^X)^T H^T W^{-1} H L^X \\ \hat{X}(0) = X_0 \\ \hat{\theta}(0) = \theta_0 \\ L^X(0) = 0 \\ U(0) = \left(\mathbb{E}(\zeta^\theta . \zeta^{\theta^T})\right)^{-1} \end{cases} \tag{4.95}$$

Cette modification dans l'estimateur nous permet d'obtenir des performances comparables au modèle linéaire. Elles sont présentées Figure 4.13 pour le problème-test complet comprenant erreurs de mesures, de discrétisation et de condition initiale en surpression de nouveau de 1000 Pa.

Il n'y a plus, pour les estimateurs non-linéaires, d'équivalence avec un problème de contrôle optimal. Le critère variationnel n'est donc plus valide. Cependant, si nous avions utilisé le même retard sur le modèle linéaire, ce critère serait alors

$$J_T(\xi) = \frac{1}{2}\xi^T U(0)\xi + \frac{1}{2}\int_{t_s}^{T} (Z - H^e \bar{X}^e(\xi))^T W^{-1}(Z - H^e \bar{X}^e(\xi))\, dt.$$

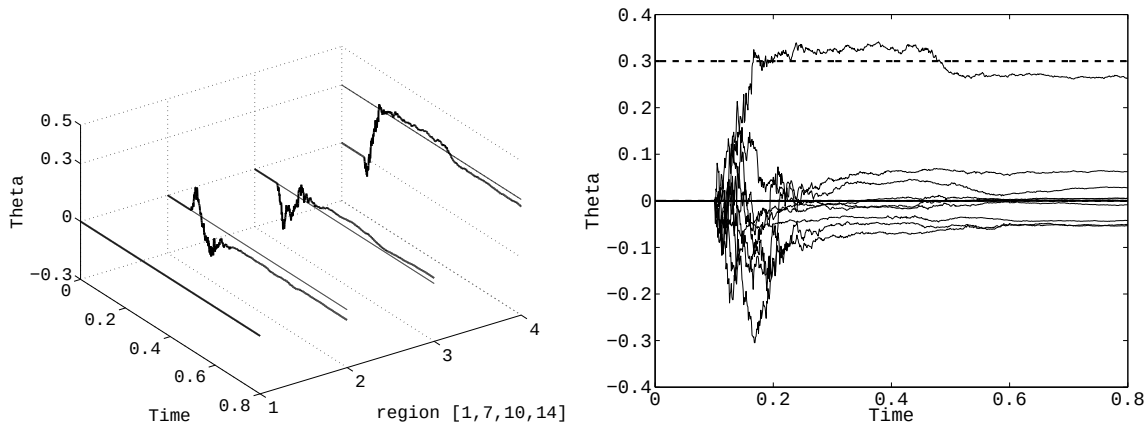


FIGURE 4.13: Estimation paramétrique pour une erreur de condition initiale en surpression de 1000 Pa. L'identification est déclenché après $t_s = 0.1s$

Or ce critère est pertinent pour éviter d'accumuler sur les premiers instants une erreur d'observation a priori fausse puisque Z est très loin de $H\bar{X}$.

REMARQUE 4.3.1 (FACTEURS D'OUBLI)
En fait, ce principe de minimisation sur de plus petites fenêtres en temps est un cas limite de l'utilisation de *facteurs d'oubli* en filtrage. Hoteit (2001) rappelle que le facteur d'oubli consiste à limiter l'impact des observations passées au travers de la dynamique de U en la modifiant par

$$\dot{U} = -\rho U + (L^x)^T H^T W^{-1} H L^x.$$

b. Contraintes paramétriques

Puisque l'estimateur s'est désormais éloigné de la commande optimale, il est possible d'envisager des variations heuristiques autour de sa formulation finale. Nous avons, entre autres, essayé d'intégrer simplement des contraintes paramétriques en « cloisonnant » l'évolution des paramètres comme suit

$$\begin{cases} \dot{\hat{\theta}} = U^{-1} L^{x^T} H^T W^{-1} (Z - H\hat{X}), & \text{if } \theta_{\min} \leq \hat{\theta} \leq \theta_{\max} \\ \dot{\hat{\theta}} = 0, \quad \text{sinon} \end{cases} \tag{4.96}$$

Au niveau de la mise en oeuvre nous avons alors traduit cette contrainte en ne prenant compte la phase de correction que pour $\theta_{\min} \leq \theta \leq \theta_{\max}$.

Dans le cas présenté Figure 4.14, nous avons repris la situation de l'échec Figure 4.12 et décidé de limiter la borne minimale pour les paramètres à 0.8 ($\tilde{\theta} = -0.2$). On remarque alors que l'estimation est nettement meilleure sans avoir recours à un décalage de l'estimation t_s. D'ailleurs Figure 4.12 nous avions aussi introduit une contrainte paramétrique $\theta_{\min} = 0.5$ pour éviter que la raideur soit singulière mais cette borne était manifestement trop grande.

REMARQUE 4.3.2 (DÉPENDANCE PARAMÉTRIQUE EXPONENTIELLE)
Une piste intéressante de gestion des paramètres notamment vis-à-vis des éventuelles bornes est de formaliser leur dépendance directement à partir de fonctions qui prennent en compte directement ces

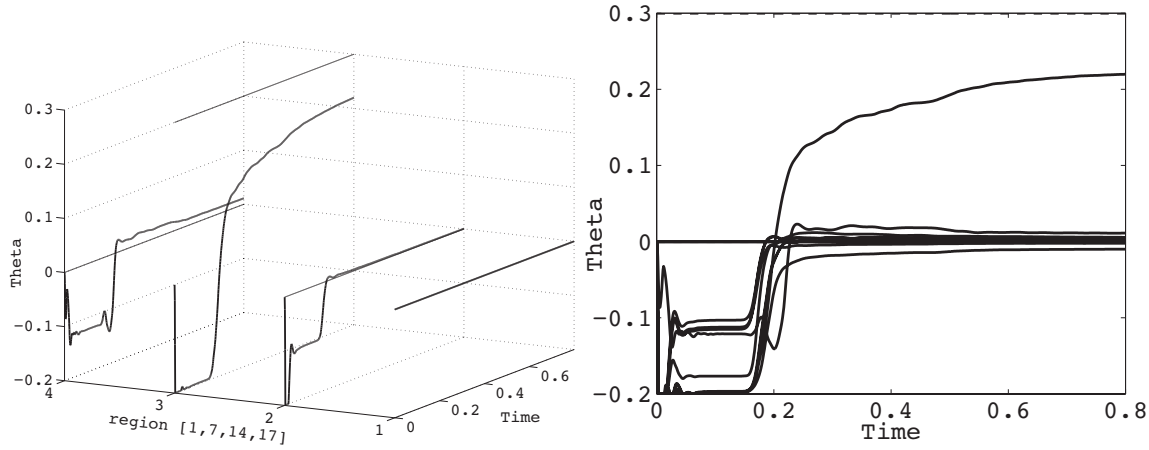


FIGURE 4.14: Exemple d'estimation avec contraintes paramétriques pour la surpression de 100 Pa et $t_s = 0$

bornes dépendance. Par exemple dans notre cas où nous cherchons une rigidité définie positive par rapport à la configuration idéale on pourrait exprimer la dépendance paramétrique sous la forme

$$K(\theta)Y = \frac{\partial K}{\partial \theta} : e^{\alpha\theta} \otimes Y,$$

avec $\alpha = \ln(10)$ typiquement. Dans ce cas, en plus d'éviter des valeurs négatives des paramètres on symétrise leur évolution par rapport à la valeur de référence. En effet il est plus adéquat de considérer que la densité de probabilité associé aux paramètres est une gaussienne sur cette échelle logarithmique, c'est-à-dire qu'il est d'égale importance d'avoir un tissu deux fois plus raide ou deux fois plus mou. Evidemment dans ce cas on n'entre plus dans le formalisme d'un modèle bilinéaire et la sensibilité paramétrique doit être calculée par rapport à cette dépendance exponentielle des paramètres.

4.3.5 Equivalence avec EKF et le filtrage non-linéaire

a. Formulation EKF

Puisque L^X suit une dynamique correspondant au tangent de la dynamique réelle, il est clair que les démonstrations présentées précédemment en linéaire permettent de vérifier que $P = L^X U^{-1} L^X$ est bien solution de l'équation de Riccati associé aux opérateurs de la dynamique tangents. Donc l'estimateur (4.92) n'est rien d'autre que le filtre de Kalman étendu (EKF) réduit dont l'extension avait déjà été présentée dans Pham et al (1997).

Cette équivalence avec le filtre EKF permet de bénéficier des résultats de convergence asymptotique introduits par Ljung (1979). Le modèle état+paramètres que nous considérons entre en effet dans la catégorie des « modèles déterministes » dans la terminologie de ces travaux avec un opérateur de transition stable et sans bruit de modèle. Il démontre alors (théorème 6.1) que l'estimation paramétrique converge presque sûrement vers un minimum de l'erreur d'observation. Reste que ce résultat asymptotique ne fournit pas d'information supplémentaire sur la vitesse de convergence. Il permet simplement d'assurer une certaine stabilité en remplaçant dans le cas non-linéaire le critère variationnel.

b. Filtre du second ordre

Puisque notre estimateur est un filtre EKF, on peut se demander s'il est possible d'appliquer des filtres de Kalman d'ordre plus élevé. Pour cela, il faut que ces filtres admettent une version réduite. Or pour notre problème bilinéaire, pour un opérateur d'observation linéaire H, le filtre de Kalman du second ordre se formule très rapidement à partir des seules opérateurs L^X et U. Puisque H est linéaire, la seule modification du filtre concerne la « prédiction » avec un terme $\frac{1}{2}\frac{\partial^2 A^e}{\partial X^2} : P^e$. On a par construction

$$P = \begin{pmatrix} L^X \\ 1 \end{pmatrix} U^{-1} \begin{pmatrix} L^{X^T} & 1 \end{pmatrix} \quad \text{et} \quad \frac{\partial^2 A^e}{\partial X^2} = \begin{pmatrix} 0 & B \\ B & 0 \end{pmatrix},$$

donc

$$\frac{\partial^2 A^e}{\partial X^2} : P^e = B : L^X U^{-1}.$$

L'estimateur devient donc

$$\dot{\hat{X}} = A\hat{X} + B : \hat{\theta} \otimes \hat{X} + R + K_X(Z - H\hat{X}) + L^X \dot{\hat{\theta}} + B : L^X U^{-1}, \tag{4.97}$$

les autres équations restant inchangées.

Bien qu'il n'y ait pas d'éléments théoriques précis pour ce filtre, on peut toujours regarder comment l'erreur d'estimation est modifiée par la correction. On reprend les mêmes étapes de calculs,

$$\dot{\eta} = (A + B\theta - K_X H)L^X\tilde{\theta} + B : \tilde{\theta} \otimes L^X\tilde{\theta} - B : L^X U^{-1} + K_X(\epsilon_h + \chi).$$

Comme $B : \tilde{\theta} \otimes L^X\tilde{\theta} = B : \tilde{\theta}^T\tilde{\theta}L^{X^T}$, on peut donc écrire

$$\dot{\eta} = (A + B\theta - K_X H)L^X\tilde{\theta} + B : (\tilde{\theta}^T\tilde{\theta}L^{X^T} - U^{-1}) + K_X(\epsilon_h + \chi).$$

Or c'est justement la vocation de U^{-1} de représenter $\mathbb{E}(\tilde{\theta}\tilde{\theta}^T)$ donc on voit qu'au niveau de l'erreur d'estimation on vise à éliminer un biais présent dans l'application directe du filtre SEEK. D'après les critères de qualité attendus d'un estimateur énumérées en introduction du premier chapitre, ce filtre (4.97) apparaît plus abouti que (4.92) pour « peu » qu'on soit capable de bien estimer la covariance paramétrique à travers P.

4.4 Cas non-linéaire et filtre SEIK

Nous venons de voir via le problème bilinéaire que, dans le cas d'une dynamique non-linéaire (ou d'observateurs non-linéaires), il était possible d'étendre la formulation de l'estimateur conjoint via la sensibilité. Mais ce faisant, nous utilisions simplement une version réduite du filtre de Kalman étendu. Cette équivalence a alors justifié la problématique de la généralisation d'autres filtres d'inspiration Kalmanienne au cas du rang réduit, ce que nous avons commencé à faire autour du filtrage de Kalman du second ordre. Cependant, si on peut ainsi résumer l'opinion de la communauté du filtrage sur le filtre du second ordre, on peut dire qu'il a rarement fait ses preuves ou du moins qu'il a donné des résultats très dépendant du contexte d'utilisation (voir notamment Simon (2006) et références pour quelques exemples).

Au contraire, le filtre UKF décrit au chapitre 1 a « nettement meilleure presse » en terme de robustesse. Notamment pour des systèmes mécaniques de faible dimension Wu et Smyth (2006) ont prouvé son efficacité pour estimer toute la gamme des paramètres de lois de comportement complexes. La difficulté reste de proposer un estimateur réduit pour ce filtre. Or la littérature n'offre pas de filtre UKF réduit se réclamant comme tel à notre connaissance (ni celle de Simon (2006)). Cependant, Pham et al (1997) a proposé une extension de son filtre SEEK sous le nom de filtre SEIK pour *Singular Evolutive Kalman Filter* où les opérteurs tangents du filtre EKF sont remplacés par une interpolation des opérateurs non-linéaires. Nous allons démontrer que **ce filtre est une formulation réduite du filtre UKF** pour des particules simplectiques au sens du chapitre 1. Pour mémoire, la publication du filtre SEIK date de 1996 sous forme de rapport technique alors que l'article fondateur pour le filtre UKF est certes de 1995 mais à été reconnu tardivement dans la communauté du filtrage (cf. notes de Simon (2006)).

4.4.1 Filtre SEIK

a. Présentation du filtre SEIK

La principale motivation de Pham et al (1997) dans la définition de son filtre est de remplacer la linéarisation opérée dans la version EKF du filtre SEEK par une interpolation qui soit exacte lorsque les opérateurs sont du second ordre. On reconnaît déjà ici le principe de la formulation UKF. Commençons par un résumé, sans démonstration, des différentes étapes de prédiction et correction du filtre SEIK.

Soit p la dimension de l'espace réduit sur lequel on veut définir le filtre SEIK (pour nous p sera la dimension de l'espace paramétrique) et soit $T \in \mathbb{M}_{p+1,p}$ la matrice définie par

$$T = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ 0 & \cdots & 0 \end{pmatrix} - \frac{1}{p+1} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

Le filtre SEIK formulé par Pham et al (1997) est défini comme suit[5]

- **Echantillonnage** :

$$\begin{cases} C_n = \sqrt{U_n^{-1}} \\ \hat{X}_{n,i}^+ = \hat{X}_n^+ + L_n C_{n,i}^T, \quad 1 \leq i \leq p+1 \end{cases} \tag{4.98a}$$

- **Prédiction** :

$$\begin{cases} \hat{X}_{n+1,i}^- = A(\hat{X}_{n,i}^+) \\ \hat{X}_{n+1}^- = \dfrac{1}{p+1} \displaystyle\sum_{1 \leq i \leq p+1} X_{n+1,i}^- \end{cases} \tag{4.98b}$$

- **Correction** :

[5] Attention, par rapport à Pham et al (1997), nous utilisons des notations inversées. Sa covariance U_n correspond à notre opérateur U_n^{-1}.

$$\begin{cases} L_{n+1} = [\hat{X}^-_{n+1,i}]T \in \mathbb{M}_{N,p} \\ Z_{n+1,i} = H(\hat{X}^-_{n+1,i}) \\ \{HL\}_{n+1} = [Z_{n+1,i}]T \\ U_{n+1} = (p+1)T^T T + \{HL\}^T_{n+1} W^{-1} \{HL\}_{n+1} \in \mathbb{M}_p \\ \hat{Z}_{n+1} = \dfrac{1}{p+1} \displaystyle\sum_{1 \le i \le p+1} Z^-_{n+1,i} \\ \hat{X}^+_{n+1} = \hat{X}^-_{n+1} + L_{n+1} U^{-1}_{n+1} \{HL\}^T_{n+1} W^{-1} (Z_{n+1} - \hat{Z}_{n+1}) \end{cases} \tag{4.98c}$$

Nous utilisons la notations $\sqrt{U_n^{-1}}$ abusivement en voulant exprimer le choix de n'importe quelle matrice C_n telle que

$$U_n^{-1} = C_n C_n^T.$$

Une décomposition de Cholesky satisfait cette condition autant que la racine carrée principale de la matrice U_n^{-1}.

