Daniel Bouffard-Landry

Algorithme novateur de contrôle d'un miroir déformable à base de ferrofluide

Mémoire présenté

à la Faculté des études supérieures et postdoctorales de l'Université Laval dans le cadre du programme de maîtrise en physique pour l'obtention du grade de Maître ès sciences (M. Sc.)

> Faculté des sciences et génie UNIVERSITÉ LAVAL QUÉBEC

> > 2012

© Daniel Bouffard-Landry, 2012

Résumé

Un nouvel algorithme de contrôle de miroir déformable à base de ferrofluide a été proposé afin de répondre à un besoin spécifique des entreprises de fabrication de lentille de précision. Il s'est avéré efficace pour améliorer la précision du contrôle en boucle ouverte par rapport à la méthode usuelle. Il a été estimé que le miroir utilisé pouvait produire une surface avec une fiabilité à $\lambda/5$, à la condition de porter une attention particulière aux instruments utilisés. Les analyses de reproductivité ont confirmé la faible sensibilité du système aux paramètres potentiellement problématiques et ont permis de déterminer une erreur maximale de 2 µm PV à la surface dans des conditions extrêmes de variations. Il a été idemontré que cette même erreur pouvait pratiquement être annulée en effectuant un entretien régulier avec calibration du miroir. Les termes non linéaires contribuant à une surface donnée ont été identifiés et la technique RIM s'est avérée efficace pour diminuer leur effet autant que désiré. Enfin, les analyses comparatives entre l'algorithme RIM et la méthode standard de contrôle en boucle ouverte ont permis de déterminer une diminution d'un facteur 2 des erreurs de contrôle par rapport aux limites physiques du système.

Abstract

A new control algorithm for a ferrofluid deformable mirror has been proposed to meet a specific need of companies manufacturing precision lens. Improvement of the accuracy of the usual open loop control algorithm was shown. It was estimated that the mirror could be used with the new technique to produce any regular surface with a reliability of λ /5, providing special attention to the instruments specifications. Repeatability analyzes have confirmed the low system sensitivity to problematic parameters and have also determined a maximum surface error of 2 µm PV in extreme conditions. It has been shown that this error could virtually be canceled only by performing regular maintenance including a calibration of the mirror. The nonlinear terms contributing to a given surface have been identified and RIM algorithm has demonstrated its possibility to reduce, as much as desired, their effects. Finally, comparative analyzes between the RIM algorithm and the standard open loop control method led to the conclusion that its usage results in a decrease by a factor 2 of control error relative to the physical limits of the system.

Remerciements

Je tiens à remercier spécifiquement Ermanno Borra, mon directeur de recherche, d'avoir mis à ma disposition tous les outils nécessaires au bon déroulement de ma maîtrise. Je tiens aussi à remercier Simon Thibault, qui possède la capacité incroyable de transformer toute problématique insolvable en une simple embuche facilement contournable et qui, par son seul exemple, réussissait à me stimuler pour travailler encore plus fort. Enfin, j'offre toute ma gratitude à Denis Brousseau, qui fut pour moi une référence tout au long de ce travail en plus d'être l'instigateur de l'idée du RIM. Nos interminables discussions ont certainement été indispensables à mon parcours et son aide, en général, fût grandement appréciée.

Je remercie aussi tous mes collègues directs qui ont pu m'aider dans différentes étapes, notamment Jocelyn Parent, qui fût pour moi l'équivalent d'un mentor. Je remercie aussi Hugues Auger pour ses compétences incroyables en tant que technicien de laboratoire et collaborateur exemplaire dans la recherche. Je remercie aussi Anna M. Ritcey pour tous ses commentaires constructifs souvent très utiles. Une pensée spéciale à Maxime Rochette avec qui j'ai partagé mon bureau et avec qui j'ai pu partager toutes mes difficultés et même mes succès. Il fut d'ailleurs d'une aide exemplaire lors de ma toute dernière expérience de laboratoire. Une pensée particulière va aussi aux individus qui ont contribué à différents niveaux dans cette recherche.

Une pensée particulière s'adresse maintenant à mes amis et ma famille, qui furent souvent une source de réconfort et qui ont fait en sorte que la maîtrise soit beaucoup plus qu'un simple travail. Vous m'avez apporté beaucoup en mettant de la vie dans mes journées. Votre présence seule a réussi à rendre cette étape de ma vie unique et inoubliable.

Enfin, il va de soi que je remercie tous les organismes subventionnaires, principalement le CRSNG et le FQRNT, sans oublier Elcan Optical Technologies de m'avoir offert une expérience unique en entreprise. Je pense spécifiquement à Stan Szapiel et Catherine Greenhalgh et toutes les autres personnes merveilleuses que j'ai eu la chance de côtoyer pendant mon séjour en Ontario.

Aux piliers de ma vie

Paul S. Landry, Raymond Landry et Roger Latreille,

Ces modèles et mentors desquels les valeurs que je porte sont grandement issues

À mes trois confidentes

Adrienne Bouffard, Ginette Savoie et Mylène Chiasson

Vos âmes pures étaient d'un réconfort inouï dans les moments plus difficiles

Tables des matières

₹ésuméi	i
Abstractii	i
Remerciementsiv	V
Dédicace	V
Γables des matièresν	i
Γable des figuresvi	i
L - Introduction	1
2 - Théorie	5
2.0 Commentaires sur la section <i>théorie</i>	5
2.1 Constitution du ferrofluide	5
2.2 Théorie du contrôle	5
2.3 Fonction d'influence, linéarisation et algorithme de contrôle	3
2.4 Mathématiques de contrôle et résiduels	1
2.4.1 Problèmes de moindres carrés (LSP et TLSP)12	2
2.4.2 Base vectorielle et résiduels14	4
2.4.3 Complément d'information	7
2.5 Sur la non-linéarité du système1	3
2.5.1 Sources externes	3
2.5.2 Calcul de la déformation de la surface	C
2.5.3 La portion non linéaire de l'équation 2.22	4
2.6 Algorithmes standards, hypothèses principales et RIM 2!	5
2.6.1 Algorithmes standards2!	5
2.6.2 Algorithme RIM	7
2.6.3 Commentaire final	9

3 - Montage et méthodologie
3.1 Montage
3.1.1 Fonctionnalités informatiques (voltages vs courants)
3.1.2 Circuit électrique 32
3.1.3 Le miroir est ses composantes
3.1.4 Montage optique
3.1.5 Analyseur de front d'onde37
3.1.6 Interféromètre Zygo et MetroPro 38
3.2 Méthodologie
3.2.1 Répétabilité des mesures 40
3.2.2 Termes non linéaires
3.2.2 a) Test de linéarité 42
3.2.2 b) Test de couplage 42
3.2.4 Test RIM

4	- Résultats et analyse	44
	4.1 Répétabilité des mesures	44
	4.1.1 Tolérancement	44
	4.1.2 Tests temporels	47
	4.2 Test de linéarité	51
	4.3 Test couplage	54
	4.4 Test RIM	58

5 - Conclusion	6	2
----------------	---	---

Annexe 1	65
Table des Acronymes	66
Bibliographie	67

Table des figures

Fig. 2.1 : Représentation schématique des nanoparticules magnétiques entourées	
de leur surfactant	5
Fig. 2.2 : Schéma d'un système en boucle fermée	6
Fig. 2.3 : Photographie montrant une vue aérienne des actuateurs	8
Fig. 2.4 : Forme typique d'une mesure de fonction d'influence	10
Fig. 2.5 : Représentation visuelle d'une FI représentée par une colonne	11
Fig. 2.6 : Représentation graphique de quelques PZ	16
Fig. 2.7 : Algorithme SOL schématisé	25
Fig. 2.8 : Algorithme SCL schématisé	26
Fig. 2.9 : Courbes d'une fonction f(x,c) quelconque montrant qu'une linéarité locale (traits rouges) permettrait de bien la représenter	27
Fig. 2.10 : Algorithme RIM schématisé	30
Fig. 3.1 : Panneau de commande du VI maison permettant de contrôler le FDM	32
Fig. 3.2 : Deux principaux panneaux de commande du logiciel HASO 3.0	33
Fig. 3.3 : Carte PowerDAQ PD2-AO-96-16	33
Fig. 3.4 : Photographie du convertisseur	34
Fig. 3.5 : Photographie du circuit électrique du miroir	34
Fig. 3.6 : Schéma du miroir	34
Fig. 3.7 : Schéma représentant la relation entre les différentes dimensions liées au	
MC	36
Fig. 3.8 : Photographie du montage optique permettant de déterminer la surface du MDF	37
Fig. 3.9 : Image représentant le concept de base du fonctionnement d'un analyseur de type Shack-Hartmann	38
Fig. 3.10 : Configuration de base d'un interféromètre de Fizeau	38
Fig. 4.1 : Graphique montrant la hauteur calculée numériquement de la surface pour des valeurs normale, maximale et minimale de courant en fonction de la distance radiale au centre du miroir	46
Fig. 4.2 : Aperçu de la surface expérimentale générée pour le test temporel	47
Fig. 4.3 : Résiduel RMS expérimentaux de la variation d'une surface en fonction du	
temps pour deux délais de mesure	48
Fig. 4.4 : Régressions polynomiales sur les résiduels RMS expérimentaux en fonction du vieillissement du FF	49

Fig. 4.5 : Coefficients de l'ordre 1 pour chacun des pixels en fonction de leur position à la surface du miroir	52
Fig. 4.6 : Coefficients de l'ordre 2 pour chacun des pixels en fonction de leur position à la surface du miroir	52
Fig. 4.7 : Configuration 1 avec tous les PPV à 1V	54
Fig. 4.8 : Configuration 2 avec les actionneurs 2, 4 et 6 à 2V	55
Fig. 4.9 : Configuration 3 avec les actionneurs 4, 5 et 7 à 2V	55
Fig. 4.10 : Résiduel expérimental entre la mesure couplée selon la configuration 1 et le témoin	56
Fig. 4.11 : Résiduel expérimental entre la mesure couplée selon la configuration 2 et le témoin	56
Fig. 4.12 : Résiduel expérimental entre la mesure couplée selon la configuration 3 et le témoin	56
Fig. 4.13 : Cible théorique utilisée pour comparer les algorithmes RIM, SCL et SOL	59
Fig. 4.14 : Résiduels expérimentaux obtenus par la méthode SCL et RIM	60

ix

1 - Introduction

Les développements de systèmes optiques permettent des prouesses technologiques qui ne cessent de repousser les limites des capacités humaines. Suffit de penser à la lunette de Galilée ou à l'invention du laser, nul ne peut douter de l'impact positif de ces inventions sur notre quotidien. Qui plus est, ce domaine est toujours florissant et actuel. Il est évident qu'il s'est segmenté en d'innombrables sous-domaines. La plus simple séparation distingue le domaine de l'éclairement (e.g. un laser) de celui de l'imagerie (e.g. un microscope).

Au fil du temps, les poussées technologiques ont permis de développer des systèmes offrant plus de possibilités tout en étant toujours plus performants. L'avènement de l'ordinateur a pour ainsi dire révolutionné les pratiques, notamment en imagerie, en permettent de simuler des systèmes complets par l'entremise des logiciels de calcul très puissants. Par contre, les éléments composant un système ou un appareil optique sont sensiblement les mêmes que ceux utilisés autrefois, mais additionnés de quelques inventions plus ou moins récentes. En effet, les miroirs et lentilles constituent encore aujourd'hui les éléments de base de tout système d'imagerie. De plus, ils partagent presque tous une caractéristique commune : ils sont conçus pour accomplir une tâche précise à un endroit précis d'un système. Par contre, quand le système est conçu de façon à permettre de faire varier une caractéristique donnée, on peut alors parler d'une *plage dynamique*, défini comme l'ensemble des possibilités de configuration (souvent un intervalle continu, mais fini). Un zoom mécanique variable et un ajustement de mise au point (focus) sont deux exemples très simples de composantes qui permettent à un objectif de caméra d'avoir une certaine plage dynamique.

La grande majorité des systèmes dynamiques sont constitués de pièces solides dont certaines sont amovibles. Ces systèmes, aussi complexes soient-ils, permettent uniquement de modifier les positionnements relatifs (de façon contrôlée) des différentes composantes. L'optique adaptative (AO) est une branche de l'optique qui se trouve au croisement des systèmes dynamiques et du contrôle numérique. Elle se distingue en misant plutôt sur la possibilité de faire varier la géométrie de certaines surfaces plutôt que leur position. Elle inclut aussi un autre aspect majeur, elle peut être utilisée de façon à « s'adapter » à un signal d'entrée variable de façon à contrôler une ou plusieurs caractéristiques du signal de sortie. Par exemple, un dispositif mécanique permettant de stabiliser des images prises à partir d'une caméra vidéo dans une voiture en mouvement serait considéré comme actif ou adaptatif, dépendamment de ses caractéristiques particulières. Cette double nomenclature est probablement un héritage historique, puisque et la frontière entre les deux définitions est floue. La nuance s'explique grossièrement par la vitesse à laquelle surviennent les variations. Les systèmes dits actifs géreront des variations « lentes », survenant à des intervalles de guelgues secondes, tandis que les systèmes d'AO feront des corrections « rapides », plusieurs fois par seconde. Le terme AO sera cependant utilisé indistinctement tout au long de ce mémoire, puisque le nom en particulier ne

Introduction

change rien à la nature des composantes ou leur utilité dans un système et ne peut apporter aucune confusion majeure.

Un survol historique des événements marquants en AO peut certainement aider à comprendre le contexte et les problématiques actuels et donc la pertinence de cette recherche. Il s'agit d'une brève histoire se déroulant sur à peine plus d'un demi-siècle, dont l'origine serait due à des besoins militaires dans les années 50. On peut retracer ce qui semble être la première proposition sur le sujet en 1953 par Babcock [1]. Cependant, le premier système complet fonctionnel n'a vu le jour que dans les années 80, étant utilisé pour corriger la turbulence atmosphérique nuisant à la qualité d'image obtenue par un télescope. D'ailleurs, l'optique adaptative a été un outil principalement utilisé en astrophysique, malgré plusieurs percées dans une multitude de domaines hétéroclites. Certaines publications font une revue historique plus exhaustive du domaine [2, 3] et plusieurs manuels de référence traitent d'optique adaptative d'un point de vue plus théorique [4, 5, 6, 7].

L'utilisation des composantes optiques adaptatives est restée assez restreinte au fil des ans pour plusieurs raisons. Tout d'abord, ces composantes ont été développées sous forme de miroir et ne pouvaient donc servir qu'en réflexion dans un montage. Il existe tout de même des exemples de composantes pouvant être utilisées en transmission [8]. Dans tous les cas, le système utilisé essaie tant bien que mal d'optimiser trois principaux facteurs : la vitesse, l'amplitude et le résiduel. Par vitesse, on sous-entend la fréquence maximale à laquelle un tel miroir peut modifier la forme de sa surface de façon contrôlée. Par amplitude, on sous-entend l'amplitude caractéristique d'une déformation pouvant être générée. La plupart des miroirs déformables ne peuvent produire des déformations supérieures à quelques micromètres, ce qui constitue d'ailleurs une des grandes faiblesses de cette technologie. Enfin, on définit comme résiduel tout ce qui traite de la différence entre une surface arbitraire désirée et la forme réelle produite. Celui-ci dépend évidemment de plusieurs facteurs comme la surface ciblée en elle-même, la géométrie et la nature du système utilisé et l'algorithme de contrôle. D'autres problèmes peuvent aussi être soulignés par rapport aux dimensions physiques d'un système désiré qui surviennent quand on tente de le reproduire à grande échelle ou de le miniaturiser.

Comme toute chose, plus un appareil offre de possibilités et plus il est abordable, plus il est probable que son utilisation se diversifie et trouve de nouveaux créneaux insoupçonnés. Malheureusement, la plupart des miroirs déformables étaient créés spécifiquement pour les télescopes tout en étant loin d'être abordables. Ceux-ci étaient constitués d'une pièce de verre qu'on déformait par points de pression afin de permettre des corrections d'aberrations de faible amplitude, avec des vitesses de quelques dizaines de kHz et des résiduels inférieurs à une fraction de longueur d'onde. Ces miroirs étaient donc limités dans les possibilités de développement : une pièce de verre ordinaire n'est pas réputée pour être déformable sur plusieurs millimètres. Les miroirs liquides, d'abord développés en raison de leur faible coût de fabrication [9], ont permis d'ouvrir la voie à de nouveaux types de miroirs offrant des possibilités uniques [10], comme la possibilité de générer des déformations de grande amplitude. Le miroir déformable à base de ferrofluide (FDM), initialement proposé par Laird et al. [11], permet ces grandes amplitudes au

2

Introduction

détriment sa vitesse d'opération. Par contre, plusieurs résultats démontrent qu'il serait envisageable d'amoindrir cet impact par différentes stratégies de contrôle [12]. Quoi qu'il en soit, plusieurs applications peuvent être envisagées avec une fréquence d'opération quasi nulle. Imaginons, par exemple, une expérience qui tenterait d'évaluer l'impact de la variation du rayon de courbure d'un miroir sur un paramètre donné. Il serait alors possible de faire l'expérience en prenant une série de mesure en avec différents miroirs de bonne qualité ayant des rayons de courbure légèrement différents. Dans ce contexte, il est clair qu'entre chaque mesure, le miroir en place doit être retiré, puis remplacé par un autre, suite auquel un alignement complet doit être effectué. D'un autre côté, il serait aussi possible d'utiliser un seul miroir déformable sans jamais devoir modifier le montage en cours d'expérience. De plus, contrairement à un miroir conventionnel, il est polyvalent. Autrement dit, il offre une plus grande possibilité de réutilisation dans un laboratoire. D'ailleurs, étant donné les amplitudes qu'il peut générer [10], il offre toute la flexibilité nécessaire au remplacement d'une grande majorité de composants optiques. L'industrie de la fabrication d'optiques de précision commence à voir le potentiel de ce genre de technologie. En effet, chaque nouvelle production doit être vérifiée par un contrôle qualité. La calibration des machines est aussi un processus long et couteux. Dans les deux cas, les outils permettant ce genre de test sont très dispendieux et très peu flexibles. Un exemple est le Computer Generated Hologram (CGH). Il est utilisé en configuration de « null testing » [13], où l'idée est d'annuler l'effet d'une composante optique à caractériser à l'aide d'une composante de référence de grande précision, mais aussi très couteux. Toute différence avec le « null » attendu sera donc attribuable à la composante sous examen. Ainsi, chaque nouveau design se voit décerner son propre CGH, dont les possibilités de réutilisation sont faibles. Les coûts associés au contrôle de la qualité ou à la caractérisation sont donc très importants et une alternative plus souple serait certainement très bénéfique.

Cette attente de l'industrie a donc motivé l'ensemble de cette recherche. La vitesse (ou fréquence d'opération) a donc été exclue des considérations, puisqu'elle ne représente pas une caractéristique recherchée dans ce cas précis. Toute l'attention était tournée vers la précision de la génération d'une surface au miroir déformable. Il est vite devenu évident que les algorithmes existants ne convenaient pas au besoin particulier. Le but ultime est d'obtenir un instrument fiable qui fournit une surface sur demande avec la meilleure précision possible. Fiable est utilisé au sens où l'on ne doit pas avoir recours à un instrument de mesure pour vérifier si la surface est générée adéquatement. Or, il est très difficile de prédire avec justesse la surface générée par un miroir déformable. Certaines techniques donnent de bonnes approximations sans toutefois assurer la précision nécessaire pour les applications industrielles si elles sont générées sans l'aide d'un appareil de mesure. La question était donc de savoir s'il était possible de trouver une nouvelle technique de contrôle, un algorithme qui permettrait d'améliorer l'exactitude d'une surface spécifique demandée au miroir. Il est important de mentionner aussi que les tests devraient porter sur les surfaces typiquement utilisées par l'industrie. Ces travaux sont donc restreints à certains types de surfaces. Celles-ci auront comme caractéristique commune de n'avoir aucune discontinuité ou variation brusque de hauteur. On dira qu'elles ont de faibles fréquences spatiales. Cette caractéristique sera développée plus clairement dans la section théorie.

Introduction

Le but premier est donc le suivant : démontrer la possibilité de contrôler précisément la surface d'un FDM sans dépendre d'un instrument de mesure en temps réel. Il importe donc de développer un nouvel algorithme et de comparer ses performances avec les méthodes existantes. Cette recherche se restreint aux fréquences temporelles et spatiales très faibles.

Comme on peut s'en douter, un nouvel algorithme a effectivement été développé et le présent mémoire relate les caractéristiques particulières de celui-ci. Pour faire référence à cette technique particulière, l'acronyme RIM (Reconstructed Influence Matrix) est utilisé tout au long de cette étude et les raisons motivant ce choix deviendront plus claires au fil de la lecture. Le document comprend quatre autres sections. Plusieurs éléments théoriques sont expliqués dans la section 2 pour permettre d'expliquer adéquatement la stratégie développée. La section suivante est naturellement dédiée à l'explication de la méthodologie utilisée dans les différentes étapes du projet qui s'est déroulé en trois phases : la première consistait à mener certaines expériences pour confirmer nos hypothèses sur les sources d'incertitudes pour les algorithmes standards. Ces hypothèses sont toutes expliquées à la section théorie du présent document. S'en est suivi une phase de simulations, afin de démontrer la faisabilité théorique du projet. Enfin, toute la viabilité de la technique ne pouvait être confirmée autrement que par des tests empiriques, forçant du même coup la tenue d'expériences en laboratoire. Les différents instruments et éléments du montage y sont donc expliqués. La quatrième section est dédiée à la présentation des divers résultats et aux analyses qui y sont rattachées. Ces deux aspects sont présentés ensemble plutôt que dans des sections distinctes pour simplifier et rendre plus fluide la lecture et faciliter la compréhension. Pour clore le tout, une section finale rapporte les différents éléments pouvant être conclus des différents travaux menés lors de cette recherche.

