



HAL
open science

Observateurs asymptotiques invariants : théorie et exemples

Silvère Bonnabel

► **To cite this version:**

Silvère Bonnabel. Observateurs asymptotiques invariants : théorie et exemples. Mathématiques [math]. École Nationale Supérieure des Mines de Paris, 2007. Français. NNT : 2007ENMP1590 . pastel-00004868

HAL Id: pastel-00004868

<https://pastel.hal.science/pastel-00004868>

Submitted on 6 Mar 2009

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

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



ED n°.... : ... (nom de l'école doctorale)

N° attribué par la bibliothèque

□□□□□□□□□□□□□□□□

T H E S E

pour obtenir le grade de
Docteur de l'Ecole des Mines de Paris
Spécialité "Mathématiques et Automatique"

présentée et soutenue publiquement par
Silvère BONNABEL

le 27/09/2007

<p>OBSERVATEURS ASYMPTOTIQUES INVARIANTS : THEORIE ET EXEMPLES</p>

Directeur de thèse : Pierre ROUCHON

Jury

M.	Ivan KUPKA	Président
M.	Yacine CHITOUR	Rapporteur
M.	Claude SAMSON	Rapporteur
M.	Philippe MARTIN	Examinateur
M.	Rodolphe SEPULCHRE	Examinateur
M.	Marc SOUCHE	Examinateur

Observateurs asymptotiques et symétries : théorie et
exemples.

SILVÈRE BONNABEL

Résumé Cette thèse a pour objet la construction d'estimateurs non-linéaires à base d'observateurs asymptotiques. Dans un premier temps nous développons un observateur destiné à estimer des concentrations en réactifs dans un réacteur de polymérisation du groupe TOTAL. Dans un deuxième temps, nous nous posons des questions d'ordre plus théorique sur l'utilisation des symétries dans la conception d'observateurs non-linéaires.

Le réacteur que nous avons considéré est un réacteur de polymérisation haute pression qui produit des polymères de type plastique composés de deux ou trois monomères. L'estimation des concentrations en certains réactifs dans les différentes zones du réacteur repose sur la modélisation de la réaction. Le modèle consiste principalement un bilan de matière, d'énergie, et l'utilisation de modèles de cinétique chimique. Ensuite via les équations du modèle et les mesures de températures et débits, on remonte aux concentrations en temps réel. L'estimateur construit est non-linéaire, et la convergence repose sur une structure triangulaire. Cet estimateur a été installé et validé sur l'unité industrielle.

La convergence de l'estimateur envisagé est indépendante du choix des unités dans lesquelles on écrit les bilans. Nous nous sommes interrogés sur la possibilité, quand on fait un observateur pour les concentrations, d'écrire des termes de correction indépendants des unités. A cet effet nous avons considéré un exemple plus académique : un réacteur chimique exothermique, pour lequel on mesure les températures et les débits, et où la cinétique chimique est d'ordre 1. On veut estimer certaines concentrations, et l'on souhaite que les propriétés de convergence soient indépendantes des unités. Cette étude a montré qu'une approche basée sur les symétries pouvait suggérer des termes de correction non-linéaires puis des changements de variables propices à l'étude de la convergence globale, pour des observateurs dont la forme est du type observateur de Luenberger ou filtre Kalman Étendu.

Ensuite, nous avons développé une méthode générale pour écrire de manière systématique les termes de corrections qui préservent les symétries. La contribution théorique principale de la thèse est de donner une méthode pour construire tous les termes de corrections non-linéaires qui préservent les symétries. On remet ensuite en perspective la notion d'erreur entre l'état et son estimée en proposant la notion d'erreur invariante. La dynamique de cette erreur invariante possède alors des propriétés fort intéressantes. En particulier, elle est indépendante de la trajectoire du système pour un système invariant à gauche sur un groupe de Lie. On applique alors cette nouvelle théorie des observateurs invariants à principalement trois exemples, un réacteur chimique pour lequel on construit un observateur globalement convergent, un exemple de voiture non-holonome où l'on construit un observateur presque globalement convergent, et un exemple emprunté au domaine de la navigation inertielle aidée par des mesures de vitesse pour lequel on obtient la convergence locale autour de toute trajectoire et tel que le comportement global de l'erreur est indépendant de la trajectoire et des entrées.

Bien que nous n'ayons abordé en détail que les situations où la dimension du groupe reste inférieure à celle de l'état, il est très naturel d'envisager de traiter avec une approche analogue des cas plus généraux. Tel est l'objet de la dernière partie de la thèse avec quatre

exemples. La synthèse d'observateur réduit pour une classe de système Lagrangien dont on mesure la position : le groupe de transformation est celui des changements de coordonnées sur l'espace de configuration. Les modèles de type Saint-Venant qui interviennent dans les modèles océanographiques : l'espace d'état est de dimension infinie car les modèles sont à base d'équations aux dérivées partielles. La fusion de données en navigation inertielle avec comme mesure une image et donc une sortie de dimension infinie. Enfin, l'estimation paramétrique d'un système quantique à deux états où la dimension du groupe est un peu plus grande que celle de l'état.

Summary This thesis is about the construction of non-linear estimators of the asymptotic observers type. We will first build an observer in order to estimate the concentrations in reactive in a polymerization reactor of TOTAL. Then we will pursue theoretical questions about the use of symmetries for the design of non-linear estimators.

The chemical reactor we worked on is a high pressure polymerization reactor which produces plastics that are polymers made of two or three monomers. The estimation of the concentrations in each reactive in the several zones of the reactor is based on a model of the reaction. The model consists of a mass balance, an energy balance, and the use of a chemical kinetics model. Thanks to the equations of the model, and to the measurements of temperature and flows, we give a real-time estimation of the concentrations. It is a non-linear estimator, and the convergence is based on contraction properties. This estimator was implemented and validated on the industrial unit.

The convergence of the estimator is independent of the choice of the units with which the balances are written. We wondered if it was always possible, when one builds an observer for the concentrations of the Luenberger or extended Kalman filter type, to write correction terms which do not depend on the units. We thus considered a more academic example : a chemical exothermic reactor, for which the temperature and flows are measured, and the chemical kinetics is of order one. We want to estimate the concentrations, and we want the convergence properties to be independent of the physical units. This study showed that an approach based on symmetries could suggest non-linear correction terms, and some change of variables which help when it comes to studying global convergence of observers of the Luenberger or extended Kalman filter type.

Then we developed a general method in order to write correction terms which systematically preserve the symmetries of the system. The principal theoretical contribution of the thesis is to give a precise method to write all the correction terms which preserve the symmetries. The notion of the error between the true state and the estimated state is re-explored via the notion of invariant state-error. The invariant state error dynamics has very interesting properties. In particular it is independent of the trajectory for a left-invariant trajectory on a Lie group. We apply the theory of invariants observers to mainly three examples, a chemical reactor for which we build a non-linear globally convergent observer, an example of a non-holonomic car for which we build an almost globally convergent ob-

server, and an example of inertial navigation assisted by velocity measurements for which we get the local convergence around any trajectory and such that the global behavior of the error does not depend neither on the trajectory nor on the inputs.

Although the theory deals with the cases where the dimension of the symmetry group is smaller than the dimension of the state, it seems very natural to use a similar method to deal with the general case. This is the topic of the last part of the thesis where we look at four examples. The synthesis of reduced observer for a class of lagrangian systems such that all the positions are measured : the transformation group consists of all the change of coordinates on the configuration space. The models of the Saint-Venant type on which the Shallow-water model are based, and which are used in oceanography : the dimension of the state space is infinite since the models use partial differential equations. The data fusion in inertial navigation for which the measurement is an image, and thus is dimension of the output is infinite. Finally, the parametric estimation of a two-states quantum system for which the dimension of the group is bigger than the dimension of the state.

Remerciements

Je remercie Ivan Kupka président du jury, Yacine Chitour et Claude Samson rapporteurs de ma thèse ainsi que les examinateurs Philippe Martin, Rodolphe Sepulchre et Marc Souche.

Je tiens à remercier aussi très chaleureusement mon directeur de thèse Pierre Rouchon pour le soin qu'il a apporté à l'encadrement du doctorat dans son ensemble. J'ai été très honoré de pouvoir travailler à ses côtés et bénéficier de ses conseils.

Je remercie François Chapelais, Jean Levine, Philippe Martin, Nicolas Petit et Laurent Praly pour leur constante disponibilité pour répondre à mes diverses questions et pour les nombreuses discussions que nous avons eues, qu'elles soient relatives au travail de thèse, ou d'ordre plus général. Je remercie également Madame Legallic qui a pu faire en sorte que tout soit le plus simple pour moi dans l'organisation des déplacements.

Je remercie encore Marc Souche pour avoir suivi le travail de thèse, fait le lien avec les divers événements liés au groupe TOTAL, et m'avoir aidé sur la partie informatique industrielle.

Je remercie beaucoup Gwenvaël Le Seac'h d'Arkema pour sa patience devant ma naïveté et mes nombreuses interrogations liées au processus, et pour sa collaboration sur la partie industrielle du travail de thèse. Je le remercie d'avoir toujours permis que je sois reçu lors des visites sur site dans des conditions optimales.