On remarque de nombreuses analogies structurelles entre les filtres SEEK et SEIK mais aussi entre filtre SEIK et UKF. En effet, SEIK comme UKF utilise un nombre restreint de particules, de l'ordre de la dimension de l'espace d'incertitude, situées à une distance de l'ordre de $\sqrt{U_n^{-1}}$. Pour comprendre l'équivalence entre les deux, il nous faut simplement analyser le rôle de T dans cet algorithme et dans la répartition particulaire

$$X^+_{n,i} = \hat{X}^+_{n-1} + L_n C_{n,i} = \hat{X}^+_{n-1} + [X_{n,i}] T C_{n,i}.$$

Pour Pham et al (1997), la matrice $T \in \mathbb{M}_{p+1,p}$ est choisie pour deux propriétés :

- Les $p+1$ lignes de T, notées $T_{,j}$, sont de somme nulle

$$\sum_{1 \le j \le p+1} T_{,j} = 0 \; ; \tag{4.99}$$

- T est de rang p

$$\text{rang}(T) = p.$$

Ainsi, les p colonnes de T forment une base du sous espace vectorielle de $\mathbb{R}^{p+1}$ des vecteurs dont la somme des composantes est nulle. Notons l la forme linéaire sur $\mathbb{R}^{p+1}$ calculant la somme des composantes

$$\forall V \in \mathbb{R}^{p+1}, \quad l(V) = \sum_{1 \le j \le p+1} V_j.$$

Dans la base canonique, $l = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix}$ et les colonnes de T forment une base de l'hyperplan de $\text{Ker}(l)$. On comprend maintenant comment Pham et al (1997) construit sa matrice à partir de la base canonique $(e_i)_{1 \le i \le p+1}$. Il choisit $e^* = \sum_{1 \le i \le p+1} e_i$ un vecteur de l'orthogonal du noyau (i.e. tout simplement l^T) et construit ses colonnes de T via

$$T_i = e_i - \frac{l(e_i)}{l(e^*)} e^*.$$

Oublions l'aspect réduit de l'algorithme SEIK pour nous focaliser sur l'espace paramétrique. Considérons alors le principe des sigma-points simplectiques pour une transformation UKF

sur un espace de dimension p (cf. Section 1.4.4). Ceux-ci sont au nombre de $p+1$ et leur choix est conditionné par leur moyenne et covariance empiriques. Pour des points simplectiques canoniques $(I_i)_{1\leq i\leq p+1}$, nous avions dit qu'il fallait donc que

$$\frac{1}{p+1}\sum_{1\leq i\leq p+1} I_i = 0,$$

et

$$\frac{1}{p+1}\sum_{1\leq i\leq p+1} I_i I_i^T = \mathbb{1} \Rightarrow \text{rang}(\sum_{1\leq i\leq p+1} I_i I_i^T) = p.$$

Ces deux conditions pour cet ensemble de points I_i (dans un espace affine donc) sont donc duales de celles sur T. Autrement dit, T **traduit les conditions satisfaites par les sigma-points** I_i **dans l'espace vectoriel des coordonnées**. Afin de vérifier définitivement cette hypothèse, nous allons donner à la section suivante une formulation matricielle du filtre UKF qui utilise des matrices analogues à T pour définir les sigma-points.

b. Formulation matricielle de la transformation *unscented*

Notre raisonnement au paragraphe précédent n'est pas complet tant qu'on ne comprend pas comment des sigma-points peuvent être fabriqués effectivement à partir de T en vérifiant non pas simplement le rang de la matrice de covariance (i.e. $\text{rang}(\sum_{1\leq i\leq p+1} I_i I_i^T) = p$) souhaité, mais ses valeurs effectives (i.e. $\frac{1}{p+1}\sum_{1\leq i\leq p+1} I_i I_i^T = \mathbb{1}$). Afin d'être parfaitement généraux et d'étendre les résultats pour différents choix de sigma-points, pas seulement simplectiques, nous introduisons les notations suivantes. Soit $(X_i)_{1\leq i\leq r}$ un ensemble de sigma-points généraux définis comme indiqué Section 1.4.4, (1.104). Soit alors $D_\alpha = \text{diag}((\alpha_i)_{1\leq i\leq r})$ et $T_\alpha \in \mathbb{M}_{r,p}$ définie par

$$\begin{cases} \sum_{1\leq i\leq r} \alpha_i (T_\alpha)_{.,i} = 0 \\ \text{rang}(T) = p \end{cases} \tag{4.100}$$

ou, de plus, T_α est le dual dans l'espace des coordonnées des contraintes définies sur les sigma-points considérés. On a immédiatement

$$X_\alpha = [X_i](\alpha_i) = [X_i]D_\alpha(1),$$

où pour simplifier les notations on utilise $(\alpha_i) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}$ et en particulier $(1) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

On démontre alors la propriété suivante,

PROPOSITION 4.4.1

$$P_\alpha = [X_i]D_\alpha T_\alpha (T_\alpha^T D_\alpha T_\alpha)^{-1} T_\alpha^T D_\alpha [X_i]^T. \tag{4.101}$$

■ ***Démonstration :*** *On a par définition de T_α,*

$$[X_i]D_\alpha T_\alpha = [X_i - X_\alpha]D_\alpha T_\alpha.$$

Ensuite, sachant que T_α traduit dans l'espace des coordonnées les contraintes sur les sigma-points, on sait que

$$\exists Q, [X_i - X_\alpha]^T = T_\alpha . Q.$$

Ces deux relations permettent alors le calcul

$$\begin{aligned}
P_\alpha &= \sum_{1\leq i\leq r} \alpha_i (X_i - X_\alpha)(X_i - X_\alpha)^T \\
&= [X_i - X_\alpha] D_\alpha [X_i - X_\alpha]^T \\
&= [X_i - X_\alpha] D_\alpha T_\alpha Q \\
&= [X_i - X_\alpha] D_\alpha T_\alpha (T_\alpha^T D_\alpha T_\alpha)^{-1} T_\alpha^T D_\alpha T_\alpha Q \\
&= [X_i - X_\alpha] D_\alpha T_\alpha (T_\alpha^T D_\alpha T_\alpha)^{-1} T_\alpha^T D_\alpha [X_i - X_\alpha]^T \\
&= [X_i] D_\alpha T_\alpha (T_\alpha^T D_\alpha T_\alpha)^{-1} T_\alpha^T D_\alpha [X_i]^T,
\end{aligned}$$

et fournissent alors une relation matricielle sur P_α à partir des matrices de sigma-points $[X_i]$ ■

Nous pouvons désormais trouver le dual de T dans l'espace des sigma-points par rapport à D_α. Examinons ce que cela donne pour les sigma-points unitaires, c'est-à-dire ceux tels que $X_\alpha = 0$ et $P_\alpha = \mathbb{1}$. La propriété 4.4.1 conduit immédiatement à choisir

$$[I_i] = (T_\alpha^T D_\alpha T_\alpha)^{-\frac{1}{2}} T_\alpha^T. \tag{4.102}$$

Inversement si on choisit les sigma-points proposés par Julier et al (2000), alors il est aisé de construire une matrice T_α convenable

$$T_\alpha = [I_i]^T,$$

à partir de sigma-points unitaires vérifiant, par définition,

$$[I_i] D_\alpha [I_i]^T = \mathbb{1}.$$

On commence maintenant à cerner le rôle de la matrice T_α dans la présentation de Pham et al (1997). Celle-ci permet d'exprimer la covariance empirique à partir d'opérations matricielles. Ces opérations laissent envisager une version réduite du filtrage UKF, ce que nous allons finalement examiner à la section suivante. Mais avant ça, fixons nous les idées définitivement sur quelques exemples de matrices T_α et $[X_i]$.

- **sigma-points simplectiques** En utilisant la matrice de Pham on peut donc construire un ensemble de points simplectiques. Ces points sont équivalents à ceux présentés aux chapitre 1 à une rotation près. En dimension 3 on peut comparer

$$[X_i]_{\text{Pham}} = \begin{pmatrix} \frac{5}{3} & -\frac{1}{3} & -\frac{1}{3} & 1 \\ \frac{1}{3} & \frac{5}{3} & -\frac{1}{3} & 1 \\ -\frac{1}{3} & -\frac{1}{3} & \frac{5}{3} & 1 \end{pmatrix} \text{ et } [X_i]_{\text{UKF}} = \begin{pmatrix} \sqrt{2} & -\sqrt{2} & 0 & 0 \\ \frac{\sqrt{2}}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & -\frac{2\sqrt{2}}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \sqrt{3} \end{pmatrix}.$$

Nous avions aussi vu que la matrice T de Pham et al (1997) est construite à partir de

$$T_i = e_i - \frac{l(e_i)}{l(e^*)} e^*.$$

où $e^* = \sum_{1\leq i\leq p+1} e_i$ est un vecteur orthogonal au noyau de l. On peut tout aussi bien choisir comme vecteur $e^* = e_n$. On obtient alors une matrice T_α

$$T_\alpha = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ -1 & \cdots & -1 \end{pmatrix},$$

dans ce cas on retrouve de façon un peu étonnante les mêmes sigma-points $[X_i]_{\text{Pham}}$.

- **sigma-points canoniques**. Il n'est pas difficile de construire la matrice T permettant l'utilisation de sigma-points canoniques. Il suffit de prendre $[X_i]^T$, à une renormalisation possible près, ou de construire l'orthogonal des contraintes vérifiées par les points. On obtient par exemple

$$T_\alpha = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix},$$

- **sigma-points étoilés** Dans ce cas, la définition des points dépend du choix des poids D_α. Cependant les contraintes permettant de définir T sont simples et donnent, entre autres,

$$T_\alpha = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ -1 & & 0 \\ & \ddots & \\ 0 & & -1 \\ 0 & \cdots & 0 \end{pmatrix}.$$

c. SEIK comme filtre UKF réduit

□ CAS SIMPLECTIQUE – Nous commençons par réinterpréter le filtre SEIK pour le choix de matrices T_α effectué par Pham et al (1997). Commençons ici par la phase, plus simple, de prédiction et considérons donc qu'on ait pu écrire la covariance a posteriori de rang p à t_n comme pour le filtre SEEK

$$P_n^+ = L_n U_n^{-1} L_n^T.$$

On définit les $p+1$ sigma-points UKF pour la phase de prédiction à partir d'un ensemble de sigma-points simplectiques unitaires

$$\hat{X}_{n,i}^+ = \hat{X}_{n-1}^+ + L_n C_n^T I_i, \quad 1 \leq i \leq p+1 \text{ avec } C_n^T C_n = U_n^{-1}.$$

On applique alors la dynamique sur ces $p+1$ sigma-points définissant

$$\hat{X}_{n+1,i}^- = A(X_{n,i}^+).$$

Ces points permettent le calcul de la moyenne et covariance empirique définies dans le filtre UKF. Utilisant les matrices D_α et T_α associées aux sigma-points simplectiques canoniques $[I_i]$,

$$\alpha_i = \frac{1}{p+1}, 1 \leq i \leq p+1, \quad D_\alpha = \text{diag}((\alpha_i)_{1\leq i\leq p+1}), \quad T_\alpha, = T$$

on a, pour la moyenne empirique

$$\begin{aligned}\hat{X}^-_{n+1} &= \frac{1}{p+1}\sum_{1\leq i\leq p+1} \hat{X}^-_{n+1,i} \\ &= [A(X^+_{n,i})].(\alpha_i),\end{aligned}$$

et pour la covariance empirique

$$\begin{aligned}P^-_{n+1} &= \sum_{1\leq i\leq p+1} \alpha_i(\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1})(\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1})^T \\ &= [\hat{X}^-_{n+1,i}]D_\alpha T_\alpha(T_\alpha^T D_\alpha T_\alpha)^{-1}T_\alpha^T D_\alpha[\hat{X}^-_{n+1,i}]^T.\end{aligned} \tag{4.103}$$

Cette dernière relation est obtenue par application de la proposition 4.4.1, car les sigma-points propagés $\hat{X}^-_{n+1}$ vérifient les mêmes contraintes que T_α, i.e. ils engendrent nécessairement un espace de rang inférieur ou égal à p.

Passons désormais à la phase de correction. Comme dans SEIK, on définit une matrice L_{n+1} suivant

$$L_{n+1} = [\hat{X}^-_{n+1,i}]D_\alpha T_\alpha,$$

où par rapport à Pham et al (1997) nous intégrons la matrice D_α directement dans L sachant qu'elle sera compensée plus tard dans notre expression de U. Le filtre UKF opère un resampling sur les particules en définissant

$$\hat{X}^*_{n+1,i} = \hat{X}^-_{n+1} + \sqrt{P^-_{n+1}}I_i$$

qui,d'après la relation (4.102), s'écrit matriciellement

$$\begin{aligned}[\hat{X}^*_{n+1,i}] &= [\hat{X}^-_{n+1}] + [\hat{X}^-_{n+1,i}]D_\alpha T_\alpha(T_\alpha^T D_\alpha T_\alpha)^{-1/2}[I_i] \\ &= [\hat{X}^-_{n+1}] + L_{n+1}(T_\alpha^T D_\alpha T_\alpha)^{-1}T_\alpha^T.\end{aligned}$$

Or avec le choix de D et T opéré par Pham et al (1997), on a

$$T(T^TT)^{-1}T^T = \begin{pmatrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix} - \frac{1}{p+1}\begin{pmatrix}1 & \cdots & 1\\ \vdots & \vdots & \vdots\\ 1 & \cdots & 1\end{pmatrix}$$

donc

$$\begin{aligned}L_{n+1}(T_\alpha^T D_\alpha T_\alpha)^{-1}T_\alpha^T &= [X_i]T(T^TT)^{-1}T^T \\ &= [\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1}],\end{aligned}$$

d'où avec ce choix

$$\hat{X}^*_{n+1,i} = \hat{X}^-_{n+1,i}.$$

Pour d'autres choix de T_α (toujours simplectiques) que T, les $\hat{X}^*_{n+1,i}$ pourraient être l'image des $\hat{X}^-_{n+1,i}$ par une transformation unitaire. En toute généralité, on calcule alors l'action de l'opérateur d'observation non-linéaire sur ces points

$$Z_{n+1,i} = H(\hat{X}^*_{n+1,i}),$$

et d'après l'algorithme UKF on obtient pour les moyennes et covariances empiriques de la phase de correction

$$\begin{aligned} P_\alpha^{\tilde{X}\tilde{Z}} &= \sum_{1\leq i\leq r} \alpha_i (\hat{X}^*_{n+1,i} - \hat{X}^-_{n+1})(Z_{n+1,i} - Z_{n+1,\alpha})^T \\ &= L_{n+1}(T_\alpha^T D_\alpha T_\alpha)^{-1}\{HL\}^T_{n+1}, \end{aligned} \tag{4.104}$$

où

$$\{HL\}_{n+1} = [H(\hat{X}^*_{n+1,i})]D_\alpha T_\alpha.$$

Par ailleurs

$$\begin{aligned} P_\alpha^{\tilde{Z}} &= W + \sum_{1\leq i\leq r} \alpha_i (Z_{n+1,i} - Z_{n+1,\alpha})(Z_{n+1,i} - Z_{n+1,\alpha})^T \\ &= W + \{HL\}_{n+1}(T_\alpha^T D_\alpha T_\alpha)^{-1}\{HL\}^T_{n+1}. \end{aligned} \tag{4.105}$$

Avec

$$\tilde{U}_n = T_\alpha^T D_\alpha T_\alpha, \tag{4.106}$$

et

$$U_{n+1} = \tilde{U}_n + \{HL\}^T_{n+1} W^{-1}\{HL\}_{n+1}, \tag{4.107}$$

on démontre aisément, en suivant la série d'expressions[6] (1.31),

$$\begin{aligned} \hat{X}^+_{n+1} &= \hat{X}^-_{n+1} + L_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1}(W + \{HL\}_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1})^{-1}(Z_{n+1} - Z_{n+1,\alpha}) \\ &= \hat{X}^-_{n+1} + L_{n+1}U_{n+1}^{-1}U_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1}(W + \{HL\}_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1})^{-1}(Z_{n+1} - Z_{n+1,\alpha}) \\ &= \hat{X}^-_{n+1} + L_{n+1}U_{n+1}^{-1}(\mathbb{1} + \{HL\}^T_{n+1}W^{-1}\{HL\}_{n+1}\tilde{U}_n^{-1})\{HL\}^T_{n+1} \\ &\qquad (W + \{HL\}_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1})^{-1}(Z_{n+1} - Z_{n+1,\alpha}) \\ &= \hat{X}^-_{n+1} + L_{n+1}U_{n+1}^{-1}\{HL\}^T_{n+1}W^{-1}(W + \{HL\}_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1}) \\ &\qquad (W + \{HL\}_{n+1}\tilde{U}_n^{-1}\{HL\}^T_{n+1})^{-1}(Z_{n+1} - Z_{n+1,\alpha}) \\ &= \hat{X}^-_{n+1} + L_{n+1}U_{n+1}^{-1}\{HL\}^T_{n+1}W^{-1}(Z_{n+1} - Z_{n+1,\alpha}). \end{aligned}$$

Le lemme d'inversion matricielle fournit finalement la covariance de correction sous la forme

$$P^+_{n+1} = L_{n+1}U_{n+1}L^T_{n+1}.$$

Cette dernière relation assurant l'hypothèse de récurrence au rang suivant, le filtre SEIK est donc exactement une version réduite du filtre UKF pour un choix de sigma-points simplectiques.