2 - Théorie

2.0 Commentaires sur la section théorie

Cette section est consacrée à l'établissement des définitions et à l'explication des éléments théoriques qui seront utilisés tout au long de ce mémoire. Bien évidemment, l'idée ici n'est pas de faire de preuves rigoureuses sur tous les éléments, puisque plusieurs approximations doivent être faites dans certains cas. Cependant, les éléments fournis ont la qualité de bien représenter notre système et d'offrir une vue d'ensemble plutôt fiable de la réalité physique. De plus, par souci de concision, il est apparu plus pertinent d'inclure uniquement certains résultats théoriques plutôt que l'ensemble de la procédure mathématique permettant d'y accéder. Ces éléments sont souvent des équations ou théorèmes bien établis pouvant être référencés à une multitude d'écrits scientifiques. Enfin, cette section inclut uniquement les éléments pertinents à la compréhension de la présente recherche. Plusieurs thèses ou mémoires portent déjà spécifiquement sur certains aspects liés aux FDM. Concrètement, tout ce qui touche la vitesse de contrôle, les équations de transfert ou l'information de nature fréquentielle n'est pas développé.

2.1 Constitution du ferrofluide

Comme le cœur de la recherche est un FDM dont l'élément central est le ferrofluide (FF), il apparaît pertinent de faire une brève description de sa constitution. Un FF est un liquide colloïdal constitué d'un fluide porteur dans lequel sont suspendues des particules magnétiques nanométriques. Ces dernières sont englobées dans une couche de surfactant permettant d'empêcher l'agglomération des particules entre elles, puisque celle-ci génère des forces de Van der Waals significativement plus grandes que les forces magnétiques responsables de l'attraction des particules entre elles. La figure 2.1 montre une représentation imagée des particules magnétiques à l'intérieur d'une couche de surfactant.



Fig. 2.1 : Représentation schématique des nanoparticules magnétiques entourées de leur surfactant

Dans le cas idéal, la concentration en particule est constante, impliquant une évaporation nulle du liquide porteur et l'agglomération inexistante, impliquant une couche de surfactant idéale ne subissant aucune dégradation. En réalité, toutes ces présomptions sont fausses et causent une variation des caractéristiques principales du FF. L'impact de telles variations sera décrit à la section 2.5.1.

2.2 Théorie du contrôle

La théorie du contrôle permet généralement de déterminer une série de procédures à suivre afin d'avoir, comme son nom l'indique, un meilleur contrôle sur un système donné. Règle générale, la théorie du contrôle devient nécessaire quand l'utilisation d'un système devient trop complexe. Pour bien comprendre, on peut simplement prendre l'exemple d'une suspension de voiture : typiquement, seuls un ressort et un fluide sont utilisés pour amortir les chocs causés par la chaussée. En contrepartie, certains systèmes plus efficaces gèrent, en temps réel, le coefficient d'amortissement pour contrer les mêmes chocs et ainsi permettre d'éliminer presque totalement les soubresauts ressentis. Ces technologies sont rarement accessibles au grand public, puisque les coûts associés seraient disproportionnés, mais peuvent convenir à des utilisateurs particuliers qui requièrent absolument de telles caractéristiques. Évidemment, l'utilisation d'un contrôle dépend directement du besoin. Avec l'industrialisation du 19^e siècle, ces besoins, liés au développement d'instruments de toute sorte, sont devenus omniprésents. Toute une sphère de recherche s'est créée dans le domaine et plusieurs livres traitent du sujet [14, 15, 16]. On peut en tirer plusieurs outils pratiques pouvant être utilisés pour un FDM. Cependant, pour le cadre des travaux présenté, il semble peu pertinent de faire une discussion sur les éléments théoriques sous-tendant certains résultats ou définitions. Cette section est plutôt réservée à donner deux définitions simples qui doivent impérativement être énoncées afin de clarifier l'ensemble du texte à suivre.

Un système de contrôle typique est composé d'éléments pouvant donner au contrôleur une certaine rétroaction. On dit alors que le système est opéré en boucle fermée. Il existe cependant une infinité de variantes de telles boucles, mais le principe directeur reste le même : l'erreur de contrôle sur la sortie (output) sert à ajuster l'entrée (entrée) selon un algorithme défini par la méthode spécifique utilisée. La figure 2.2 permet de visualiser un système en boucle fermée de façon schématisée.



Fig. 2.2 : Schéma d'un système en boucle fermée (Les fonctions X, G Y et H sont arbitraires et sont présentés à titre d'exemple)

On peut voir que le schéma contient des fonctions dépendant d'une certaine variable « s ». C'est la conséquence directe de l'utilisation omniprésente de la transformation de Laplace en théorie du contrôle, permettant ultimement de prédire la sortie « S » en fonction de l'entrée « E », soit S(E). Globalement, chaque « boîte » est une fonction de transfert représentant la modification mathématique du signal lors de son « passage » dans la boîte. Selon le schéma de la figure 2.2, X(s) serait la transformée de Laplace du signal d'entrée, G(s) la fonction de transfert du signal, Y(s) la transformée inverse de la sortie et H(s) la fonction représentant la rétroaction. Dans l'espace temporel, les multiplications seraient plutôt des convolutions, soit des intégrales plus ou moins sympathiques, voire impossibles à intégrer analytiquement. Pour éliminer toute confusion, il est important de mettre le présent problème en contexte. La théorie du contrôle sert principalement aux systèmes dynamiques, c'est-à-dire les systèmes pour lesquels on veut déterminer une fonction de sortie (dépendante du temps) liée directement au signal d'entrée (aussi dépendant du temps). Mathématiquement, on écrirait la fonction de sortie comme y(f(t)), où y serait une fonction composée des fonctions x, g et h. Par contre, il a été mentionné que l'aspect temporel (ou fréquentiel) du système n'allait pas être abordé. Notre problème se résume donc à un système pouvant être représenté par la fonction y(x). La théorie du contrôle ne sert donc, dans le présent cas, que de prétexte pour définir certains aspects théoriques, mais pourrait certainement être utilisée dans une autre approche.

Maintenant que la définition de la boucle fermée est donnée, il est aisé de lui opposer une définition concernant les systèmes sans rétroaction. On dit qu'ils sont en boucle ouverte. La représentation d'un tel système serait identique à la figure 2.2, une fois l'élément H(s) enlevé. Il peut sembler impertinent d'avoir des systèmes de *contrôle* en boucle ouverte, mais celle-ci permet notamment de déterminer certaines équations pour simuler des systèmes complexes (comme la fonction de transfert) ou très difficilement représentables mathématiquement (dans l'espace temporel).

Les définitions données jusqu'à maintenant sont très générales et ne comportent aucun élément spécifique des FDM. De plus, elles sont issues de la théorie du contrôle où la variable temporelle est omniprésente, puisqu'elle traite de l'évolution du système dans le temps. Bien que les équations régissant l'évolution temporelle soient utiles dans certaines circonstances, elles semblent moins appropriées dans des circonstances où l'intérêt est concentré sur l'état final uniquement. Il est donc nécessaire d'utiliser des définitions plus appropriées au présent problème, dans lesquelles les termes « ouvert » et « fermé » sont liés uniquement à la présence (ou l'absence) de rétroaction. On définit *Standard Open Loop* (SOL) et *Standard Closed Loop* (SCL) comme des méthodes numériques (ou algorithmes) permettant le contrôle d'un paramètre donné, à savoir la surface du miroir. Il faut comprendre que le terme standard est utilisé dans un sens utile plus que formel. On doit l'interpréter comme « techniques les plus couramment utilisées ». Il ne fait aucunement allusion au schéma présenté en figure 2.2.

2.3 Fonction d'influence, linéarisation et algorithme de contrôle

La section précédente donnait plusieurs éléments théoriques très généraux qui n'étaient jamais liés au miroir utilisé pour cette recherche. Le mot algorithme a été mentionné à plusieurs reprises sans pour autant en décrire les rouages. Afin d'avoir une explication claire des différentes particularités de l'algorithme en question, il est important d'y aller par étape sous peine de perdre énormément de clarté. Ainsi, la première étape qui vient intuitivement à l'esprit est l'explication du processus de génération d'une surface et la détermination des outils disponibles pour se faire. Tout d'abord, sans rentrer dans les détails du montage, il faut savoir que le miroir est composé de 91 actionneurs uniformément distribués dans une configuration hexagonale (voir figure 2.3). Il arrive parfois que certains actionneurs soient inutilisés. En effet, puisque l'effet d'un actionneur se mesure par réflexion, l'utilisation d'un faisceau lumineux de petit diamètre rend inutiles les actionneurs en périphérie. Les raisons poussant à l'utilisation d'un tel faisceau sont variées, mais la principale est que le système de mesure peut résoudre de bien plus grandes amplitudes au miroir si on restreint la taille de la pupille.



Fig. 2.3: Photographie montrant une vue aérienne des actuateurs (L'interface FF/air est parallèle au circuit imprimé se lequel s'attachent les actionneurs)

Ces actionneurs se trouvent physiquement sous un récipient contenant 1 mm de FF un peu plus large que la région contenant l'ensemble des actionneurs (note : la section 3 est consacrée, entre autres, à l'explication détaillée des éléments du montage). Ces actionneurs sont de simples bobines alimentées par un courant allant de -0,2 A à 0,2 A, générant un champ de quelques Gauss à la surface du FF. Comme il sera expliqué plus loin, ce champ permet de déformer la surface localement au-dessus de l'actionneur en créant une dépression ou une petite bosse de quelques microns en amplitude et de dimensions latérales de l'ordre de la distance entre deux actionneurs voisins, soit quelques millimètres. Une mesure d'une telle déformation est appelée *fonction d'influence* (FI). Lors de la conception du miroir, une composante a été conçue afin de linéariser la réponse d'un actionneur (son fonctionnement ne sera pas expliqué avant la section 2.4), mais il est tout de même possible de se référer à [10, 17]. Faisons pour l'instant l'hypothèse que le système est parfaitement linéaire. Cette hypothèse implique premièrement

que l'influence simultanée de deux actionneurs est égale à leur influence individuelle additionnée. Cela implique aussi que si l'on connait l'influence d'un actionneur alimenté par un courant donné, on peut prédire exactement l'influence du même actionneur pour un courant différent en multipliant la fonction d'influence par le ratio des deux courants (un courant 2 fois moins fort produira une fonction d'influence avec exactement 2 fois moins d'amplitude). Ainsi, en déterminant un ensemble de courants $\{I_i\} : \{i_1, i_2, ..., i_N\}$, injectés dans chacun des *N* actionneurs utilisés, on peut prédire la surface du miroir, en faisant une combinaison linéaire des FI mesurées (en supposant les erreurs de mesure nulles) dont les coefficients seraient les ratios de courant prémentionnés. Mathématiquement, si \mathcal{G}_i représente une FI mesurée avec un courant injecté M_i, la combinaison linéaire sera donnée par :

$$\sum_{i=1}^{N} \frac{I_i}{M_i} * \mathcal{G}_i$$
(2.1)

On peut facilement alléger le tout en divisant les FI mesurées par le courant auxquels elles correspondent et les identifier par le symbole $\mathcal{F}_i = \mathcal{Q}_i / M_i$. On dit de ces nouvelles FI qu'elles sont *normalisées*. D'ailleurs, toute référence subséquente aux FI dans ce document sous-entendra celles qui sont normalisées (sauf avis contraire). Ainsi, la combinaison linéaire se réduit à la somme des FI multipliés par leur courant respectif.

$$\sum_{i=1}^{N} I_i * \mathcal{F}_i \tag{2.2}$$

Enfin, il faut savoir deux choses importantes concernant les fonctions d'influence. La première est que, pour des besoins de simulation, il faut définir une fonction d'influence comme étant une fonction gaussienne en trois dimensions, centrée sur l'actionneur selon :

$$\mathcal{F}_i = A * e^{-\frac{(R-R_0)^2}{2\sigma^2}},\tag{2.3}$$

dont les paramètres A (amplitude), σ (déviation standard) et R₀ (centre de la FI) peuvent être déterminés expérimentalement. Utiliser ces fonctions donne une excellente approximation du système (voir figure 2.4) et a l'avantage d'être une fonction déterminée en tout point de l'espace.

Par contre, il est en général plus simple d'utiliser directement la mesure plutôt que la fonction donnée en 2.3, puisque les paramètres doivent être extirpés à partir d'un « fit » gaussien. Cette opération requiert du temps en plus de représenter un résultat contenant au moins autant d'incertitudes que la mesure elle-même. De plus, cette fonction semble varier au fil du temps au fur et à mesure que le FF se dégrade, il faudrait donc continuellement recommencer le processus de calibration. Ainsi, l'utilisation de la mesure semble être le choix le plus simple, mais implique un nombre fini de points et donc des FI discrétisées. On peut cependant gérer facilement et simplement cette information par le biais de matrices de données. En effet, les fonctions d'influence peuvent être évaluées par plusieurs outils qui ont tous une caractéristique commune :

la mesure est une matrice finie de points. Nous n'avons donc pas une fonction continue, mais bien un ensemble de mesures discrètes. Par simplicité, on nomme ces points des pixels, indépendamment du système utilisé. Il va de soi que, plus l'instrument fournit un grand nombre de pixels, plus la FI s'approche de la fonction réelle continue représentant la réalité. La contrainte de travailler dans un espace pixélisé, les équations doivent être modifiées. Il est alors possible d'utiliser une représentation matricielle pour les fonctions d'influence.



Fig. 2.4 : Forme typique d'une mesure de fonction d'influence

Afin de bien comprendre cette nouvelle notation, supposons qu'un instrument fournit des mesures selon une matrice de dimensions l-par-k, pour un total de m=l·k pixels. Il est aisé de numéroter individuellement chaque pixel de 1 à m. La linéarité du système nous permet de prétendre que la valeur d'amplitude de chaque pixel est donnée par la combinaison linéaire des FI correspondant au même pixel. Autrement dit, la valeur « W_j » du j^e pixel est donnée par la combinaison linéaire des FI combinaison des N FI par :

$$W_{j} = \sum_{i=1}^{N} I_{i} * F_{j}^{j}, \qquad (2.4)$$

où l_i est le courant injecté dans le i^e actionneur et F^j_i est la valeur d'amplitude de la fonction d'influence du i^e actionneur, mesurée au j^e pixel. Définissant une matrice « H », contenant toutes les Fl, on peut écrire W=H*I où H est une matrice m-par-N (m étant le nombre de pixels) et où l est un vecteur colonne de dimensions N-par-1. On en déduit que les mesures des fonctions d'influence doivent être redimensionnées en des vecteurs colonne selon un isomorphisme¹ quelconque et rangées dans un ordre prédéterminé. La façon la plus simple est de concaténer les k colonnes d'une mesure en une seule colonne et d'ordonner celles-ci dans H par un numéro unique qu'on assigne à chaque actionneur. La figure 2.5 montre visuellement ce qui vient d'être expliqué.

¹ Un isomorphisme est une transformation bijective reliant chacun des éléments d'un ensemble à un seul élément d'un second ensemble et qui permet de retourner au même élément de l'ensemble initial (et lui seul) par la transformation inverse.

Au final, on peut déterminer toute surface W à partir d'un vecteur courant I, à condition que la matrice H ait été adéquatement construite, par un simple produit matriciel :

$$W = H * I \tag{2.5}$$

Cette équation toute simple n'est cependant pas d'une grande utilité sous cette forme. En effet, le vrai problème est plutôt l'inverse de cette équation. Dans la plupart des cas, il ne s'agit pas de déterminer la surface produite par un ensemble de courant, puisque de toute façon elle peut aisément être mesurée. Il faut plutôt trouver une façon d'inverser l'équation pour obtenir une façon de calculer un vecteur courant permettant de générer une surface cible donnée. Cette inversion n'est pas aussi simple qu'il y parait et nécessite quelques outils mathématiques pour y parvenir. L'explication de ceux-ci constitue le corps de la prochaine section.



Fig. 2.5 : Représentation visuelle d'une FI représentée par une colonne

2.4 Mathématiques de contrôle et résiduels

Comme il a été fait mention à la fin de la dernière section, la vraie problématique est de trouver une méthode pour calculer un courant qui génère une surface cible W. Il faut donc isoler « I » dans l'équation 2.5. Cependant, il ne suffit pas de diviser par H de chaque côté de l'égalité ou de multiplier par l'inverse. En effet, en général H est une matrice (m-par-n) où m (le nombre de pixels²) est significativement plus grand que n (le nombre d'actionneurs, soit 91 au maximum). Par contre, on peut remarquer que l'équation est algébriquement équivalente à un système de m équations à n variables. Bien qu'on n'ait aucune certitude que les m équations soient toutes indépendantes, on peut tout de même statuer qu'en général, le système est surdéterminé et qu'il n'existe aucune solution à 2.5.

² Varie selon l'instrument mais est de l'ordre de quelques milliers à quelques centaines de milliers.

2.4.1 Problèmes de moindres carrés (LSP et TLSP)

C'est ici qu'entrent en jeu les problèmes de moindres carrés. Les acronymes LSP et TLSP viennent de l'anglais *Least Square Problem* et *Total Lest Square Problem*. Le premier s'énonce comme suit :

Soit la matrice « H » (m-par-n) avec m > n et des vecteurs « I » (n-par-1) et « W » (m-par-1). On cherche la solution « I » qui minimise

$$|H * I - W||,$$
 (2.6)

où l'opérateur $\|\cdot\|$ est la longueur euclidienne (norme) habituelle : $\|\xi\| = \xi^T \xi$. Ainsi, on voit que $\|H * I - W\| \ge 0$ et, comme mentionné préalablement, il n'existe en général aucune solution à l'équation 2.5 qui permettrait d'obtenir $\|H * I - W\| = 0$. Par contre, on peut poser un problème équivalent en supposant qu'il existe une erreur δ sur le vecteur W, pouvant être minimisée et telle qu'il existe une solution « l' » à l'équation :

$$H * I' = W + \delta \tag{2.7}$$

On dit alors que $W + \delta \in range(H)$. Le fait de rajouter une erreur à W peut sembler être une mauvaise idée, puisqu'en général W est une cible *idéale*, donc une valeur *posée*. En contrepartie, la matrice H est construite à partir de mesures expérimentales qui elles, doivent clairement contenir des erreurs. Le problème devrait donc être de trouver une solution l' dans

$$(H + \varepsilon) * I' = W, \tag{2.8}$$

tel que ε est minimisé. On se rend vite compte que cette approche est mauvaise puisqu'il est impossible de reproduire n'importe quel W par combinaison linéaire des N FI et ce, indépendamment des erreurs de mesures dans celles-ci, étant donné qu'elles ne forment pas une base vectorielle. Ainsi, l'équation 2.8 imposerait un « ε » qui ne servirait pas uniquement à considérer les erreurs de mesure des FI, mais serait plutôt utilisée pour les *déformer* afin de permettre une solution l'. Considérant donc que les erreurs de mesures sont généralement très faibles, l'équation 2.7 semble être une bonne approche pour trouver « l' » si on ne considère pas « δ » comme une incertitude, mais comme une légère perturbation d'une équation sans solution. Enfin, comme mentionné plus haut, la solution au problème exposé dans les équations 2.6 ou 2.7 est très bien connue et il est démontré qu'il existe un « l' » unique qui minimise l'équation 2.6 si rang(H) = n, mais qu'il en existe une infinité dans le cas rang(H) < n. Par contre, il n'existe toujours qu'une seule solution (donnée par l'équation 2.9) qui minimise à la fois l'équation 2.6 et sa propre norme.

$$I' = pinv(H) * W \tag{2.9}$$

Ici, pinv(H) est le pseudo-inverse (PI) de la matrice H (m-par-n). La compréhension approfondie du PI peut être utile dans certaines utilisations. En effet, le cœur même de cette solution repose sur une décomposition aux valeurs singulières (SVD). La SVD s'énonce comme suit :

Un nombre réel σ est valeur singulière d'une matrice M (opérant sur un champ K et de dimensions n-par-m) si et seulement si il existe un vecteur unitaire u dans K^m et v dans Kⁿ tel que $Mu = \sigma v$ et $M^* v = \sigma u$. Dans toute SVD, $M = U \Sigma V^*$ où Σ est la matrice diagonale contenant les « p » valeurs singulières³, en comptant les dégénéressances. Cette matrice est unique à une permutation près. Les matrices U et V sont les matrices contenant les vecteurs singuliers (dits gauche et droit) et sont unitaires. Ces éléments peuvent être vus plus en détail dans [18]. L'information utile provient des valeurs singulières en elles-mêmes. Si l'ordre de grandeur de la plus grande valeur singulière est très différent de celui de la valeur minimale, la matrice sera définie comme mal conditionnée et mènera certainement à des erreurs computationnelles significatives. Prenons un exemple simple : on sait qu'un ordinateur traite les opérations mathématiques sous forme binaire et que chaque valeur est enregistrée sur un nombre fini de bits. Ainsi, la simple opération d'additionner 2⁴⁰+2⁻²³ donnera simplement 2⁴⁰ puisque la mémoire est forcée d'arrondir tout ce qui dépasse le nombre maximal de bits. En poussant d'avantage l'idée, il est aisé sur papier d'établir que ((10⁴⁰+10⁻²³)-10⁴⁰)*10²³=1, alors qu'un ordinateur donnerait 0. Pour déterminer si une matrice est mal conditionnée, il suffit de comparer les valeurs singulières maximale et minimale. Plus leur ratio est près de l'unité, meilleur est le « conditionnement » de la matrice. Concrètement, les valeurs singulières d'une matrice de FI sont associées aux modes fondamentaux du miroir. L'ensemble de ces modes forme une base orthonormée sur laquelle on peut décomposer toute surface réalisable, mais où les valeurs singulières servent de valeurs de pondération. Ainsi, une valeur singulière trop faible cause non seulement des erreurs computationnelles, mais indique aussi que la présence du mode fondamental associé sera toujours négligeable. Ces informations peuvent être utilisées pour manuellement imposer une valeur singulière nulle à ces modes de façon à ne garder que les éléments significatifs. De plus, il n'est généralement pas surprenant de constater que les valeurs singulières élevées sont liées à des modes fondamentaux de faible fréquence spatiale et inversement pour les faibles valeurs. On peut donc aisément évaluer la fréquence spatiale de coupure du système via cette décomposition. Inversement, un des modes fondamentaux sera un terme de piston, impossible à mesurer avec un analyseur de front d'onde. Ce mode sera donc généralement filtré aussi.