Mes remerciements iront aussi à Pascal Schaefer pour son aide très bienvenue quant à la fiabilisation du programme d'estimation d'un point de vue informatique industrielle.

Je remercie Philippe Canin et Jacqueline Jeanjean du site de Carling.

Je remercie enfin Pascal Fourcade et Philippe Lagardère pour leurs utiles aides ponctuelles.

Table des matières

1	Introduction	1
I	Capteur logiciel pour un réacteur de polymérisation	5
2	Réacteur HP EDA à Carling	9
2.1	Description du procédé	9
2.2	Modèle dynamique	12
2.3	Conclusion : modèle dynamique retenu	16
3	Observateur	21
3.1	Principe de l'observateur	22
3.1.1	Cas Homopolymère	22
3.1.2	Cas copolymère	23
3.1.3	Cas terpolymère	26
3.1.4	Stabilité avec la recycle	28
3.1.5	Récapitulatif des mesures et des grandeurs estimées	30
3.2	Résultats	31
4	Conclusion	37
II	Observateurs et symétries	39
5	Un réacteur exothermique	41
5.1	Symétries	42
5.2	Convergence globale	45
6	Théorie des observateurs invariants	49
6.1	Introduction	49
6.2	Systèmes invariants, observateurs et erreurs invariants	51
6.2.1	Système invariant et sortie compatible	51

6.2.2	Hypothèses de base	52
6.2.3	Méthode du repère mobile, champs de vecteurs invariants, coordonnées de base et de fibre	52
6.2.4	Caractérisation des pré-observateurs invariants	56
6.2.5	Pre-observateurs invariants : une méthode constructive	61
6.2.6	Erreur d'état invariante et convergence	61
6.3	Observateurs invariants sur un groupe de Lie	62
6.4	Observateurs invariants sur un groupe de Lie dans un cas particulier	65
6.5	Conclusion	70
7	Exemples	71
7.1	La voiture non-holonome	72
7.2	Navigation inertielle avec mesure de vitesse	76
7.2.1	Un pré-observateur invariant	79
7.2.2	Simulations	83
8	Luenberger, Kalman et invariance	91
8.1	Point d'équilibre	92
8.2	Trajectoires permanentes	92
8.3	Observateur invariant et filtre de Kalman étendu	96
8.3.1	Introduction	96
8.3.2	Bruits et symétries	98
8.3.3	Filtre de Kalman étendu invariant à gauche	99
8.3.4	Exemple	104
8.3.5	Simulations	107
8.3.6	Conclusion	107
9	Observateurs invariants en temps discret	111
9.1	Systèmes invariants, sorties compatibles et observateurs invariants	111
9.2	Construction et caractérisation des observateurs invariants discrets	112
9.3	Erreur d'état invariante et convergence	113
III	Extensions	115
10	Système quantique	119
10.1	Dynamique et invariances	120
10.1.1	Dynamique	120
10.1.2	Invariances	121
10.2	Un observateur invariant	121
10.2.1	Interprétation géométrique avec la sphère de Bloch	122

10.2.2	Diverses dimension et normalisation	123
10.3	Convergence	124
10.4	Simulations	130
10.5	Conclusion	131
11	Nudging invariant	133
11.1	Modèle Shallow Water	133
11.2	Invariance par permutation d'indices	134
11.3	Invariances des équations	135
11.4	Nudging invariant	135
11.5	Etude du système linéarisé	137
12	Vidéo en navigation inertielle	141
12.1	Hypothèses	141
12.2	Modèle cinématique	142
12.3	Environnement infiniment loin et calage des gyroscopes	144
12.4	Conclusion	148
13	Un observateur réduit intrinsèque	149
13.1	Introduction	149
13.2	Un observateur réduit intrinsèque	151
13.2.1	Lemme	152
13.2.2	L'observateur réduit	153
13.2.3	Cas où la courbure est positive : simulations sur la sphère	155
14	Conclusion	159
IV	Annexes	163
A	Cinétique de polymérisation	165
A.1	Justification théoriques des modèles de cinétique chimique	165
A.2	Justification expérimentale des paramètres des modèles	167
B	Quaternions	169
C	Observateur et équation d'erreur autonome : Un exemple	171

Chapitre 1

Introduction

L'objet de ce mémoire est l'estimation de quantités non directement mesurées par des observateurs asymptotiques (capteurs logiciels). Pour les systèmes linéaires observables, il existe des méthodes de construction d'observateur générales avec les travaux de Kalman et Bucy [30] et aussi ceux de Luenberger [35]. En effet on considère le système linéaire de la forme

$$\frac{d}{dt}x = Ax + Bu, \quad y = Cx$$

où $x \in \mathbb{R}^n$ est l'état du système, $u \in \mathbb{R}^m$ les entrées, et $y \in \mathbb{R}^p$ la sortie ; A est une matrice $n \times n$, B une matrice $n \times m$, et C une matrice $p \times n$. Les observateurs du type filtre de Kalman, ou Luenberger s'écrivent sous la forme

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu + L(C\hat{x} - y)$$

L'observateur consiste en une copie de la dynamique et un terme de correction $L(C\hat{x} - y)$ qui dépend des mesures et est nul quand $\hat{x} = x$. Le filtre de Kalman [30] donne un réglage automatique du gain $L(t)$ qui dépend alors du temps t . Si le système est observable, on peut choisir aussi le gain L constant de manière à ce que les valeurs propres de $A + LC$ soient à partie réelle strictement négative, et l'observateur fournit une estimation de x dans le sens où $\hat{x} - x \rightarrow 0$ exponentiellement quand $t \rightarrow \infty$ puisque l'on a

$$\frac{d}{dt}(\hat{x} - x) = (A + LC)(\hat{x} - x)$$

Pour les systèmes généraux fondés sur un modèle non-linéaire sous forme d'état avec les mesures données par une loi de sortie, qui s'écrivent donc sous la forme générale

$$\frac{d}{dt}x = f(x, u), \quad y = h(x)$$

où f et h sont des applications régulières, on ne dispose pas à l'heure actuelle de méthode systématique pour construire des observateurs asymptotiques comme c'est le cas pour les systèmes linéaires observables. Cependant, pour certains types de non linéarités, des méthodes existent. Pour les systèmes dits linéarisables par injection de sortie [27] on se ramène au cas linéaire où les méthodes usuelles s'appliquent. Pour les systèmes ayant une forme normale d'observabilité uniforme de structure triangulaire dans les bonnes coordonnées, les observateurs à grands gains sont bien adaptés [23]. A notre connaissance, il n'existe pas de travaux théoriques qui, de manière systématique, exploitent les symétries pour la synthèse d'observateurs asymptotiques. L'une des contributions de cette thèse est de combler cette lacune pour les systèmes invariants par rapport à l'action d'un groupe de symétries. Schématiquement, ce travail contient une partie appliquée, où l'on décrit une étude, effectuée pour le groupe TOTAL, qui consistait à élaborer un capteur logiciel pour un réacteur de polymérisation, et une partie théorique, où l'on explore les liens entre observateurs et symétries.

Le réacteur TOTAL considéré est un réacteur chimique de polymérisation Haute-Pression, situé sur le site de Carling en Moselle. Il produit des polymères du type polyéthylène qui sont des matières plastiques. Certains polymères produits sont issus d'une polymérisation de l'éthylène avec un ou deux autres monomères. Ce sont des plastiques spécifiques, dont certains ne sont produits nulle part ailleurs. Dans ce cas, il y a dans le réacteur plusieurs espèces chimiques. Les opérateurs pilotent l'unité et contrôlent la production notamment via l'injection des divers réactifs. Notre but est de faire un capteur logiciel qui permette d'estimer la concentration en chacun des réactifs à plusieurs endroits du réacteur. Cela permet de donner aux opérateurs une image interne de la réaction et ainsi de les aider à piloter l'unité. La démarche classique que nous avons adoptée pour "remonter" aux compositions internes à partir des mesures de débits, de pression, et de température a été de construire un observateur. Pour cela nous avons établi un modèle simple de la réaction, qui s'appuie sur un bilan de matière, et un bilan thermique. Le modèle est complété par un modèle de cinétique chimique. En injectant la mesure temps réel de température dans le modèle, et en simulant les équations (observateur avec injection de sortie) on obtient un système qui converge et fournit donc une estimation des compositions internes en temps réel. On peut déduire de cette estimation la composition du polymère final, qui est une donnée "qualité" sur le produit final. L'observateur a été implémenté sur site industriel.