□ CAS GÉNÉRAL – Nous proposons maintenant à partir de cette équivalence méthodologique une extension du résultat de Pham et al (1997) pour d'autres choix de sigma-points.

[6] Voir aussi notre démonstration du lemme d'inversion matricielle en introduction du chapitre 1

Autrement dit, nous revoyons la présentation ci-dessus pour les matrices T_α associées, par exemple, aux sigma-points canoniques ou étoilés. Or, à la différence des sigma-points simplectiques, ces sigma-points vérifient des contraintes supplémentaires dans leur définition pour générer un espace de rang p à partir de $m > p+1$ points. Ainsi, en reprenant la démonstration ci-dessus, on constate que la relation (4.103) ne tient plus pour un choix de sigma-points quelconque car les points propagés ne vérifient plus les contraintes imposées par T permettant l'application de la proposition 4.4.1. Notamment on imagine aisément que des points étoilés au départ ne le restent pas après propagation et définissent alors un espace de rang supérieur à p.

Cependant si nous cherchons toujours une matrice de rang p, il est possible de projeter cette covariance empirique sur la meilleure matrice de rang p pour une certaine norme. Si cette dernière est la norme de Frobenius, cette opération revient simplement à extraire la SVD à p composantes de la matrice P_α. Examinons comment réaliser cela sur la matrice de covariance a priori. On cherche donc la SVD de

$$P_\alpha = [\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1}] D_\alpha [\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1}]^T \in \mathbb{M}_N(\mathbb{R}),$$

de rang maximum r. Posons $V = [\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1}] D_\alpha^{\frac{1}{2}}$, on démontre qu'il suffit pour cela de diagonaliser la grammienne associée

$$G_\alpha = V^T V \in \mathbb{M}_r(\mathbb{R}).$$

En effet soit la diagonalisation de G_α

$$G_\alpha = U_r \Sigma_r U_r^T, \text{ avec } U_r^T U_r = \mathbb{1},$$

et les valeurs propres dans Σ_r ordonnées de manière décroissante. Alors, les vecteurs propres normalisés de $P_\alpha = VV^T$ sont les colonnes de $VU_r\Sigma_r^{-\frac{1}{2}}$ car

$$\begin{cases} VV^T(VU_r) = VU_r\Sigma_r \\ \Sigma_r^{-\frac{1}{2}} U_r^T V^T V U_r \Sigma_r^{-\frac{1}{2}} = \mathbb{1} \end{cases}$$

d'où finalement

$$P_\alpha = (VU_r\Sigma_r^{-\frac{1}{2}})\Sigma_r(\Sigma_r^{-\frac{1}{2}} U_r^T V^T) = VU_rU_r^TV^T.$$

Cette structure nous permet alors de définir la SVD à p en définissant la restriction U_p de U_r ne conservant que les p premières colonnes,

$$P_{\text{SVD}} = VU_pU_p^TV^T.$$

Cette matrice de rang p écrite sous forme factorisée, permet alors simplement de régénérer r particules qui de nouveau vérifieront les contraintes imposées par le filtre UKF. Pour cela en effet, il suffit d'utiliser les r sigma-points unitaires de notre choix et de définir

$$\begin{aligned} [\hat{X}^*_{n+1,i}] &= [\hat{X}^-_{n+1}] + VU_p[I_i] \\ &= [\hat{X}^-_{n+1}] + [\hat{X}^-_{n+1,i} - \hat{X}^-_{n+1}] D_\alpha^{1/2} U_p[I_i], \end{aligned}$$

correspondant, par définition, à une matrice de covariance a priori

$$P^-_{n+1} \stackrel{def}{=} [\hat{X}^*_{n+1,i}] D_\alpha T_\alpha (T_\alpha^T D_\alpha T_\alpha)^{-1} T_\alpha^T D_\alpha [\hat{X}^*_{n+1,i}]^T.$$

Les nouveaux points $\hat{X}^*_{n+1,i}$ doivent de nouveau être propagés par les opérateurs d'observation. En reprenant les notations du filtre UKF Section 1.4.4, nous pouvons écrire le filtre sous la forme

$$K = P_\alpha^{\tilde{X}\tilde{Z}}(P_\alpha^{\tilde{Z}})^{-1}.$$

Cet opérateur est calculable dans notre cas en utilisant le lemme d'inversion matricielle. Pour mener à bien les calculs, on utilise des notations condensées

$$[\tilde{X}] = [X^*_{n+1,i} - \hat{X}^-_{n+1}], \quad [\tilde{Z}] = [Z_{n+1,i} - Z_{n+1,\alpha}].$$

On a alors

$$\begin{aligned} K &= [\tilde{X}]D_\alpha[\tilde{Z}]^T(W + [\tilde{Z}]D_\alpha[\tilde{Z}]^T)^{-1} \\ &= [\tilde{X}]D_\alpha[\tilde{Z}]^T\Big(W^{-1} - W^{-1}[\tilde{Z}](D_\alpha^{-1} + [\tilde{Z}]^TW^{-1}[\tilde{Z}])^{-1}[\tilde{Z}]^TW^{-1}\Big) \\ &= [\tilde{X}]D_\alpha\Big(\mathbb{1}_r - [\tilde{Z}]^TW^{-1}[\tilde{Z}](D_\alpha^{-1} + [\tilde{Z}]^TW^{-1}[\tilde{Z}])^{-1}\Big)[\tilde{Z}]^TW^{-1}. \end{aligned}$$

Posons alors

$$D_{\mathfrak{m}} = [\tilde{Z}]^TW^{-1}[\tilde{Z}] \in M_r(\mathbb{R})$$

qui, à la différence de $P_\alpha^{\tilde{Z}}$, est une matrice calculable car sa dimension est égale au nombre de particules choisi. On peut alors écrire

$$K = [\tilde{X}]D_\alpha(\mathbb{1} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1})[\tilde{Z}]^TW^{-1}. \tag{4.108}$$

En reprenant la démonstration de la proposition 4.4.1, on peut aussi écrire ce gain sous la forme

$$K = L_{n+1}(T_\alpha^TD_\alpha T_\alpha)^{-1}T_\alpha^T(\mathbb{1} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1})[\tilde{Z}]^TW^{-1}, \tag{4.109}$$

où cette fois

$$L_{n+1} = [\hat{X}^*_{n+1,i}]D_\alpha T_\alpha,$$

à partir des r particules $\hat{X}^*_{n+1,i}$ correctement échantillonnées. En revanche, pour le terme de droite de (4.108), on ne peut pas utiliser la matrice $\{HL\}$ combinée à T_α car les sigma-points propagés par l'opérateur d'observation ne vérifient pas les contraintes d'échantillonnages imposées par T_α. Une fois le gain défini, il reste à calculer la covariance a posteriori qui va nous servir ensuite au rééchantillonnage. Cette covariance doit donc vérifier la propriété de rang réduit égal à p. On a

$$\begin{aligned} P^+_{n+1} &= P^-_{n+1} - P_\alpha^{\tilde{X}\tilde{Z}}(P_\alpha^{\tilde{Z}})^{-1}(P_\alpha^{\tilde{X}\tilde{Z}})^T \\ &= P^-_{n+1} - [\tilde{X}]D_\alpha(D_{\mathfrak{m}} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1}D_{\mathfrak{m}})D_\alpha[\tilde{X}]^T \\ &= P^-_{n+1} - L_{n+1}(T_\alpha^TD_\alpha T_\alpha)^{-1}T_\alpha^TD_\alpha(D_{\mathfrak{m}} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1}D_{\mathfrak{m}})D_\alpha T_\alpha(T_\alpha^TD_\alpha T_\alpha)^{-1}L^T_{n+1}. \end{aligned}$$

L'avantage de la dernière formulation est qu'elle permet de nouveau d'écrire

$$P^+_{n+1} = L_{n+1}U^{-1}_{n+1}L^T_{n+1}$$

où, toujours avec le lemme d'inversion matricielle,

$$\begin{aligned} U^{-1}_{n+1} &= (T_\alpha^TD_\alpha T_\alpha)^{-1} - (T_\alpha^TD_\alpha T_\alpha)^{-1}T_\alpha^TD_\alpha(D_{\mathfrak{m}} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1}D_{\mathfrak{m}})D_\alpha T_\alpha(T_\alpha^TD_\alpha T_\alpha)^{-1} \\ &= (T_\alpha^TD_\alpha T_\alpha + T_\alpha^TD_\alpha(D_{\mathfrak{m}} - D_{\mathfrak{m}}(D_\alpha^{-1} + D_{\mathfrak{m}})^{-1}D_{\mathfrak{m}})^{-1})D_\alpha T_\alpha)^{-1} \\ &= \Big(T_\alpha^TD_\alpha T_\alpha + T_\alpha^TD_\alpha\big((D_\alpha + D_\alpha D_{\mathfrak{m}}^{-1}D_\alpha)^{-1} - (T_\alpha^TD_\alpha T_\alpha)^{-1}\big)D_\alpha T_\alpha\Big)^{-1}, \end{aligned}$$

soit

$$U_{n+1} = T_\alpha^T D_\alpha T_\alpha + T_\alpha^T D_\alpha \big((D_\alpha + D_\alpha D_{\mathfrak{m}}^{-1} D_\alpha)^{-1} - (T_\alpha^T D_\alpha T_\alpha)^{-1}\big) D_\alpha T_\alpha \tag{4.110}$$

Une telle formulation $P_{n+1}^+ = L_{n+1} U_{n+1}^{-1} L_{n+1}^T$ signifie que la covariance a posteriori est déjà de rang p. La cause en est l'utilisation de points $\hat{X}_{n+1,i}^*$ échantillonnés correctement qui, dans la covariance, apparaissent directement, sans avoir été propagés par un opérateur non-linéaire. Cette formulation assure donc qu'on peut continuer la récurrence pour la prédiction suivante.

Cette formulation **généralise** donc celle proposée par Pham et al (1997) pour n'importe quel choix de sigma-points. Cependant sa mise ne oeuvre impose un resampling supplémentaire et un calcul de covariance un peu plus complexe. Cette complexité accrue pourra potentiellement trouver une justification en pratique si :

- la qualité de l'interpolation de certains opérateurs dynamiques (ou d'observation) non-linéaire bénéficie d'un type spécifique de conformation des particules ;
- certaines lois de probabilité sont mieux représentées par un échantillonnage plus riche (voir Julier et al (2000)).

4.4.2 Application à l'estimation paramétrique

Nous utilisons ce filtre UKF réduit pour étendre nos précédents résultats d'estimation paramétrique. Ainsi, le principe est inchangé par rapport au début du chapitre : utiliser le filtrage d'état « physique » pour limiter la covariance à l'espace paramétrique puis utiliser un filtre réduit sur le premier estimateur. On rappelle que les algorithmes que nous venons de voir sont donc appliquées à la variable $X^e = \left(\begin{smallmatrix} X \\ \theta \end{smallmatrix}\right)$.

a. Formulation

Nous nous limitons à la formulation utilisant des points simplectiques et formulons ainsi la version non-linéaire de notre estimateur conjoint. La particularité de cette formulation vient de la dynamique nulle vérifiée par θ.

- **Echantillonnage** :

$$\begin{cases} C_n = \sqrt{U_n^{-1}} \\ \hat{X}_{n,i}^+ = \hat{X}_n^+ + L_n^X C_n^T I_i, \quad 1 \le i \le p+1 \\ \hat{\theta}_{n,i}^+ = \hat{\theta}_n^+ + L_n^\theta C_n^T I_i, \quad 1 \le i \le p+1 \end{cases} \tag{4.111a}$$

- **Prédiction** :

$$\begin{cases} \hat{X}_{n+1,i}^- = A(\hat{X}_{n,i}^+, \hat{\theta}_{n,i}^+) + K_X(Z_n - H(\hat{X}_{n,i}^+)) \\ \hat{X}_{n+1}^- = \hat{X}_{n+1,\alpha}^- \\ \hat{\theta}_{n+1}^- = \hat{\theta}_{n+1}^+ \end{cases} \tag{4.111b}$$

- **Correction** :

$$\begin{cases} L^X_{n+1} = [\hat{X}^-_{n+1,i}]D_\alpha T_\alpha \in \mathbb{M}_{N,p} \\ L^\theta_{n+1} = [\theta^-_{n+1,i}]D_\alpha T_\alpha \\ Z_{n+1,i} = H(X^-_{n+1,i}) \\ \{HL^X\}_{n+1} = [Z_{n+1,i}]D_\alpha T_\alpha \\ U_{n+1} = (T_\alpha)^T D_\alpha T_\alpha + \{HL^X\}^T_{n+1} W^{-1}\{HL^X\}_{n+1} \in \mathbb{M}_p \\ \hat{X}^+_{n+1} = \hat{X}^-_{n+1} + L^X_{n+1}U^{-1}_{n+1}\{HL^X\}^T_{n+1}W^{-1}(Z_{n+1} - \hat{Z}_{n+1,\alpha}) \\ \hat{\theta}^+_{n+1} = \hat{\theta}^-_{n+1} + L^\theta_{n+1}U^{-1}_{n+1}\{HL^X\}^T_{n+1}W^{-1}(Z_{n+1} - \hat{Z}_{n+1,\alpha}) \end{cases} \tag{4.111c}$$

Par rapport au filtre paramétrique SEEK, nous voyons apparaître en plus des termes non-linéaires un nouvel opérateur L^θ. Afin de l'interpréter, l'idéal est de formuler cet estimateur dans le cadre linéaire et de le comparer au filtre SEEK, ce que nous faisons au paragraphe suivant.

b. Equivalence pour les dynamiques linéaires

Reprenons une dynamique linéaire avec une dépendance paramétrique linéaire. De plus, on suppose les opérateurs autonomes pour simplifier la formulation sans perte de généralité :

$$\bar{X}_{n+1} = A_K\bar{X}_n + B\bar{\theta}_n + R.$$

De même l'opérateur d'observation est supposé linéaire, et indépendant du temps. L'estimateur SEIK ci-dessus se simplifie considérablement puisque l'action des opérateur linéaires ne modifie pas la répartition autour de la moyenne de l'échantillon. Pour les prédictions on obtient donc

$$\begin{cases} \hat{X}^-_{n+1} = A_K\hat{X}^+_n + B\hat{\theta}^+_n + R \\ \hat{\theta}^-_{n+1} = \hat{\theta}^+_{n+1} \end{cases}$$

Puis pour les corrections

$$\begin{cases} \hat{X}^+_{n+1} = \hat{X}^-_{n+1} + L^X_{n+1}U^{-1}_{n+1}(L^X_{n+1})^T H^T W^{-1}(Z_{n+1} - H\hat{X}^-_{n+1}) \\ \hat{\theta}^+_{n+1} = \hat{\theta}^-_{n+1} + L^\theta_{n+1}U^{-1}_{n+1}(L^X_{n+1})^T H^T W^{-1}(Z_{n+1} - \hat{H}X^-_{n+1}) \end{cases}$$

où (L^X_n) et (L^θ_n) et (U_n) n'ont a priori pas la même dynamique pour le SEIK que pour le SEEK. On rappelle d'ailleurs que dans SEEK, il n'y avait pas de (L^θ_n) ou plutôt que celui-ci était constant égal à 1. Dans le cas SEIK, on a désormais

$$\begin{aligned} L^X_{n+1} &= [X^-_{n+1,i}]D_\alpha T_\alpha \\ &= [A_K\hat{X}^+_{n,i} + B\hat{\theta}^+_{n,i}]D_\alpha T_\alpha \\ &= A_K\hat{X}^+_n(1)^T D_\alpha T_\alpha + AL^X_n C^T_n[I_i]D_\alpha T_\alpha + B\hat{\theta}^+_n(1)^T D_\alpha T_\alpha + BL^\theta_n C^T_n[I_i]D_\alpha T_\alpha \\ &= (A_K L^X_n + BL^\theta_n)C^T_n[I_i]D_\alpha T_\alpha, \end{aligned}$$

donc

$$L^X_{n+1} = (A_K L^X_n + BL^\theta_n)C^T_n[I_i]D_\alpha T_\alpha, \tag{4.112}$$

et de même

$$L^\theta_{n+1} = L^\theta_n C^T_n[I_i]D_\alpha T_\alpha = L^\theta_n C^T_n[I_i]D_\alpha T_\alpha. \tag{4.113}$$

Enfin, on a

$$U_{n+1} = (T_\alpha)^T D_\alpha T_\alpha + ({L^X}_{n+1})^T H^T W^{-1} H L^X_{n+1}.$$

Ces structures sont similaires à celle du filtre SEEK si on se rappelle que celui-ci à été défini via une décomposition de type

$$P_n^+ = L_n^T U_n^{-1} L_n,$$

qui n'est pas unique. Notamment rien n'empêche d'opérer une décomposition du type

$$P_n^+ = L_n^T \sqrt{U_n}^{-T} D_n^T D_n \sqrt{U_n}^{-1} L_n,$$

pour toute racine carrée de U_n et toute matrice unitaire $D_n \in \mathcal{O}_n(\mathbb{R})$.