Pour plus d'information sur les éléments théoriques mentionnés, il est possible de se référer à [18, 19]. On peut donc revenir au cœur du problème : le PI. Malgré tous les éléments apportés, le PI se résume grossièrement à une matrice très similaire à une matrice inverse. En effet, dans le cas où toutes les valeurs singulières sont non nulles et où la matrice est bien conditionnée, ce qui est généralement les cas, pinv(H)*H=I_n, où I_n est la matrice identité (n-par-n) permettant ainsi d'isoler le vecteur courant dans l'équation 2.5. L'analogie s'arrête ici puisque H'*H \neq I si au moins une valeur singulière est nulle. La solution reste néanmoins celle de l'équation 2.9, représentant l'algorithme SOL.

Il existe un problème plus général que le LSP qui peut s'appliquer à la situation actuelle, à savoir le TLSP. L'idée est la même que pour le LSP, mais les erreurs sont considérées à la fois sur W

³ Ici, p = min(m, n).

et H. De façon non formelle, on peut combiner les équations 2.7 et 2.8 pour chercher une solution l' qui minimise δ et ϵ dans :

$$(H + \varepsilon) * I' = W + \delta$$
(2.10)

Ce problème n'est cependant pas à l'abri des problèmes mentionnés dans le paragraphe suivant l'équation 2.8. Malgré tout, l'énoncé formel devient :

$$\min_{\varepsilon,\delta} \|(\varepsilon|\delta)\|_F \tag{2.11}$$

assujetti à :

$$W + \delta \in range(H + \varepsilon) \tag{2.12}$$

Il faut « lire » à l'équation 2.11 : minimiser la norme de Frobenius de la matrice augmentée⁴ $\epsilon | \delta$. Cette norme se calcule par la somme des carrés de tous les éléments de la matrice. Encore une fois, il existe une solution unique qui minimise à la fois l'équation 2.11 et sa propre norme. Cependant, après plusieurs comparaisons expérimentales entre les deux solutions (menant pratiquement à la même solution) et en conséquence que l'équation 2.9 est plus simple à traiter que la solution des équations 2.11 et 2.12, il a été déterminé d'utiliser la première approche. Pour constater la « complexité » du TLSP, on peut consulter [19].

On peut finalement mentionner que les deux approches présentées peuvent être légèrement plus générales et inclure des matrices de pondération (weighting matrices) qui peuvent permettre de considérer certains actionneurs où certains pixels comme étant plus dominants ou plus importants dans le calcul. Les erreurs sont donc davantage distribuées sur les pixels ou actionneurs « moins significatifs ». Cependant, on ne trouve aucune raison évidente qui pourrait justifier l'utilisation de telles matrices pour le système concerné.

2.4.2 Base vectorielle et résiduels

Il est maintenant temps d'expliquer un peu plus ce qui est entendu par « mesure », les méthodes associées et leur raison d'être. Le but n'est pas de s'attarder sur ce dernier point pour l'instant puisqu'une section entière est réservée pour l'explication du montage et des méthodes. L'idée ici est plutôt de donner les outils mathématiques nécessaires à la compréhension des éléments à venir. Il va de soi qu'un minimum d'information sur le montage doit être connu *a priori* pour établir la base des outils utilisés, puisque ce sont eux qui s'adaptent à la réalité physique et non l'inverse. Le besoin est donc de trouver une méthode de mesure de la surface du miroir. Il va de soi qu'une méthode par contact (e.g. une sonde) est inadéquate sur un liquide. Cependant, il est relativement aisé de mesurer un front d'onde après sa réflexion sur le miroir et déterminer la forme de la surface en analysant la déformation opérée sur celui-ci, considérant sa forme initiale

⁴ La matrice $\varepsilon | \delta$ est la matrice ε à laquelle on rajoute le vecteur δ à la n+1^e colonne.

connue. Il est aisé de démontrer que si le front d'onde initial est plan, la mesure de la forme du front d'onde correspond à deux fois la surface du miroir.

Décrire mathématiquement un front d'onde devient donc une nécessité. De plus, la très grande majorité des montages optiques utilisent des faisceaux de section circulaire. Les coordonnées cylindriques sont donc « naturelles » pour cette description. La réalité physique d'un faisceau impose certaines contraintes qui doivent être considérées dans l'établissement d'une fonction mathématique. En effet, aucune discontinuité ne peut exister dans une fonction qui décrirait parfaitement un front d'onde *réel* à un moment donné. Une telle fonction vectorielle sera donc forcément analytique. Cette constatation peut sembler anodine, mais en fait elle assure l'existence d'une base vectorielle orthonormée unique pouvant décrire tout faisceau. La seule condition est d'effectuer une transformation de similarité pour que la fonction soit décrite sur un disque unitaire⁵ (de rayon = 1). Autrement dit, peu importe le rayon réel du faisceau, il est toujours possible de le normaliser de façon à décrire le front d'onde en fonction d'une variable r allant de 0 à 1.

Cette base vectorielle orthogonale de rayon unitaire nous est donnée par les polynômes de Zernike (PZ). Il existe deux types de PZ, qu'on distingue par la parité de la fonction génératrice. Les PZ pairs sont donnés, selon la convention utilisée par J. C. Wyant [20], par :

$$Z_n^m = R_n^m(r) * \cos(m\theta) \tag{2.13}$$

et les impairs par :

$$Z_n^m = R_n^m(r) * \sin(m\theta) \tag{2.14}$$

où n \ge m, r est la distance radiale et θ l'angle azimutal.

De plus, pour n-m pair :

$$R_n^m = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} r^{(n-2k)}$$
(2.15)

et donne exactement 0 pour n-m impair.

De ces équations, on peut voir plusieurs choses. Premièrement, le paramètre m est le seul à avoir une influence azimutale tandis que n est associé principalement à r. Les coefficients de chaque terme dépendent à la fois de m et n. Enfin, pour un n donné on génère une série polynômes en faisant varier m, mais un polynôme sur deux est nul.

⁵ La même idée est utilisée dans la définition des fonctions trigonométrique définies sur un cercle de rayon unitaire.

Pour vérifier l'orthogonalité de ces polynômes, on peut utiliser :

$$\int_{r,\varphi} Z_n^m Z_{n'}^{m'} r dr d\varphi = \frac{(1+\delta_{m,0})\pi}{2(n+1)} \delta_{m,m'} \delta_{n,n'}$$
(2.16)

où δ est le delta de Dirac. Les valeurs des r et φ sont limitées aux points du disque unitaire. Quelques-uns des PZ sont présentés à la figure 2.6.



Fig. 2.6 : Représentation graphique de quelques PZ

L'élément principal à constater est que plus m et n sont élevés, plus les variations d'amplitude sont rapprochées. On dit alors que les figures présentent des fréquences spatiales élevées. On comprend donc aisément pourquoi un système à 91 actionneurs ne pourra jamais produire le PZ Z_{100}^{56} . De façon plus réaliste, il est aussi aisé de comprendre une des sources principales d'erreur lors de la génération d'une surface arbitraire. Si on la décompose sur la base des PZ et que la contribution des termes de hautes fréquences est importante, la génération sera systématiquement de mauvaise qualité. La seule façon d'améliorer ce point est de reconfigurer le miroir pour augmenter la fréquence spatiale des actionneurs. Mais il va de soi qu'il y a une limite à la miniaturisation et aux capacités computationnelles. Un système « parfait » aurait une infinité d'actionneurs infiniment minces.

Ces définitions mènent directement au concept de résiduel. Celui-ci est défini comme la matrice donnée par la différence entre une surface cible et celle mesurée. On se sert cependant souvent d'outils comme la valeur RMS (Root Mean Square) ou le PV (Peak-to-Valley) de ladite matrice pour définir le résiduel. Ces mesures sont bien évidemment imparfaites et peuvent cacher des biais, mais donnent en général un bon aperçu global de la situation. Ainsi, en analysant les résiduels d'une surface, il faut toujours garder en tête qu'une partie provient de termes à hautes fréquences et que seulement la portion provenant des fréquences faibles peut réellement être améliorée. C'est d'ailleurs pourquoi certains chercheurs utilisent différents filtres fréquentiels sur leurs surfaces pour ne garder que l'information utile, mais ces techniques viennent avec leur lot de complications et leur utilisation ne semble pas justifiée dans le présent contexte.

2.4.3 Complément d'information

La base orthogonale de PZ est un outil utile pour définir la surface du FDM. De plus, il est possible d'évaluer rapidement la « faisabilité » d'une surface donnée en la décomposant sur la base des PZ et en évaluant l'importance de ses composantes à haute fréquence. On pourrait alors croire que plus on diminue la fréquence, plus on diminue l'erreur résiduelle. Cette diminution s'effectue jusqu'à, finalement, complètement disparaitre si on tente de reproduire Z_0^0 . Or, il y a une subtilité à considérer. Chaque actionneur produit une déformation très près d'une gaussienne en trois dimensions. Si on tente de décomposer un tel front d'onde dans la base des PZ, on se rend vite compte qu'il est nécessaire d'utiliser l'ensemble des polynômes. Chaque FI se définit comme :

$$F_{i} = \sum_{i,m,n} A_{m,n}^{i} Z_{m}^{n}$$
(2.17)

où $A_{m,n}^i$ est le coefficient de Z_m^n pour la i^e FI.

Ainsi, s'il existe une combinaison linéaire des FI pour générer Z_0^0 , elle s'écrit :

$$Z_0^0 = \sum_i c_i \sum_{i,m,n} A_{m,n}^i Z_m^n = \sum_{i,m,n} c_i A_{0,0}^i Z_0^0 + \sum_{\substack{i,m,n \\ m,n \neq 0,0}} c_i A_{m,n}^i Z_m^n$$
(2.18)

Il y a donc une double condition à respecter pour reproduire un front d'onde exempt de résiduel :

$$\sum_{i,m,n} c_i A^i_{0,0} = 1 \qquad \text{et} \qquad \sum_{\substack{i,m,n \\ m,n \neq 0,0}} c_i A^i_{m,n} Z^n_m = 0$$
(2.19)

Bien qu'il existe une infinité de solutions à la première condition, aucune d'elle n'est, en général, simultanément solution de la deuxième condition, sauf si tous les $Z_m^n = 0$. Heureusement, les termes d'ordres élevés dans les FI d'un FDM sont souvent d'amplitudes extrêmement faibles, voire nulles. De plus, les besoins de l'industrie sont, comme il était mentionné en introduction, limités à des surfaces ayant des fréquences spatiales relativement basses. Le potentiel est donc élevé dans les circonstances.

Finalement, l'ensemble des éléments concernant les mathématiques de contrôle permet clairement d'établir qu'il existera toujours une limite physique au système, même si toutes les mesures étaient idéales et qu'il n'existait aucune non-linéarité dans le processus réel. De plus, certains tests ont démontré (résultats à être publiés) que l'utilisation des fonctions gaussiennes idéales dans un système parfaitement linéaire, n'incluant aucune erreur de mesure, permet de prédire les résultats obtenus par l'algorithme en boucle fermée avec une précision surprenante. C'est donc dire que cet algorithme donne environ la limite physique du système. Basé sur cette constatation, il semble donc important de comparer le nouvel algorithme avec celui à améliorer, mais aussi de le comparer avec la boucle fermée, afin d'avoir une base commune sur laquelle s'appuyer.

2.5 Sur la non-linéarité du système

Toute la théorie développée jusqu'à présent repose sur la linéarité du système utilisé. Cette dernière a été démontrée sans équivoque dans [10]. Il apparait intéressant de se pencher sur les causes de non-linéarité dans le système, afin de déterminer des méthodes pour les éliminer ou amoindrir leur effet.

2.5.1 Sources externes

On entend par source externe tout apport, force, énergie ou travail transféré par un phénomène ou une chose qui n'est pas partie intégrante du miroir. Quelques causes ont été identifiées et, dans certains cas, des actions ont été prises pour régler ou amoindrir le problème. Ces causes peuvent toutes être classées dans la même catégorie, soit celle des turbulences environnementales. En effet, le système est extrêmement sensible aux vibrations environnantes, aux déplacements d'air dans le laboratoire et principalement aux variations de température⁶. L'impact des vibrations est clairement le plus subtil. Sa présence est surtout notable lorsqu'on utilise un FF de faible viscosité. Dans ce cas précis, le fait de marcher autour du montage, de s'appuyer sur la table optique ou quelque geste équivalent fait systématiquement apparaître du mouvement à la surface du miroir. On peut aisément minimiser cet effet en augmentant la viscosité du FF et en isolant le système des vibrations (e.g. par l'entremise d'une table à suspension pneumatique). Concernant les courants d'air, le seul geste concret qui peut être fait est de tenter de confiner le miroir dans une enceinte sans toutefois bloquer le parcours optique. Ainsi, une boîte de plexiglas a été construite pour contenir le miroir. Seul un trou pour le faisceau est présent et la circulation d'air est énormément amoindrie. Enfin, le principal problème provient de l'échauffement du miroir. En effet, plusieurs bobines sont actives en permanence et dissipent de la chaleur par effet Joules (typiquement quelques Watts). Cette chaleur crée un gradient de température, donc un mouvement dans le fluide (voir [21]), responsable de la déformation de la surface de plusieurs microns. Pour régler une grande partie du problème, il s'est avéré suffisant d'« aérer » le système, c'est-à-dire de laisser passer l'air aux endroits qui semblaient plutôt clos.

⁶ Ces éléments ont été énumérés en ordre croissant d'importance.

Pour amoindrir encore davantage ces effets, il est possible de prendre plusieurs fois la même mesure et d'en faire la moyenne. En effet, puisqu'ils varient dans le temps de façon chaotique selon une statistique semi-aléatoire, moyenner une série de mesures est équivalent à extraire en bonne partie la composante statique du miroir et donc élimine pratiquement toutes les contributions susmentionnées.

Du point de vue statique, d'autres effets sont présents. En effet, la formation d'un ménisque dans le récipient donne une légère courbure à la surface. Celle-ci est extrêmement faible puisque le choix des dimensions du contenant a été fait en connaissance de cause. Il y a aussi le composant mentionné rapidement en section 2.2 qui linéarise la réponse d'un actionneur. On le nomme bobine de Maxwell (MC), en honneur au célèbre James C. Maxwell, puisque son fonctionnement est expliqué dans son traité [22]. Un MC sert théoriquement à produire un champ magnétique de direction et d'amplitudes constantes sur un plan de l'espace (l'utilité d'un tel système est expliquée en détail à la section suivante). Cependant, certains calculs démontrent que le système réel crée bel et bien un champ de direction constante, mais d'amplitude légèrement variable, causant ainsi une déformation statique au repos. La dernière déformation statique notable n'est pas réellement physique ou directement liée à la surface. Elle s'explique par le fait que la mesure de la surface se fait par un moyen détourné utilisant la réflexion d'un front d'onde sur le miroir. Ce même front d'onde doit parcourir un trajet optique, traverser des lentilles et subir une multitude de transformations (le détail du montage sera donné en section 3). Bien que les équipements soient de bonne qualité, ils introduisent tout de même une certaine erreur statique dans les mesures. Heureusement, les déformations statiques sont très faciles à éliminer, à condition que leur amplitude ne vienne pas compromettre le parcours optique. En effet, les systèmes de mesure de front d'onde viennent généralement avec une option pour référencer la surface. Référencer la surface signifie qu'à toute mesure est soustraite une référence, permettant ainsi de voir uniquement la « variation » de la surface, éliminant du même coup toute forme d'erreur statique indésirable. Cet outil peut aussi être utilisé pour faire certaines mesures spécifiques, comme les tests de couplage⁷. En effet, il est possible de référencer la déformation générée par un groupe d'actionneurs et d'analyser le comportement d'un actionneur dans ces nouvelles conditions.

Toutes ces « formalités » servent principalement à isoler l'effet réel des actionneurs sur la surface du miroir afin de permettre une étude précise de leur comportement, menant ainsi à un contrôle accru du système. Dans les faits, si le référencement cause un problème technique au niveau du processus, il est toujours possible d'ajuster les courants pour simultanément générer une surface voulue et rendre plane la surface au repos (par superposition linéaire selon l'hypothèse de linéarité). Encore une fois, l'amplitude à corriger doit être relativement faible, puisque la plage de courants utilisables change et pourrait diminuer la plage dynamique du miroir.

Il y a cependant un dernier élément à considérer pour tenir compte de tous les phénomènes pouvant nuire à l'acuité de génération d'une surface désirée. On peut constater son

⁷ On dit qu'il y a couplage s'il y a présence d'une dépendance inter-actionneurs.

effet seulement sur des périodes de temps supérieures à plusieurs jours. Certains indices laissent croire qu'il pourrait s'agir d'un effet linéaire pouvant donc être inclus dans le modèle, mais aucune preuve formelle n'indique qu'il n'induit pas aussi des variations non linéaires. Il est question ici de l'effet du temps sur le FF. Comme mentionné en section 2.1, on sait de façon certaine que le liquide porteur du FF s'évapore graduellement et qu'il y a formation lente d'agglomérats au fil de la dégradation du surfactant. Ces variations ont un effet notable sur les propriétés du fluide et causent notamment une diminution marquée de l'amplitude des fonctions d'influence. Tout porte à croire que l'effet ne s'arrête pas là. En effet, un simple facteur d'échelle ne change en rien la forme relative d'une fonction d'influence. Cependant, une légère modification sur la forme a été remarquée après comparaison entre fonctions d'influences prises à des temps très éloignés ayant été multipliés par un facteur d'échelle leur donnant le même PV. Voilà pourquoi la grande majorité des expériences se sont déroulées dans les heures suivant la mesure des FI. Dans le cas hypothétique où un produit fini serait développé pour l'industrie, une telle calibration pourrait très aisément être automatisée et prendre tout au plus quelques minutes ne compromettant donc pas la viabilité du projet.

Finalement, concernant les erreurs pouvant provenir d'une modification du montage comme changer une lentille pour une autre équivalente, mais de meilleure qualité, changer complètement le FF ou toute interaction qui ne devrait théoriquement « rien changer » à la configuration initiale -, il a été établi qu'il était plus judicieux de reprendre les mesures de FI et de calibrer de nouveau le système. Certains résultats de simulations qui seront montrés en section 4 montrent que de telles modifications génèrent des erreurs qui sont probablement minimes et sans conséquence. Cependant, puisque les propriétés du FF évoluent dans le temps, une stratégie allant dans le sens de « le plus souvent calibré sera le mieux » permet d'éviter bien des problèmes potentiels et, de surcroit, aide à la minimisation des erreurs de contrôle imputables à la dégradation du FF.

2.5.2 Calcul de la déformation de la surface

On peut tirer plusieurs éléments utiles à l'élaboration de la méthode RIM à partir de certaines équations théoriques liés à la dynamique des fluides. On peut déjà mentionner que celles-ci sont issues d'une fastidieuse démarche mathématique où plusieurs approximations ont été faites et où plusieurs éléments ont été posés. Elles sont donc, tout au plus, une bonne approximation de la réalité. Ces équations ne seront cependant pas utilisées dans l'algorithme final, mais mettent en lumière les éléments non linéaires, lesquels doivent être « compensés ».

La première équation (probablement la plus utile) découle des équations très générales de *Navier-Stokes* sur la dynamique des fluides, tout comme une équation d'électromagnétisme classique découle systématiquement d'une combinaison savante des équations de Maxwell. Par souci de clarté et de concision, seuls les résultats désirés seront fournis, soit trois formes distinctes, mais pratiquement équivalentes de l'équation qui relie la hauteur relative (Δ h) d'un

point de la surface en fonction de la densité de flux magnétique B dans le vide⁸ générée par les différentes bobines, aux points de l'interface FF/air. Pour un développement exhaustif des équations, on peut se référer à [23, 24, 25, 26, 27].