Pour implémenter le capteur logiciel, il a fallu choisir les unités dans lesquelles on écrit les équations de l'observateur, et adapter la structure non-linéaire des modèles (de cinétique) à ce choix. Ces adaptations sont des opérations mathématiques sur le modèle indépendantes de la physique du système, et l'on voulait que la convergence de l'observateur ne soit pas affectée par le choix des unités - alors que c'est le cas en général pour les observateurs du type filtre de Kalman, filtre de Kalman étendu, observateurs de Luenberger. Nous avons donc repris les questions posées et les résultats de [1], et nous avons considéré le problème de la construction d'observateurs indépendants du choix des unités pour les

réacteurs chimiques exothermiques. Puis nous avons considéré plus généralement le problème de la construction d'observateurs qui respectent les invariances (ou symétries) des systèmes. Nous avons alors développé et complété la théorie des observateurs invariants qui était suggérée dans [1]. On appelle observateurs invariants les observateurs qui respectent les symétries. La contribution théorique principale de la thèse est de compléter la méthode de construction des observateurs, puis de caractériser les observateurs invariants -ce qui permet d'avoir une méthode systématique en plusieurs étapes de construction de tous les observateurs invariants possibles; et enfin de remettre en perspective la notion d'erreur d'état quand le système présente des symétries. Il est apparu dans les exemples qu'on a traités qu'utiliser les symétries dans la manière dont on définit l'erreur simplifiait notablement l'étude de la convergence et le réglage des gains.

Nous avons cherché à appliquer la théorie à des exemples qui proviennent de l'ingénierie. La classe des systèmes mécaniques se prête particulièrement bien au problème car on bénéficie en général des invariances galiléennes. C'est la raison pour laquelle nous sommes intéressés aux systèmes mécaniques. Quand la dimension du groupe de symétries est la même que celle du système, on peut donner un sens intrinsèque à l'erreur d'estimation sur l'état. Nous avons obtenu pour un système non-linéaire, et dont les entrées dépendent du temps - problème qui provient du domaine de la navigation inertielle assistée par mesure de vitesse - une équation d'erreur autonome, et donc un comportement de l'erreur d'estimation indépendant de la trajectoire de l'état et aussi des entrées.

Nous proposons une théorie des observateurs invariants pour une classe importante de systèmes qui présentent des symétries : l'espace d'état et des entrées y sont de dimension finie, et la somme des dimensions y est plus grande que le nombre de degrés de liberté du groupe de symétries. Nous avons essayé cependant de confronter cette idée d'observateurs invariants à des exemples qui sortent de cette classe. Il apparaît que la notion d'observateur invariant reste un guide pour aborder une classe plus large d'exemples, et notamment les systèmes décrits par des équations aux dérivées partielles, où l'état est de dimension infinie.

Le manuscrit comporte trois parties :

- Dans la première partie, on rendra compte du travail effectué pour le groupe TOTAL, à savoir la construction d'un capteur logiciel pour un réacteur chimique industriel.
- Dans la deuxième partie, on développera la théorie des observateurs invariants et on l'appliquera à des exemples issus de la chimie et de la mécanique.
- Dans la troisième partie, on abordera la construction d'observateurs invariants pour divers problèmes d'estimation qui ne rentrent pas dans le cadre de la théorie développée en seconde partie (groupe de symétries de dimension infinie, système gouverné par des équations aux dérivées partielles). On montrera l'intérêt de prendre en compte les symétries pour l'estimation et la fusion de données.

Première partie

Capteur logiciel pour un réacteur de polymérisation

Dans cette partie nous présentons le travail qui a été réalisé pour TOTAL PETROCHEMICALS. L'unité concernée est l'unité polyéthylène haute pression de l'usine de Carling. Nous nous intéressons en particulier au réacteur où l'éthylène (mélangé à quelques autres monomères) est transformé en polymère du type matières plastiques. Nous le modélisons comme une cascade de réacteurs exothermiques parfaitement agités. L'unité est pourvue de plusieurs capteurs qui mesurent en particulier certains débits et les températures. Notre travail a consisté à réaliser un capteur logiciel, qui s'appuie sur un modèle basé sur les lois de conservation de la physique, et qui à partir des mesures de débit et de température, permet de remonter en temps réel aux compositions internes. Il donne ainsi une image interne de la réaction. On peut alors en déduire en particulier la composition (et donc la qualité) du polymère formé. L'algorithme a été installé sur site. Il permet d'améliorer la conduite de l'unité, puisqu'il fournit aux opérateurs des indications quantitatives sur la réaction. Ainsi, on n'a pas besoin d'attendre les mesures laboratoire (tardives) pour corriger les paramètres opératoires de façon à produire la qualité recherchée. Le travail a été mené en collaboration avec :

Gwenvael LE SEAC'H	ARKEMA Mont
Pascal SCHAEFER	TOTAL Lyon
Marc SOUCHE	TOTAL Lyon
Thomas ROUSSEL	Groupement de Recherche de Lacq
Jacqueline JEANJEAN	TOTAL PETROCHEMICALS Carling

De nombreuses visites sur site ont été effectuées : à Carling, où se situe le réacteur, pour interagir avec les exploitants du site et installer l'algorithme d'estimation. Egalement à Lacq où est situé le Groupement Recherche de Lacq (GRL) pour ce qui concerne la validité du modèle de cinétique chimique.

Dans le chapitre 2, on présente l'unité, le réacteur, et le procédé. On présente le modèle du réacteur, qui est basé sur les lois de conservations de l'énergie et de la matière, et qu'on a souhaité être le plus simple possible. Il fait aussi intervenir un modèle de cinétique chimique qui prend différentes formes selon les types de qualités produites. Ces descriptions sont soumises à des impératifs de confidentialité qui limitent leur précision.

Dans le chapitre 3, on présente l'observateur. On donne une preuve de convergence pour certains grades. On présente finalement les résultats. On ne peut valider qu'indirectement le modèle, en recoupant les résultats fournis par l'observateur, avec les mesures d'une part de production massique de polymère, et d'autre part de composition du polymère.

Pour la modélisation du réacteur on s'est appuyé sur les modélisations usuelles pour les cascades de réacteurs parfaitement agités (par exemple [54, 5]). Pour le modèle cinétique, on s'est appuyé sur [50, 41, 25]. De nombreuses publications sont consacrées à la détermination expérimentale des coefficients du modèle cinétique en fonction des différents types de plastiques produits ([8, 36, 16, 17, 19, 18]...).

Chapitre 2

Réacteur HP EDA à Carling

Le réacteur considéré est du type autoclave Haute Pression (2000 bars environ). Le réacteur est dit autoclave car il est homogénéisé, zone par zone, en composition et température par agitation et fonctionne en mode "adiabatique", c'est-à-dire qu'il n'y a pas d'échange avec des zones d'eau comme en tubulaire. La réaction est exothermique et le mélange réactionnel est refroidi par de l'éthylène frais injecté en divers endroits.

Les développements ci-dessous concernant les bilans enthalpiques et massiques s'appuient sur un rapport à diffusion restreinte de N. Aghannan, C. Lambert, et N. Petit pour TOTAL daté du 8 janvier 2003.

2.1 Description du procédé

L'unité polyéthylène de Carling comporte trois lignes de production. Nous nous intéressons à la ligne 41 qui produit du polyéthylène Haute Pression, et qui est la seule à produire également des plastiques qui sont les lotryls et lotaders. Ils sont regroupés sous le nom de Ethylène Dérivés Acryliques (EDA).

- Les lotryls sont des copolymères de l'éthylène avec de l'Acrylate d'Ethyle, de Butyle ou de Méthyle.
- Les lotaders sont des terpolymères de l'éthylène avec un des précédents acrylates, et de l'Anhydride Maléique (MAH), ou du Méthacrylate de Glycidile (MAGLY).

La gamme comprend une quinzaine de qualités. La qualité est déterminée par la composition du polymère en chacun des monomères, la densité du polymère, et l'indice de fluidité (IF) du polymère (appelée aussi Melt Index) qui est fonction de la longueur des chaînes de polymère. La production de la ligne 41 est de plusieurs kilo-tonnes par an.

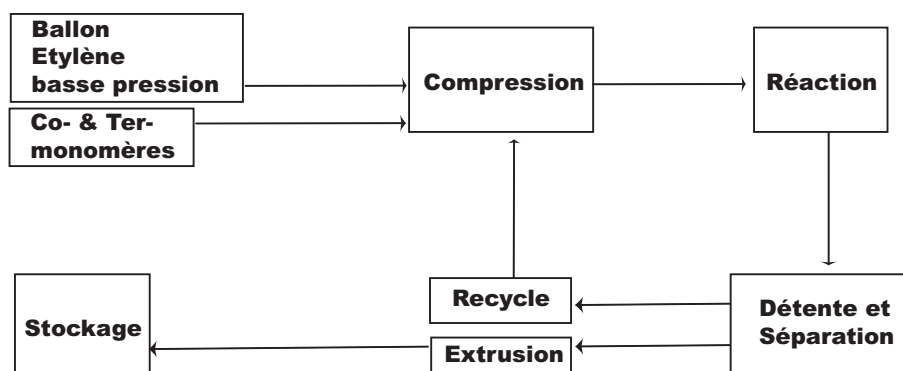


FIG. 2.1 – Procédé autoclave PE EDA Carling

Le procédé

L'éthylène après avoir été compressé, et mélangé aux autres monomères, est injecté en divers endroits du réacteur sous forme de nappes. La partie fluide du mélange réactionnel est monophasique et à l'état supercritique. A la sortie du réacteur, le mélange est détendu et les particules de polymère "précipitent". La part du mélange qui n'a pas réagi est recyclée et ré-injectée en entrée du réacteur.