Afin de justifier cette idée, il est d'ailleurs intéressant de regarder comment se comporte le système de l'erreur pour le filtre SEIK. L'idée est alors d'effectuer le même type de changement de variable (4.36) que pour l'algorithme SEEK

$$\eta_n = \tilde{X}_n^- - L_n^X ({L^\theta}_n)^{-1} \tilde{\theta}_n, \tag{4.114}$$

à la différence donc que nous y insérons l'inverse de L^θ. Or les expressions (4.112) et (4.113) nous permettent immédiatement de simplifier le terme en $C_n^T[I_i] D_\alpha T_\alpha$

$$\eta_{n+1} = \tilde{X}_{n+1}^- - (A_K L_n^X ({L^\theta}_n)^{-1} + B)\tilde{\theta}_{n+1},$$

d'où on tire, tous calculs faits, la même dynamique qu'en (4.60)

$$\eta_{n+1} = A_K \eta_n,$$

et pour $({L^\theta}_n U^{-1} {L^\theta}_n)\tilde{\theta}_n$, une expression similaire à (4.61) multipliée par U_n.

c. Résultats numériques

Nous avons testé l'algorithme UKF réduit avec sigma-points simplectiques (donc SEIK) avec la version non-linéaire décrite Section 3.4.7 de notre problème-test. Nous cherchons donc toujours les paramètres de précontrainte qui cette fois cependant sont à la fois au second membre de la dynamique mais aussi dans la raideur. En effet la précontrainte s'écrit désormais

$$\mathcal{P}^{\text{ext}} = \sum_{1 \le i \le 17} \int_{\Omega_i^{AHA}} \theta_i \int_\Omega \sigma_0 w(\underline{x}, t)\ \mathrm{tr}(\ \mathrm{d}_{\underline{y}} \underline{\underline{e}} \cdot \underline{v})\, d\Omega = \delta Y^T \cdot B(Y)\theta. \tag{4.115}$$

De nouveau, tous les paramètres sains sont laissés à 1 alors que la région AHA 14 est infarcie avec $\theta_{14} = 0.5$. Enfin les mesures sont des vitesses volumiques.

Nous profitons de cette section de résultats pour insister sur les avantages de l'algorithme SEIK par rapport à SEEK en termes d'implémentation. En effet celui-ci ne nécessite pas de calcul d'opérateurs tangents, notamment par rapport aux paramètres. En revanche, pour chaque particule, on doit résoudre un problème non-linéaire donc une boucle de Newton pendant la prédiction alors que, pour SEEK, le tangent est par définition linéaire. En conséquence, la complexité en temps est plus grande mais l'algorithme est bien plus simple à paralléliser puisque à chaque étape de prédiction chaque boucle de Newton par particule peut être résolue indépendamment.

□ TEST ALGORITHMIQUE – Nous commençons par une évaluation de l'algorithme d'estimation à partir de données synthétiques parfaites. On rappelle que l'incertitude se limite donc exclusivement aux paramètres. les performances sont présentées Figure 4.15 et montrent la performance théorique de l'estimateur.

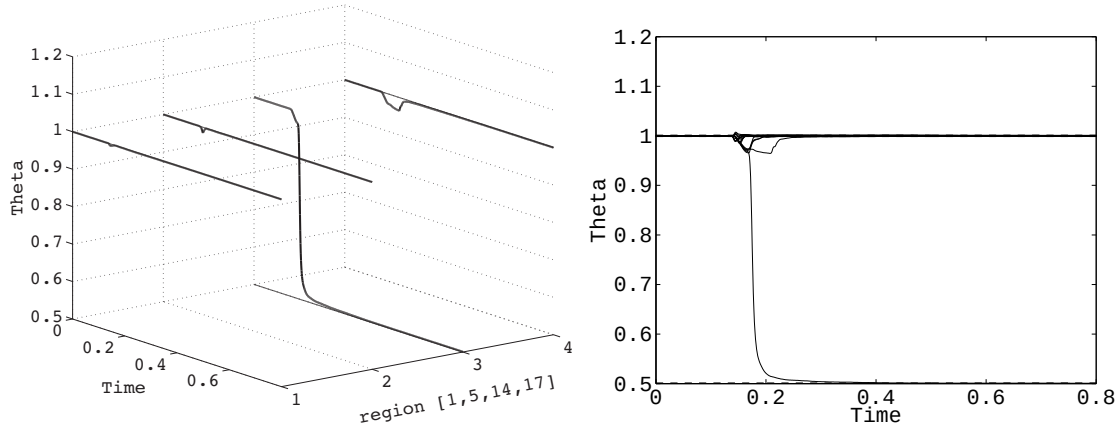


FIGURE 4.15: Test algorithmique de l'identification non-linéaire « en conditions idéales »

□ CONFIGURATION BRUITÉE – Nous pouvons alors tester notre algorithme dans des conditions de bruit réalistes. Le bruit de mesure est supposé blanc à 10% de la vitesse maximale

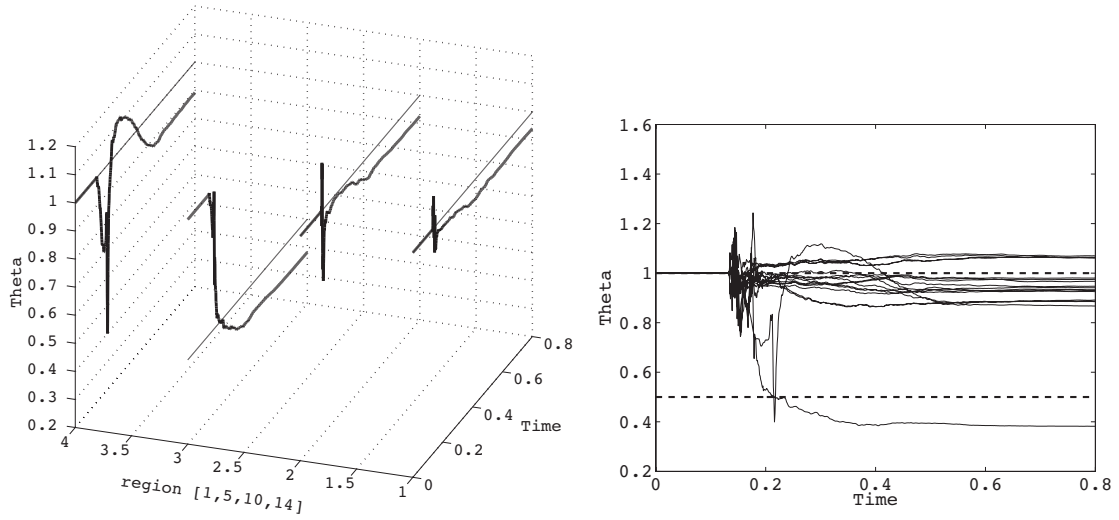


FIGURE 4.16: Estimation conjointe pour le problème non-linéaire avec une erreur de condition initiale générée par une surpression de 1000 Pa

comme précédemment et l'erreur de condition initiale correspond à la surpression de 1000 Pa. L'estimation est déclenchée à $t_s = 0.1s$.

Les résultats apparaissent Figure 4.16 et sont suffisants pour la détection de la zone infarcie. Pour autant, ils sont moins bons qu'en linéaire et l'erreur d'estimation finale de l'ordre de 10% correspond plus aux résultats observés en bilinéaire. Là encore c'est le couplage entre dépendance paramétrique et dépendance de l'état qui défavorise l'estimation.

□ ESTIMATION LINÉAIRE SUR MODÈLE NON-LINÉAIRE – Malgré la relative faible complexité de l'estimateur non-linéaire, on peut imaginer que certains aimeraient se limiter à un estimateur linéaire pour toutes les simplifications numériques qu'il apporte. Dans ce cas, on introduit une erreur de modèle qu'on aimerait « négligeable » par rapport à l'objectif de détection de zone infarcie. Afin de tester ce type de procédé, nous nous sommes livrés, avec P.E. Ellefsen et S. Deumont[7], à l'exercice numérique suivant :

1. Simuler avec le modèle non-linéaire utilisé ci-dessus pour différentes valeurs de σ_0.

2. Estimer comme à la Section 4.2.6 en supposant le modèle linéarisé autour de 0 (petits déplacements).

On rappelle que, en non-linéaire, respecter les ordres de grandeurs des volumes éjectés conduit à choisir la contractilité $\sigma_0 = 4.5\ 10^4$Pa, alors qu'en linéaire il suffisait de prendre 3.10^4Pa. L'écart entre la déformation générée par les deux types de modèles pour le même σ_0 est donc conséquent pour le système cardiaque, et diminue pour une contrainte active moindre dans le tissu (jusqu'à la limite des petits déplacements). Cet écart est mesuré sur la dilatation globale de l'organe mais aussi sur ses déformations locales. En particulier, nous avions mentionné le fait que l'orientation des fibres engendre un phénomène de torsion que les médecins appellent le twist (entre 10 et 15 degrés). Or, on sait (voir illustration Figure 4.17) que les mouvement de rotation sont pris en compte de façon très différente entre un modèle linéaire et un modèle non-linéaire, d'où d'ailleurs des dilatations plus importantes pour le modèle linéaire à même σ_0.

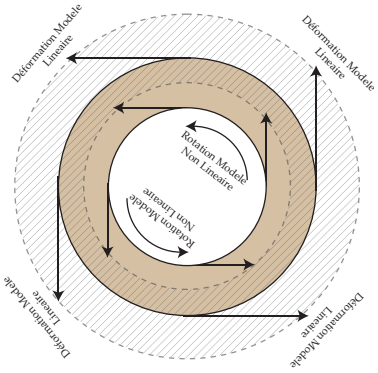


FIGURE 4.17: Mouvement de rotation d'ensemble pour le modèle non-linéaire et l'équivalent produit par le modèle linéaire.

Les résultats obtenus par P.E. Ellefsen et S. Deumont sont présentés Figure 4.18 pour différentes valeurs de σ_0 correspondant à différentes valeurs de twist mesuré. Les erreurs se limitent à l'erreur de modèle et l'erreur paramétrique. Déjà pour $\sigma_0 = 3.10^4$, le résultat d'estimation linéaire ne permet plus la détection de la zone infarcie. Le seuil limite de validité est dans ce cas-test $\sigma_0 = 2.10^4$, pour un twist de 5 degrés (au lieu des 12 degrés physiologiques) et un volume éjecté de de 30 mL (au lieu de 70 mL). Cet exemple tend à prouver que, pour les dé-

[7]étudiants en troisième année de l'École Polytechnique, lors de leur module « Travaux de Modélisation et Simulations »

formations attendues dans le cœur, il n'est donc pas envisageable de réaliser une identification paramétrique à base de modèles linéaires.

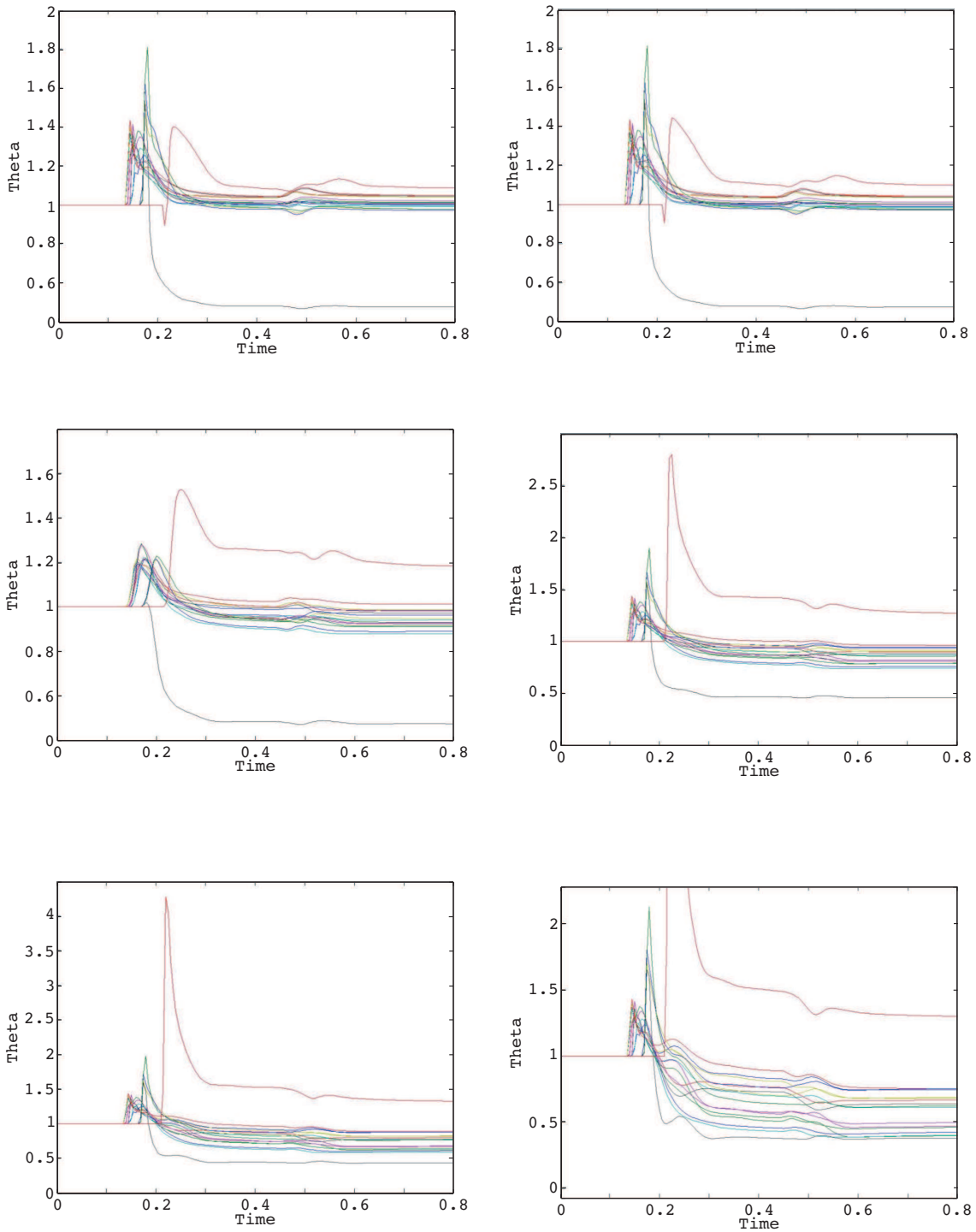


FIGURE 4.18: Résultats d'identification correspondant à des simulations directes pour différentes valeurs de σ_0. Le twist mesuré est respectivement de 0.05, 0.5, 4.7, 10, 19.5 et 32 degrés.