Ainsi, on obtient la relation suivante :

$$\Delta h = \frac{(\mu_r - 1)}{2\mu_r \mu_0 \rho g} \Big[\Big(B_n^2(x, y) - B_n^2(x_o, y_o) \Big) + \mu_r \Big(B_t^2(x, y) - B_t^2(x_o, y_o) \Big) \Big]$$
(2.20)

où μ_r est la permittivité relative du FF, μ_0 la permittivité du vide, ρ la densité massique du FF, g l'accélération gravitationnelle, B_n^2 la composante de densité de flux magnétique normale à l'interface, B_t^2 sa composante tangentielle et (x₀, y₀) un point quelconque de la surface. Il est important de conserver cette notion de hauteur relative, sous peine de faire face à des absurdités. D'ailleurs, on trouve principalement dans la littérature l'équation :

$$h = \frac{(\mu_r - 1)}{2\mu_r \mu_0 \rho g} [B_n^2(x, y) + \mu_r B_t^2(x, y)]$$
(2.21)

Cette équation sous-entend cependant qu'on a soustrait une jauge arbitraire à tous les points, de telle sorte qu'au moins un point de la surface se trouve à une hauteur nulle. Sans cette présomption, un champ uniforme $\vec{B} = cst \cdot \hat{a}_n$ donnerait une hauteur h non nulle à tout le liquide, l'obligeant donc à être en lévitation. Cependant, la seule chose qui nous intéresse de l'équation 2.20 (au sens où on peut en tirer de l'information pertinente) est la portion non constante, soit l'équation 2.21. Pour les mêmes raisons, on pourrait simplement écrire :

$$h \propto [B_n^2(x, y) + \mu_r B_t^2(x, y)],$$
 (2.22)

Les composantes B_n et B_t du champ total \vec{B} ne peut provenir que de trois sources dont les contributions respectives sont simplement additionnées. Une partie du champ est générée par le MC (avec les composantes β_n et β_t), une autre par l'ensemble des actionneurs (de composantes $b_n = \sum_i b_n^i$ et $b_t = \sum_i b_t^i$) et la dernière du champ magnétique terrestre (de composantes T_n et T_t). On peut faire l'expansion du terme de droite de l'expression 2.22 pour obtenir :

$$(\beta_n + T_n)^2 + 2(\beta_n + T_n)(b_n) + (b_n)^2 + \mu_r [(\beta_t + T_t)^2 + 2(\beta_t + T_t)(b_t) + (b_t)^2]$$
(2.23)

⁸ Le terme *densité de flux magnétique* est utilisé pour se distinguer de *l'intensité de champ magnétique* (souvent identifié comme le champ H). Cependant, puisque ce dernier n'apparait dans aucune équation, le champ B sera identifié par l'expression *champ magnétique* tout simplement.

Maintenant, en considérant que $(\beta + T) \gg b$ et que les termes $(\beta_n + T_n)^2$ et $(\beta_t + T_t)^2$ peuvent être référencés⁹, l'équation 2.22 peut simplement se réécrire comme étant :

$$h \propto (\beta_n + T_n)(b_n) + \mu_r(\beta_t + T_t)(b_t)$$
(2.24)

On obtient ainsi une relation linéaire entre la forme de la surface et le champ généré par les actionneurs avec une erreur donnée par $(b_n)^2 + \mu_r (b_t)^2$. De façon relative, cette erreur peut être très faible en comparaison avec la déformation totale, ce qui rend le système « très linéaire ». C'est d'ailleurs un des points principaux démontrés dans [10]. Cependant, dans un contexte de fabrication de lentilles de précision, l'erreur relative a peu de signification. L'erreur absolue est la seule mesure qui compte. C'est pourquoi négliger $(b_n)^2 + \mu_r (b_t)^2$ pose un problème fondamental, étant donné que cette erreur mesurée expérimentalement lors de la génération d'une surface cible est, en général, trop élevée pour les standards de l'industrie.

On peut pousser le développement un peu plus avant puisque le système est réellement contrôlé à partir d'un ensemble de courants, non pas de champs. Il faut donc s'assurer que 2.24 est aussi linéaire en fonction des courants. En fait, la question à se poser est la suivante : est-ce que les composantes d'un champ généré par une bobine (étant donné que toute composante du miroir alimentée par un courant est nécessairement une bobine) suivent une relation $b \propto I$ pour tous les points de l'espace? La réponse à cette question est oui, si on fait la simple supposition qu'une bobine peut être représentée par un ensemble de spires circulaires. Elles sont donc toutes disjointes (un cercle parfait revient nécessairement sur lui-même), faites d'un fil de section infiniment mince (sans quoi il faudrait décomposer ledit fil en une infinité de sous-fils tous collés les uns sur les autres), toutes parcourues par le même courant I et distribuées spatialement de façon à reproduire le plus fidèlement possible la configuration exacte. De cette façon, on peut simplement additionner les contributions individuelles de chacune et mettre I en évidence dans cette somme (la dernière partie de l'énoncé est vraie si et seulement si l'équation analytique d'une spire unique est une fonction linéaire en I). On peut effectivement démontrer (voir [28]) que cette équation est donnée par :

$$\vec{B}(r,z) = \frac{\mu_0 I k}{4\pi \sqrt{ar^3}} [P_r \cdot \hat{a}_r + P_z \cdot \hat{a}_z]$$
(2.25 a)

$$P_r = -(z-h)\left(K - \frac{2-k^2}{2(1-k^2)}E\right)$$
(2.25 b)

$$P_{z} = \left(K + \frac{k^{2}(r+a) - 2r}{2r(1-k^{2})}E\right)$$
(2.25 c)

⁹ Cela est rendu possible puisque, lors d'une utilisation normale du miroir, β_n et β_t sont des paramètres posés et constants générant des déformations statiques pouvant donc être référencés.

$$K(k) = \int_0^{2\pi} \frac{d\alpha}{\sqrt{1 - k^2 \sin^2 \alpha}}$$
 (2.25 d)

$$E(k) = \int_0^{2\pi} \sqrt{1 - k^2 \sin^2 \alpha} \, d\alpha$$
 (2.25 e)

$$k(r,z) = \sqrt{\frac{4ar}{(r+a)^2 + (z-h)^2}}$$
(2.25 f)

Les équations sont données en coordonnées cylindriques et nous permettent de voir rapidement qu'il n'y a aucune composante dans la direction angulaire ni de dépendance sur l'angle (qui serait venu briser la symétrie). Ainsi, toute l'information sur le champ dans l'espace tridimensionnel est contenue dans seulement deux dimensions. Concernant les différents paramètres à l'intérieur des équations (2.25 a) à (2.25 f), tous les symboles préalablement définis ont la même définition pendant que r et z deviennent les coordonnées de l'espace bidimensionnel, h le rayon et a la hauteur de la spire considérée par rapport à l'origine¹⁰ et finalement K et E sont respectivement les intégrales elliptiques exactes des premier et second ordres. Cette série d'équations peut sembler extrêmement compliquée à gérer, mais les seuls éléments requis sont r, z, a et h, en plus d'une façon d'obtenir les valeurs de K et E en fonction de la valeur de k calculée. Ces dernières sont fournies par des tables ou tout logiciel de calculs mathématiques suffisamment avancé (e.g. MatLab). Comme mentionné précédemment, ces équations sont exactes dans les conditions établies et démontrent formellement que le champ varie linéairement avec le courant injecté.

Avant de passer à la section suivante, il peut être important de présenter les ordres de grandeur en jeu dans l'équation 2.23. En utilisant les équations 2.25 a) à f), on peut établir que le terme dominant $A = 2(\beta_n + T_n)(b_n) + 2\mu_r(\beta_t + T_t)(b_t)$ n'est pas véritablement dominant devant $B = (b_n)^2 + \mu_r (b_t)^2$, principalement aux limites d'utilisation d'un actionneur. En effet, en utilisant un seul actionneur avec son courant maximal, on peut calculer la grandeur du ratio B/A, soit |B/A|. Ce qu'on constate est que |B/A| vaut tout de même environ 0,2. De plus, à l'endroit précis où la composante tangentielle du champ est à son maximum, le ratio |B/A| monte autour de 1,7, ce qui démontre que les termes négligés ne sont pas si négligeables après tout. Cependant, il est important de pondérer cette dernière information venant contredire toutes les hypothèses faites jusqu'à maintenant. Premièrement, l'erreur est proportionnelle au carré du facteur de diminution de courant. Ainsi, sur la majeure partie de la plage dynamique, le ratio devient effectivement très faible et négligeable. De plus, le ratio maximal correspond à des points où la valeur de la fonction d'influence est faible. Ainsi, le terme non linéaire agit faiblement sur la forme générale de la fonction d'influence. En effet, en effectuant plutôt le ratio entre B et la valeur maximale de A définie comme A₀, on obtient une erreur inférieure à 15% de la hauteur maximale, et ce, dans le cas le plus extrême. Même si cette valeur semble tout de même élevée, il sera démontré que l'algorithme proposé réussit à outrepasser cette difficulté.

¹⁰ L'origine est positionnée en plein centre du contenant de FF exactement à l'interface FF / air.

2.5.3 La portion non linéaire de l'équation 2.22

À la section précédente, la méthode de linéarisation a été expliquée en détail. Un problème sérieux a été soulevé concernant le terme non linéaire négligé $(b_n)^2 + \mu_r (b_t)^2$. Il semble donc important d'analyser plus en profondeur ce dernier afin de comprendre la nature des contributions non linéaires. On peut partir d'une composante générale b², puisque les indices spécifiques ne changent rien au développement mathématique et que les deux termes sont complètement indépendants.

On a :

$$b^{2} = \left(\sum_{i} b_{i}\right)^{2} = \sum_{i} b_{i}^{2} + 2 \cdot \sum_{i} b_{i} \cdot \left(\sum_{j \neq i} b_{j}\right)$$
(2.26)

Il y a donc deux composantes non linéaires distinctes à considérer. Le premier des deux termes indique que la relation reliant la surface à la variable b_i (ou, de façon équivalente, au courant associé I_i) est quadratique. L'erreur est donc de plus en plus significative au fur et à mesure que le courant injecté est éloigné du courant auquel la FI a été mesurée. On dit de cette erreur que c'est une erreur d'échelle. Aucun choix de courant ne permet de l'éliminer. Seul un courant au centre de la plage permise (donc égal à 0) minimise l'erreur d'échelle moyenne pour un courant aléatoire donné. Cependant, le centre de la plage est aussi l'endroit où le ratio signal/bruit (SNR) est le moins bon. La stratégie est donc généralement, avec la méthode standard, de mesurer les FI autour du quart de la plage de courant accessible et de les diviser par ce même courant pour les normaliser.

Il y a ensuite le second terme non linéaire. Pour un actionneur donné, son effet varie linéairement avec le courant injecté dans celui-ci. Cependant, il dépend aussi de tous les courants injectés dans les autres actionneurs. On dit ainsi que sa réponse est couplée. Maintenant, considérant la distance entre les actionneurs, l'intensité du champ magnétique généré par chacun et les symétries dans le positionnement des actionneurs, on peut émettre l'hypothèse (qui sera vérifiée plus loin) que s'il y a présence décelable de couplage, celui-ci sera probablement limité à l'interaction avec les plus proches voisins (PPV) uniquement¹¹. On peut aussi supposer que l'importance de ce terme non linéaire sera plus faible que l'autre. Enfin, son amplitude peut être approximée en n'utilisant que le courant moyen de couplage, donc indépendamment de la configuration particulière de ceux-ci¹². Ce terme varierait donc aussi linéairement avec ce nouveau paramètre. De plus, le nombre de degrés de liberté du système serait ainsi limité, permettant l'élaboration d'une méthode relativement simple, ne devant considérer que la fonction d'influence, le facteur d'échelle et le paramètre de couplage moyen.

¹¹ L'amplitude du champ créé par un dipôle magnétique va environ comme 1/r³.

¹² Le terme de couplage étant lui-même extrêmement faible, il est probable que la différence serait difficilement détectable entre deux configurations distinctes valeur moyenne égale.

2.6 Algorithmes standards, hypothèses principales et RIM

Tous les morceaux sont maintenant en place pour définir la méthode RIM. Il est important de se souvenir que les résultats obtenus par RIM seront ultimement comparés à ceux obtenus par deux autres méthodes dites « standards ». Il apparaît donc pertinent de décrire brièvement ces deux algorithmes puisqu'ils seront aussi utilisés.

2.6.1 Algorithmes standards

Les deux algorithmes sont intimement liés l'un à l'autre. Comme mentionné préalablement, le SOL n'est qu'une étape particulière d'un algorithme plus complet, le SCL. En guise d'introduction, une analogie avec l'algèbre de base s'impose. Quand on mesure une FI, on le fait à une valeur de courant donnée. On effectue ensuite une opération de normalisation. Par analogie, imaginons une mesure unique « y » qui varie en fonction d'un « x », soit y=y(x). Si la courbe y(x) est linéaire, suivant l'équation y(x) = mx, alors sa mesure « normalisée » (au sens défini préalablement) sera la valeur y(x) mesurée à x = 1 donc y(1). Il est tout aussi valable de mesurer y(x₀) = mx₀ et diviser cette mesure par x₀ pour obtenir y(x₀)/x₀ = m = y(1). C'est qu'en fait, la FI normalisée est en quelque sorte la pente de la droite, qui se calcule généralement par a = $\Delta y/\Delta x$, n'impliquant pas que Δx soit obligatoirement 1. Maintenant, pour déterminer quel paramètre x_c donnera une valeur cible y_c, il suffit de résoudre l'équation y_c = y(1)*x_c. On trouve donc facilement x_c = [y(1)]⁻¹y_c. Il est donc clair que la même logique est cachée derrière la procédure montrée à la section 2.4 menant à : H*I=W => I=pinv(H)*W, considérant que pinv(H) est l'équivalent de H⁻¹. On peut comprendre visuellement la démarche à l'aide de la figure 2.7.



Fig. 2.7 : Algorithme SOL schématisé

Cependant, pour que cette stratégie soit bonne, il faut que les FI soient linéaires et donne une mesure nulle à I=O, puisque y(x) = mx+bm, b=O dans l'analogie. Or, basé sur les éléments de la section précédente, on sait en réalité que la relation entre les FI et les courants n'est pas parfaitement linéaire et que la mesure prise à I=O n'est pas parfaitement nulle puisqu'il y a présence de bruit et d'erreurs multiples (même après un référencement et/ou une moyenne). Le courant I_c ne génèrera pas exactement la surface cible, mais sera une bonne approximation de la valeur recherchée. En fait, en injectant ce courant dans les actionneurs, la mesure correspondante

pourrait être vue comme $W_c + \delta$ où δ serait l'erreur sur la prédiction. Ici s'arrête l'algorithme du SOL. Cet algorithme produira toujours une erreur δ , qui sera d'autant plus grande qu'il y a présence d'erreurs d'échelle et de couplage dans la figure générée. L'erreur δ sera aussi dépendante des erreurs contenues dans les FI et donc dans H.

De son côté, l'algorithme SCL est venu de l'idée d'utiliser δ pour corriger I_c calculé en SOL. En effet, si on suppose qu'il existe un vecteur courant optimisé I_o, I_c peut s'écrire I_c = I_o + I_{\delta}. On peut donc voir dans H*(I_o + I_{\delta}) = W_c + δ que l'erreur sur la surface est en fait une erreur sur le courant. Or, il est très facile de trouver I_{\delta} qui génère δ , du moins une bonne approximation : on utilise encore une fois I_{\delta} = pinv(H)*W_{\delta}. Nous avons donc une meilleure approximation I_{c2} de I_o obtenue par I_{c2} = I_c - I_{\delta}. Encore une fois, I_{c2} ne sera pas exactement I_o, mais plutôt I_o + δ_2 . En supposant que $\delta_{k+1} < \delta_k \forall k$, il sera toujours possible de répéter l'opération jusqu'à la précision souhaitée. On se rend compte qu'en réalité, le courant calculé finit par osciller autour de I_o, mais que l'erreur associée correspond à une erreur fondamentale due à la géométrie du système. On dira alors, par abus de langage, que l'algorithme a convergé. L'algorithme complet présenté à la figure 2.8.



Fig. 2.8 : Algorithme SCL schématisé

Cette procédure n'est pas affectée par les erreurs d'échelle ou de couplage, ni par les erreurs de mesure des FI. Le seul élément pouvant être affecté par ceux-ci est le nombre d'itérations nécessaire à la convergence de la boucle fermée.

2.6.2 Algorithme RIM

À ce point-ci, développer un algorithme améliorant le SOL, mais n'utilisant pas de boucle fermée quelle qu'elle soit, semble être une tâche difficile. Il est clair qu'il importe de tenir compte de l'erreur d'échelle et du couplage. Cependant, toute l'information doit être recueillie dans une seule séance de calibration. De plus, la technique doit être suffisamment simple pour éviter de faire une trop grande quantité de mesures ou même de calculs, puisque l'algorithme, bien que bon sur papier, pourrait être affecté par l'ajout trop important d'erreurs computationnelles. Il ne faut pas oublier que le but ultime est d'améliorer la méthode standard et non de simplement inventer une nouvelle méthode. Enfin, on sait que le classement du data en matrices constitue une approche efficace pour le traitement de données, bien que les opérations requises pour extraire l'information recherchée, à savoir le vecteur courant « optimisé », doivent uniquement être linéaires, en raison de la nature de l'algèbre matricielle et vectorielle.

La méthode RIM semble avoir tout le potentiel requis pour respecter ces conditions. L'idée générale est d'utiliser le LSP (comme le fait si bien l'algorithme SOL) sur une « approximation de linéarité locale ». Autrement dit, s'il est possible de mesurer, *a priori*, une FI pour chaque actionneur dans une condition presque identique à celle requise pour générer la surface, un calcul au voisinage de ces FI devrait permettre de trouver un vecteur courant contenant très peu d'erreurs, puisque la réponse de chaque actionneur est fortement linéaire.

En effet, l'hypothèse que les FI sont mesurées dans les bonnes conditions de couplage et mesurées à un courant qui ne nécessitera pas un grand facteur d'échelle implique que les deux principales sources d'erreur sont pratiquement annulées. L'analogie avec la fonction simple f(x) peut permettre de visualiser ce qui vient d'être expliqué. Supposons donc une fonction un peu plus générale notée f(x,c) qui représenterait toute mesure prise à une certaine valeur du paramètre x, subissant une certaine influence externe (analogue au couplage moyen) c. La figure 2.9 est un exemple imagé montrant la « linéarité locale ».



Fig. 2.9 : Courbes d'une fonction f(x,c) quelconque montrant qu'une linéarité locale (traits rouges) permettrait de bien la représenter
Théorie

L'idée de base est donc posée. Il reste cependant une panoplie de problèmes à surmonter pour parvenir à une méthode viable. Déjà, on peut se demander pourquoi la fonction f(x,c) n'a pas été tracée en graphique tridimensionnel, plutôt que d'être représentée par une superposition de différentes courbes représentant des valeurs discrètes de c. Ce choix vient naturellement avec le LSP. En effet, le problème final doit se réduire à une matrice (H) reliée à un vecteur cible W représentant des amplitudes. Le LSP sert d'outil pour extrapoler (ou interpoler au besoin) la « droite continue » représentant l'amplitude en fonction des courants. Le couplage, pour sa part, ne peut être considéré que par une série de mesures discrètes. Il y a aussi un problème beaucoup plus fondamental à régler : comment faire pour connaître a priori la bonne « localité » à choisir pour effectuer le calcul. En effet, la surface cible ne donne absolument aucun indice sur le vecteur courant recherché et la méthode en boucle fermée a déjà été exclue. En fait, ce qui a été exclu des possibilités n'est pas la boucle en soi ou toute boucle en général, mais le besoin de devoir mesurer la surface pendant l'utilisation du miroir. Il est donc permis de faire un simple calcul du vecteur courant avec la méthode SOL. Comme il a déjà été mentionné, le vecteur ainsi calculé est déjà une excellente approximation du vecteur courant optimisé recherché. On peut ainsi se servir de cette information pour déduire les conditions d'amplitude et de couplage approximatives d'où faire l'extrapolation selon la méthode RIM.

Il ne reste plus qu'à trouver comment se servir de ces valeurs approximatives de courant et des couplages moyens associés (calculables directement) pour développer un algorithme viable pour la suite. En fait, la solution envisagée est de construire, pendant la phase de calibration, une table de valeurs (lookup table ou simplement LT) pour chaque actionneur, contenant les mesures de plusieurs FI mesurées à des intervalles de courants et de conditions de couplage réguliers prédéterminés. On doit cependant s'assurer qu'elle contient un ensemble de mesures normalisées afin de pouvoir réaliser la toute première étape du procédé. Ainsi, si on trouve, par exemple, une valeur de courant de 8.4¹³ et que la LT contient une mesure à 8 et une autre à 9, toutes deux mesurées à une condition de couplage tout près de celle calculée, il suffit de trouver l'équation matricielle passant par ces deux vecteurs et le tour est joué. Pour chaque actionneur, on pourra calculer un vecteur « pente » (équivalent à une FI normalisée contenue dans H) donné par :

$$\bar{a} = \frac{(h^+ - h^-)}{(I^+ - I^-)}$$
(2.27)

Les signes + et – signifient respectivement à la valeur supérieure et inférieure de courant la plus près de celui calculé avec SOL. La « droite » locale ne passe cependant plus nécessairement par 0. La relation pour chaque actionneur a plutôt la forme w = āI + ē. Il faut donc considérer un « vecteur à l'origine » ē qui devra être additionné à āI et ce, pour chaque actionneur. On l'obtient simplement par :

$$\bar{e} = \frac{(h^{-}I^{+} - h^{+}I^{-})}{(I^{+} - I^{-})}$$
(2.28)

¹³ Ces chiffres sont totalement arbitraires et sans unité, ils servent uniquement à l'exemple.