Le réacteur

Il est le siège de la réaction de polymérisation et constitue le coeur du procédé. Il est assimilé à un ensemble de plusieurs réacteurs parfaitement agités dans chaque zone et à la suite les uns des autres en série. La température est mesurée, et elle est contrôlée par l'injection de catalyseur (régulateur proportionnel/integral (PI)). Les seuls échanges d'énergie se font via les flux de matière entrée-sortie (par abus de langage on dit le réacteur "adiabatique"). La réaction est exothermique. Le mélange réactionnel est refroidi par les nappes qui contiennent de l'éthylène frais et éventuellement des comonomères, termonomères, et additifs.

Mesures et données physico-chimiques

Les mesures et données relatives à une zone et utiles à la modélisation dynamique de la réaction sont résumées dans le tableau 2.1. On pourra se référer à la figure 2.1. Problème

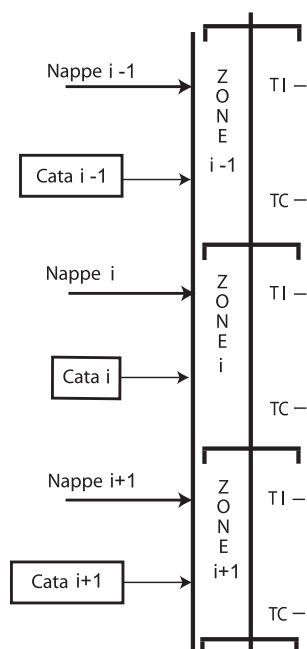


FIG. 2.2 – Schéma simplifié de 3 zones consécutives du réacteur

Zone i		
<i>Notation</i>	<i>Description</i>	<i>unités</i>
P^i	Pression	Bar
T^i	Température réacteur	°C
T_e^i	Température Nappe i	°C
F_e^i	Débit Nappe i	kg/H
F_{Cata}	Debit de catalyseur	kg/H
$X_{E,e}$	Fraction massique éthylène dans les Nappes	
$X_{C,e}$	Fraction massique comonomère dans les Nappes	
$X_{T,e}$	Fraction massique termonomère dans les Nappes	
ρ_E	Densité de l'éthylène HP	
ρ_C	Densité du comonomère HP	
ρ_T	Densité du termonomère HP	
V^i	Volume de la zone	L

TAB. 2.1 – Mesures et données accessibles relatives à la zone i (nappes=alimentation, HP=haute pression)

2.2 Modèle dynamique

Le modèle est basé sur les lois de conservation de la matière (bilan massique), et de l'énergie (bilan enthalpique). La température étant régulée rapidement, et variant peu, le bilan enthalpique correspond au régime quasi-statique. Dans les cas où l'on fabrique des copolymères et des terpolymères, on utilise un modèle cinétique fondé sur les ratios des vitesses de réaction, et validé expérimentalement.

Le but est d'avoir un modèle dynamique non-linéaire qui incorpore ces données structurales sur le procédé pour ensuite en déduire (par un calcul en temps réel) à partir des mesures de débits et températures, les compositions à l'intérieur du réacteur. On ne dispose pas de capteur de composition à de telles pressions.

Le bilan massique, et le bilan enthalpique sont effectués pour chaque zone du réacteur. La précision des mesures est limitée par la difficulté de mesurer avec exactitude des débits en Haute Pression. Les approximations de modélisation conduisent à des erreurs qui sont a priori petites devant celles qui découlent des imprécisions de mesure au niveau des débits.

Bilan massique

On se place dans une zone particulière. On suppose qu'un seul type de catalyseur est actif dans la zone donnée. Si A désigne l'un des réactifs (éthylène E ou comonomère C ou monomère T), PE le polymère et $Cata$ le catalyseur encore actif dans une zone, un bilan de matière pour la zone i du réacteur s'écrit :

$$M^i \frac{d}{dt} X_A^i = F_e^i X_{A,e}^i + F_{out}^{i-1} X_A^{i-1} - F_{out}^i X_A^i - R_{A,mass}^i \quad (2.1)$$

$$M^i \frac{d}{dt} X_{PE}^i = F_{out}^{i-1} X_{PE}^{i-1} - F_{out}^i X_{PE}^i + R_{PE,mass}^i \quad (2.2)$$

$$M^i \frac{d}{dt} X_{Cata}^i = F_{Cata}^i + F_{out}^{i-1} X_{Cata}^{i-1} - F_{out}^i X_{Cata}^i - R_{Cata,mass}^i \quad (2.3)$$

où X_A^i , X_{PE}^i et X_{Cata}^i désignent la fraction massique de réactif A , de polymère et de catalyseur (supposés uniformes) dans la zone i et si ρ_{HP} est la densité du mélange à la température T^i et pression P^i

$$M^i = \rho_{HP} V^i$$

Les termes du type $R_{A,mass}^i$ représentent la masse de réactif A qui disparaît dans la réaction par unité de temps. On néglige la quantité de catalyseur dans la zone. Le débit massique de sortie de zone en régime permanent est alors donné par

$$\begin{aligned} F_{out}^0 &= 0 \\ F_{out}^i &= F_{out}^{i-1} + F_e^i \end{aligned}$$

Le volume de chaque zone est constant, on a donc obtenu un système différentiel qui représente la dynamique des fractions massiques en réactifs, polymère et catalyseur actif dans la zone considérée.

Bilan enthalpique

Les hypothèses pour le réacteur sont d'une part qu'il n'y a pas d'échange thermique avec l'extérieur et d'autre part que chaque zone est parfaitement agité et les températures et concentrations sont uniformes dans une zone. Un bilan enthalpique pour la zone i du réacteur (basé sur la conservation de l'énergie) s'écrit alors :

$$\begin{aligned} \left\{ M_{PE}^i C_{PPE} + \sum_A M_A^i C_{PA} \right\} \frac{d}{dt} T^i &= F_e^i \left\{ \sum_A X_{Ae}^i C_{PA} \right\} (T^{e^i} - T^i) \\ &+ F_{out}^{i-1} \left\{ X_{PE}^{i-1} C_{PPE} + \sum_A X_A^{i-1} C_{PA} \right\} (T^{i-1} - T^i) \quad (2.4) \\ &+ \sum_A R_{A, mass}^i \Delta H_A(T^i) \end{aligned}$$

où $\Delta H_A(T^i)$ représente la chaleur de polymérisation massique du réactif A (et donc $\sum_A R_{A, mass}^i \Delta H_A(T^i)$ la chaleur produite par la réaction par unité de temps), X_{Ae}^i la fraction massique de réactif A dans la nappe i qui arrive dans la zone i , X_A^i est la fraction massique en réactif A dans la zone i (supposée uniforme), les C_{PA} sont les capacités calorifiques massiques moyennes et tous les autres termes ont été définis plus haut. On fait l'hypothèse que toutes les capacités calorifiques sont toutes égales et valent 0.6 J/kg/K . Ces équations représentent la dynamique de la température de la zone considérée.

Si l'on suppose maintenant la température de chaque zone parfaitement contrôlée et stable, qu'alors on néglige les variations temporelles de température on obtient finalement le système algébrique suivant :

$$\begin{aligned} F_e^i \left\{ \sum_A X_{Ae}^i C_{PA} \right\} (T^i - T^{e^i}) + F_{out}^{i-1} \left\{ X_{PE}^{i-1} C_{PPE} + \sum_A X_A^{i-1} C_{PA} \right\} (T^i - T^{i-1}) \quad (2.5) \\ = \sum_A R_{A, mass}^i \Delta H_A(T^i) \end{aligned}$$

Débits massiques de production

Les modèles de cinétique chimique classiques qui trouvent leur justification en annexe (pour les deux premiers cas) nous donnent les relations suivantes pour les termes de production massiques, dans le cas où l'on produit du polyéthylène (homopolymérisation), ou des copolymères, ou encore des terpolymères.

Cas Homopolymère :

On considère le modèle classique suivant ([50, 41]). Il est confirmé expérimentalement ([19, 18]) dans les conditions propres aux réacteur.

$$\begin{aligned} R_{E,mol}^i &= V^i k_o \exp\left(\frac{-E_o}{RT^i}\right) \sqrt{[Cata]^i} [E]^i \\ R_{Cata,mol}^i &= V^i k_d \exp\left(\frac{-E_d}{RT^i}\right) [Cata]^i \end{aligned} \quad (2.6)$$

où E désigne l'éthylène, et où les constantes (k_d, E_d) et (k_o, E_o) doivent être déterminées expérimentalement pour chaque type de catalyseur.