4.5 Perspectives autour du système cardiaque

4.5.1 Estimation pour *des paramètres dynamiques*

L'expression (4.114) pour l'algorithme SEIK, nous laisse imaginer que le $L^\theta = \mathbb{1}$ de l'estimateur paramétrique SEEK n'était pas limitant. Autrement dit rien n'empêchait à L^θ d'avoir une dynamique, ce qui dans le cadre SEEK signifie que θ lui même pourrait avoir une dynamique propre. Nous proposons comme dernier résultat (ici uniquement méthodologique) de généraliser ce que nous venons de voir dans un tel cadre.

En automatique, le type de changement de variable que nous avons utilisé est formellement, en particulier pour (4.114), classique. En fait, si son utilisation ici peut être originale pour l'estimation adaptative sur des systèmes de grandes tailles, l'idée de réduire le rang de l'espace d'état existe depuis les observateurs de Luenberger. En effet, les observateurs réduits de Luenberger (voir Bonnans et Rouchon (2005)) consistent aux aussi à limiter le filtrage sur une partie du système en assurant la convergence de l'erreur par ce type de changement de variable. Or ce genre d'observateur fonctionne sur des systèmes dont la structure est très générale, et pas seulement pour une partie de l'état de dynamique nulle. Le terme de paramètres est alors quelque peu impropre puisque si ces grandeurs ont désormais une dynamique, elles doivent être appelées variables d'état. Comme cette situation apparaît en complément de ce que nous venons de présenter et annonce les perspectives de la section suivante en termes d'applications, nous choisissons de l'appeler estimation paramétrique généralisée.

a. Formulation en linéaire

Sur la dynamique linéaire considérée jusqu'à présent nous choisissons tout d'abord d'ajouter une dynamique linéaire autonome à θ,

$$\begin{cases} \dot{X} = AX + B\theta + R \\ \dot{\theta} = C\theta + S \\ X(0) = X_0 + \zeta^X \\ \theta(0) = \theta_0 + \zeta^\theta \end{cases} \tag{4.116}$$

avec des observations toujours sous la forme $Z = HX + \chi$. On propose l'observateur suivant

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B\theta + R + K_X(Z - H\hat{X}) + L^X U^{-1} {L^X}^T H^T W^{-1}(Z - H\hat{X}) \\ \dot{\hat{\theta}} = C\hat{\theta} + S + L^\theta U^{-1} {L^X}^T H^T W^{-1}(Z - H\hat{X}) \\ \hat{X}(0) = X_0 \\ \hat{\theta}(0) = \theta_0 \end{cases} \tag{4.117}$$

où

$$\begin{cases} \dot{L}^X = (A - K_X H)L^X + BL^\theta \\ \dot{L}^\theta = CL^\theta \\ \dot{U} = {L^X}^T H^T W^{-1} H L^X \\ L^X(0) = 0 \\ L^\theta(0) = \mathbb{1} \\ U(0) = \mathbb{E}(\zeta^\theta {\zeta^\theta}^T)^{-1} \end{cases} \tag{4.118}$$

Pour cet estimateur, le système de l'erreur s'analyse par un changement de variable correspondant aux observateurs réduits de Luenberger

$$\eta = \tilde{X} - L^X (L^\theta)^{-1} \tilde{\theta}.$$

Le système différentiel homogène vérifié par ce système s'écrit en négligeant les différents bruits de mesures et de discrétisation

$$\begin{aligned} \dot{\eta} &= \dot{\tilde{X}} - \dot{L}^X (L^\theta)^{-1}\tilde{\theta} + L^X (L^\theta)^{-1} \dot{L}^\theta (L^\theta)^{-1}\tilde{\theta} - L^X (L^\theta)^{-1}\dot{\tilde{\theta}} \\ &= (A - K_X H)\tilde{X} + B\tilde{\theta} - L^X U^{-1} {L^X}^T H^T W^{-1} H \tilde{X} - ((A - K_X H)L^X + BL^\theta)(L^\theta)^{-1}\tilde{\theta} \\ &\quad + L^X (L^\theta)^{-1} C\tilde{\theta} - L^X (L^\theta)^{-1}(C\tilde{\theta} - L^\theta U^{-1} L^T H^T W^{-1} H \tilde{X}) \\ &= (A - K_X H)(\tilde{X} - L^X (L^\theta)^{-1}\tilde{\theta}) \\ &= (A - K_X H)\eta. \end{aligned} \tag{4.119}$$

Puis pour θ, on calcule la dynamique de

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}}((L^\theta)^{-1}\tilde{\theta}) &= -(L^\theta)^{-1}\dot{L}^\theta (L^\theta)^{-1}\tilde{\theta} + (L^\theta)^{-1} C\tilde{\theta} - (L^\theta)^{-1} L^\theta U^{-1} L^X H^T W^{-1} H \tilde{X} \\ &= U^{-1} L^X H^T W^{-1} H L^X ((L^\theta)^{-1}\tilde{\theta}) - U^{-1} L^X H^T W^{-1} H L^X (L^\theta)^{-1}\eta, \end{aligned} \tag{4.120}$$

soit, pour $(L^\theta)^{-1}\tilde{\theta}$, une dynamique connue constituée d'une partie homogène dépendant des conditions d'observabilité et un second membre forcé par la dynamique exponentiellement stable de η.

Le schéma en temps associé à ce filtre est alors directement obtenu de la version discrète du filtre SEEK. Cet algorithme de type prédiction-correction vérifie alors la version discrète des équations (4.119),(4.120) pour le système de l'erreur $(\tilde{X}_n - L^X_n (L^\theta_n)^{-1}\tilde{\theta}_n,\ (L^\theta_n)^{-1}\tilde{\theta}_n)$. Cette propriété assure alors la stabilité du schéma.

b. Formulation en non-linéaire

Cette fois la dynamique est

$$\begin{cases} \dot{X} = A(X, \theta) \\ \dot{\theta} = f(\theta) \\ X(0) = X_0 + \zeta^X \\ \theta(0) = \theta_0 + \zeta^\theta \end{cases} \tag{4.121}$$

pour des opérateurs d'observations soit linéaires $Z = HX + \chi$ soit de la forme $D(Z, X) = \chi$. Il suffit alors de remplacer la formulation SEEK par la formulation SEIK, s'étendant à partir

des expressions précédentes pour tous les types d'opérateur d'observation. Les équations sont directement étendues à partir de (4.111) en ajoutant un schéma stable sur la dynamique de θ. On rappelle que ce dernier est de toute façon nécessaire à la résolution du problème direct (i.e. sans prise en compte des données). Ce système vérifie alors la dynamique d'erreur linéarisée (4.119),(4.120) en version discrète que l'on pourrait d'ailleurs aussi écrire en continu sous la forme proposée par Sarkka (2007). Celle-ci, bien qu'un peu artificielle, permet en continu de réaliser le même type de linéarisation que Section 4.4.2.b..

4.5.2 Implications pour l'estimation cardiaque

Les paragraphes qui suivent sont donc des remarques et des perspectives à partir des résultats énoncés. Nous avons essayer d'imaginer comment les résultats déjà obtenus laissent envisager des outils pour la configuration générale de l'estimation cardiaque. Cette dernière nécessite en effet de prendre en compte pleinement les deux autres catégories de variables d'état abordées au chapitre 2 que sont les variables de pression et les variables électriques.

a. Estimation de conditions aux limites

L'estimation de conditions aux limites est centrale dans de nombreuses applications en mécanique et particulièrement en mécanique cardiaque. Nous avons vu en effet, Section 2.3.5, combien il était délicat de définir des conditions aux limites raisonnables pour le cœur isolé. Nous avons alors imaginé des conditions aux limites élastiques en (2.42) conduisant à une dynamique de la forme

$$M\ddot{Y} + (C + C_{\mathrm{BC}}(\theta))\dot{Y} + (K + K_{\mathrm{BC}}(\theta))Y = F,$$

avec θ paramétrant (linéairement) les coefficients de raideur et viscosité des conditions aux limites. Le paramétrage peut alors être estimé comme pour n'importe quel opérateur de raideur ou de viscosité. Naturellement, ceci présuppose que le choix de ce modèle de conditions aux limites représente correctement la réalité, comme pour tout problème d'assimilation de données.

b. Variables de pression

Quand on peut disposer de mesures de pression, nous avons vu qu'un observateur de Luenberger était très simple à réaliser puisqu'il suffisait d'introduire un terme en $\gamma(P - P_{\mathrm{m}})$ dans la dynamique de la pression artérielle. Cette idée présentée comme perspective Section 3.5.8 pourrait de plus être généralisée à des variables de pression solutions d'un modèle d'EDP. Dans ce cas on s'engagerait sur l'observation et le contrôle de systèmes d'EDP couplés type fluide-structure. Resterait de tout façon la difficulté d'obtenir de telle mesures en pratique (sauf à utiliser des pressions périphériques associées à des modèles) et à pouvoir gérer les changement de phase. Notamment la phase isovolumétrique où la pression est un multiplicateur de Lagrange resterait un enjeu.

Ce qu'on peut retenir de ce chapitre est que lorsqu'on n'a pas d'idée sur un filtre « physique », il est toujours possible d'envisager de faire un filtre de Kalman en deuxième étage

d'un premier estimateur de Luenberger. La condition *sine qua non* reste la dimension du système sur lequel on a limité le rang de la covariance. Or, les variables de pression du modèle cardiaque présenté au chapitre 2 sont en faible nombre et paraissent donc bien adaptés à un observateur adaptatif. En fait ce raisonnement n'est pas complètement exact puisque la dynamique déplacements-pression n'a pas la structure (4.116). Même dans le cas où les valves sont ouvertes, en simplifiant le modèle de valve, on obtient le système (3.108) dont formellement on peut écrire le linéarisé sous la forme

$$\begin{cases} \dot{X} = AX + BP + R \\ \dot{P} = CP - B^T X + S \end{cases} \tag{4.122}$$

Rien n'empêche a proiri de définir le même type d'estimateur de Luenberger réduit pour ce type de problème. On obtient dans ce cas en linéaire

$$\begin{cases} \dot{\hat{X}} = A\hat{X} + B\hat{P} + R + K_X(Z - H\hat{X}) + L^X U^{-1} L^{X^T} H^T W^{-1}(Z - H\hat{X}) \\ \dot{\hat{P}} = C\hat{P} - B^T \hat{X} S + L^P U^{-1} L^{X^T} H^T W^{-1}(Z - H\hat{X}) \end{cases} \tag{4.123}$$

avec

$$\begin{cases} \dot{L}^X = (A - K_X H)L^X + BL^P \\ \dot{L}^P = CL^P - B^T L^X \end{cases} \tag{4.124}$$

Dans ce cas, le changement de variable $(\tilde{X},\, \tilde{P}) \to (\eta = \tilde{X} - L^X(L^P)^{-1}\tilde{P},\, \tilde{P})$ trigonalise toujours l'erreur. Cependant la dynamique de η devient après simplifications

$$\dot{\eta} = (A - K_X H - L^X(L^P)^{-1}B^T)\eta,$$

et ne suit donc plus la dynamique de l'estimateur d'état. Donc on ne bénéficie plus des résultats précédents et la dépendance de la dynamique de η en L^X et L^P rend l'étude plus difficile. Cet estimateur doit donc être analysé en détail et il n'est d'ailleurs pas clair que le changement de variable soit le bon, ni même l'estimateur. Nous avions ainsi imaginé trigonaliser le système (4.122) sur la partie pression afin de se ramener au problème précédent. Or même si cette trigonalisation est toujours possible algébriquement, les variables résultantes n'ont pas forcément d'interprétation physique. Notamment cette manipulation pourrait ne pas être robuste au raffinement de la discrétisation. Ce problème reste donc à traiter. Notons enfin que l'estimateur (4.123) est défini en linéaire mais que la version SEIK permet de proposer en temps discret une généralisation en non-linéaire.

c. Extension autour de l'activation électrique

Si nous cherchons un exemple de système où les variables d'état suivent une dynamique « triangulaire », il faut que la variable θ de (4.116) soit une entrée du système mécanique. Or, la simulation cardiaque fournit directement ce type d'exemple avec l'entrée électrique u. Cependant, la variable électrique est une variable distribuée au même titre que les déplacements et les vitesses, donc sa dimension impose plutôt d'imaginer un observateur de Luenberger sur ce champ. Un tel observateur doit alors être conçu pour un système de réaction-diffusion et à partir des mesures disponibles.

Si nous nous limitons aux données électriques, nous avons vu que les mesures sont soit des ECGs (ou des mesures à la surface de thorax) soit des potentiels à la surface du cœur pour des méthodes invasives. Dans le deuxième cas, la position du problème de l'estimation est alors très proche de ce que nous avons décrit en mécanique si ce n'est que le système est parabolique avec un terme de réaction non-linéaire (fondamental pour la propagation du front). Les choses se compliquent si on doit passer par une méthode non-invasive où les observations sont plus indirectes et le problème inverse de l'obtention des données à la surface du coeur mal posé (i.e. opérateur de diffusion entre la surface du coeur et le thorax). Ce type de reconstruction appelé imagerie électrocardiographie (voir Rudy et Oster (1992)) a été mis en oeuvre avec des vestes comportant 250 électrodes au lieu des 12 dérivations ECG. On pourra se reporter à Moreau-Villégier (2005) pour une discussion détaillée de ces différentes techniques. Elle-même propose d'ailleurs dans le cas de mesures invasives des méthodes variationnelles de reconstruction focalisées sur l'estimation non pas du champ de potentiel mais des temps de dépolarisation. Cette idée pourrait d'ailleurs suffire si l'objectif se limite à l'estimation mécanique car l'amplitude de la dépolarisation est en multiple de σ_0 dans la loi de contraction et donc paramètrable globalement avec lui.