Théorie

Ainsi, en construisant la matrice A (ou de façon équivalente en reconstruisant H, d'où le nom de la technique) avec des colonnes correspondant aux vecteurs ā, ainsi qu'une matrice E de colonnes ē, on peut écrire une équation similaire à 2.5 qui représente le système localement. En effet, on obtient maintenant la surface W par :

$$W = A * I + E * (i) \tag{2.29}$$

Le terme (:) dans 2.29 est un artéfact mathématique permettant d'additionner toutes les colonnes de E entre elles. Il correspond donc à un vecteur de mêmes dimensions que I, ayant ses entrées toutes égales à 1. Ainsi, définissant $\tilde{E} = E * (:)$ et $\tilde{W} = W - \tilde{E}$, l'équation 2.29 devient simplement :

$$\widetilde{W} = A * I \tag{2.30}$$

L'équation 2.30 a donc exactement la forme recherchée et la solution I est, encore une fois, obtenue par le LSP :

$$I = pinv(A) * \widetilde{W} \tag{2.31}$$

La méthode respecte donc toutes les contraintes préalablement mentionnées, et reste relativement simple. La LT nécessaire à l'obtention d'une bonne précision demeure arbitraire, mais on peut s'attendre à ce que, plus elle contiendra de données, plus la méthode sera fiable. Il peut aussi être mentionné que la démarche suivie pour déterminer l'équation 2.27 impliquait de choisir les mesures correspondant au paramètre de couplage le plus près. On pourrait prétendre qu'il serait possible de faire le même genre de procédé qui a permis de découler 2.28, mais dans la dimension du coefficient de couplage. On aurait alors trouvé une « pente » correspondant approximativement à la moyenne de \bar{a}^+ et \bar{a}^- qui serait calculée avec les mesures correspondant respectivement aux paramètres c^+ et c^- , diminuant ainsi davantage l'erreur sur le vecteur courant calculé. Or, cette avenue ne permet pas de trouver une équation équivalente à 2.28 et E devient en quelque sorte arbitraire.

La méthode décrite représente donc la meilleure approche qui a pu être imaginée pour respecter les critères établis au tout début de cette section. La méthodologie détaillée se résume à au schéma de la figure 2.10.

2.6.3 Commentaire final

Pour les besoins de cette recherche, aucune LT complète n'a été construite. Une telle tâche demande un processus automatisé qui n'était pas disponible alors. Seules les mesures utiles ont été mesurées *après* avoir calculé le premier courant par la méthode SOL. Ces mesures étaient prises « comme si » elles avaient été extraites de la LT, ne changeant donc rien à la viabilité des

Théorie

expériences.



Fig. 2.10 : Algorithme RIM schématisé

3 - Montage et méthodologie

3.1 Montage

Cette sous-section est réservée à l'explication du montage utilisé. On y trouve d'abord le détail concernant les différents outils informatiques utilisés, autant au niveau programmes que des codes utilisés. Ensuite, le circuit électrique reliant l'ordinateur aux actionneurs du miroir, ses différentes composantes et leur fonctionnement sont expliqués. S'ensuit une section exclusivement dédiée au MDF, avant d'aboutir sur une autre consacrée au détail du montage optique et enfin, sur une dernière portant sur l'analyseur de front d'onde.

3.1.1 Fonctionnalités informatiques (voltages vs courants)

Le montage est organisé de telle sorte que, une fois tous les ajustements effectués, l'utilisateur n'interagit qu'avec l'ordinateur. En cette matière, seulement deux outils travaillant simultanément sont nécessaires. Un premier contrôle le miroir pendant qu'un autre effectue des acquisitions de fronts d'onde. Si certains calculs doivent être effectués pour déterminer les paramètres de contrôle à partir de certaines mesures de fronts d'onde, un troisième outil, en l'occurrence le logiciel MatLab R2009b, sert d'intermédiaire. Ce dernier n'est cependant pas essentiel au fonctionnement primaire du montage. Un quatrième logiciel contrôle la source de courant alimentant le MC, mais son utilisation est limitée à la portion *ajustements* d'une expérience.

Le miroir est composé de 91 actionneurs (n'étant rien d'autre que des enroulements de fils) devant être contrôlés de façon indépendante. La stratégie adoptée a donc été de se doter d'une PowerDAQ PD2-AO-96-16, soit une carte branchée directement à l'ordinateur avec 96 ports indépendants pouvant fournir une sortie analogique 16 bits sur la plage [-10 V, 10 V]¹⁴. La section 3.1.2 explique comment et pourquoi ces tensions sont transformées en courants, à leur tour amplifiés. Le choix du logiciel de commande s'est arrêté sur LabView, étant apparu comme la meilleure option disponible. Un instrument virtuel (VI) maison a donc été créé, permettant principalement de commander indépendamment les 91 sorties, de charger un vecteur courant à partir d'un document *.txt* enregistré sur l'ordinateur et d'enregistrer dans ce même format

¹⁴ Il peut sembler surprenant de voir une plage en [V] alors que la théorie était développée autour d'un vecteur-courant. Or, le circuit électrique sert exclusivement à transformer la plage [-10V, 10V] en [-200mA, 200mA]. La sortie donnée en [V] correspond donc à une valeur unique de commande pour l'actionneur, donnée en [A]. Le terme *courant* est donc utilisé au sens général de *signal*, indépendamment de l'unité à laquelle il fait spécifiquement référence. Dans ce contexte, un *courant* peut donc indistinctement être donné en [V] ou en [A], mais sa nature sera spécifiée en cas d'ambiguïté majeure.

l'ensemble des valeurs envoyées dans le circuit au moment de la sauvegarde. La figure 3.1 montre à quoi ressemble le panneau de commande.



Fig. 3.1 : Panneau de commande du VI maison permettant de contrôler le FDM

De son côté, la portion acquisition est assurée par le logiciel de contrôle *HASO 3.0* fourni par *Imagine Optics*, fabricant de l'analyseur *HASO HR 44* utilisé¹⁵. Il va de soi qu'il offre une multitude d'options répondant à différents besoins et qu'il est inutile d'en faire l'étalage exhaustif. Il existe, entre autres, une multitude d'options de calibration, de reconstitution de fronts d'onde, d'affichage et de sauvegarde. Certaines d'entre elles seront expliquées à la section 3.1.5. On peut obtenir un aperçu du logiciel à la figure 3.2, montrant les deux principaux panneaux de commande et la variété des options disponibles.

3.1.2 Circuit électrique

Comme mentionné précédemment, mis à part le fait qu'il achemine les signaux de l'ordinateur aux actionneurs, le circuit électrique ne sert qu'à convertir le signal généré de la plage [-10V, 10V] à la plage [-200 mV, 200 mV]. En effet, chaque sortie de la carte est branchée à un convertisseur-amplificateur, qui est à son tour connecté aux bornes d'un actionneur. Il est nécessaire de passer par ce convertisseur, puisque la carte est conçue pour servir de source de tension et opérer à très faibles courants. En contrepartie, les actionneurs nécessitent une source de courant dépassant les spécifications de la carte. En effet, le calibre du fil utilisé peut-

¹⁵ Analyseur de type Shack-Hartmann



Fig. 3.2 : Deux principaux panneaux de commande du logiciel HASO 3.0

fonctionner à des courants de plusieurs dixièmes d'ampères. Pour éviter de les briser ou de les faire surchauffer, il a été décidé de bloquer le courant maximal à 200mA. Physiquement, le circuit est constitué de trois compartiments compacts (la carte, le convertisseur et le miroir) reliés entre eux par des câbles plats (flat-cables). Le convertisseur sera uniquement présenté comme une simple *boîte noire* puisque sa conception a été assurée par un sous-traitant spécialisé. Finalement, pour couvrir entièrement la portion électrique du système, il importe de mentionner brièvement que le MC est branché à sa propre source stabilisée en courant (brièvement mentionnée en section 3.1.1). Ce circuit n'est donc formé que du fil et de la source. Son utilisation est expliquée à la section 3.1.3. Les figures 3.3 à 3.5 montrent les différentes composantes électriques du système.



Fig. 3.3 : Carte PowerDAQ PD2-AO-96-16



Fig. 3.4 : Photographie du convertisseur



Fig. 3.5 : Photographie du circuit électrique du miroir

3.1.3 Le miroir est ses composantes

Le miroir est en fait un assemblage modulaire de plusieurs morceaux. Comme on peut le voir à la figure 3.6, il est constitué d'une plaque, à laquelle sont fixés les actionneurs. En observant la figure 2.3, on comprend que la plaquette en question est un circuit imprimé permettant d'intégrer les 182 fils et 6 connecteurs dans un disque de 10 cm de diamètre. Cet élément est placé sous un contenant suspendu par le MC et à l'intérieur duquel se trouve le FF. Toutes les dimensions ont été calculées pour que la surface du FF et le plan horizontal de symétrie du MC soient confondus.



Fig. 3.6 : Schéma du miroir (modification d'une image tirée de [10])

Montage et méthodologie

Il est à noter que la couche de FF a une épaisseur de 1 mm et que cette épaisseur n'a pas été établie arbitrairement. En effet, une couche trop épaisse de FF donne une surface au repos très instable, attribuable principalement aux courants d'air et à la fluidité du liquide. De plus, augmenter l'épaisseur signifie s'éloigner des actionneurs. L'amplitude des FI en est donc significativement affectée. En contrepartie, une couche trop mince souffre de deux problématiques. D'abord, il se peut que le contenant ne soit pas parfaitement au niveau. Le liquide, pour sa part, le sera certainement. Avec suffisamment de FF, cela ne cause aucun problème puisque seule l'interface est importante pour les déformations (à un certain niveau). Cependant, si la couche est trop mince, il se peut que le fond « ressorte » du FF, causant ainsi un « trou » en périphérie du front d'onde. L'autre phénomène néfaste est ce qui est couramment appelé « interaction avec le fond ». En effet, si la couche est trop mince, des forces initialement négligeables des équations de Navier-Stokes doivent dorénavant être considérées, changeant ainsi drastiquement la dynamique du système. Le choix de l'épaisseur de 1 mm s'impose donc comme le meilleur compromis possible et est le résultat d'une série d'essais et d'erreurs.

Au niveau des actionneurs, la configuration et les dimensions ont aussi été longuement réfléchies. Encore une fois, si on se réfère à la figure 2.3, on voit qu'ils sont distribués selon une structure hexagonale. Cette dernière a été choisie, car, si l'espace entre les actionneurs est réduit à zéro, la configuration devient hexagonale compacte. Elle représente la distribution la plus compacte possible de disques égaux sur une surface donnée, donnant ainsi une densité maximale de points de contrôle et diminuant le résiduel apparaissant dans l'équation 2.7. Cependant, il est fait mention à la section 2.5.3 qu'il y a présence d'un terme de couplage qui dépend probablement du champ moyen des PPV. Ainsi, plus les actionneurs sont distants les uns des autres, moins il y a de couplage. Cette configuration représente, encore une fois, un compromis judicieux entre des paramètres à optimiser. Il faut aussi considérer le côté pratique résultant de l'augmentation du diamètre de la pupille avec un même nombre d'actionneurs. En effet, certaines applications, comme le null testing, utilisent des faisceaux de grands diamètres. Avoir un miroir aux dimensions similaires devient donc un atout. Dans ce cas précis, les actionneurs ont 2 mm de diamètre et sont séparés de 5 mm centre-à-centre. Le faisceau fait environ 50 mm de diamètre, permettant techniquement de percevoir l'effet des 91 actionneurs sur une mesure. Cependant, puisque les mesures doivent être prises à la main et vu la quantité de mesures à prendre, il a été décidé d'utiliser uniquement les 19 premiers actionneurs. Un masque numérique viendra éliminer les points de mesures dépassant la pupille souhaitée.

Enfin, le MC est le dernier morceau formant le miroir. Ses dimensions sont déterminées en fonction de ratios très précis ayant pour effet de rendre le champ à la surface du miroir constant jusqu'à la dérivée de 6^e ordre. La figure 3.7 montre ses principales caractéristiques.

Le MC est censé être constitué de trois boucles parfaitement circulaires de fil infiniment mince. De plus, les deux boucles de rayon $\sqrt{4/7} R$ doivent être parcourues d'un courant exactement égal à 49/64 de celui circulant dans la boucle centrale. Le montage réel est évidemment légèrement différent. Les fils sont plutôt enroulés afin de permettre la diminution du courant à injecter et le ratio 49/64 correspond au ratio du nombre d'enroulements. De ce fait, les trois boucles peuvent

Montage et méthodologie

être reliées entre elles et être parcourue du même courant. Les dimensions finies des enroulements doivent cependant être considérées dans des simulations de précision du champ généré.



Fig. 3.7 : Schéma représentant la relation entre les différentes dimensions liées au MC

La source stabilisée en courant alimentant le MC est fabriquée par *BK Precision*, modèle 1786B. Elle vient avec un logiciel de contrôle permettant de modifier les paramètres directement à l'écran. Cette dernière permet de générer les courants envoyés au MC (nécessairement limités à 1A) avec une précision de ±4mA, selon les spécifications du fabricant.

3.1.4 Montage optique

Le montage optique est relativement simple. L'idée est de générer un front d'onde aussi plan possible et de rayon suffisamment large pour éclairer la totalité de la pupille du miroir. Le faisceau réfléchi est redimensionné par un système afocal pour ajuster le diamètre du faisceau à la pupille d'entrée de l'analyseur de front d'onde. La figure 3.9 montre un schéma du montage.

Pour obtenir le faisceau désiré, une diode couplée dans une fibre génère un front d'onde quasi sphérique. On utilise un iris pour ne conserver que la portion d'éclairement uniforme du faisceau. Celui-ci est collimé par la lentille f1, placée exactement à une distance focale de la sortie de la fibre. Ensuite, ce faisceau passe par un premier système afocal de grandissement f3/f2, constitué des lentilles du même nom. Le faisceau doit ensuite être réfléchi par un miroir plat à 45° sur la surface horizontale du miroir. Le faisceau réfléchi passe par un second système afocal (M = f4/f3) par le système formé des lentilles f3 et f4 et le *beam splitter*. Le faisceau collimé entre ensuite dans l'analyseur de fronts d'onde qui fait la mesure indirecte de la surface. Dans le montage présenté à la figure 3.9, les lentilles f1 et f3 ont des focales de 100 mm, tandis que la lentille f2 a une focale de 600 mm. La distance séparant l'iris de la lentille f1 est de 100 mm, celle séparant f1 et f2 est de 700 mm, celle entre f2 et le FDM est de 600 mm, celle entre f2 et f3 est de 700 mm et finalement, celle entre f3 et l'analyseur de front-d'onde est de 100 mm.

Les erreurs de fabrications, d'alignement et de positionnement causent des erreurs statiques. Celles-ci peuvent aisément être référencées, pourvu que leur présence ne représente pas une erreur trop grande, à l'aide de l'outil réservé à cet effet dans le logiciel HASO 3.0.



Fig. 3.8 : Photographie du montage optique permettant de déterminer la surface du MDF

3.1.5 Analyseur de front d'onde

Un analyseur de front d'onde de type Shack-Hartmann (SH) est constitué principalement d'une matrice de microlentilles et d'un capteur CCD. L'image de la figure 3.10 est tirée du manuel de présentation d'Imagine-Optics dans la section expliquant le fonctionnement de l'analyseur. L'idée derrière la matrice de microlentilles est de décomposer le front d'onde en plusieurs souspupilles. Chacune de ces sous-pupilles passe dans une lentille. Un front d'onde idéal produit un *spot* correspondant à la tâche d'Airy générée par le système. Celle-ci est cependant centrée sur l'axe optique. Si au contraire, le front d'onde est courbé, le faisceau correspondant à la souspupille convergera en un point légèrement hors axe. De ce déplacement, il est possible de déduire la pente locale de faisceau incident et donc de reconstruire sa totalité.

La mesure représente donc plus ou moins les composantes du gradient de la surface et la reconstruction est donc, par analogie, l'intégrale de la mesure. Cet énoncé n'a pas pour but d'expliquer comment l'algorithme fonctionne, simplement de mettre en lumière un point important : la reconstruction est valide, à une constante arbitraire près. Cette constante est déterminée de telle sorte que l'amplitude moyenne soit égale à 0. Autrement dit, si on mesure une gaussienne de différentes amplitudes, le « 0 » varie en fonction de celle-ci. Il faudra donc en tenir compte dans certains tests. D'un autre côté, la reconstruction dépend également du bruit sur la CCD ainsi que de la base polynomiale utilisée. En effet, concernant le dernier point, il faut savoir qu'un SH peut être utilisé en mode « modal », où la surface mesurée est donnée comme une combinaison linéaire des PZ. Bien évidemment, on doit utiliser un nombre fini de PZ et chaque reconstruction tronque une partie de la surface correspondante aux termes d'ordres plus élevés.





Fig. 3.9 : Image représentant le concept de base du fonctionnement d'un analyseur de type Shack-Hartmann

Le danger avec un tel système est que si le front d'onde est trop aberré, il est possible que certains *spots* se confondent sur la CCD et que le logiciel soit incapable de le reconstruire partiellement ou globalement. Il faut donc porter une attention particulière aux pentes locales générées à la surface du miroir. Bien qu'il existe des algorithmes permettant d'étendre la plage d'utilisation du SH, nous limiterons les amplitudes (donc les pentes locales) afin de pouvoir utiliser le logiciel fourni avec l'instrument. Chaque mesure obtenue contient donc 1936 données correspondant à la matrice de 44x44 microlentilles. Il est aussi possible, au besoin, de connaître l'intensité lumineuse à chacun de ses points de mesure.

3.1.6 Interféromètre Zygo et MetroPro

Un interféromètre de type *Fizeau* produit par Zygo sera aussi utilisé. Le schéma de la figure 3.11permet d'entrée de jeu de saisir le concept de base.



Fig. 3.10 : Configuration de base d'un interféromètre de Fizeau. Tiré de [29].

Le principe fondamental repose sur la création d'une interférence entre le front d'onde réfléchi à la sortie de la surface de référence et la surface à tester (Reference surface et Surface under test, figure 3.11). Un algorithme analyse ensuite le patron d'interférence produit au plan d'observation (Observation plane, figure 3.11) et reconstruit la forme du front d'onde et donc la forme du miroir de façon indirecte. En effet, la forme du front d'onde aura exactement la même forme que la surface, mais l'amplitude sera doublée par l'effet de la réflexion. Toute une série de manœuvres de calibration est nécessaire à l'obtention de mesures adéquates avec ce système, mais leur

Montage et méthodologie

explication détaillée est inutile. Elles visent principalement l'alignement des différents éléments en jeu.

Le logiciel de contrôle est fourni par le fabricant et est connu sous le nom de MetroPro. Les principaux paramètres auxquels il faut porter une attention particulière pendant les mesures sont la référence utilisée et les zones non définies par une mesure donnée. En effet, on peut enregistrer une référence qui sera automatiquement soustraite de chaque mesure pour n'obtenir, dans le cas présent, que l'influence d'un actionneur. Le problème principal rencontré provient des zones qu'on peut qualifier de non définies. En effet, si les franges sont trop serrées ou l'illumination relative inadéquate, il se peut que l'algorithme arrive à reconstruire le front d'onde, mais seulement en partie. Certains pixels n'auront donc aucune valeur d'amplitude associée. Le principal ennemi est la densité de frange qui augmente avec les variations brusques d'amplitude à la surface du miroir. Il faut généralement limiter les amplitudes à quelques microns, soit environ le tiers central de la plage dynamique du miroir.

Cet appareil est extrêmement précis, probablement même le meilleur, pour mesurer la surface totale du miroir, mais est toutefois très lent (environ une mesure à toutes les 2-3 secondes au maximum). La limitation en amplitude pourrait aussi être un problème puisque son utilisation sera réservée à la détection des effets non linéaires, d'amplitudes très faibles, particulièrement à l'intérieur de la plage dynamique préalablement mentionnée.

3.2 Méthodologie

Cette section est réservée pour donner le détail concernant les différentes stratégies employées lors des différentes étapes de la recherche. On y trouve uniquement les éléments utiles au mémoire et non des perspectives d'avenir ou des possibilités de recherches ultérieures. On peut rappeler ici que le but principal est de démontrer la viabilité de l'algorithme RIM dans un contexte industriel. Certains aspects doivent donc être démontrés même s'ils n'entrent pas directement dans l'algorithme. En effet, l'industrie voudra certainement savoir, entre autres, à quel point le système est sensible à la configuration initiale. On voudra connaître la répétabilité des tests, savoir quels sont les paramètres sensibles du montage, quel est l'impact sur la précision du scénario catastrophe (*worst-case scenario*) et à quelle fréquence il est nécessaire de refaire une calibration. D'un autre côté, il sera important de quantifier les variables présentées dans la section théorie pour connaître l'ordre de grandeur des éléments en jeu. Enfin, il sera impératif de vérifier si la méthode proposée améliore effectivement la technique SOL.