Cas Copolymère :

On traite ici le cas de l'éthylène avec un acrylate. On considère le modèle suivant ([50, 41]). Il est confirmé par expérimentalement ([16, 17, 36]) comme étant une bonne approximation :

$$\frac{R_{E,mol}^i}{R_{C,mol}^i} = \frac{[E]^i r_E [E]^i + [C]^i}{[C]^i [E]^i + r_C [C]^i} \quad (2.7)$$

$$R_{Cata,mol}^i = V^i k_d \exp\left(\frac{-E_d}{RT^i}\right) [Cata]^i$$

où E désigne l'éthylène, C le comonomère et où les constantes (k_d, E_d) et (r_e, r_A) doivent être déterminées expérimentalement pour chaque type de catalyseur et d'acrylate. Certains des paramètres de polymérisation r_e, r_a ont été pour certains mesurés par M. Buback [16, 17] de l'institut de physique chimique de Göttingen, ou par G. Luft [36] de l'institut de technologie chimique de Darmstadt. Nous avons dans le cas échéant repris ces valeurs. Les autres ont été tabulés en s'appuyant sur la connaissance que l'on a du procédé et des réactivités puis validés sur le réacteur pilote ARKEMA du Groupement Recherche de Lacq (GRL). Les équations sont peu sensibles aux erreurs sur ces paramètres et le modèle a été validé expérimentalement pour tous les types de co-polymères au GRL (voir annexe A).

Cas Terpolymère :

On traite ici le cas de trois monomères. On considère alors la formule suivante, dite équation de terpolymérisation d'Alfrey et Goldfinger (voir [25, 36]), qui est une généralisation de (2.7) et qui donne le ratio entre deux termes quelconques de production massique

parmi ceux de l'éthylène, comonomère, et termonomère :

$$\begin{aligned}
 & R_{E,mol}^i : R_{C,mol}^i : R_{T,mol}^i \\
 & = [E] \left(\frac{[E]}{r_{31}r_{21}} + \frac{[C]}{r_{21}r_{32}} + \frac{[T]}{r_{31}r_{23}} \right) \left([E] + \frac{[C]}{r_{12}} + \frac{[T]}{r_{13}} \right) \\
 & : [E] \left(\frac{[E]}{r_{12}r_{31}} + \frac{[C]}{r_{12}r_{32}} + \frac{[T]}{r_{32}r_{13}} \right) \left(\frac{[E]}{r_{21}} + [C] + \frac{[T]}{r_{23}} \right) \\
 & : [E] \left(\frac{[E]}{r_{13}r_{21}} + \frac{[C]}{r_{13}r_{23}} + \frac{[T]}{r_{13}r_{23}} \right) \left(\frac{[E]}{r_{31}} + \frac{[C]}{r_{32}} + [T] \right)
 \end{aligned} \tag{2.8}$$

où E désigne l'éthylène, et C le comonomère, T le termonomère et où les constantes r_{ij} sont les paramètres de polymérisation croisés et doivent être déterminées expérimentalement pour chaque couple de monomères.

Recycle

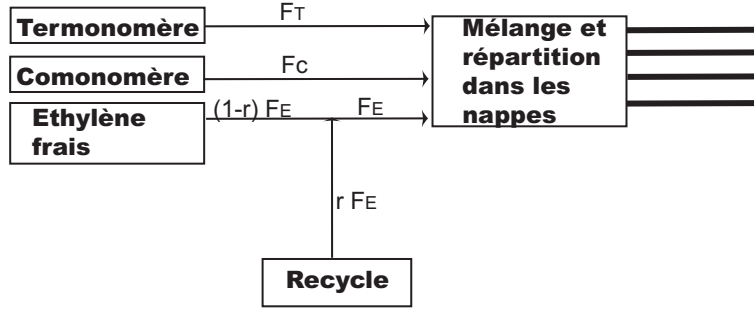


FIG. 2.3 – Schéma simplifié de la préparation du mélange en entrée de réacteur

Les fractions massiques dans les nappes $X_{E,e}$, $X_{C,e}$ et $X_{T,e}$ sont en fait calculées à partir de la connaissance des débits massiques d'entrée en éthylène, comonomère, et termonomère ($F_E(t)$, $F_C(t)$, $F_T(t)$). Ces débits sont mélangés et ensuite redistribué dans les nappes d'alimentation de chacune des zones. On a donc

$$\begin{aligned}
 X_{E,e}(t) &= \frac{F_E(t)}{F_E(t) + F_C(t) + F_T(t)} \\
 X_{C,e}(t) &= \frac{F_C(t)}{F_E(t) + F_C(t) + F_T(t)} \\
 X_{T,e}(t) &= \frac{F_T(t)}{F_E(t) + F_C(t) + F_T(t)}
 \end{aligned} \tag{2.9}$$

Temps de retour et composition du gaz de recycle

On appelle n l'indice de la dernière zone. Le mélange en sortie de dernière zone est composé d'une quantité $F_{out}^n X_E^n$ (kg/s) d'éthylène, d'une quantité $F_{out}^n X_C^n$ (kg/s) de copolymère et d'une quantité $F_{out}^n X_T^n$ (kg/s) de termonomère. X_C^n et X_T^n sont très petits, car les acrylates (comonomères) et le MAGLY et MAH (termonomères) sont injectés en plus faible quantité et sont plus réactifs que l'éthylène. Ainsi on a plus de 95% d'éthylène dans le gaz qui recycle :

$$\frac{X_E^n}{X_E^n + X_C^n + X_T^n} > 0.95$$

C'est la raison pour laquelle dans un premier temps on peut négliger la recycle et considérer que le gaz de recycle n'est composé que d'éthylène. Après détente, le polymère est séparé de la phase gazeuse. Cette dernière est recyclée vers l'entrée du réacteur. Soit V_{retour} le volume total du chemin de retour du gaz vers l'entrée, ρ_{retour} la densité moyenne du gaz dans le chemin de retour, alors le temps de recycle est donné par

$$\tau_r = \frac{\rho_{retour}}{V_{retour} F_{out}^n}$$

Il est de l'ordre de quelques minutes et doit être calculé pour chaque qualité produite.

Calcul des fractions massiques en entrée

Le gaz provenant de la recycle est mélangé à de l'éthylène frais et est composé à plus de 95% d'éthylène. Le flux F_E est en fait composé d'une quantité $0 < r < 1$ en proportion (massique) de gaz provenant de la recycle et d'une quantité $1 - r$ d'éthylène frais, comme indiqué sur la figure 2.3. Pour tenir compte de la recycle, il faut en fait modifier le calcul (2.9) des fractions massiques dans les nappes (donc les proportions en entrée) de la façon suivante :

$$X_{E,e}(t) = \left(r \frac{X_E^n(t - \tau_r)}{X_E^n(t - \tau_r) + X_C^n(t - \tau_r) + X_T^n(t - \tau_r)} + (1 - r) \right) \frac{F_E(t)}{F_E(t) + F_C(t) + F_T(t)}$$

$$X_{C,e}(t) = r \frac{X_C^n(t - \tau_r)}{X_E^n(t - \tau_r) + X_C^n(t - \tau_r) + X_T^n(t - \tau_r)} \frac{F_E(t)}{F_E(t) + F_C(t) + F_T(t)} + \frac{F_C(t)}{F_E(t) + F_C(t) + F_T(t)}$$

$$X_{T,e}(t) = r \frac{X_T^n(t - \tau_r)}{X_E^n(t - \tau_r) + X_C^n(t - \tau_r) + X_T^n(t - \tau_r)} \frac{F_E(t)}{F_E(t) + F_C(t) + F_T(t)} + \frac{F_T(t)}{F_E(t) + F_C(t) + F_T(t)}$$

2.3 Conclusion : modèle dynamique retenu

On donne les équations retenues pour la dynamique du réacteur dans les cas où il y a un, deux, ou trois monomères. Dans le cas homopolymère, le bilan enthalpique donne

le taux de conversion et le modèle cinétique permet de relier la quantité d'éthylène qui réagit à la quantité de catalyseur dans réacteur. Dans les deux autres cas, le catalyseur se décompose selon une loi de cinétique chimique d'ordre 1. On ne relie pas la quantité de catalyseur à la réaction. Le bilan enthalpique donne la quantité totale d'énergie libérée par la réaction, c'est à dire le nombre global d'ouverture de double liaison par unité de temps. Le modèle cinétique donne alors les "clés de répartition" entre les diverses voies possibles de polymérisation.