Toujours est-il que le problème de l'estimation électrique est souvent traité de façon découplée de la contraction mécanique. Or l'électrique étant la source de la contraction, l'imagerie cardiaque apporte clairement une information pour l'identification des grandeurs électriques. Autrement dit, comme u est une entrée de la mécanique dans le problème direct, cette dernière devient une source d'identifiabilité de l'activation dans le problème inverse de l'estimation. Donc si on imagine un observateur de Luenberger pour les variables électriques, il faudrait que celui-ci puisse être couplé à notre estimateur mécanique et en bénéficier. Nous constatons donc que, **hormis la dimension du problème électrique**, nous sommes aussi confrontés à un problème d'estimation couplé sous la forme (4.116), (4.117). Afin de contourner la difficulté de la taille du problème nous pouvons envisager utiliser une base réduite sur la variable électrique. Il existe en effet de nombreuses méthodes cherchant à réduire la dimension d'un problème EDP en généralisant la notion de modes linéaires. On peut notamment citer les méthodes POD Kunisch et Volkwein (2002). Le principe est de décomposer les opérateurs dynamiques sur une base de Galerkin de rang faible construite à partir d'un ensemble de solutions précalculées. En pratique d'ailleurs, on continue de construire les opérateurs sur une base éléments finis usuelle avant de les projeter sur la base réduite pour accélérer la résolution au niveau de l'inversion des opérateurs de transition. Soit V le vecteur de degrés de libertés éléments finis discrétisant la variable électrique u suivant une équation de réaction diffusion de la forme

$$M_{\text{el}}\dot{V} + D_{\text{el}}V = R_{\text{el}}(V)$$

où M_{el} correspond à une norme L^2, D_{el} est l'opérateur de diffusion et $R_{\text{el}}(V)$ est le terme de réaction non-linéaire. Soit $[\Psi_i^{\text{POD}}] \in \mathbb{M}_p(\mathbb{R})$ (p petit) la matrice de passage des POD vers la base élément finis utilisée et $\theta_u \in \mathbb{R}^p$ les coefficients décomposant V dans la base POD, on a

$$V = [\Psi_i^{\text{POD}}]\theta_u$$

donc

$$M_{\text{el}}^{\text{POD}}\dot{\hat{\theta}}_u + D_{\text{el}}^{\text{POD}}\theta_u = [\Psi_i^{\text{POD}}]^T R_{\text{el}}^{\text{POD}}([\Psi_i^{\text{POD}}]\theta_u) \tag{4.125}$$

avec

$$M_{\text{el}}^{\text{POD}} = [\Psi_i^{\text{POD}}]^T M_{\text{el}}[\Psi_i^{\text{POD}}], \quad D_{\text{el}}^{\text{POD}} = [\Psi_i^{\text{POD}}]^T D_{\text{el}}[\Psi_i^{\text{POD}}]$$

L'équation (4.125) permet donc d'utiliser **directement** la version non-linéaire de notre filtre paramétrique généralisé (4.117). L'avantage d'une telle formulation est qu'on est sûr de pouvoir imaginer sur ces variables un filtre (en l'occurrence un filtre EKF ou UKF) pour n'importe quel type d'observation. De plus, nous avions noté Remarque 4.2.2 que nous pouvions utiliser d'autres opérateurs d'observation sur le deuxième étage de l'estimation. Cette remarque prend tout son sens ici puisque sur la partie électrique réduite on peut alors utiliser, en plus des observations cinématiques tirées de l'image, les mesures d'ECG ou de potentiel.

$$\begin{cases} \dot{\hat{X}} = A(\hat{X}, V) + K_X(Z - H\hat{X}) + L^X(L^\theta)^{-1}\dot{\theta}_u \\ \dot{\hat{\theta}}_u = f(\hat{\theta}_u) + L^\theta U^{-1} {L^X}^T H^T W^{-1}(Z - H\hat{X}) + L^\theta U^{-1} L^\theta (H_{\text{el}}^{\text{POD}})^T W_{\text{el}}^{-1}(Z_{\text{el}} - H_{\text{el}}^{\text{POD}}\theta_u) \\ \dot{U} = {L^X}^T H^T W^{-1} H L^X + {L^\theta}^T H^T W_{\text{el}}^{-1} H L^\theta \\ V = [\Psi_i^{\text{POD}}]\theta_u \end{cases} \tag{4.126}$$

avec les dynamiques L^X et L^θ (et en conséquence U) du filtre SEEK ou SEIK, et par exemple pour les dérivations ECG

$$Z_{\text{el}} = T_{\text{ECG}}[\Psi_i^{\text{POD}}]\theta_u$$

où T_{ECG} est la matrice de transfert du potentiel électrique à la surface du coeur vers le thorax. On remarquera que vis-à-vis de variable électrique initiale V, l'estimateur (4.126) est une forme de filtrage réduit où la base de décomposition de la covariance est fixée sur la base POD (voir des variantes des filtres SEEK et SEIK sur base fixe dans Hoteit (2001)).

Reste à étudier la performance des méthodes POD en électrophysiolgie. Elles le sont déjà largement en mécanique (solide et fluide) avec des résultats pratiques de réduction très concluants et ont déjà donné lieu à des résultats préliminaires en mécanique cardiaque Pham (2007). Pour des systèmes de réaction diffusion en chimie, Shvartsman et al (2000) ont aussi montré la performance de méthodes POD avec, de plus, la volonté, comme ici, de pouvoir définir des observateurs et stabiliseurs.

Dans un premier temps, ce type de formulation pourrait d'ailleurs être testé sur un système parabolique classique où la réduction de modèle a déjà largement fait ses preuves. Le système cardiaque ne présente pas ce type de couplage entre le problème mécanique et un problème parabolique mais ces systèmes se rencontrent souvent en génie civil pour des problèmes de thermoélasticité (voir Nassiopoulos (2008)). De tels systèmes permettraient de tester en profondeur l'estimateur (4.126) ou des couplage à imaginer avec deux observateurs de Luenberger inspirés du contrôle dans des configurations linéaires (petits déplacements).

Conclusion et perspectives

The concept of a virtual physical human is a sophisticated computer modelling tool, which compares observations of an individual patient and relates them to a vast dataset of observations of others with similar symptoms and known conditions. By processing all this information, the model can simulate the likely reaction of the individual patient to possible treatments or interventions

— The STEP Consortium, *Roadmap to the VPH (Virtual Physiological Human)*

Les mathématiciens sont comme les français : quoi que vous leur disiez ils le traduisent dans leur propre langue et le transforment en quelque chose de totalement différent.

— Johann Wolfgang von Goethe

L'assimilation de données est une discipline qui rencontre de plus en plus de succès dans un nombre croissant de sciences appliquées. Ainsi, le travail de cette thèse avait pour vocation de mettre à profit ces méthodes pour les besoins associés à la modélisation cardiaque, principalement l'identification paramétrique. Rapidement nous avons pourtant pris conscience que, même si l'incertitude principale concerne les paramètres, l'état lui-même devait être estimé, d'où ce souci constant de traiter le problème d'estimation conjointement sur l'état et les paramètres. Les méthodes d'assimilation de données introduites sont donc initialement adaptées au système cardiaque, et nous espérons que les résultats obtenus s'avéreront utiles dans ce cadre. Cependant les caractéristiques mécaniques de ce dernier sont très exigeantes (loi de comportement commandée multi-échelles, non-linéarités fortes, couplages avec des variables de pression, contraintes isovolumiques des cavités) et la complexité même de la configuration étudiée nous assure alors que les résultats méthodologiques présentés dépassent largement le cadre de l'estimation cardiaque pour pouvoir être valorisés dans le cadre de l'estimation en mécanique en général, notamment dans des configurations d'asservissement visuel ou d'estimation à base d'images (cf. Hild et Roux (2006)), ou encore de détection de défauts sur des ouvrages de génie civil... De même, tous les systèmes hyperboliques du second ordre sont potentiellement concernés par les résultats proposés, notamment les problèmes d'acoustique ou les phénomènes vibratoires en général. Plus généralement, nous pensons que le filtrage a un avenir dans les problèmes d'évolution, pour peu qu'on bénéficie de résultats de contrôle

afin d'imaginer de nouveaux types de filtres de Luenberger adaptés à la physique considérée. Nous en avons d'ailleurs donné un premier exemple avec l'équation de la chaleur (parabolique). Une fois ce filtre créé, l'identification des paramètres s'étendra directement à partir des méthodes proposées au chapitre 4. D'une certaine manière, nous pouvons dire que les résultats du chapitre 4 peuvent être compris sous la forme : Pour toute EDP, si vous avez un filtre de Luenberger stabilisant l'état sur une trajectoire observée, alors vous pouvez conjointement identifier les paramètres de manière efficace.

Nous ne voulions pourtant pas simplement nous limiter à revisiter des méthodes classiques pour justifier l'intérêt des méthodes d'estimation en élastodynamique et biomécanique. Ainsi, nous nous sommes orientés résolument vers le filtrage par rapport aux méthodes variationnelles plus classiques même si, a posteriori, nous avons en plus assuré une certaine efficacité à nos filtres en terme de performance, de stabilité générale, de complexité algorithmique et de mise en œuvre dans des codes dédiés. Nous espérons donc que nos résultats sont originaux notamment sur la base des éléments suivants :
Chapitre 3 :

- Proposition d'un estimateur de Luenberger comme outil d'approximation numérique d'un système. En particulier, il apparaît pertinent d'utiliser les données (raisonnablement bruitées) dès qu'elles sont disponibles plutôt qu'une approximation numérique classique, même à connaissance parfaite de la condition initiale et du modèle.
- Obtention de filtres de Luenberger originaux dans le cas de données de type déplacements (données Lagrangiennes) ou positions (données Eulériennes). Le choix d'introduire une forme de norme H^1 dans l'adjoint de l'opérateur d'observation à partir d'un relèvement harmonique (statique mais à la volée) des données nous apporte une efficacité similaire à l'utilisation (plus classique chez les ingénieurs) de vitesses.
- Généralisation non-linéaire de ces filtres. En particulier pour des observations très non-linéaires sous la forme de contours dans des séquences d'images, notre modification des classiques forces-image comme correction, non pas dans le principe des travaux virtuels, mais dans l'équation sur la vitesse comme dérivée du déplacement, nous a permis d'assurer la convergence de l'estimateur, au moins en terme de linéarisé.

Chapitre 4 :

- Obtention de l'équivalence (sous condition d'absence d'erreur de modèle autre que l'incertitude paramétrique) entre le filtre adaptatif proposé par Zhang et Clavel (2001) et le filtre SEEK de Pham et al (1997). Cette équivalence permet de relier le filtre adaptatif à un critère moindres carrés et de proposer un schéma numérique stable. Inversement, elle permet l'analyse d'erreur d'estimation du filtre SEEK par le classique changement de variable $(\tilde{X}, \tilde{\theta}) \to (\eta, \tilde{\theta})$ qui apparaissait déjà dans les théories autour du filtrage réduit de Luenberger (cf. Luenberger (1971) ; Bonnans et Rouchon (2005)).
- Extensions multiples du filtre adaptatif initial avec des versions d'ordre 2, des contraintes paramétriques, un déclenchement retardé mais surtout une version H^∞ pour laquelle nous avons pu mener l'analyse de convergence.

- Démonstration du fait que le filtre SEIK proposé par Pham et al (1997) est une version réduite du filtre UKF pour sigma-points simplectiques. Nous avons alors généralisé le formalisme au cas d'autres choix de sigma-points. Comme pour des systèmes mécaniques de faible dimension (cf. Wu et Smyth (2006) ; Mariani et Ghisi (2007)), ce filtre non-linéaire se révèle efficace, à la fois en termes de résultat et de mise en oeuvre numérique (absence de différentiation).

Ces résultats méthodologiques originaux ont été accompagnés de « preuves de concept » à partir d'exemples numériques que nous avons souhaités les plus pertinents possibles. Ceci explique largement le très gros effort autour du code de simulation direct, mais aussi sur la compréhension des différents types d'observations disponibles en cardiologie. En particulier le stage de master de O. Talcolth (cf. Talcoth (2007)), sur la génération d'échographies synthétiques et l'analyse d'erreur pour le flot optique à partir de données cardiaques synthétiques, a déjà montré combien la modélisation et le traitement d'images s'apportaient mutuellement.

Il reste cependant beaucoup à faire pour envisager de créer des outils véritablement efficaces sur données réelles.

- **Côté modélisation**, on imagine notamment une gestion plus fine (après tous les efforts consentis au niveau microscopique) du couplage entre l'entrée électrique nécessaire à la mécanique et les potentiels intra- et extra-cellulaire fournis classiquement par les modélisations électriques. De plus, les aspects énergétiques au sens physiologique (apport en oxygène) sont sans doute déterminants pour de nombreuses pathologies. Enfin, de nombreux médecins pensent que les aspects de modélisation géométrique sont très importants dans la contraction notamment pour le ventricule droit (piliers, bandes modératrices (voir Netter (1969)), tissus conjonctifs entourant les valves...).
  Conjointement aux efforts croissants de modélisation, il faudra nécessairement obtenir des résultats de validation de cette modélisation sur des sous-systèmes, notamment avec des expériences de Hill, voire sur des cellules cardiaques.

- **Côté traitement d'images**, même si nous nous sommes efforcés de proposer des estimateurs à partir d'images, il apparaît déjà que le potentiel des tags pour l'observabilité du système mécanique est considérable. Or à ce jour, les mesures ne sont pas encore entièrement 3D et de nombreux outils de traitement restent à imaginer. Je pense d'ailleurs qu'il faut peut-être se résoudre à ne pas espérer trop longtemps obtenir des déplacements 3D comme mesures, mais plutôt à s'adapter à la spécificité actuelle des observations d'IRM taggé sous formes de séries de plans en mouvement.

- Enfin l'estimation sur données réelles est inenvisageable si nous ne pouvons pas gérer conjointement les variables mécaniques, les paramètres, mais aussi les variables de pressions et l'entrée électrique, si possible à partir de données non invasives (pression aux doigts, ECG...). Nous avons déjà évoqué ces perspectives, reste maintenant à démontrer leur potentiel.

Pour conclure, comme nous avions commencé, sur le cadre général scientifique , cette thèse est née, je crois, de la conviction de D. Chapelle, P. Le Tallec et M. Sorine que dans la modélisation cardiaque, il ne pouvait y avoir d'impact applicatif sans que la modélisation soit intimement liée à l'assimilation de données hétérogènes disponibles après examen d'un patient. J'espère que ce travail a pu servir cette conviction que l'assimilation de données permet en effet comme dans d'autres sciences de l'ingénieur d'adapter la trajectoire d'un modèle générique à la configuration spécifique du patient, notamment dans le souci d'identifier des pathologies éventuelles. Mais plus encore dans cette discipline confrontée au vivant, elle permet déjà de justifier le pouvoir prédictif du modèle. En effet si les méthodes d'assimilation de données assurent sur des modèles et des observations synthétiques réalistes que l'on est capable d'estimer l'état du système au sens large (trajectoire et paramètres), alors on détient un outil formidable pour évaluer ensuite la qualité d'un modèle à prédire potentiellement tel ou tel type de pathologie. Insistons finalement une fois de plus sur le fait que l'objectif n'est pas, en modélisation pour le vivant, de « supporter » toute l'envergure de la prédiction médicale, mais plutôt d'apporter au médecin, sur un diagnostic en cours, des outils quantitatifs de diagnostic et de pronostic qui lui permettront d'affiner la thérapeutique à suivre. Ainsi la notion de « bon modèle » ne se justifie que par rapport à l'application ciblée et en fonction des données disponibles, et c'est bien l'assimilation de données qui unifie ce triptyque modélisation, observations et applications. La citation du *Virtual Physiological Human* qui ouvre cette conclusion laisse entrevoir de vastes perspectives pour la modélisation en médecine. Notre espoir est que les travaux et la réflexion ici présentés offriront un axe de recherche pertinent pour accompagner cet élan scientifique prometteur.