Afin de répondre à ces questions, une simulation devra d'abord être entreprise pour vérifier théoriquement l'effet de certains paramètres *sensibles*. Ce genre de simulation est couramment appelé *test de tolérancement* (ou test de robustesse). L'approche sera de se concentrer uniquement sur les paramètres propices aux changements pouvant survenir lors d'un entretien régulier ou d'une calibration. Les paramètres « non variables » seront de toute façon référencés. Ainsi, en faisant l'hypothèse que toutes les pièces en dehors du contenant à FF sont

Montage et méthodologie

fixes, les seuls paramètres potentiellement variables sont la hauteur du FF dans le contenant (pouvant survenir lors du remplacement du FF) et les valeurs exactes de courants injectés (pouvant provenir d'une légère instabilité des sources utilisées). De façon complémentaire, on devra aussi s'assurer de la constance du système dans le temps, ou du moins un intervalle raisonnable de constance. Pour ce faire, il suffira de définir un vecteur courant qui sera injecté jours après jour. Chaque surface ainsi générée sera mesurée et permettra de suivre l'évolution des variations temporelles du système. Une fois ces éléments démontrés, il faudra quantifier les variables issues des équations théoriques 2.20 et 2.25. Les outils de simulation de la première expérience seront alors très utiles. Ils permettront, entre autres, de quantifier l'erreur d'échelle et l'erreur de couplage. Enfin, la partie cruciale sera celle où il faudra comparer la méthode RIM aux méthodes SOL et SCL. L'hypothèse est que la méthode RIM devrait avoir une précision intermédiaire. Il sera tout de même intéressant d'évaluer la différence de précision entre cette technique et le SCL, censée représenter l'erreur minimale physiquement atteignable.

3.2.1 Répétabilité des mesures

La première étape est le test de tolérancement. Il suffit de simuler le système, en utilisant les équations 2.20 et 2.25, et d'évaluer la forme de la surface théorique pour différentes valeurs¹⁶. La simulation du système impose d'identifier toutes les spires, leur diamètre, leur courant et leur position selon un repère dont l'origine est posée en plein centre de l'interface air/FF. Il faut ensuite calculer l'apport individuel de chacune au champ total, et ce, en un point quelconque de l'espace. Ce champ n'est en fait que la somme de toutes les contributions. Une fois ces valeurs obtenues, il est possible de calculer la forme de la surface générée.

Concrètement, il est possible de décomposer l'interface FF/air en une grille de points ayant des coordonnées spatiales bien définies. De plus, il est possible de calculer indépendamment les champs du MC et d'un actionneur sans même tenir compte des courants respectifs. En effet, selon l'équation 2.25, l'opération revient à calculer une certaine quantité multipliant le courant I. Mathématiquement, l'équation 2.25 a) peut se réécrire comme la somme des contributions de chaque spire, multipliée le courant commun l :

$$\vec{B}(r,z) = I * \vec{P} \tag{3.1a}$$

$$\vec{P} = \sum_{i} \left(\frac{\mu_0 k}{4\pi \sqrt{ar^3}} [P_r \cdot \hat{a}_r + P_z \cdot \hat{a}_z] \right)_i$$
(3.1 b)

L'indice i à l'extérieur de la parenthèse dans l'expression 3.1 b) signifie que le calcul s'effectue selon les paramètres particuliers de la i^e spire de courant.

¹⁶ Tous les calculs seront effectués avec MatLab et nécessiteront la création d'un code maison pour éviter les erreurs et obtenir des résultats dans un temps raisonnable. Seul l'actionneur central sera considéré et l'erreur sera majorée pour tenir compte d'une potentielle contribution des faibles erreurs de couplage dans des circonstances plus générales.

Le champ total généré à la fois par un actionneur (indices a) et le MC (indices C) sera donné par :

$$\vec{B} = I_C * \overrightarrow{P_C} + I_a * \overrightarrow{P_a} + \vec{T}$$
(3.2)

T est le champ magnétique terrestre. Pour les comparaisons entre différentes valeurs de courant ou de position, la valeur T n'est pas nécessaire puisqu'elle disparait dans la soustraction (on suppose le champ terrestre extrêmement constant et uniforme à la surface du FF). On utilise enfin l'équation résultante 3.2 pour calculer la surface théorique à l'aide de l'équation 2.20.

Le test sur la variation de l'épaisseur de liquide peut se faire en changeant les coordonnées de la grille de points préalablement définie. Le test sur les courants se fait simplement en les faisant varier dans l'équation 3.2 (puis en recalculant la surface avec l'équation 2.20). Comme la précision sur la hauteur est d'environ 0,1 mm et celle des courants, inférieure à 0.01A, le test sera effectué sur un intervalle de \pm 0.2 mm et \pm 1 % autour des valeurs normales de hauteur et de courants. Ces valeurs sont respectivement posées à 0,0 mm, 0,8A (MC) et 0,2 A (actionneur).

Du côté des mesures temporelles, une certaine surface doit être générée et mesurée quotidiennement avec un même vecteur courant. Entre les différentes mesures, le système devrait idéalement ne pas être utilisé et absolument aucune composante ne doit être touchée. Certaines observations laissent croire que deux facteurs causent les variations temporelles et qu'ils sont liés à la dégradation du FF. Il y a d'abord la dégradation du surfactant qui peut potentiellement être responsable d'une agglomération graduelle des particules en suspension qui finissent éventuellement par se déposer au fond du contenant. Il y a aussi l'évaporation lente du liquide porteur qui semble être accentuée pendant les utilisations intensives du système. La perte de liquide porteur semble principalement affecter la viscosité du FF qui, à son tour, semble influencer le temps de réponse du système sans toutefois affecter les amplitudes.

Deux types de mesures sont nécessaires : une mesure prise rapidement après l'injection des courants et une autre après un temps suffisant pour que le liquide soit certainement à l'équilibre. Par expérience, il semble qu'un court délai de 0,5 s et un long de 10 s soient raisonnables pour les buts fixés. Un VI particulier devra être créé à cet effet permettant de commander le miroir et de déclencher la mesure par un signal *trigger* après un délai donné. Afin de vérifier l'hypothèse de l'évaporation accélérée lors d'une utilisation intensive, une journée sera réservée vers la fin de l'expérience pour pousser le système en mode intensif sur une période de 24 heures. Il est extrêmement ardu d'émettre une quelconque hypothèse concernant l'évolution des résiduels au fil du temps. On peut néanmoins s'attendre à des comportements distincts pour les deux délais.

3.2.2 Termes non linéaires

À la section 2, il a été démontré que les équations utilisées prédisaient deux termes non linéaires. Ces deux termes étaient, par hypothèse, la première cause d'imprécision de la méthode SOL. Il semble donc majeur de vérifier qu'ils contribuent effectivement de façon significative aux erreurs de contrôle. Pour se faire, il suffit de déterminer la proportion non linéaire d'une FI (liée à l'erreur d'échelle) et de mesurer clairement un effet de couplage.

3.2.2 a) Test de linéarité

Il n'existe qu'une seule façon de vérifier la linéarité globale de la réponse d'un actionneur : prendre une mesure à intervalle régulier pour un actionneur donné et déterminer si la FI se déforme proportionnellement au courant injecté. La littérature montre souvent la confirmation de la linéarité d'un système par la simple mesure du PV en fonction du courant. Celle-ci permet cependant l'introduction de plusieurs biais dans l'évaluation de la relation, notamment par certaines valeurs qui apparaissent régulièrement en périphérie d'une mesure et qui, visuellement, sont clairement des données aberrantes. Dans le même ordre d'idée, il est aussi très ardu de détecter toute forme d'erreur systématique qui pourrait potentiellement apparaître. Enfin, l'information spécifique à chaque point de mesure est perdue.

Afin d'aller chercher un maximum d'information, une régression quadratique sera effectuée sur chacun des points de mesure en couvrant entièrement la plage dynamique des actionneurs. Il faut donc s'attendre à de grandes amplitudes, ce qui justifie l'emploi du SH. Il est cependant possible que, malgré cette précaution, cet instrument ne puisse pas mesurer adéquatement les FI de grande amplitude. Afin d'empêcher ce genre de situation, la valeur de courant circulant dans le MC sera ajustée à la baisse (au besoin) pour permettre l'étude de la pleine plage. Une mesure sera effectuée sur la plage [-10 V, 10 V] à chaque intervalle de 1 (sauf OV). Chaque régression sera ainsi basée sur 20 données et retournera les coefficients de l'équation quadratique générale pour chaque point de mesure et ceux-ci seront compilés dans trois matrices distinctes qui permettront d'obtenir un aperçu visuel de la situation.

L'objectif spécifique est ici uniquement de démontrer la présence d'un terme quadratique et son amplitude. Dans le cas idéal, il serait possible de corroborer les mesures expérimentales et les prédictions qui peuvent être faites à partir des équations de la section 2.

3.2.2 b) Test de couplage

Le test de couplage requiert une meilleure précision de mesure que le test de linéarité. Ainsi, un interféromètre de type Zygo sera utilisé. Les courants injectés doivent demeurer très faibles pour permettre les mesures, conséquence des capacités de mesure de l'appareil. L'idée ici est de mesurer une FI non normalisée de l'actionneur central dans différentes conditions de couplage. Il faut tester à la fois s'il y a présence de couplage et si son impact varie en fonction de la

Montage et méthodologie

configuration exacte des actionneurs. Ainsi, 4 configurations seront analysées : la première avec tous les actionneurs éteints (pour avoir une référence), la deuxième avec tous les actionneurs à 1V, la troisième avec un actionneur sur deux à 2V (afin d'assurer la même intensité de champ en plein centre) et finalement, la quatrième avec simplement 1V distribués asymétriquement. Pour n'obtenir que la FI, il faudra évidemment référencer l'influence de la configuration de couplage.

3.2.4 Test RIM

Cette étape est sans doute celle qui est la plus déterminante de toutes. Elle reste tout de même extrêmement simple à expliquer. D'ailleurs, la section 2.6 résume déjà très bien la méthode appliquée. Le seul élément à définir est la surface cible. Il est évident qu'il est impossible de vérifier la viabilité du système pour absolument toutes les configurations possibles. De plus, si une surface cible est trop arbitraire, il se pourrait que ce cas particulier soit problématique pour l'algorithme, menant vers des conclusions erronées, sans toutefois qu'il soit possible de s'en apercevoir. Une surface continue et relativement lisse semble donc être un bon point de départ. Or, il se trouve que c'est exactement le genre de surface qui nous intéresse : les surfaces typiques de lentilles industrielles. Une simple combinaison linéaire de quelques PZ suffit. Cette combinaison s'obtient très aisément par des logiciels comme MatLab. La cible ainsi calculée peut aisément être utilisée par les trois méthodes présentées. Chacune de ces méthodes formera une surface qui sera mesurée. De chacune de ces mesures sera soustraite la surface cible et, enfin, de ce résiduel sera calculé le résiduel RMS, ce qui réduira la comparaison des techniques à une comparaison de trois valeurs. Évidemment, plus faible sera la valeur RMS, plus précise sera la méthode.

4.1 Répétabilité des mesures

Comme il a été mentionné à plusieurs reprises, le souci premier était l'utilisation du FDM en milieu industriel. La première devait naturellement être de vérifier que le système offre une certaine fiabilité, qu'il est possible de l'utiliser de façon intensive et que d'une fois à l'autre, les résultats concordent. Toute entreprise qui utilise un nouvel instrument a besoin de plus d'un mode d'emploi, il faut aussi connaître le plan d'entretien, les limitations, les aspects problématiques. Il est donc important de déterminer une plage temporelle pour laquelle l'instrument est réputé « constant » et à la suite de laquelle un entretien devra avoir lieu. Il faut donc aussi s'assurer que le système reviendra à son état initial après celui-ci. Deux tests ont été effectués et les résultats sont présentés ici. Ceux traitant de la dégradation dans le temps sont nommés *tests temporels* tandis que ceux concernant la répétabilité d'un montage après un entretien sont englobés sous le titre *tolérancement*.

4.1.1 Tolérancement

Pour le test de tolérancement, deux paramètres variables ont été isolés. Comme la position relative des pièces mécaniques est invariante dans le temps, seuls les courants injectés et la quantité de FF peuvent varier d'un montage à l'autre. L'étude a été menée sur un seul actionneur, mais les faibles variations potentielles d'un actionneur à l'autre ont été compensées par une large surestimation des erreurs. Il y a donc trois paramètres indépendants à vérifier : l'erreur sur le courant injecté dans l'actionneur, l'erreur sur le courant injecté dans le MC et l'erreur sur la quantité de FF (ou, de façon équivalente, l'erreur sur la hauteur de FF). Cependant, la construction mécanique en elle-même comporte des erreurs de positionnement relatif qui font que l'erreur réelle de positionnement de l'actionneur par rapport à l'interface air/FF n'est pas nécessairement la même que celle entre le MC et cette même interface. Pour considérer l'ensemble des paramètres, il faut donc distinguer deux erreurs liées à la hauteur du FF. On peut ainsi facilement imaginer qu'une FI typique peut être décrite par une fonction mathématique quelconque qui dépend des trois paramètres spatiaux et des quatre paramètres à étudier. En se servant de la symétrie du système, on peut cependant réduire le tout à une fonction à cinq variables : $h(r, \Delta i, \Delta I, \Delta b, \Delta B)$. Les lettres en majuscules sont associées au MC et leurs équivalents minuscules à l'actionneur. Le paramètre r est simplement la distance radiale du centre du miroir et donc du centre de l'actionneur. La fonction représente donc simplement la hauteur de la surface. Représenter cette dernière graphiquement est impossible puisqu'il faudrait un espace à cinq dimensions. Une méthode simple et efficace peut simplement être de trouver les cas extrêmes ou worst case scenarios qui peuvent être comparés au cas idéal.

Il faut cependant borner adéquatement les différentes valeurs envisageables, sinon l'évaluation perd tout son sens. Cependant, puisque la presque totalité des instruments sont

construits « maison », le point de départ concernant les incertitudes ne peut être qu'estimé. Par exemple, nous savons que la carte PowerDAQ PD2-AO-96-16, présentée à la figure 3.3 et qui contrôle l'actionneur, génère une sortie numérique avec une résolution de 16 bits, soit une résolution de 6x10⁻⁶A une fois convertie en courant dans la plage [-0,2 A, 0,2 A]. Cependant, le panneau de contrôle, commandé en volts sur la plage [-10 V, 10 V], est rarement utilisé avec des entrées numériques aux chiffres significatifs allant au-delà de 0,01V, ce qui correspond à une « erreur » (humaine) correspondante de ±0,2 mA, soit plus de 300 fois plus élevée que la résolution elle-même. Aux fins d'analyse, cette erreur a été majorée à ±1 mA. Il n'y a aucune raison de croire que cette erreur n'englobe pas toute autre source non considérée. Concernant la source au MC, comme il a été mentionné à la section 3.1.3, le courant est contrôlé avec une incertitude de ± 0,4 mA. Encore une fois, les tests ont été effectués sur une incertitude majorée de \pm 0,8 mA, soit le double. Enfin, concernant la hauteur de FF, il est possible de se baser sur l'erreur sur la quantité. La quantité est mesurée en grammes, et la hauteur est déterminée par la masse volumique du FF fournie par le fournisseur $(1,21 \text{ g/cm}^3)$ et par l'équation du volume d'un cylindre : $V = h\pi r^2$. On connaît évidemment le rayon du contenant, soit 3,80 ± 0,01 cm. De son côté, l'incertitude sur une mesure de la masse est de \pm 0,1 g. Combinées, elles correspondent à \pm 30 μ m sur la hauteur. Cette incertitude a été majorée à \pm 100 μ m pour les besoins du test. Concernant les valeurs dites idéales, elles ont été posées de façon à, elles aussi, déterminer l'erreur maximale. Ainsi, comme mentionné à la section 3.2.1, le courant idéal de l'actionneur a été posé au maximum de sa plage dynamique, soit 0,2 A, et le MC à la valeur de 0,8 A. Pour les besoins, on fait varier r de 0 à 4 mm par bonds de 0,1 mm.

Le paramètre le plus critique a été, sans équivoque, la précision sur le courant au MC, suivi du courant à l'actionneur, puis de l'erreur relative de hauteur de l'interface par rapport à l'actionneur. L'erreur équivalente liée au MC a été mesurée comme étant nulle en raison du nombre de décimales limité associé aux calculs effectués par MatLab. Le graphique montrant les courbes de hauteurs maximale, minimale et idéale est montré à la figure 4.1. Comme le calcul de la hauteur est effectué à partir de l'équation 2.20 et que ce même calcul est déterminé à une constante près, il a été décidé de ramener la hauteur des trois courbes pour le point r=4mm à 0, afin de mieux voir la différence relative.

On peut donc constater que l'erreur sur la surface ne dépassera jamais 2µm. Ce chiffre peut, d'amblée, sembler élevé, mais il faut garder en mémoire que toutes les variables sensibles ont été simultanément exagérées dans le même sens. Il est évident que ce scénario est plus qu'improbable, puisque généralement les erreurs ne sont pas corrélées. De plus, celles-ci ont été volontairement exagérées pour déterminer l'erreur extrême qu'il est techniquement impossible de dépasser. Cette évaluation repose cependant sur une hypothèse non vérifiée, à savoir si les équations utilisées représentaient bien le système réel. Une conclusion peut tout de même être tirée de cette analyse : il n'y a pas de raison physique de croire que le système est sensible aux erreurs de préparation. Autrement dit, deux techniciens compétents sur deux systèmes distincts devraient techniquement obtenir des résultats très similaires, indépendamment de leurs erreurs de manipulation respective. De plus, ces erreurs proviennent spécifiquement de changements

opérés durant un entretien régulier ou lors du remplacement de certaines pièces mécaniques du système.



Fig. 4.1 : Graphique montrant la hauteur calculée numériquement de la surface pour des valeurs normale, maximale et minimale de courant en fonction de la distance radiale au centre du miroir

Il va de soi que rien n'empêche une nouvelle calibration du système. Ce faisant, les erreurs se trouvent incluses à l'intérieur des mesures et l'algorithme de contrôle en tient compte pendant l'utilisation normale du miroir. Changer la constante de proportionnalité d'une équation linéaire ne change en rien la linéarité en soi. Pour une cible donnée, les valeurs calculées seront certes différentes, mais dans le contexte proposé, le seul intérêt réel est de pouvoir générer une surface voulue avec le FDM. Le désagrément principal causé par une différente calibration vient du fait qu'un ancien vecteur courant calculé ne pourra plus être utilisé de façon fiable. Cependant, cet exercice a clairement démontré qu'ils peuvent encore être utiles en tant que bonne approximation.

Dans un contexte où ces variations deviennent problématiques, il est possible de choisir une source de courant offrant une meilleure précision. En effet, la source utilisée avait une plage dynamique s'étalant de 0 à 3 A avec une précision d'environ 5 mA. Le simple fait de choisir, par exemple, une source 0-1 A avec une précision de 0,5 mA permet une forte diminution des erreurs potentielles. De plus, il est plus que probable que les erreurs réelles soient de l'ordre de quelques dizaines de microns, puisque l'influence à 0,2 A est généralement plus près de 18 microns que de

22 et les incertitudes sont systématiquement très inférieures à celles proposées. On en conclut qu'une erreur de mesure a peu de chance de provenir de la configuration physique du montage.

4.1.2 Tests temporels

Certains tests ont aussi été effectués sur les effets du temps sur la répétabilité d'une surface. En effet, il est clair que le temps dégrade les particules et cause leur agglomération. Il est tout aussi clair que le liquide porteur s'évapore tranquillement, faisant donc varier la hauteur du liquide. Le point qui doit être établi est le temps caractéristique permettant une dégradation suffisante pour causer des erreurs significatives de répétabilité. Autrement dit, il faut déterminer la fréquence à laquelle un entretien est requis. Une surface continue relativement lisse a donc été mesurée à intervalle régulier pendant une période de 40 jours. Chaque mesure était comparée à la première et la valeur RMS de la différence entre les deux mesures était calculée. Cette figure telle que mesurée au premier jour est présentée à la figure 4.2.



Fig. 4.2 : Aperçu de la surface expérimentale générée pour le test temporel

Il avait déjà été observé par le passé que le temps semblait rendre le FF plus visqueux, notamment en augmentant significativement son temps de réponse. L'hypothèse émise alors était qu'une partie du liquide porteur s'était évaporée, entrainant une augmentation de la concentration du FF en particules et indirectement de sa viscosité. Cependant, l'impact majeur de la viscosité survient au niveau du temps de stabilisation de la surface après injection des courants. En temps normal, le FF EHF1, produit par Ferrotech, a une viscosité de 6cP et un temps de réaction caractéristique d'environ 100 ms. Après quelques mois, ce temps caractéristique semblait s'étirer à plusieurs secondes (autour de 5). Il a donc été décidé de mesurer l'effet sur la surface pour une mesure prise rapidement après l'injection des courants, soit 0,5 s et pour une autre prise après la stabilisation de la surface, soit 10 s pour déterminer distinctement les variations dues au temps de stabilisation et celles dues uniquement aux variations de nature plus « mécanique » du FF. Ces paramètres mécaniques sont principalement la concentration en particule et le dépôt causé par les agglomérats au fond du contenant.

La figure cible est une simple superposition linéaire de différents PZ. Selon la convention de Wyant [20], la surface désirée¹⁷ génère un front d'onde réfléchi qui se décompose en trois PZ : Z_2^2 , Z_3^1 et Z_3^3 , de coefficients respectifs 3,6, 1,1 et 1,9 [µm]. Ce genre de surface est très représentative des surfaces typiquement « difficiles » à tester : amplitude trop élevée pour une mesure directe avec la plupart des instruments de précision, surface continue relativement lisse et suffisamment distincte d'une surface sphérique.

Les résiduels RMS en fonction du temps sont présentés à la figure 4.3. On note un comportement radicalement distinct entre les deux séries de mesure.