Cas Homopolymère :

Si l'on rassemble les équations (2.1) et (2.6) on obtient le modèle suivant pour la dynamique des compositions (fractions massiques)

$$M^i \frac{d}{dt} X_E^i = F_e^i + F_{out}^{i-1} X_E^{i-1} - F_{out}^i X_E^i - \mathcal{M}_E V^i k_o \exp\left(\frac{-E_o}{RT^i}\right) \sqrt{[Cata]^i} [E]^i \quad (2.10)$$

$$\frac{d}{dt} [Cata]^i = V^i k_d \exp\left(\frac{-E_d}{RT^i}\right) [Cata]^i \quad (2.11)$$

où $[Cata]$ et $[E]$ sont les concentrations molaires en catalyseur et éthylène. On ne connaît pas les coefficients k_d et k_o avec exactitude, et on obtient de très mauvais résultats en simulant ce système. Cependant, on peut récrire le bilan de l'éthylène en intégrant l'hypothèse de quasi-stationnarité en température (cf (2.5)) :

$$M^i \frac{d}{dt} X_E^i = F_e^i + F_{out}^{i-1} X_E^{i-1} - F_{out}^i X_E^i - R^i / \Delta H_E(T) \quad (2.12)$$

où

$$R^i = Cp F_{out}^i(t)(T^{in}(t) - T^i(t))$$

Cas Copolymère :

Si l'on rassemble les équations (2.1), (2.5), (2.7), on obtient les bilans de monomère et de comonomère s'écrivent ainsi :

$$M^i \frac{d}{dt} X_E^i = F_{out}^i(t)(X_{m,E,e}^i - X_E^i) - X_E \mathcal{M}_E (r_{e,mass} X_E + X_C) R^i \quad (2.13)$$

$$M^i \frac{d}{dt} X_C = F_{out}^i(t)(X_{m,C,e}^i - X_C^i) - X_C \mathcal{M}_C (X_E + r_{C,mass} X_C) R^i \quad (2.14)$$

où l'on note \mathcal{M}_A la masse molaire du réactif A (avec $A \in \{E, C\}$) et $X_{m,A,e}^i$ la fraction massique moyenne en réactif A qui entre dans la zone i définie par :

$$X_{m,A,e}^i = \frac{F_e^i X_{A,e}^i + F_{out}^{i-1} X_A^{i-1}}{F_e^i + F_{out}^{i-1}}$$

et R^i donné par

$$R^i = \frac{Cp (F_{out}^i T^i - F_e^i T_e^i - F_{out}^{i-1} T^{i-1})}{\mathcal{M}_C X_C (X_E + r_{C,mass} X_C) \Delta H_C(T) + \mathcal{M}_E X_E (r_{e,mass} X_E + X_C) \Delta H_E(T)}$$

Les paramètres de polymérisation massique sont calculés à partir des paramètres de polymérisation molaires de la manière suivante :

$$r_{E,mass} = \frac{\mathcal{M}_C}{\mathcal{M}_E} r_{E,mol}$$

$$r_{C,mass} = \frac{\mathcal{M}_E}{\mathcal{M}_C} r_{C,mol}$$

Cas Terpolymère :

Si l'on rassemble les équations (2.1), (2.5), (2.8), les bilans matières des trois monomères s'écrivent :

$$M^i \frac{d}{dt} X_E^i = F_{out}^i(t) (X_{m,E,e}^i - X_E^i) - X_E^i \left(\frac{X_E^i}{\mathcal{M}_E r_{31} r_{21}} + \frac{X_C^i}{\mathcal{M}_C r_{21} r_{32}} + \frac{X_T^i}{\mathcal{M}_T r_{31} r_{23}} \right) \left(\frac{X_E^i}{\mathcal{M}_E} + \frac{X_C^i}{\mathcal{M}_C r_{12}} + \frac{X_T^i}{\mathcal{M}_T r_{13}} \right) \frac{R^i}{D^i}$$

$$M^i \frac{d}{dt} X_C^i = F_{out}^i(t) (X_{m,C,e}^i - X_C^i) - X_C^i \left(\frac{X_E^i}{\mathcal{M}_E r_{12} r_{31}} + \frac{X_C^i}{\mathcal{M}_C r_{12} r_{32}} + \frac{X_T^i}{\mathcal{M}_T r_{32} r_{13}} \right) \left(\frac{X_E^i}{\mathcal{M}_E r_{21}} + \frac{X_C^i}{\mathcal{M}_C} + \frac{X_T^i}{\mathcal{M}_T r_{23}} \right) \frac{R^i}{D^i}$$

$$M^i \frac{d}{dt} X_T^i = F_{out}^i(t) (X_{m,T,e}^i T^i - X_T^i) - X_T^i \left(\frac{X_E^i}{\mathcal{M}_E r_{13} r_{21}} + \frac{X_C^i}{\mathcal{M}_C r_{23} r_{12}} + \frac{X_T^i}{\mathcal{M}_T r_{13} r_{23}} \right) \left(\frac{X_E^i}{\mathcal{M}_E r_{31}} + \frac{X_C^i}{\mathcal{M}_C r_{32}} + \frac{X_T^i}{\mathcal{M}_T} \right) \frac{R^i}{D^i}$$

où pour chaque réactif A parmi les trois monomères

$$X_{m,A,e}^i = \frac{F_e^i X_{A,e}^i + F_{out}^{i-1} X_A^{i-1}}{F_e^i + F_{out}^{i-1}}$$

et avec

$$D^i = X_E^i \left(\frac{X_E^i}{\mathcal{M}_E r_{31} r_{21}} + \frac{X_C^i}{\mathcal{M}_C r_{21} r_{32}} + \frac{X_T^i}{\mathcal{M}_T r_{31} r_{23}} \right) \left(\frac{X_E^i}{\mathcal{M}_E} + \frac{X_C^i}{\mathcal{M}_C r_{12}} + \frac{X_T^i}{\mathcal{M}_T r_{13}} \right) \Delta H_E(T)$$

$$+ X_C^i \left(\frac{X_E^i}{\mathcal{M}_E r_{12} r_{31}} + \frac{X_C^i}{\mathcal{M}_C r_{12} r_{32}} + \frac{X_T^i}{\mathcal{M}_T r_{32} r_{13}} \right) \left(\frac{X_E^i}{\mathcal{M}_E r_{21}} + \frac{X_C^i}{\mathcal{M}_C} + \frac{X_T^i}{\mathcal{M}_T r_{23}} \right) \Delta H_C(T)$$

$$+ X_T^i \left(\frac{X_E^i}{\mathcal{M}_E r_{13} r_{21}} + \frac{X_C^i}{\mathcal{M}_C r_{23} r_{12}} + \frac{X_T^i}{\mathcal{M}_T r_{13} r_{23}} \right) \left(\frac{X_E^i}{\mathcal{M}_E r_{31}} + \frac{X_C^i}{\mathcal{M}_C r_{32}} + \frac{X_T^i}{\mathcal{M}_T} \right) \Delta H_T(T)$$

et

$$R^i = C_p (F_{out}^i T^i - F_e^i T_e^i - F_{out}^{i-1} T^{i-1})$$

Chapitre 3

Observateur

Dans le chapitre 2, on a établi un modèle dynamique du réacteur zone par zone, basé sur un bilan de matière et un bilan enthalpique. Les termes de production massique qui apparaissent dans le bilan massique sont formulés à l'aide d'un modèle de cinétique chimique pour les trois types de polymères produits : homopolymère (un seul monomère en entrée), copolymère (deux monomères en entrée) et terpolymère (trois monomères en entrée).

Les mesures dont on dispose, et qui sont reliées au modèle sont

- Les débits d'entrée (certains volumiques et d'autres massiques)
- La température des entrées et température du réacteur
- La pression du gaz en entrée et la pression dans réacteur

Les pressions n'apparaissent qu'implicitement dans le modèle car c'est à partir d'elles et de la température qu'on calcule les densités et qu'on convertit tous les débits en débits massiques.

On souhaite estimer en ligne les compositions dans chacune des zones. L'intérêt pour l'opérateur est d'avoir une image interne zone par zone afin de mieux piloter le réacteur dans son ensemble. A partir de l'estimation des compositions on remonte alors simplement à la composition du polymère qu'on fabrique dans chaque zone, ainsi qu'aux consommations spécifiques en catalyseur par zone (ratio catalyseur injecté dans la zone par quantité de polymère formé). Cette dernière information donne une indication de stabilité car la réaction est très exothermique. Le réacteur peut s'emballer s'il y a "trop" de catalyseur en solution : c'est le phénomène de décomposition typique dans les réacteurs polyéthylène haute pression (voir [55]).

On considère dans un premier temps qu'il n'y a pas de recycle. Cela revient à dire que tous les réactifs autres que l'éthylène ont réagi à la sortie de la dernière zone, et donc que le gaz qui recycle n'est composé que d'éthylène.

3.1 Principe de l'observateur

Les mesures et données dont on dispose sont résumées dans le tableau 2.1, page 11. On prend pour état les compositions. L'observateur est une simple copie des bilans dans lesquels on injecte les mesures de température et de débit.