Bibliographie

AHA/ACC/SNM. 1992, «Standardization of cardiac tomographic imaging», *Circulation*, vol. 86, pp. 338–339. 71

Angelini, E. et O. Gérard. 2006, «Review of myocardial motion estimation methods from optical flow tracking on ultrasound data», dans *Proc. of the 28th IEEE EMBS Annual International Conference*. 147

Astorino, M., J.-F. Gerbeau, O. Pantz et K. Traoré. 2008, «Fluid-structure interaction and multibody contact. Application to aortic valves», *Comp. Meth. Appl. Mech. Engng.* To appear. 100

Augenstein, K. F., B. W. Cowan, I. J. LeGrice, P. M. F. Nielsen et A. A. Young. 2005, «Method and apparatus for soft tissue material parameter estimation using tissue tagged magnetic resonance imaging», *J. Biomech. Eng.*, vol. 127, no. 1, pp. 148–157. 246

Axel, L., A. Montillo et D. Kim. 2005, «Tagged magnetic resonance imaging of the heart : a survey», *Med. Image Anal.*, vol. 9, p. 376. 139

Başar, T. 2002, «Paradigms for robustness in controller and filter designs», *J. Macroeconomic Dynamics*. 53

Başar, T. et P. Bernhard. 1995, *H^∞-Optimal Control and Related Minimax Design Problems - A Dynamic Game Approach*, 2nde édition, Birkhaüser. 53

Baerentzen, J. et H. Aanaes. 2005, «Signed distance computation using the angle weighted pseudo-normal», *IEEE Trans. Visual. Comput. Graph.*, vol. 11, no. 3, pp. 243–253. 116, 231

Banks, H. T. et R. H. Fabiano. 1998, «Approximation issues for applications in optimal control and parameter estimation», dans *Modelling and computation for applications in mathematics, science, and engineering (Evanston, IL, 1996)*, Numer. Math. Sci. Comput., Oxford Univ. Press, New York, pp. 141–165. 195

Banks, H. T. et K. Kunisch. 1989, *Estimation techniques for distributed parameter systems*, *Systems & Control : Foundations & Applications*, vol. 1, Birkhäuser Boston Inc., Boston, MA. 8

Bardos, C., G. Lebeau et J. Rauch. 1992, «Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary», *SIAM J. Control Optim.*, vol. 30, no. 5, pp. 1024–1065, ISSN 0363-0129. 185

Bathe, K. J. 1996, *Finite Element Procedure*, Prentice-Hall. 124, 126

Benoit-Cattin, H., G. Collewet, B. Belaroussi, S. Saint-Jalmes et C. Odet. 2005, «The simri project : a versatile and interactive MRI simulator», *Journal of Magnetic Resonance*, vol. 173, no. 97–115. 137

Bensoussan, A. 1971, *Filtrage optimal des systèmes linéaires*, Dunod. 7, 8, 29, 207

von Berg, J., H. Barschdorf, T. Blaffert, S. Kabus et C. Lorenz. 2007, «Surface based cardiac and respiratory motion extraction for pulmonary structures from multi-phase ct», dans *Medical Imaging 2007 : Physiology, Function, and Structure from Medical Images*, vol. 6511, édité par A. Manduca et X. Hu, SPIE, p. 65110Y, doi :10.1117/12.709395. URL `http://link.aip.org/link/?PSI/6511/65110Y/1`. 229

Bernadou, M., P. L. George, A. Hassim, P. Joly, P. Laug, B. Muller, A. Perronnet, E. Saltel, D. Steer, G. Vanderborck et M. Vidrascu. 1988, *MODULEF : une bibliothèque modulaire d'éléments finis*, 2nde édition, Institut National de Recherche en Informatique et en Automatique (INRIA), Rocquencourt. 166

Bestel, J. 2000, *Modèle différentiel de la contraction musculaire contrôlée. Application au système Modèle différentiel de la contraction musculaire contrôlée. Application au système cardio-vasculaire*, Thèse, University Paris IX Dauphine. 92, 93, 156

Bestel, J., F. Clément et M. Sorine. 2001, «A biomechanical model of muscle contraction.», dans *Lecture Notes in Computer Science*, vol. 2208, Eds W.J. Niessen and M.A. Viergever, Springer-Verlag. 94

Blum, J., F. X. Le Dimet et I. M. Navon. 2008, «Data assimilation for geophysical fluids», dans *Handbook of Numerical Analysis : Computational Methods for the Atmosphere and the Oceans*, édité par R. Temam et J. Tribbia, Elsevier. 8, 65, 161, 243

Bonnans, F. et P. Rouchon. 2005, *Commande et optimisation de systèmes dynamiques*, Editions de l'Ecole Polytechnique. 299, 306

Bonnans, J. F., J. C. Gilbert, C. Lemaréchal et C. A. Sagastizábal. 2006, *Numerical optimization : Theoretical and practical aspects*, 2nde édition, Universitext, Springer-Verlag, Berlin. 162

Boulakia, M., M. A. Fernández, J.-F. Gerbeau et N. Zemzemi. 2007, «Towards the numerical simulation of electrocardiograms», dans *Functional Imaging and Modeling of the Heart*, édité par F. Sachse et G. Seemann, no. 4466 dans Lecture Notes in Computer Science, Springer Verlag, pp. 240–249. 107

Boulakia, M., M. A. Fernández, J.-F. Gerbeau et N. Zemzemi. 2008, «A coupled system of pdes and odes arising in electrocardiograms modelling», *Applied Math. Res. Exp.* To appear. 107, 108, 131

Bourquin, F. 1995, «A numerical controllability test for distributed systems», *J. Struct. Contr.*, pp. 5—23. 176

Bourquin, F. 1998, «Control of flexible structures : Control theory and approximation issues», dans *CIMNE*, Barcelona. 176

Bui, H. D. 1994, *Inverse Problems in the Mechanics of Materials : An Introduction*, CRC Press. 246

Burq, N. et P. Gérard. 2002, «Contrôle optimal des equations aux derivees partielles», Cours de l'Ecole Polytechnique. 184, 185

Caillerie, D., M. Mourad et A. Raoult. 2003, «Cell-to-muscle homogenization. Application to a constitutive law for the myocardium», *ESAIM Math. Model. Numer. Anal.*, vol. 37, no. 4, pp. 681–698. 99

Casella, G. et R. L. Berger. 2002, *Statistical Inference*, 2nde édition, Duxbury Press. 14

Cazorla, O., A. Lacampagne, J. Fauconnier et G. Vassort. 2003, «A ca2+ antagonist with length-dependent ca2+ -sensitizing properties in cardiac myocytes», *Brit. J. Pharmaco*, vol. 139, no. 99–108. 5, 154

Chapelle, D., F. Clément, F. Génot, P. Le Tallec, M. Sorine et J. Urquiza. 2001, «A physiologically-based model for the active cardiac muscle.», dans *Lectures Notes in Computer Science*, vol. 2230, Eds T. Katila, I.E. Magnin, P. Clarysse, J. Montagnat, J. Nenonen, Springer-Verlag. 97, 130

Chapelle, D., M. A. Fernàndez, J.-F. Gerbeau, P. Moireau, J. Sainte-Marie et N. Zenzemi. 2008, «A 3D model for the electromechanical activity of the heart», In preparation. 9, 11, 107, 131, 147, 156

Chapelle, D., P. Moireau et P. Le Tallec. 2009, «Robust filtering for joint state-parameter estimation in distributed mechanical systems», *DCDS–A*, vol. 23, no. 1/2, pp. 65–84. A special issue Dedicated to Ta-Tsien Li on the Occasion of his 70th Birthday. 10, 11, 178, 268

Chapelle, D., P. Moireau, M. Sermesant, M. A. Fernández et H. Delingette. 2006, «Le Cœur Numérique», Film. Réalisation A. Paouri, Production INRIA Multimedia. 9, 12

Chavent, G. 2008, «Non linear least squares for inverse problems», In redaction. 8, 44, 163

Ciarlet, P. G. 1988, *Mathematical Elasticity, Vol. I : Three-Dimensional Elasticity*, Studies in Mathematics and its Applications, North Holland. 90, 185

Ciarlet, P. G. et G. Geymonat. 1982, «Sur les lois de comportement en élasticité non linéaire», *C.R.A.S, Série II*, vol. 295, pp. 423–426. 89

Cohen, I., L. D. Cohen et N. Ayache. 1991, «Using deformable surfaces to segment 3-D images and infer differential structures.», *Image Understanding*, vol. 53, no. 2, pp. 211–218. 228

Collet, M., V. Walter et P. Delobelle. 2003, «Active damping of a micro-cantilever piezo-composite beam», *J. Sound Vibration*, vol. 260, no. 3, pp. 453–476. 190

Colli Franzone, P., L. Guerri et C. Viganotti. 1980, «Relating surface to epicardial potentials with direct and inverse calculations», dans *Simulation of systems '79 (Proc. Ninth IMacs Congress, Sorrento, 1979)*, North-Holland, Amsterdam, pp. 665–674. 4

Coron, J. M. 2007, *Control and Nonlinearity*, vol. 136, Mathematical Surveys and Monographs. 7, 171, 233, 249

Cottet, G.-H. et E. Maitre. 2006, «A level-set method for fluid-structure interactions with immersed surfaces», *Math. Model. Meth. Appl. Sci.*, vol. 16, no. 3, pp. 415–438. 154

Cox, S. et E. Zuazua. 1994, «The rate at which energy decays in a damped string», *Comm. Partial Differential Equations*, vol. 19, no. 1-2, pp. 213–243. 186, 187, 199, 211

Cox, S. et E. Zuazua. 1995, «The rate at which energy decays in a string damped at one end», *Indiana Univ. Math. J.*, vol. 44, no. 2, pp. 545–573. 187, 211

Crépeau, E. et M. Sorine. 2007, «A reduced model of pulsatile flow in an arterial compartment», *Chaos, Solitons Fractals*, vol. 34, no. 2, pp. 594–605. 152, 239

Curtain, R. F. et H. Zwart. 1995, *An introduction to infinite-dimensional linear systems theory*, *Texts in Applied Mathematics*, vol. 21, Springer-Verlag, New York. 186

Declerck, J., T. Denney, C. Oztürk, W. O'Dell et E. R. McVeigh. 1999, «LV motion reconstruction from tagged MRI : a comparison», dans *International Society of Magnetic Resonance in Medicine*, Philadelphia, PA, USA, p. 1283. 139

Diniz Dos Santos, N. M., J.-F. Gerbeau et J.-F. Bourgat. 2008, «A partitioned fluid-structure algorithm for elastic thin valves with contact», *Comp. Meth. Appl. Mech. Engng.*, vol. 197, pp. 1750–1761. 100

Djabella, K. et M. Sorine. 2005, «Differential model of the excitation-contraction coupling in a cardiac cell for multicycle simulations», dans *EMBEC'05*, vol. 11, Prague, pp. 4185–4190. 108

Duan, Q., E. Angelini, S. Homma et A. Laine. 2006, «Tracking the endocardium using optical flow along iso-value curves», dans *Proc. of IEEE EMBS Annual International Conference*, pp. 707–710. 147

Duan, Q., P. Moireau, E. Angelini et D. Chapelle. 2007, «Simulation of 3D ultrasound with a realistic electro-mechanical model of the heart», dans *Proceedings of Functional Imaging and Modeling of the Heart 2007 (FIMH'07)*, *Lecture Notes in Computer Science (LNCS)*, vol. 4466, pp. 463–470. 9, 12, 144, 145, 146, 147

Ervedoza, S. et E. Zuazua. 2008, «Uniformly exponentially stable approximations for a class of damped systems», To be published. 197, 199

Evensen, G. 2007, *Data Assimilation – The Ensemble Kalman Filter*, Springer Verlag. 65

Faurre, P. et M. Robin. 1984, *Éléments d'automatique*, Dunod. 49, 52

Fenton, F. et A. Karma. 1998, «Vortex dynamics in three-dimensional continuous ortex dynamics in three-dimensional continuous myocardium with fiber rotation : Filament instability and fibrillation», *Chaos*, vol. 8, no. 20. 108

Fillard, P., V. Arsigny, X. Pennec et N. Ayache. 2007, «Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics», *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1472–1482. 115, 119

Fleming, H. W. 1997, «Deterministic nonlinear filtering», *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série*, vol. 25, no. 3-4, pp. 435–454. 48

Fleming, H. W. et R. W. Rischel. 1975, *Deterministic and Stochastic Optimal Control*, Springer-Verlag. 48, 49

Frey, P. 2001, «`YAMS`. A fully automatic adaptive isotropic surfacere meshing procedure», Rap. Tech. 0252, INRIA. 112

Frey, P. et P. L. George. 2008, *Mesh Generation. Application to finite elements*, 2nde édition, Wiley. 112

Fung, Y. 1993, *Biomechanics : Mechanical Properties of Living Tissues*, 2nde édition, Springer. 5

George, P. L., F. Hecht et E. Saltel. 1990, «Fully automatic mesh generator for 3d domains of any shape», *Impact of Comp. in Sci. ans Eng.*,, vol. 2, pp. 187–218. 112

Gohberg, I. C. et M. G. Kreĭn. 1969, *Introduction to the theory of linear nonselfadjoint operators*, Translated from the Russian by A. Feinstein. Translations of Mathematical Monographs, Vol. 18, American Mathematical Society, Providence, R.I. 187

Guesmia, A. 2000, *Contributions à la controlabilité exacte et la stabilisation des systèmes d'évolution*, Thèse, Université Louis Pasteur, Strasbourg. 185

Haddad, R., I. E. Magnin et P. Clarysse. 2007, «A new fully-digital anthropomorphic and dynamic thorax/heart model», dans *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, Lyon. 137

Hahn, W. 1967, *Stability of Motion*, Springer Verlag. 169, 171

Hassibi, B., A. H. Sayed et T. Kaylath. 1999, *Indefinite-quadratic Estimation and Control : A Unified Approach to H^2 and H^∞ Theories*, SIAM. 32, 45

Hauret, P. et P. Le Tallec. 2006, «Energy-controlling time integration methods for nonlinear elasticity», *Comput. Meth. Appl. Mech. Eng.*, vol. 195, no. 4890–4916. 127

Hild, F. et S. Roux. 2006, «Digital image correlation : from displacement measurement to identification of elastic properties – a review», *Strain*, vol. 42, no. 2, pp. 69–80(16). 244, 305

Hill, T. L. 1975, «Theoretical formalism for the sliding filament model of contraction, part ii», *Prog. Biophys. molec. Biol.* 154

Hill, V. A. 1938, «The heat of shortening and the dynamic constants of muscle», dans *Proc. Roy. Soc. London*, pp. 126–136. 97

Hoteit, I. 2001, *Filtres de Kalman réduits et efficaces pour l'assimilation de données en océanographie*, Thèse, Université J. Fourier, Grenoble. 8, 255, 281, 304

Hunter, P. J. 1975, *Finite Element Analysis of Cardiac Muscle Mechanics*, Thèse, University of Oxford. 4, 5

Hunter, P. J., A. D. McCulloch et H. E. D. ter Keurs. 1998, «Modelling the mechanical properties of cardiac muscle», *Progr. Biophys. Mol. Biol.*, vol. 69, pp. 289–331. 99

Huxley, A. F. 1957, «Muscle structure and theories of contraction», dans *Progress in biophysics and biological chemistry*, vol. 7, Pergamon press, pp. 255–318. 92, 93

Jazwinsky, A. H. 1970, *Stochastic processes and filtering theory*, Academic Press. 58

Julier, S., J. Uhlmann et H. Durrant-Whyte. 1995, «A new approach for filtering nonlinear systems», dans *American Control Conference*, pp. 1628–1632. 10, 59

Julier, S., J. Uhlmann et H. Durrant-Whyte. 2000, «A new method for the nonlinear transformation of means and covariances in filter and estimators», *IEEE Transactions on Autonatic Control*, vol. 45, no. 3, pp. 447–482. 60, 62, 287, 293

Julier, and J. K. Uhlmann, S. J. 2002a, «Reduced Sigma Point Filters for the Propagation of Means and Covariances through Nonlinear Transformations», dans *Proc. of IEEE Am. Contr. Conf.*, IEEE, Anchorage AK, USA, pp. 887–892. 63

Julier, and J. K. Uhlmann, S. J. 2002b, «The Scaled Unscented Transformation», dans *Proc. of IEEE Am. Contr. Conf.*, IEEE, Anchorage AK, USA, pp. 4555–4559. 63

Kalman, R. et R. Bucy. 1961, «New results in linear filtering and prediction theory», *Trans. ASME J. Basic. Eng.*, vol. 83, pp. 95–108. 31

Kato, T. 1995, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin. Reprint of the 1980 edition. 195

Khalil, H. K. 1992, *Nonlinear systems*, Macmillan Publishing Company, New York, ISBN 0-02-363541-X. 171

Komornik, V. 1994, *Exact controllability and stabilization. The multiplier method*, RAM : Research in Applied Mathematics, Masson, Paris. 185

Kunisch, K. et S. Volkwein. 2002, «Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics», *SIAM J. Numer. Anal.*, vol. 40, no. 2, pp. 492–515. 303

Laleg, T. M., E. Crépeau et M. Sorine. 2007, «eparation of arterial pressure into a nonlinear superposition of solitary waves and a windkessel flow», *Biomed. Signal. Process. Contr.*, vol. 2, pp. 163–170. 152, 239

Lasiecka, I. et R. Triggiani. 2003, «$L_2(\Sigma)$-regularity of the boundary to boundary operator B^*L for hyperbolic and Petrowski PDEs», *Abstr. Appl. Anal.*, , no. 19, pp. 1061–1139. 176, 177, 207

Le Tallec, P. 1994, «Numerical methods for nonlinear three-dimensional elasticity», dans *Handbook of Numerical Analysis*, vol. 3, édité par P. G. Ciarlet et J.-L. Lions, Elsevier. 88, 96

Le Tallec, P. 2000, *Introduction à la Dynamique des Structures, Cours de l'Ecole Polytechnique.*, Ellipses. 188

Lefebvre, T., H. Bruyninckx et J. De Schuller. 2002, «Comment on "a new method for the nonlinear transformation of means and covariances in filters and estimators" [and authors' reply]», *IEEE Trans. Automat. Contr.*, vol. 47, no. 8, pp. 1406– 1409. 60, 62

Legland, F. 2002, «Filtre de kalman et modèles de markov cachés», Master's Lectures Notes. 16, 65