Plusieurs éléments peuvent être exposés à partir de ce simple graphique. Il faut d'abord savoir que les courbes ont été artificiellement juxtaposées de façon à ce que les deux mesures au jour 3 (donc après un vieillissement de 2 jours) soient identiquement nulles. En effet, chaque mesure est comparée avec sa mesure initiale respective. De plus, deux mesures distinctes possèdent toujours

¹⁷ La décomposition est approximative est effectuée sur la mesure au jour 1.

un certain résiduel moyen de quelques centaines de microns, indépendant du phénomène de dégradation. Cependant, ce résiduel est normalement constant d'une mesure à l'autre, ce qui permet de soustraire sa valeur des valeurs calculées. Se faisant, on isole en quelque sorte la variation de la forme. De plus, l'expérience cachait une seconde expérience qui pouvait être opérée en simultanée. En effet, pendant les 40 jours, le système n'a été mis en marche que quelques minutes par jour, et ce, uniquement pour prendre les mesures. Cependant, un délai de 24 heures a été utilisé pour vérifier l'effet du vieillissement dans une utilisation « intensive ». Du jour 34 à 35 (vieillissement de 33 à 34 jours sur le graphique), le miroir fonctionnait en continu, générant la surface cible. Le but était de vérifier si une utilisation intensive, similaire à celle qu'on doit s'attendre dans une entreprise, accélérerait le processus de dégradation. Pour vraiment constater l'effet, une régression polynomiale a été opérée sur les points mesurés entre 2 et 33 jours de vieillissement et extrapolée pour les jours subséquents. Ces régressions sont présentées à la figure 4.4.





Les meilleures régressions polynomiales (*best fit*) ont été obtenues et la courbe pour les données avec délai de 0,5 s est un polynôme d'ordre 4, comparativement à la régression sur les points avec délai de 10 s donnant un polynôme d'ordre 1. On constate très rapidement le « décalage » entre la courbe à l'ordre 4 et les points après 34 jours de vieillissement. Ce comportement n'est cependant pas apparent pour l'autre série de données. De plus, on voit qu'il est possible de faire corréler ces points décalés simplement par translation. On peut donc en déduire qu'ils ont vieilli

prématurément. Les résiduels constatés auraient dû être observés 1,6 jour plus tard, représentant donc un vieillissement 2,6 fois plus rapide.

Pour expliquer pourquoi l'effet apparaît seulement pour les mesures avec un court délai, il faut revenir à un niveau un peu plus fondamental. Comme il a déjà été mentionné, deux phénomènes simultanés viennent modifier le comportement du FF. Le premier est la quantité du liquide porteur et le second est la dégradation du surfactant causant l'agglomération des particules. Il faut aussi savoir que l'utilisation continue mène à des pertes en chaleur se traduisant en un échauffement du FF (de quelques degrés Celsius). Les observations nous indiquent donc que l'échauffement accélère l'évaporation du liquide porteur, mais n'agit pas de façon détectable sur la dégradation du surfactant. De plus, on peut aussi déduire que l'augmentation de la concentration du FF en particules n'influence pas la dégradation. En effet, si le liquide porteur s'évapore, la concentration en particules augmente et les collisions entre-elles font de même. On pourrait alors croire qu'elles pourraient causer une dégradation accélérée proportionnelle à leur concentration. Mais les données contredisent cette hypothèse. Comment alors expliquer les comportements respectifs des deux séries de mesures?

L'explication du comportement des mesures prises avec un délai de 10 s ne peut se faire mathématiquement, ni se baser sur des faits démontrés. Cependant, on peut émettre l'hypothèse selon laquelle la dégradation du surfactant d'une particule s'effectue de façon constante. Cette dégradation cause une agglomération qui interfère en modifiant le champ apparent à l'interface, modifiant ainsi linéairement l'amplitude de la déformation. De plus, la mesure est prise dans une position d'équilibre statique, puisque la phase dynamique ne dure que quelques dixièmes de seconde. Voilà pourquoi la viscosité, principal responsable de l'augmentation du temps de réponse, n'influence pas les résiduels pour cette série de données. L'effet est cependant bien présent pour les mesures à 0,5 s de délai. En effet, le temps de stabilisation nominal est autour de 0,1 s, mais peut augmenter jusqu'à quelques secondes à l'intérieur d'un délai de vieillissement de 3 mois. Il est donc normal de voir ses effets sur une période de 40 jours. En effet, si le temps de réponse augmente, la surface prend de plus en plus de temps pour se stabiliser et la mesure se trouve à être prise pendant son état transitoire entre deux états stables. Au fil des jours, la même mesure « s'éloigne » temporellement de l'état stable, ce qui explique pourquoi la variation augmente significativement plus rapidement pour cette série de données. La variation est donc due à une combinaison de la dégradation du surfactant et de l'augmentation de la viscosité, mais ce dernier a un effet beaucoup plus notable si on compare les deux courbes de résiduels.

D'un point de vue pratique, on peut se demander à quelle fréquence il serait raisonnable d'effectuer un entretien, soit de vidanger le FF et en remettre du frais, afin de garder les mesures comparables d'une journée à l'autre. La réponse à cette question ne peut se faire en dehors d'un contexte spécifique. Il faut connaître les conditions d'expérimentations, à savoir si le miroir est utilisé de façon dynamique ou statique, de façon intensive ou non, les besoins en termes de précision de surface, etc. Dans un cas typique où la précision attendue est de l'ordre d'un dixième de longueur d'onde, soit environ 70 nm avec un laser HeNe, dans des conditions d'utilisations restreintes, mais avec une utilisation dynamique, un entretien hebdomadaire est suffisant.

50

Cependant, la plupart des entreprises ont une utilisation intensive de leurs instruments. Comme mentionné ci-haut, cette utilisation a pour effet d'accélérer le vieillissement d'un facteur 2,6. Le délai raisonnable tombe alors à environ 3 jours. Avec un processus automatisé de calibration, un tel processus peut être effectué en moins de 15 minutes. Même s'il doit être fait de façon bihebdomadaire, il ne semble pas être une embuche majeure aux opérations commerciales normales. De plus, la calibration régulière ne peut être que bénéfique en général puisqu'elle permet d'éliminer toute autre source d'erreurs statiques qui pourraient survenir en cours de route.

4.2 Test de linéarité

Le test de linéarité est un des plus simples à réaliser. Il suffit d'étudier individuellement l'évolution de chaque pixel pour une série de mesure. En effet, les mesures obtenues au SH sont en réalité des matrices de 1936 valeurs (voir figure 4.2 à titre d'exemple). Sur une série de mesure, chaque pixel évolue, selon l'hypothèse, suivant sa propre équation quadratique. Chaque régression quadratique retourne trois valeurs distinctes représentant les trois coefficients de l'équation quadratique générale : $a_0 + a_1 x + a_2 x^2$. En rangeant ces coefficients dans trois matrices distinctes, aux mêmes positions que les pixels associés, on peut définir trois nouvelles matrices A0, A1 et A2 qui contiennent l'entièreté de l'information recueillie. La matrice A0 est cependant inutile, puisqu'elle contient principalement l'erreur statique moyenne sur chaque mesure, erreur qui devrait normalement disparaître par référencement. De plus, elle n'apporte aucune information physique ou utile au problème et devrait, selon les équations de la section 2, donner identiquement 0 partout. D'ailleurs, il faut garder en mémoire que le but principal de cette expérience est uniquement de démontrer l'existence d'une composante de second ordre et non pas d'obtenir des valeurs exactes permettant des calculs précis éventuels. C'est d'ailleurs pourquoi un simple analyseur de front d'onde de type SH a été utilisé, plutôt qu'un instrument de haute précision qui aurait limité la plage dynamique des mesures.

Dans le cas présent, les mesures ont été prises par intervalle de 1V sur la plage dynamique complète du FDM, soit de -10 V à 10 V. Pour se faire, il a cependant été nécessaire de diminuer le courant au MC à 0,3 A, puisque l'instrument n'était pas en mesure de reconstruire adéquatement le front d'onde passé ±4 V. Un seul actionneur a été sollicité (l'actionneur central) afin d'empêcher toute forme de couplage qui aurait contaminé les données. Les matrices A1 et A2 obtenues par régression sur les mesures entre -10 V et 8 V¹⁸ sont présentées dans les figures 4.5 et 4.6.

51

¹⁸ Les mesures à 9 et 10V ont dû être rejetées, car elles présentaient trop de données aberrantes. L'explication est fournie dans la discussion.



Fig. 4.5 : Coefficients de l'ordre 1 pour chacun des pixels en fonction de leur position à la surface du miroir

L'ordre 1 est ni plus ni moins la FI de l'actionneur central. On peut voir que celui-ci était décentré dans la pupille pendant la mesure. La déformation est de l'ordre de 0,55 μ m/V, totalisant 5,5 V à pleine amplitude.



Fig. 4.6 : Coefficients de l'ordre 2 pour chacun des pixels en fonction de leur position à la surface du miroir

On peut constater la présence d'un effet d'ordre 2 spécifiquement centré sur l'actionneur utilisé. Pour s'en convaincre, il suffit de comparer les figures 4.6 et 4.5 et constater que les deux régions d'intérêt coïncident spatialement. Cet effet non linéaire contribue à la déformation totale à raison d'environ 4,5 nm/V² PV, soit 0,45 μ m à pleine amplitude.

La première chose à constater est qu'il y a bel et bien présence d'un terme à l'ordre 2. De plus, ce dernier peut contribuer jusqu'à 2/3 λ pour une longueur d'onde de 632,8 nm. Ainsi, il est non négligeable si on veut contrôler la surface avec une précision de l'ordre de λ /5. On peut cependant se servir de l'amplitude PV du coefficient d'ordre 2 pour déterminer l'erreur équivalente générée par la méthode RIM¹⁹. En effet, en prenant des mesures régulièrement espacées de $\delta = 1$ [V], l'erreur d'interpolation est $\Delta = a\epsilon - a\epsilon^2$ où a est l'amplitude PV du coefficient d'ordre 2. Puisque $|\epsilon| \leq \delta$ (s'il s'avère que $|\epsilon| > \delta$, la méthode impose de sélectionner une nouvelle plage), il en résulte que $\Delta \leq \frac{a}{4}$, soit environ 1,1 nm au maximum. La régression polynomiale confirme donc l'hypothèse selon laquelle la méthode améliore la génération d'une surface. De plus, selon l'équation de l'annexe 1, il est possible de diminuer δ , soit l'espacement entre les mesures, afin de réduire l'erreur d'échelle à une valeur aussi minime que souhaité.

Concernant les données obtenues, il est clair qu'une meilleure résolution pour les mesures aurait été souhaitable. En effet, la forme du coefficient a de la figure 4.6 est clairement souséchantillonnée. Cependant, il faut garder en mémoire que le seul but de cet exercice était de déterminer un ordre de grandeur et qu'il est raisonnable de croire que ces mesures soient largement suffisantes. Il est cependant difficile déterminer si les coefficients de signe négatif distribués en anneau autour des pixels centrés sur l'actionneur représentent bel et bien la réalité. En effet, selon les équations de la section 2.5, le coefficient d'ordre 2 devrait normalement être une somme de termes au carré multiplié par une quantité positive. Ainsi, la matrice A2 des coefficients d'ordre 2 aurait, selon la théorie, dû être strictement positive. En considérant que plusieurs approximations sont faites et que plusieurs valeurs sont posées pour aboutir à ces équations, il n'est pas surprenant de voir ce genre de différence entre les prédictions théoriques et expérimentales. De toutes les approximations effectuées dans le développement mathématique de ces équations, une apparaît comme étant la cause la plus probable de cette divergence. En effet, le développement repose sur la supposition que la normale est invariante, alors que la réalité est tout autre. De plus, en raison de la forme d'une FI, la variation d'orientation des vecteurs normaux est maximale dans un anneau centré sur l'actionneur, corroborant ainsi les valeurs obtenues.

Enfin, il a été mentionné que le courant au MC avait dû être diminué à 0,3 A pour permettre les mesures. On voit aussi que les régressions ont été faites sur la plage [-10 V, 8 V] plutôt que sur l'ensemble des mesures. Concernant le premier point, il n'influence en rien les coefficients d'ordre 2, puisque ceux-ci dépendent uniquement du courant dans l'actionneur. Pour ce qui est de la régression polynomiale, il a été nécessaire d'exclure les mesures à 9 et 10V, puisqu'elles présentaient trop de bruit et venaient compromettre la qualité des résultats. On peut simplement expliquer ce bruit par l'utilisation prolongée du miroir. En effet, les premières mesures ont été effectuées immédiatement après avoir mis le système en marche et il a été mentionné que le miroir est sensible aux variations de température qui sont nécessairement présentes lors de la mise en marche. Il a été décidé de conserver ces mesures malgré tout, puisque les premières étaient de très bonne qualité.

¹⁹ Voir Annexe 1

4.3 Test couplage

Le test de couplage visait, comme le test de linéarité, à déterminer si un effet de couplage était mesurable et, le cas échéant, quelle est l'amplitude de sa contribution à la surface. En parallèle, il fallait aussi vérifier l'impact de la configuration des éléments de couplage, à savoir s'il était possible de se fier uniquement à une influence moyenne, plutôt que sur une configuration spécifique. Comme il a été mentionné à la section 3.22 b), trois configurations différentes ont été testées. Les figures 4.7 à 4.9 illustrent l'état des actionneurs dans les différentes conditions de couplage.

Dans les trois cas, on regarde l'effet du couplage des PPV sur l'actionneur central (gris). Les choix des valeurs de voltage sont déterminés de façon à produire exactement le même champ au point central de l'actionneur 1²⁰. En effet, la symétrie de la configuration est suffisante pour nous convaincre que le champ doit être identique au point central. Comme ce dernier est proportionnel au courant, l'utilisation de la moitié des actionneurs doit être compensée par un par un doublage des courants d'alimentation. De façon parallèle, on peut constater que les figures 4.7 à 4.9 ne présentent que les 19 premiers actionneurs, alors que le miroir en contient 91. Ce choix est délibéré et, puisqu'aucun des 72 actionneurs manquants n'était utilisé, il vise uniquement la clarté de la présentation.



Fig. 4.7 : Configuration 1 avec tous les PPV à 1V

²⁰ Concernant la numérotation des actionneurs, on peut visualiser la configuration comme une série d'hexagones concentriques considérés comme des *niveaux*. Ainsi, la numérotation s'effectue successivement sur chaque niveau, en partant de l'actionneur central et, pour chacun d'eux, en sens antihoraire à partir de l'actionneur le plus à droite (orienté à 3 h), reproduisant ainsi la logique de la convention généralement utilisée pour les coordonnées polaires.



Fig. 4.8 : Configuration 2 avec les actionneurs 2, 4 et 6 à 2V



Fig. 4.9 : Configuration 3 avec les actionneurs 4, 5 et 7 à 2V

Pour la prise de données, une première mesure référencée de l'influence de l'actionneur 1 a été prise sans couplage afin de servir de témoin. Les analyses préliminaires semblaient indiquer que l'effet de couplage serait d'amplitude très faible et pouvait donc être difficile à identifier par les mesures avec l'analyseur SH. Cette réalité explique le choix de l'interféromètre Zygo pour ces mesures et par conséquent, la limitation à de faibles courants. L'influence mesurable maximale a été obtenue en injectant 2V à l'actionneur central. Les résiduels entre les mesures avec couplage et le témoin sont présentés aux figures 4.10 à 4.12.



Fig. 4.10 : Résiduel expérimental entre la mesure couplée selon la configuration 1 et le témoin



Fig. 4.11 : Résiduel expérimental entre la mesure couplée selon la configuration 2 et le témoin



Fig. 4.12 : Résiduel expérimental entre la mesure couplée selon la configuration 3 et le témoin

À la première observation, on constate immédiatement que les résiduels identifient clairement la position des PPV actifs. Les configurations des figures 4.7 à 4.9 sont très facilement identifiables dans les résiduels. On constate néanmoins que le miroir n'était pas parfaitement orienté au moment des mesures. Il était orienté selon un axe faisant un angle d'environ 20° par rapport à l'axe habituel. Pour ce qui est des dimensions, les différents pixels illustrés sont ceux du détecteur CCD inclus dans le ZYGO. Une méthode simpliste a été utilisée pour déterminer la largeur correspondante sur la surface du miroir : un faisceau d'environ 50 mm de diamètre était imagé sur environ 140x140 pixels. On pouvait donc en déduire un ratio de 0,35 mm/pixel. Ainsi, les mesures présentées correspondent environ à 1 cm² de la surface du miroir.

Maintenant, pour ce qui est des formes observées, il est évident qu'elles ne corroborent pas les hypothèses de départ. Déjà, on voit clairement que l'influence du couplage est extrêmement dépendante de la configuration des PPV. On voit aussi que le résiduel est négatif sur toute la zone de couplage, alors que les équations utilisées prédisaient un résiduel positif. Ensuite, on s'attendait à ce que l'impact soit plus important aux endroits où le champ de l'actionneur central est maximal. Ainsi, le résiduel obtenu aurait, selon cette hypothèse, dû présenter une certaine forme gaussienne, ce qui n'est assurément pas le cas. Enfin, les mesures montrent des contributions par couplage de l'ordre de 250 nm pour seulement 2V. À plein rendement, si le résiduel est linéaire comme le prédisent les équations, cette même erreur serait augmentée à 1 micron, soit autour de 1,5 λ pour un laser HeNe.

Le résiduel négatif est certainement déconcertant, puisqu'aucun élément théorique simple ne peut être utilisé pour expliquer les observations. Le seul outil disponible est le raisonnement physique. La première constatation à faire est que l'effet du couplage de l'actionneur central sur les PPV n'a jamais été considéré. Comme ceux-ci subissent tous le même, il est raisonnable d'envisager qu'un résiduel présentant leur forme soit détecté. Ensuite, on peut interpréter la déformation de la surface au repos comme une reconfiguration spatiale des particules magnétiques. De plus, il est raisonnable de croire que plus l'amplitude de la déformation augmente, plus il devient difficile de déplacer les particules manquantes nécessaires à la formation de la configuration finale. Autrement dit, l'énergie requise pour déplacer une particule augmente au fur et à mesure que son énergie potentielle augmente. Les PPV augmentent ainsi l'énergie potentielle moyenne des particules disponibles pour la déformation centrale. De façon équivalente, l'actionneur central fait augmenter l'énergie potentielle des particules ayant initialement servi à former la déformation causée par les PPV. Ainsi, la surface à l'équilibre subit une diminution généralisée de son amplitude finale, ce qui pourrait expliquer en grande partie les résiduels observés. Il reste que, encore une fois, l'intérêt n'est pas de déterminer les équations analytiques permettant de prédire la surface à l'équilibre, il est seulement d'identifier les contributions non linéaires et leur ordre de grandeur.

Finalement, on peut constater que l'erreur de couplage est plus importante encore que l'erreur d'échelle. Cependant, il est clair que l'obtention de mesures près des conditions de

couplage désirées, servant à la méthode RIM, diminuerait de beaucoup cet effet. Heureusement, la cible utilisée pour la méthode RIM est une cible calculée analytiquement. Par contre, il devient très ardu de construire a priori une LT pouvant servir à toutes les situations envisageables. Il faut donc changer légèrement l'approche. En effet, la LT imaginée devait être construite pendant la phase de calibration précédant l'utilisation du miroir. Or, comme il a été mentionné à la section 4.1, il est nécessaire d'effectuer ces calibrations aussi souvent que possible. Ainsi, il est raisonnable de croire qu'à chaque nouvelle production, une entreprise voudra calibrer de nouveau ses instruments. Ses besoins ponctuels sont donc toujours spécifiques à une surface ou une pièce donnée. Nul besoin de construire une LT complète à chaque fois. Il suffit donc de modifier l'algorithme de calibration automatique préalablement déterminé. Plutôt que de faire une seule phase de mesure et construire une LT complète pour ensuite effectuer une seule phase de calcul, il suffit d'intercaler ces mêmes mesures et calculs dans un ordre approprié, et ce, sans alourdir le procédé. L'idée devient alors de mesurer les FI de chaque actionneur afin de calculer les courants approximatifs. Ceux-ci permettent alors de déterminer les conditions presque exactes de couplage et donc de déterminer les mesures restantes (et surtout les conditions correspondantes) nécessaires à la construction de la RIM. Elles sont ainsi effectuées dans une deuxième phase de mesures et utilisées pour calculer le vecteur courant final recherché. Cette approche est, non seulement, plus précise que celle initialement imaginée, elle permet aussi d'alléger grandement le procédé.

4.4 Test RIM

Le test RIM constitue l'étape ultime pour démontrer la viabilité de l'algorithme proposé. Bien que très simple, cette simple expérience constitue le cœur même de tout le travail. L'idée s'est donc limitée à tenter de produire avec précision une surface cible selon les trois algorithmes présentés à la section 2.6. Les mesures des trois surfaces ont été comparées à la surface cible et la valeur RMS de ces résiduels a été calculée. La cible utilisée est présentée à la figure 4.13.

Cette cible est uniquement celle utilisée pour la section 4.1, avec une symétrie par rapport aux coordonnées en x. Elle représente le genre de défi que rencontrerait n'importe quel fabricant d'optiques de précision, autant par sa forme asymétrique que par son amplitude. Les trois algorithmes ont été utilisés pour des fins de comparaison. Il faut cependant noter que la conclusion de la section 4.3 était déjà sue et donc, l'algorithme amélioré a été utilisé dans ce casci. Les résultats sont présentés à la figure 4.13. Il faut noter que seuls les 19 premiers actionneurs ont été utilisés.