Le bilan thermique stationnaire (la température étant régulée rapidement, on la suppose constante) donne la cinétique globale : nous n'utilisons de la modélisation cinétique que les ratios entre les divers termes de production massique. Les ratios dépendent des compositions via le modèle cinétique. La cinétique globale est alors vue ici comme un fonction du temps et le modèle des ratios de réaction est une fonction non-linéaire des compositions. Les équations différentielles de bilan des diverses espèces s'écrivent alors sous la forme $\frac{d}{dt}X = f(X, t)$ où la dépendance en temps de f vient des mesures en ligne de débits et de températures. L'observateur est alors une simple copie de ce système $\frac{d}{dt}\hat{x} = f(\hat{x}, t)$. On montre dans les deux cas (homopolymérisation, copolymérisation) que ce système dynamique qui dépend du temps oublie sa condition initiale : c'est un filtre non-linéaire asymptotiquement stable, dont l'état donne une estimation des compositions au sein du réacteur. En effet si l'on prend deux trajectoires du système $\frac{d}{dt}X_1 = f(X_1, t)$ et $\frac{d}{dt}X_2 = f(X_2, t)$ telles que $X_1(0) \neq X_2(0)$, on a $X_2(t) - X_1(t) \rightarrow 0$ quand $t \rightarrow \infty$. Par conséquent on a bien que l'erreur d'estimaion $\Delta x = \hat{x} - x \rightarrow 0$ quand $t \rightarrow \infty$. On utilise le même filtre dans le cas terpolymérisation, bien nous ne disposons pas de preuve de convergence complète. On a néanmoins des indications numériques fortes de convergence. Cet estimateur donne une estimation très réaliste des compositions internes, de celle du polymère et de sa quantité. Sa validation ne peut être qu'indirecte via les mesures en laboratoire de la composition des échantillons de polymères produits en sortie de l'unité.

3.1.1 Cas Homopolymère

Dans ce cas, il n'y a que deux espèces chimiques dans le réacteur : de l'éthylène et du polymère (polyéthylène). On remplace dans (2.12), X_E^i par son estimée \hat{X}_E^i qui obéit à la dynamique suivante :

$$M^i \frac{d}{dt} \hat{X}_E^i = F_c^i + F_{out}^{i-1} \hat{X}_E^{i-1} - F_{out}^i \hat{X}_E^i - R^i(t) / \Delta H_E(T)$$

avec

$$R^i(t) = Cp F_{out}^i(t) (T^{in}(t) - T^i)$$

C'est un filtre stable car dans la zone i la composition de la zone précédente $\hat{X}_E^{i-1}(t)$ est vue comme une fonction du temps uniquement, et la dynamique s'écrit $\frac{d}{dt} \hat{X}_E^i = f(\hat{X}_E^i, t)$ avec $\frac{\partial}{\partial x} f(x, t) = \frac{F_{out}^i}{M^i} \leq a < 0$ où a est constant. Ainsi \hat{X}_E^i converge exponentiellement vers X_E^i avec une constante de temps plus courte que $1/a$. Une fois que l'observateur fournit une estimation correcte de la composition dans la zone du dessus $M^i \frac{d}{dt} \Delta X_E^i = -F_{out}^i \Delta X_E^i$ où $\Delta X_E^i = \hat{X}_E^i - X_E^i$ est l'erreur d'estimation.

3.1.2 Cas copolymère

Nous allons montrer que quelque soit l'initialisation $\hat{X}_E(0)$ et $\hat{X}_C(0)$ on a bien que les erreurs $\hat{X}_E - X_E$ et $\hat{X}_C - X_C$ tendent vers 0. On ne va présenter l'observateur que sur la première zone du réacteur. Le principe est similaire pour les zones suivantes disposées en cascade les unes derrière les autres. Seuls changent les entrées : les températures, les débits d'entrée, et la composition du mélange en entrée de zone dépendent de la zone précédente. Soient \hat{X}_E et \hat{X}_C les fractions massiques estimées en éthylène et en acrylates pour la première zone du réacteur. On reprend le bilan dynamique pour les compositions dans le cas copolymère (2.13)-(2.14). On choisit \hat{X}_E et \hat{X}_C arbitraires à l'instant $t = 0$ solutions de

$$\begin{aligned} M \frac{d}{dt} \hat{X}_E &= F_e(t)(X_{m,E,e} - \hat{X}_E) - \hat{X}_E \mathcal{M}_E (r_{e,mass} \hat{X}_E + \hat{X}_C) R(\hat{X}_E, \hat{X}_C, t) \\ M \frac{d}{dt} \hat{X}_C &= F_e(t)(X_{m,C,e} - \hat{X}_C) - \hat{X}_C \mathcal{M}_C (\hat{X}_E + r_{C,mass} \hat{X}_C) R(\hat{X}_E, \hat{X}_C, t) \end{aligned}$$

$$R(\hat{X}_E, \hat{X}_C, t) = \frac{C_p F_e (T - T_e)}{\mathcal{M}_C \hat{X}_C (\hat{X}_E + r_{C,mass} \hat{X}_C) \Delta H_C(T) + \mathcal{M}_E \hat{X}_E (r_{e,mass} \hat{X}_E + \hat{X}_C) \Delta H_E(T)}$$

Cela correspond bien au modèle car en zone 1 on a $F_e = F_{out}$. Nous allons prouver la convergence des trajectoires indépendamment de leur condition initiale. Elle repose sur deux propriétés :

- Les trajectoires sont bornées : \hat{X}_E et \hat{X}_C restent toujours entre 0 et 1. En effet le champ est strictement rentrant au bord de ce domaine. Qui plus est, les compositions si elles sont initialisées entre 0 et 1 deviennent inférieures aux compositions d'entrée $X_{m,E,e}$ et $X_{m,C,e}$ en temps fini. En effet la dérivée de \hat{X}_E (resp. \hat{X}_C) est plus petite qu'un réel strictement négatif pour $\hat{X}_E > X_{m,E,e}$ (resp. $\hat{X}_C > X_{m,C,e}$).
- Le système, par un changement de variable approprié, se met sous forme triangulaire

$$\frac{d}{dt} X_1 = f_1(X_1, t), \quad \frac{d}{dt} X_2 = f_2(X_1, X_2, t)$$

avec f_1 décroissante par rapport à X_1 et f_2 décroissante par rapport à X_2 . Comme les pentes sont strictement plus petites qu'un réel strictement négatif, on a convergence exponentielle des estimations \hat{X}_e et \hat{X}_c vers X_E et X_C , respectivement.

Etude de la convergence

En notant

$$R(X_E, X_C, t) = \frac{Mg(t)}{\mathcal{M}_C X_C (X_E + r_{C,mass} X_C) \Delta H_C(T) + \mathcal{M}_E X_E (r_{e,mass} X_E + X_C) \Delta H_E(T)}$$

on obtient avec, K , g , et u fonctions du temps t et indépendantes des compositions internes

$$\begin{aligned} \frac{d}{dt}X_E &= u(t) - K(t)X_E \\ &\quad - g(t) \frac{\mathcal{M}_E X_E (r_{e,mass} X_E + X_C)}{\mathcal{M}_C X_C (X_E + r_{C,mass} X_C) \Delta H_C(T) + \mathcal{M}_E X_E (r_{e,mass} X_E + X_C) \Delta H_E(T)} \\ &= u(t) - K(t)X_E - \frac{g(t)}{\Delta H_E(T)} \frac{1}{1 + \frac{\mathcal{M}_C \Delta H_C(T)}{\mathcal{M}_E \Delta H_E(T)} \frac{X_C (X_E + r_{C,mass} X_C)}{X_E (X_C + r_{e,mass} X_E)}} \end{aligned}$$

et de manière similaire

$$\frac{d}{dt}X_C = v(t) - K(t)X_C - \frac{g(t)}{\Delta H_C(T)} \frac{1}{1 + \frac{\mathcal{M}_E \Delta H_E(T)}{\mathcal{M}_C \Delta H_C(T)} \frac{X_E (X_C + r_{e,mass} X_E)}{X_C (X_E + r_{C,mass} X_C)}}$$

Soient

$$\begin{aligned} \varphi : z &\mapsto \frac{\mathcal{M}_C \Delta H_C(T)}{\mathcal{M}_E \Delta H_E(T)} \frac{z(1 + r_{C,mass} z)}{(z + r_{e,mass})} \\ \psi : z &\mapsto \frac{\mathcal{M}_E \Delta H_E(T)}{\mathcal{M}_C \Delta H_C(T)} \frac{z(1 + r_{E,mass} z)}{(z + r_{C,mass})} \end{aligned}$$

On remarque qu'à l'aide de $u, v, K, g_E, g_C \geq 0$ la dynamique s'écrit

$$\frac{d}{dt}X_E = u(t) - K(t)X_E - \frac{g(t)}{\Delta H_E(T)} \frac{1}{1 + \varphi\left(\frac{X_C}{X_E}\right)} \quad (3.1)$$

$$\frac{d}{dt}X_C = v(t) - K(t)X_C - \frac{g(t)}{\Delta H_C(T)} \frac{1}{1 + \psi\left(\frac{X_E}{X_C}\right)} \quad (3.2)$$

On fait l'hypothèse que les enthalpies de réaction ΔH_E et ΔH_C sont constantes et on considère le changement de variable :

$$\begin{aligned} \eta &= \Delta H_E X_E + \Delta H_C X_C \\ \mu &= \Delta H_E X_E - \Delta H_C X_C \end{aligned}$$

La dynamique sur η est stable :

$$\frac{d}{dt}\eta = \Delta H_E u(t) + \Delta H_C v(t) - K(t)\eta - g(t)$$

car si l'on pose $\frac{d}{dt}\eta = f(\eta, t)$ on a bien $\frac{\partial f}{\partial \eta} = -K(t) \leq a < 0$ où $a > 0$ est constant.