Lions, J.-L. 1968, *Contrôle optimal de systemes gouvernés des équations aux derivées partielles*, Dunod. 7

Lions, J.-L. 1988, *Contrôlabilité exacte, perturbations et stabilisation de systèmes distribués. Tome 1, Recherches en Mathématiques Appliquées*, vol. 8, Masson, Paris, ISBN 2-225-81477-5. Avec appendices de E. Zuazua, C. Bardos, G. Lebeau and J. Rauch. 185

Lipster, R. S. 2008, «Orthogonal projection. conditional expectation in orthogonal projection. conditional expectation in the wide sense», Lecture Notes for the course Stochastic Processes. 16, 17, 18

Ljung, L. 1979, «Asymptotic behavior of the extended kalman filter as parameter estimator for linear systems», *IEEE Trans. Automat. Contr.*, vol. AC-24, no. 1, pp. 36–50. 282

Lorensen, W. E. et H. E. Cline. 1987, «Marching cubes : A high resolution 3d surface construction algorithm», *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169. 115

Luenberger, D. G. 1963, *Determining the State of a Linear with Observers of Low Dynamic Order*, Thèse, Stanford University. 170

Luenberger, D. G. 1971, «An introduction to observers», *IEEE Transactions on Automatic Control*, vol. 16, pp. 596–602. 306

Luo, C. H. et Y. Rudy. 1994, «A model of the ventricular cardiac action ptentiel. depolarisation, repolarisation,and their interaction», *Circ. Res.*, , no. 68, pp. 1071–1096. 108

M., P., M. Sermesant, C. Chung, G. Liu, E. V. McVeigh, E. Crystal et G. Wright. 2007, «An experimental framework to validate 3D models of cardiac electrophysiology via optical imaging and MRI», dans *Proceedings of Functional Imaging and Modeling of the Heart 2007 (FIMH'07), Lecture Notes in Computer Science (LNCS)*, vol. 4466, pp. 100–109. 153

Mallet, V. et B. Sportisse. 2006, «Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations : An ensemble approach applied to ozone modeling», *J. Geophys. Res.*, vol. 111. 29

Mariani, S. et A. Ghisi. 2007, «Unscented kalman filtering for nonlinear structural dynamics», *Nonlinear Dynamics*, vol. 49, no. 1–2, pp. 131–150. 59, 307

Marsden, J. E. et T. J. R. Hughes. 1983, *Mathematical Foundations of Elasticity*, Dover Publications, New York. 67, 179

McVeigh, E. V., O. Faris, D. Ennis, P. Helm et F. Evans. 2001, «Measurement of ventricular wall motion, epicardial electrical mapping, and myocardial fiber angles in the same heart», dans *Functional Imaging and Modeling of the Heart (FIMH'01)*, no. 2230 dans Lecture Notes in Computer Science (LNCS), Springer, pp. 72–82. 153

Meunier, J. et M. Bertrand. 1995, «Echographic image mean gray level changes with tissue dynamics : A system-based model study», *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 4, pp. 403–410. 144, 145

Micu, S. et E. Zuazua. 2004, «An introduction to the controllability of partial differential equations», dans *Quelques questions de théorie du contrôle*, édité par T. Sari, Collection Travaux en Cours Hermann, pp. 69–157. 175

Mitchell, C. C. et D. G. Schaeffer. 2003, «A two-current model for the dynamics of cardiac membrane», *Bulletin Math. Bio.*, , no. 65, pp. 767–793. 108

Moireau, P., D. Chapelle et P. Le Tallec. 2008a, «Filtering for distributed mechanical systems using position measurements : Perspective in medical imaging», Accepted for publication in Inverse Probl. 10, 11, 215

Moireau, P., D. Chapelle et P. Le Tallec. 2008b, «Joint state and parameter estimation for distributed mechanical systems», *Computer Methods in Applied Mechanics and Engineering*, vol. 197, pp. 659–677. 10, 11, 178, 251, 254

Montillo, A., D. N. Metaxas et L. Axel. 2002, «Automated segmentation of the left and right ventricles in 4D cardiac SPAMM images», dans *MICCAI '02 : Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part I*, Springer-Verlag, London, UK, pp. 620–633. 228

Moreau-Villégier, V. 2005, *Méthodes variationnelles et séquentielles pour l'étude de la contraction cardiaque*, Thèse, Université de Nice-Sophia Antipolis-UFR sciences. 153, 303

Mortensen, R. E. 1968, «Maximum likelihood recursive nonlinear filtering», *J. Optim. Theory Appl.*, , no. 2, pp. 386–394. 48, 51

Mourad, A. 2003, *Description topologique de l'architecture fibreuse et modélisation mécanique du myocarde*, Thèse, Ecole doctorale MSTI de l'INPG. 71

Nassiopoulos, A. 2008, *Identification rapide de la température dans les structures du génie civil*, Thèse, Ecole Nationale des Ponts et Chaussées. 304

Netter, F. H. 1969, *The Netter Collection of Medical Illustrations, Heart : Volume 5*, Saunders Elsevier. 68, 74, 108, 307

Pan, L., J. L. Prince, J. A. C. Lima et N. O. Osman. 2005, «Fast tracking of cardiac motion using 3D-HARP», dans *IEEE Trans. on Biom. Eng.*, vol. 52. 139

Panfilov, A. V. et A. T. Winfree. 1985, «Dynamical simulations of twisted scroll rings in three-dimensional excitable media», *Phys. D*, vol. 17, no. 3, pp. 323–330. 4

Papademetris, X., A. J. Sinusas, D. P. Dione et J. S. Duncan. 2004, «Estimation of 3D left ventricular deformation from echocardiography», *Med. Image Anal.*, vol. 8, pp. 285–294. 228

Penrose, R. 1955, «A generalized inverse for matrices.», dans *Proc. Cambridge Phil. Soc.*, vol. 51, pp. 406–413. 16

Peskin, C. S. 1982, «The Fluid Dynamics of Heart Valves : Experimental, Theoretical, and Computational Methods», *Annual Review of Fluid Mechanics*, vol. 14, pp. 235–259. 4

Peters, J., O. Ecabert, C. Meyer, H. Schramm, R. Kneser, A. Groth et J. Weese. 2007, «Automatic whole segmentation in static magnetic resonance image volumes», dans *Proc. MICCAI 2007*. 139

Petersen, S., V. Peto, M. Rayner, J. Leal, R. Luengo-Fernandez et A. Gray. 2005, *European cardiovascular disease statistics*, BHF, London, UK. 2

Peyrat, J. M., M. Sermesant, X. Pennec, H. Delingette, C. Xu, E. V. McVeigh et N. Ayache. 2007, «A computational framework for the statistical analysis of cardiac diffusion tensors : Application to a small database of canine hearts», *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1500–1514. 115, 121, 122, 123

Peyré, G. et L. Cohen. 2005, «Geodesic computations for fast and accurate surface remeshing and parameterization», dans *Progress in Nonlinear Differential Equations and Their Applications*, vol. 63, pp. 157–171. 119

Pham, D. T., J. Verron et M. C. Roubeaud. 1997, «A singular evolutive interpolated kalman filter for data assimilation in oceanography», *J. Marine Systems*, vol. 16, pp. 323–341. 8, 10, 251, 252, 255, 263, 282, 284, 285, 287, 288, 289, 290, 293, 306, 307

Pham, K. 2007, *Réduction de modèle pour la simulation cardiaque*, Rapport de Master, ENS Cachan - Paris 6. 304

Pioletti, D. P., L. R. Rakotomanana, J. F. Benvenuti et P. F. Leyvraz. 1998, «Viscoelastic constitutive law in large deformations : application to human knee ligaments and tendons», *J. Biomech.*, vol. 31, pp. 753–757. 92

Pons, J. P., F. Ségonne, D. Boissonnat, L. Rineau, M. Yvinec et R. Keriven. 2007, «High-quality consistent meshing of multi-label datasets», dans *Information Processing in Medical Imaging*, pp. 198–210. 142, 240

Preumont, A. 2002, *Vibration Control of Active Structures, An Introduction*, 2nde édition, Kluwer Academic Publishers. 190, 192, 193

Pustoc'h, A., J. Ohayon, Y. Usson, A. Kamgoue et P. Tracqui. 2005, «An integrative model of the self-sustained oscillating contractions of cardiac myocytes,», *Acta Biotheor.*, vol. 53, no. 4, pp. 277–293. 154

Raoult, A. 2009, «Symmetry groups in nonlinear elasticity - an exercise in vintage mathematics», *Comm. Pure Appl. Math.*, vol. 8, no. 1, pp. 435–456. In redaction. 91

Roder Tcheugoué Tébou, L. et E. Zuazua. 2007, «Uniform boundary stabilization of the finite difference space discretization of the 1-d wave equation», *Adv. Comput. Math.*, vol. 26, pp. 337–365. 195, 199

Rouchdy, Y., J. Pousin, J. Schaerer et P. Clarysse. 2007, «A nonlinear elastic deformable template for soft structure segmentation : application to the heart segmentation in MRI», *Inverse Problems*, vol. 23, no. 3, pp. 1017–1035. 229

Rudy, Y. et H. S. Oster. 1992, «The electrocardiographic inverse problem.», *Crit. Rev. Biomed. Eng.*, vol. 20, no. 1-2, pp. 25–45. 303

Sachse, F. B. 2004, *Computational Cardiology, Modeling of Anatomy, Electrophysiology, and Mechanics*, *Lecture Notes in Computer Science*, vol. 2966, Springer. 99

Sachse, F. B., R. Frech, C. D. Werner et O. Dossel. 1999, «A model based approach to assignment of myocardial fibre orientation», *Computers in Cardiology 1999*, pp. 145–148. 119

Sainte-Marie, J., D. Chapelle, R. Cimrman et M. Sorine. 2006, «Modeling and estimation of the cardiac electromechanical activity», *Comp. & Struct.*, vol. 84, pp. 1743–1759. 105, 130, 156

Sarkka, S. 2007, «On unscented kalman filtering for state estimation of continuous-time nonlinear systems», *IEEE Trans. Automat. Contr.*, vol. 52, no. 9, pp. 1631–1641. 301

Schmid, H. et P. J. Hunter. 2006, «Biomechanical modelling at the molecular, cellular, and tissue levels», Lectures Notes – 14th CISM-IUTAM Summer School. 99, 117

Sermesant, M. 2003, *Modèle électromécanique du coeur pour l'analyse d'image et la simulation*, Thèse, Université de Nice-Sophia Antipolis-UFR sciences, Ecole doctorale STIC. 138, 229, 241

Sermesant, M., P. Moireau, O. Camara, J. Sainte-Marie, R. Andriantsimiavona, R. Cimrman, D. L. Hill, D. Chapelle et R. Razavi. 2006, «Cardiac function estimation from MRI using a heart model and data assimilation : Advances and difficulties», *Medical Image Analysis*, vol. 10, no. 4, pp. 642–656. 11

Sermesant, M., K. Rhode, A. Anjorin, S. Hegde, G. Sanchez-Ortiz, D. Rueckert, P. Lambiase, C. Bucknall, D. Hill et R. Razavi. 2004, «Simulation of the electromechanical activity of the heart using XMR interventional imaging», dans *Medical Image Computing and Computer-Assisted Intervention (MICCAI'04)*, *Lecture Notes in Computer Science (LNCS)*, vol. 3217, Springer, pp. 786–794. 153

Shubov, M. A. 1997, «Spectral operators generated by damped hyperbolic equations», *Integr. Equat. Operat. Theor.*, vol. 28, pp. 358–372. 186

Shung, K. 1993, *Ultrasonic Scattering in Biological Tissues*, CRC Press, Boca Raton. 143

Shvartsman, S. Y., C. Theodoropoulos, R. Rico-Martínez, I. G. Kevrekidis, E. S. Titi et T. J. Mountziaris. 2000, «Order reduction for nonlinear dynamic models of distributed reacting systems», *J. Process. Contr.* 304

Simon, D. 2006, *Optimal State Estimation : Kalman, H^∞, and Nonlinear Approaches*, Wiley-Interscience. 58, 59, 283, 284

Smith, N. P., D. P. Nickerson, E. J. Crampin et P. J. Hunter. 2000, «Computational mechanics of the heart. from tissue structure to ventricular function», *J. of Elasticity*, vol. 61, no. 1, pp. 113–141. 70, 71

Stergiopulos, N., B. E. Westerhof et N. Westerhof. 1999, «Total arterial inertance as the fourth element of the windkessel model», *Am. J. Physiol.*, vol. 276, pp. H81–H88. 103

Streeter, D. D. 1979, «Gross morphology and fiber geometry of the heart», dans *Handbook Physiology. The cardiovascular system*, édité par R. Berne, N. Sperelakis et S. Geiger, Am. Phys. Soc. Williams Wilkins, Baltimore. 71

Sundnes, J., G. T. Lines, X. Cai, B. F. Nielsen, K.-A. Mardal et A. Tveito. 2006, *Computing the electrical activity in the heart*, *Monographs in Computational Science and Engineering*, vol. 1, Springer-Verlag, Berlin. 107, 110

Talagrand, O. 2003, «Bayesian estimation. optimal interpolation. statistical linear estimation», dans *Data Assimilation for the Earth System*, Advanced Study Institute (NATO), Acquafredda di Maratea, Italy, May-June 2002, Kluwer Academic Publisher, Dordrecht, The Netherlands, pp. 21–35. 23

Talcoth, O. 2007, *Movement Estimation in Simuated 3D Ultrasound Images*, Rapport de Master, Ecole Polytechnique. 9, 144, 307

ten Tusscher, K. H. W. J., D. Noble, P. J. Noble et A. V. Panfilov. 2004, «A model for human ventricular tissue.», *Am J. Physiol. Heart. Circ. Physiol.*, vol. 286, pp. 1573–1589. 108

Tisseur, F. et K. Meerbergen. 2001, «The quadratic eigenvalue problem», *SIAM Review*, vol. 43, no. 2, pp. 235–286. 187

Tong, S. et P. Shi. 2006, «Cardiac motion recovery : Continuous dynamics, discrete measurements, and optimal estimation», dans *MICCAI*, pp. 744–751. 246

Truesdell, C. et W. Noll. 1965, «The nonlinear field theories of mechanics», dans *Handbuch der Physik*, vol. III/3, édité par S. Flügge, Springer Verlag. 88

Usyk, T. P., R. Mazhari et A. D. McCulloch. 2000, «Effect of laminar orthotropic myofiber architecture on regional stress and strain in the canine left ventricle.», *J. Elasticity*, vol. 61, pp. 143–164. 71

Vallet, B., E. Angelini et A. Laine. 2006, «Variational segmentation framework in prolate spheroidal coordinates for 3d real-time echocardiography», dans *Proceedings of SPIE Volume : 6144, Medical Imaging 2006 : Image Processing,*, pp. 1370–1380. 117, 147

Wesseling, K. 1985, «Method and device for controlling the cuff pressure in measuring the blood pressure in a finger by means of photo-electric plethysmograph», United State Patent. 152

Wiener, N. 1949, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*, The Technology Press of the Massachusetts Institute of Technology,, Cambridge, MA. 8

Wu, M. et A. W. Smyth. 2006, «Application of the unscented Kalman filter for real-time nonlinear structural system identification», *Struct. Contr. Health. Monit.*, vol. 14, no. 7, pp. 971–990. 59, 284, 307

Yosida, K. 1995, *Functional analysis*, Classics in Mathematics, Springer Verlag, Berlin. 184

Zahalac, G. H. 1981, «A distribution-moment approximation for kinetic theories of muscular contraction», *Mathematical Biosciences*, vol. 114, pp. 55–89. 94

Zhang, Q. 2002, «Adaptive observer for mimo linear time varying systems,», *IEEE Trans. Automat. Contr.*, vol. 3, pp. 525–529. 255, 257, 260, 262, 263, 272

Zhang, Q. et A. Clavel. 2001, «Adaptive observer with exponential forgetting factor for linear time varying systems», dans *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, vol. 4, pp. 3886–3891. 254, 257, 276, 306

Zhang, Q. et A. Xu. 2001, «Global adaptive observer for a class of nonlinear systems», dans *Proc. of the 40th IEEE Conf. on Decision and Control*, vol. 4, pp. 3360–3365. 262