Il semble à première vue que la méthode SOL a été négligée. Cependant, il faut se rappeler que la méthode SCL est simplement constituée de plusieurs itérations SOL. Ainsi, sa première mesure n'est nulle autre que la valeur SOL recherchée. D'un autre côté, la méthode SCL fournit plusieurs valeurs RMS et ne semble pas converger directement vers une valeur unique.



Fig. 4.13 : Cible théorique utilisée pour comparer les algorithmes RIM, SCL et SOL

Les résultats de la cinquième itération ont été présentés afin de montrer que la valeur de la quatrième itération était minimale. À partir de ce point, l'algorithme fait osciller la valeur RMS autour d'un palier et a peu de chance de trouver un nouveau seuil minimal, même avec plusieurs dizaines d'itérations supplémentaires. C'est pourquoi généralement la méthode s'arrête exactement au moment où une itération supplémentaire fait augmenter la valeur RMS. Dans le cas présent, la valeur de résiduel associé à l'algorithme SCL est d'environ 0,38 µm, une amélioration notable par rapport au 0,5 µm associé au SOL. La valeur de 0,45 µm obtenue pour la méthode RIM est donc une amélioration notable, tout en restant une technique en boucle ouverte, ayant réduit pratiquement de moitié l'erreur associée à un algorithme « ouvert ».

Comme il a été spécifié dans la portion théorie, la méthode SCL permet de s'approcher très près des limites physiques du système. Ainsi, il apparait surprenant que la limite physique observée soit si élevée. En effet, 0,38 μ m représente environ $\lambda/2$ à 682 nm, valeur insuffisante pour la plupart des utilisations qui requièrent souvent des résiduels inférieurs à $\lambda/5$, voire $\lambda/10$. Or, le problème observé du bruit dans les mesures d'un instrument donné. En effet, le bruit considéré est celui observé après un référencement avec un appareil. Pour un même miroir, un Zygo présentera un bruit de l'ordre de 50 nm PV, alors que le SH d'Imagine Optics présentera plutôt 100 nm PV. Il faut comprendre que dans ce contexte, le bruit n'a rien à voir avec l'appareil en soi. En effet, un Zygo fait mieux que 50nm PV en général. Le bruit vient plutôt des variations à la surface du liquide qui génèrent une erreur dans les mesures. Ces mêmes variations peuvent ensuite causer des variations plus ou moins grandes d'erreurs dans les algorithmes de reconstruction. On peut aisément constater leur présence en comparant plusieurs mesures de la même surface au miroir. La variabilité des résultats est beaucoup plus importante que pour la mesure d'une surface solide. Cette contribution peut difficilement être soustraite des mesures, puisque l'effet n'est pas constant. Il ne peut pas non plus être moyenné aisément puisqu'il varie trop lentement. Le fait de soustraire une référence à chaque mesure aide certainement, mais si la mesure est prise trop longtemps après la référence, il arrive que la mesure référencée contienne autant des contributions non désirées que la mesure brute. Ainsi, au moment des mesures, un nouveau SH avait été installé pour ses capacités uniques de mesures à fréquence élevée. Les caractéristiques de bases sont toutes extrêmement similaires à celui d'Imagine Optics, mais le bruit des mesures tournait autour de 600 nm PV. Une telle amplitude est difficilement explicable, mais il est clair qu'elle perturbe grandement la valeur RMS minimale obtenue et explique pourquoi l'erreur doit être imputée à l'instrument plutôt qu'au miroir en tant que tel. En contrepartie, elle n'influe pas sur la conclusion selon laquelle la méthode RIM améliore le SOL, puisque toutes les mesures contiennent la même erreur systématique. Ainsi, le comportement relatif reste inchangé et les conclusions tirées peuvent être considérées comme étant fiables.



Fig. 4.14 : Résiduels expérimentaux obtenus par la méthode SCL et RIM

Il y a un dernier point qui mérite d'être mentionné pour clore l'analyse de la technique. Pour des raisons indéterminées, probablement liées à la forte présence de bruit, il a été nécessaire de recommencer l'expérience plusieurs fois. En effet, les premières tentatives étaient faites machinalement, sans trop se soucier justement du bruit sur les mesures. Il a donc fallu user de patience pour en arriver à prendre des mesures de bruit minimal. Pour se faire, il a fallu espacer celles-ci suffisamment longtemps pour que le miroir ne subisse jamais les effets de l'échauffement et chaque mesure était subjectivement choisie lorsqu'elle « semblait » présenter peu de bruit. Il est clair qu'une telle subjectivité est à proscrire pour un technicien d'entreprise, mais il apparait évident qu'un bon choix d'instrument règlerait ce problème. Il reste néanmoins que le seul paramètre arbitraire a été le niveau de bruit et, puisque celui-ci est associé à l'instrument, il est possible de conclure que le FDM utilisé possède la capacité de produire une surface en boucle ouverte, moyennant l'utilisation de la méthode RIM, présentant des résiduels significativement inférieurs à ceux de la boucle ouverte standard. De plus, comme le SOL est fiable généralement

autour de $\lambda/3$ RMS et le SCL à $\lambda/20$ RMS, il est raisonnable de penser que la technique RIM pourrait représenter une fiabilité de $\lambda/5$ RMS, ce qui représente une amélioration notable et suffisante pour intéresser certains fabricants de lentilles de précision.

5 – Conclusion

Le but principal des présents travaux était d'améliorer les techniques en place pour le contrôle du miroir. Le tout devait être fait pour un créneau spécifique de l'optique, à savoir la caractérisation de surfaces par null testing. L'intérêt principal pour cette amélioration touchait principalement les entreprises de fabrication de lentilles de précision. En effet, ceux-ci disposent de peu d'outils permettant de faire les caractérisations requises pour assurer la qualité de leur produit. De plus, les options disponibles sont couteuses, peu versatiles et occasionnent souvent des délais supplémentaires dans le processus fastidieux de conception et production d'une nouvelle composante. La méthode proposée visait principalement à améliorer une méthode de contrôle utilisée pour les FDM de façon à rendre ceux-ci à la fois compétitifs et polyvalents pour plusieurs applications. De façon parallèle, une multitude de vérifications devaient être faites afin de vérifier les hypothèses de départ et surtout de baliser les limites de l'instrument et ses caractéristiques spécifiques. Ainsi, le premier aspect étudié fut la répétabilité des mesures en fonctions de plusieurs paramètres, notamment le temps et les variations physiques de l'instrument au fil des manipulations. Ces dernières pouvaient être limitées aux variations de hauteur de FF utilisé, dues soit à l'évaporation naturelle du liquide porteur le constituant ou à l'imprécision sur la quantité mesurée lors d'un entretien régulier et aux légères variations de courant pouvant survenir entre deux sources distinctes. Ensuite, afin de vérifier les résultats théoriques prédisant la présence de termes non linéaires potentiellement dommageables pour l'acuité des surfaces générées par un processus idéalement linéaire, certaines analyses ont porté sur la détection et la caractérisation de ceux-ci. Il importait de confirmer leur présence et dans quelles mesures ils pouvaient affecter la méthode standard. Ces termes étaient identifiés comme l'erreur d'échelle et l'erreur de couplage et pouvaient être étudiés individuellement. Enfin, la méthode RIM a été confrontée aux méthodes SOL et SCL. La première représente la technique à améliorer et la seconde plus ou moins la limite physique du système.

Les conclusions tirées des tests de répétabilité sont pour le moins rassurantes. Les deux facteurs d'intérêt, les variations de courants et la hauteur de FF dans le miroir, ont été largement exagérés afin de le soumettre à des conditions extrêmes et pouvoir déterminer une limite maximale d'erreurs de répétabilité, tout en permettant d'identifier les variables les plus sensibles, plus susceptibles de causer des problèmes. Du côté des variables physiques, les erreurs typiques étaient déterminées soit à partir d'une valeur nominale fournie par un fabricant, soit à partir d'une estimation raisonnable appuyée sur une utilisation répétée des instruments. L'erreur de courants aux actionneurs a été majorée d'un facteur 5 et celle du courant au MC d'un facteur 2. De son côté, l'erreur sur la hauteur de FF a été majorée d'un facteur 3. Le principal facteur sensible a été identifié comme étant l'erreur de courant au MC. Malgré ces exagérations atypiques de la réalité, la combinaison simultanée de toutes les erreurs maximales a permis de déterminer qu'une surface sera toujours produite à $\pm 2 \,\mu$ m PV. De plus, cette valeur peut facilement être pratiquement annulée par une calibration régulière du FDM. Du côté temporel, une même surface a été générée à répétition sur une période de 40 jours. Il a été possible de déterminer d'entrée de jeu qu'un

Conclusion

entretien hebdomadaire est suffisant pour assurer une répétabilité acceptable des mesures dans un cas d'utilisation modérée. Pour une utilisation soutenue, cet entretien devait plutôt être effectué sur une base bihebdomadaire. De plus, deux manifestations distinctes de la dégradation du FF ont pu être observées. La première, liée à l'évaporation du liquide porteur, a été observée uniquement dans une utilisation à fréquence élevée du miroir. Il a été déduit que cette évaporation causait l'augmentation de la viscosité du liquide et par conséquent du temps de réponse caractéristique. Il a aussi été démontré que cette augmentation causait une erreur plus importante à haute fréquence que la dégradation des particules magnétiques et que l'effet était accentué par une utilisation intensive du miroir causant une dégradation 2,6 fois plus rapide que celle observée lors d'une utilisation modérée. Les mesures à faibles fréquences ont pu mettre en lumière l'effet isolé de la dégradation des particules magnétiques, plus spécifiquement leur agglomération par dégradation graduelle du surfactant les protégeant. Cet effet s'est avéré indépendant de l'intensité d'utilisation du FDM. On conclut de l'ensemble de ces tests qu'un entretien bihebdomadaire suivi d'une calibration suffisait à assurer la guasi-absence d'erreurs liées aux composantes physiques. Un bon choix d'appareils précis aide aussi grandement à réduire les risques d'erreurs.

De son côté, l'analyse des contributions non linéaires a mis au jour certaines difficultés imprévisibles. Les hypothèses étaient tirées d'équations elles-mêmes issues d'une longue démarche mathématique, truffée d'approximations et de considérations variées. Celles-ci se résumaient au final à deux équations. Une première servait à calculer le champ magnétique généré par le miroir et l'autre à déterminer la surface résultante de ces champs. Le traitement algébrique de ces équations a permis de prédire la présence de deux termes non linéaires. Le premier, l'erreur d'échelle, représentait un terme de faible amplitude et généralement négligé proportionnel au carré du courant injecté dans un actionneur. L'appellation « erreur d'échelle » met en évidence la conséquence inévitable de l'annulation d'un faible terme d'ordre 2 : plus la plage d'extrapolation est large, plus sa contribution est importante. L'autre terme présent représentait l'erreur de couplage, causée par la dépendance de l'influence d'un actionneur sur l'état de ses voisins. Dans les deux cas, l'objectif visait à démontrer leur présence dans les mesures et à caractériser leur amplitude maximale. Le terme représentant l'erreur d'échelle a été détecté et représentait une contribution maximale de 2/3 λ PV d'amplitude dans la méthode SOL. La démonstration a été faite que cette contribution pouvait être réduite à une valeur arbitrairement petite à l'aide d'un paramètre libre faisant partie intégrante de l'algorithme RIM : l'intervalle de mesure. Le terme de couplage s'est avéré plus problématique. Il a d'ailleurs contribué à forcer la modification de l'algorithme initialement proposé, mais celle-ci c'est avéré être une amélioration notable, simplifiant les opérations et minimisant les mesures supplémentaires requises. La première proposition d'algorithme RIM requérait une quantité phénoménale de mesures pour éliminer une contribution d'environ 1,5 λ à l'erreur. La deuxième permettait à la fois de diminuer la quantité de mesures à celles requises uniquement, plutôt que la construction d'une LT complète tout en permettant aussi de minimiser l'erreur à une valeur arbitrairement petite. Deux autres observations inattendues sont venues remettre en question la démarche théorique utilisée. La forme et l'amplitude du terme de couplage étaient radicalement différentes des attentes. Dans le

63
Conclusion

premier cas, l'erreur était due au fait de n'avoir pas considéré le couplage réciproque de l'actionneur sur ses PPV. Dans l'autre, une hypothèse portant sur l'énergie requise pour déplacer les particules magnétiques semblait raisonnable pour expliquer les amplitudes négatives du terme de couplage, plutôt que positives tel qu'attendu. Ces tests ont ainsi permis de solidifier l'assise théorique tout en permettant d'améliorer grandement l'algorithme proposé.

Les considérations énoncées ont été extrêmement utiles pour démontrer la viabilité de la méthode en tant qu'alternative à la méthode SOL. En effet, la surface générée présentait une erreur résiduelle pratiquement à mi-chemin entre la méthode SOL et la méthode SCL, considérée comme la limite physique du système. Un bémol a dû être ajouté concernant le choix de l'instrument. En effet, il importe de choisir un instrument offrant une capacité de mesure exempte, le plus possible, de bruit tel que décrit dans l'analyse. Omettre ce fait compromet grandement la viabilité du système. Il a été estimé qu'une fiabilité à $\lambda/5$ était aisément envisageable pour ce genre de configuration.

En définitive, la méthode proposée semble présenter une série d'avantages tout en restant une méthode en boucle ouverte. Elle requiert cependant certains calculs numériques qui peuvent aisément être inclus dans un logiciel automatisé. Pour des résultats optimaux, le choix des instruments est crucial et un entretien régulier est nécessaire. Les caractéristiques uniques de la méthode permettent d'offrir une nouvelle option viable aux entreprises concernées. Certains aspects restent probablement à être optimisés, comme l'augmentation de la viscosité du FF qui pourrait hypothétiquement diminuer les résiduels au détriment des amplitudes de déformation, mais une optimisation correspondant aux besoins spécifiques d'un montage pourrait certainement améliorer les résultats. La méthode est aussi suffisamment générale pour pouvoir s'utiliser avec un système complètement différent d'un FDM, pourvu qu'il soit possible de dégager un comportement pratiquement linéaire sur un certain intervalle de valeurs.

Annexe 1

Annexe 1

Théorème : La méthode RIM appliquée à une fonction quadratique $f(x) = ax^2 + bx + c$ entre deux points x_0 et $x_0 + \delta$, donne une erreur $\Delta = g(x_0 + \epsilon) - f(x_0 + \epsilon) = a\delta\epsilon - a\epsilon^2$, où g(x) est la fonction linéaire créée par la méthode.

Preuve: g est définie comme étant : g(x) = mx + k, où m est donnée par $\frac{f(x_0+\delta)-f(x_0)}{(x_0+\delta)-(x_0)}$.

Or,

$$f(x_0 + \delta) = a(x_0 + \delta)^2 + b(x_0 + \delta) + c = a(x_0)^2 + (2a\delta + b)x_0 + (a\delta^2 + b\delta + c)$$

et

$$f(x_0) = a(x_0)^2 + bx_0 + c$$

Ainsi,

$$m = \frac{1}{\delta}(2a\delta x_0 + a\delta^2 + b\delta) = 2ax_0 + a\delta + b$$

De plus, la méthode impose que $g(x_0) = f(x_0)$. Cette égalité permet de fixer k dans g(x):

$$(2ax_0 + a\delta + b)x_0 + k = a(x_0)^2 + bx_0 + c$$

d'où on tire :

$$k = -a(x_0)^2 - a\delta x_0 + c$$

On réécrit donc :

$$g(x) = (2ax_0 + a\delta + b)x - a(x_0)^2 - a\delta x_0 + c$$

Enfin, l'erreur Δ se trouve directement par $g(x_0 + \epsilon) - f(x_0 + \epsilon)$:

$$\Delta = \left[(2ax_0 + a\delta + b)(x_0 + \epsilon) - a(x_0)^2 - a\delta x_0 + c \right] - \left[a(x_0 + \epsilon)^2 + b(x_0 + \epsilon) + c \right]$$
$$\Delta = a\delta\epsilon - a\epsilon^2$$

Ce qui complète à preuve •

Table des Acronymes

AO : Adaptive Optic (Optique adaptative)

CGH : Computer Generated Hologram (Hologramme généré par ordinateur)

FDM : Ferrofluid Deformable Miror (Miroir déformable à base de ferrofluide)

FF : Ferrofluide

FI : Fonction d'Influence

LSP : Least Square Problem (Problème des moindres carrés)

LT : Lookup table (Table de reference)

MC : Maxwell Coil (Bobine de Maxwell)

PPV: Plus Proche Voisin

PZ : Polynôme de Zernike

RIM: Reconstructed Influence Matrix (Matrice d'influence reconstruite)

SCL : Standard Closed Loop (Boucle fermée usuelle)

SH : Shack-Hartmann

SNR : Signal-Noise Ratio (Ratio signal sur bruit)

SOL : Standard Open Loop (Boucle ouverte usuelle)

TLSP : Total Least Square Problem (Problème complet du moindre carré)

Bibliographie

- H. W. Babcock, «The Possibility of Compensating Astronomical Seeing,» Publications of the Astronomical Society of the Pacific, vol. 65, n° 386, p. 229, 1953.
- [2] J. W. Hardy, «Active optics : a new technology for the control of light,» Proc. IEEE 66, pp. 651-697, 1978.
- [3] H. W. Babcock, «Adaptive optics revisited,» Science 249, pp. 253-257, 1990..
- [4] R. K. Tyson, Principles of Adaptive Optics, Boston: Academic Press, 1991.
- [5] J. W. Hardy, Adaptive Optics for Astronomical Telescopes, Oxford University Press, 1998.
- [6] D. M. Alloin et J. M. Mariotti, Adaptive Optics for Astronomy, Kluwer Academic Publishers, 1994.
- [7] R. K. Tyson, Adaptive Optics Engineering Handbook, Marcel Dekker, 1999.
- [8] X. Hu and al., «Ionic liquid based variable focus lenses,» Soft Matter, vol. 7, n° 13, pp. 5941-5943, 2011.
- [9] E. F. Borra, «The liquid-mirror telescope as a viable astronomical tool,» *Journal of the Royal* . *Astronomical Society of Canada*, vol. 76, n° 4, pp. 245-256, 1982.
- [10] D. Brousseau, E. F. Borra, M. Rochette et D. B. Landry, «Linearization of the response of a 91actuator,» Optic express, vol. 18, n° 8, pp. 8239-8250, 2010.
- [11] P. Laird, E. F. Borra, R. Bergamesco, J. Gingras, L. Truong et A. Ritcey, «Deformable mirrors based on magnetic liquids,» chez *Proceedings of SPIE 5490*, 2004.
- [12] J. Parent, «Dynamique des miroirs déformables liquides à base de ferrofluides,» Université Laval, Québec, 2008.
- [13] M. Wang et S. Thibault, «CGH null design and fabrication for CFH telescope simulator,» chez Proceedings of SPIE 5252, 2003.
- [14] J. Van de Vegte, Feedback Control Systems, 3rd Edition, Prentice-Hall, 1993.
- [15] G. J. Silva, A. Datta et S. R. Bhattacharyya, PID Controllers for Time-Delay Systems, Boston: Birkhäuser, 2005.

- [16] K. J. Astrom et T. Hagglund, PID Controllers : Theory, Design, and Tuning, 2nd Edition, Instrument Society of America, 1995.
- [17] A. Iqbal et J. B. Amara, «Modeling of a Magnetic-Fluid Deformable Mirror for Retinal Imaging Adaptive,» International Journal of Optomechatronics, vol. 1, n° 2, pp. 180-208, 2007.
- [18] R. A. Horn et C. A. Johnson, Matrix analysis, New York: Cambridge University Press, 1985.
- [19] G. H. Golub et C. F. Van Loan, «An analysis of the total least square problem,» SIAM Journal of numerical analysis, vol. 17, n° 6, pp. 883-893, 1980.
- [20] J. C. Wyant et C. Creath, «Basic Wavefront Aberration Theory for Optical Metrology,» chez APPLIED OPTICS AND OPTICAL ENGINEERING, VOL. XI, vol. XI, Academic Press inc., 1992, pp. 1-53.
- [21] C. E. Nanjundappa et I. S. Shivakumara, «Effect of Velocity and Temperature,» Journal of Heat transfer, vol. 130, n° 10, 2008.
- [22] J. C. Maxwell, Treatise on Electricity and Magnetism, Oxford: The Clarendon Press, 1873.
- [23] M. Rioux, «Miroirs déformables à base de ferrofluide,» Université Laval, Québec, 2006.
- [24] M. d'Aquino, G. Miano, C. Serpico, W. Zamboni et G. Coppola, «Forces in Magnetic Fluids Subject to Stationary Megnetic Fields,» *IEEE Transactions on magnetics*, vol. 39, n° 5, pp. 2657-2659, 2003.
- [25] B. R. Munson, D. F. Young et T. H. Okiishi, Fundamentals of Fluid Mechanics, Hoboken: Wiley, 2002.
- [26] R. E. Rosensweig, Ferrohydrodynamics, New-York: Dover, 1997.
- [27] S. Bobbio, Electrodynamics of Materials, San Diego: Academic Press, 2000.
- [28] K. Kuns, «Calculation of magnetic field inside plasma chamber,» 2007. [En ligne]. Available: http://plasmalab.pbworks.com/w/file/fetch/17178415/bfield.pdf. [Accès le 04 avril 2012].
- [29] D. Malacara, M. Servín et Z. Malacara, Interferogram Analysis for Optical Testing, 2nd Edition, Boca Raton: Taylor & Francis Group, 2005.
- [30] C. L. Lawson et R. J. Hanson, Solving least squares problems, Englewood Cliffs: Prentice-Hall, 1974.