Cela veut dire notamment que puisque l'observateur consiste en une simple copie de la dynamique, en appelant $\hat{\eta} = \Delta H_E \hat{X}_E + \Delta H_C \hat{X}_C$, η et $\hat{\eta}$ vérifient les mêmes équations différentielles, et si on appelle $\Delta\eta = \hat{\eta} - \eta$ on a

$$\frac{d}{dt}\Delta\eta = -K(t)\Delta\eta$$

et donc la composante de l'erreur d'estimation sur η converge bien vers zéro.

Quant à μ on a

$$\frac{d}{dt}\mu = \Delta H_E u(t) - \Delta H_C v(t) - K(t)\mu - g(t)\frac{1}{1+\varphi(z)} + g(t)\frac{1}{1+\frac{1}{\varphi(z)}}$$

où l'on note $z = \frac{X_E}{X_C}$ et on a utilisé la relation $\varphi(1/z)\psi(z) = 1$.

On veut montrer qu'à η bloqué le membre de droite est une fonction décroissante de μ . Il suffit pour cela de montrer que

$$g(t)\frac{1}{1+\varphi(z)}$$

est une fonction croissante de μ . Cela est vrai puisque $z = \frac{\Delta H_E}{\Delta H_C} \frac{\eta - \mu}{\eta + \mu}$ est une fonction clairement décroissante de μ et φ est une fonction toujours croissante (sa dérivée est du signe de $2r_E z + z^2$). Ainsi si on écrit que $\frac{d}{dt}\mu = h(\eta, \mu, t)$ on a $\frac{\partial h}{\partial \mu} \leq -K(t) \leq b$ avec $b < 0$ constant.

Cela veut dire notamment que si $\hat{\eta} = \eta$ au cours du temps, on a la convergence de $\Delta\mu = \hat{\mu} - \mu$ vers zéro.

Conclusion de l'étude

Nous avons montré que quelque soit l'initialisation $\hat{X}_E(0)$ et $\hat{X}_C(0)$ on a bien que $\hat{X}_E - X_E$ et $\hat{X}_C - X_C$ tendent vers 0, puisque (\hat{X}_E, \hat{X}_C) et (X_E, X_C) vérifient les mêmes équations et peuvent être ramenés par changement de variable à η et μ . On a mis le système sous la forme triangulaire

$$\frac{d}{dt}\eta = f(\eta, t), \quad \frac{d}{dt}\mu = h(\mu, \eta, t)$$

avec f décroissante par rapport à η et de dérivée plus petite qu'un réel strictement négatif, h bornée et uniformément continue par rapport à η (car $\varphi > 0$), décroissante par rapport à μ et de dérivée strictement plus petite qu'un réel négatif. Les solutions sont uniformément bornées. Le système converge donc en deux temps. D'abord η converge et oublie sa condition initiale (toutes les solutions de la première équation différentielle ci-dessus $\frac{d}{dt}\eta = f(\eta, t)$ convergent les unes vers les autres, y compris η correspondant au vrai système et $\hat{\eta}$ correspondant à l'observateur, puisqu'elles sont deux solutions du même système qui "oublie" sa condition initiale). On considère alors que $\hat{\eta} = \eta$. On a vu alors que μ oublie

sa condition initiale. En pratique, on observe que les estimées de composition convergent rapidement avec une constante de temps plus rapide que le temps de séjour dans la zone considérée.

3.1.3 Cas terpolymère

Dans le cas terpolymère, on procède de la même manière pour construire l'observateur, mais nous n'avons pas de preuve de la convergence. Nous avons seulement certaines garanties sur les trajectoires. En effet l'observateur s'écrit maintenant :

$$\begin{aligned}
M^i \frac{d}{dt} \hat{X}_E^i &= F_{out}^i(t) (X_{m,E,e}^i - \hat{X}_E^i) \\
&\quad - \hat{X}_E^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{31} r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{21} r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{31} r_{23}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_E} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{12}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{13}}} \right) \frac{R(t)}{H(\hat{X}_E, \hat{X}_C, \hat{X}_T)} \\
M^i \frac{d}{dt} \hat{X}_C^i &= F_{out}^i(t) (X_{m,C,e}^i - \hat{X}_C^i) \\
&\quad - \hat{X}_C^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{12} r_{31}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{12} r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{32} r_{13}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_C} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{23}}} \right) \frac{R(t)}{H(\hat{X}_E, \hat{X}_C, \hat{X}_T)} \\
M^i \frac{d}{dt} \hat{X}_T^i &= F_{out}^i(t) (X_{m,T,e}^i - \hat{X}_T^i) \\
&\quad - \hat{X}_T^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{13} r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{23} r_{12}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{13} r_{23}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{31}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_T} \right) \frac{R(t)}{H(\hat{X}_E, \hat{X}_C, \hat{X}_T)}
\end{aligned}$$

avec

$$\begin{aligned}
H(\hat{X}_E, \hat{X}_C, \hat{X}_T) &= \hat{X}_E^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{31} r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{21} r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{31} r_{23}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_E} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{12}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{13}}} \right) \Delta H_E(T) \\
&\quad + \hat{X}_C^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{12} r_{31}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{12} r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{32} r_{13}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_C} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{23}}} \right) \Delta H_C(T) \\
&\quad + \hat{X}_T^i \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{13} r_{21}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{23} r_{12}}} + \frac{\hat{X}_T^i}{\mathcal{M}_{T r_{13} r_{23}}} \right) \left(\frac{\hat{X}_E^i}{\mathcal{M}_{E r_{31}}} + \frac{\hat{X}_C^i}{\mathcal{M}_{C r_{32}}} + \frac{\hat{X}_T^i}{\mathcal{M}_T} \right) \Delta H_T(T)
\end{aligned}$$

et

$$R(t) = Cp (F_{out}^i T^i - F_e^i T_e^i - F_{out}^{i-1} T^{i-1})$$

Les trajectoires présentent deux propriétés intéressantes :

Les trajectoires restent bornées

Si elles sont initialisées dans le parallélépipède $(\hat{X}_E, \hat{X}_C, \hat{X}_T) \in [0, X_{m,E,e}^i] \times [0, X_{m,C,e}^i] \times [0, X_{m,T,e}^i]$, elles y restent pour tout temps t positif. En effet la dynamique peut s'écrire sous la forme :

$$\begin{aligned} M^i \frac{d}{dt} \hat{X}_E^i &= F_{out}^i(t)(X_{m,E,e}^i - \hat{X}_E^i) - \hat{X}_E^i Q_1(\hat{X}_E^i, \hat{X}_C^i, \hat{X}_T^i, t) \\ M^i \frac{d}{dt} \hat{X}_C^i &= F_{out}^i(t)(X_{m,C,e}^i - \hat{X}_C^i) - \hat{X}_C^i Q_2(\hat{X}_E^i, \hat{X}_C^i, \hat{X}_T^i, t) \\ M^i \frac{d}{dt} \hat{X}_T^i &= F_{out}^i(t)(X_{m,T,e}^i - \hat{X}_T^i) - \hat{X}_T^i Q_3(\hat{X}_E^i, \hat{X}_C^i, \hat{X}_T^i, t) \end{aligned}$$

où les $Q_i(\hat{X}_E^i, \hat{X}_C^i, \hat{X}_T^i, t)$ sont des quantités positives pour $\hat{X}_E^i, \hat{X}_C^i, \hat{X}_T^i$ positifs. En effet le champ est strictement rentrant aux bords du parallélépipède $[0, X_{m,E,e}^i] \times [0, X_{m,C,e}^i] \times [0, X_{m,T,e}^i]$.

Plus généralement on peut affirmer que les trajectoires si elles sont initialisées dans le cube $[0, 1]^3$ y restent pour tout temps positif.

Les trajectoires du système dans \mathbb{R}^3 convergent vers l'intérieur d'un triangle

En effet si l'on note η la variable composite

$$\eta = \Delta H_E(T) \hat{X}_E + \Delta H_C(T) \hat{X}_C + \Delta H_T(T) \hat{X}_T$$

Sa dynamique s'écrit

$$\frac{d}{dt} \eta = F_{out}^i(t)(\Delta H_E(T) X_{m,E,e}^i + \Delta H_C(T) X_{m,C,e}^i + \Delta H_T(T) X_{m,T,e}^i) - F_{out}^i(t) \eta - R(t) \quad (3.3)$$

On fait l'hypothèse que les enthalpies ci-dessus $\Delta H_E, \Delta H_C, \Delta H_T$ sont constantes (car la température est parfaitement contrôlée). La dynamique de η peut se ré-écrire

$$\frac{d}{dt} \eta = f(t) - K(t) \eta$$

où $K(t) < a < 0$ et a est un réel constant. La dynamique sur η est exponentiellement stable. Donc η converge indépendamment de sa condition initiale. Les trajectoires convergent donc vers le sous-ensemble dépendant du temps

$$\{(\hat{X}_E, \hat{X}_C, \hat{X}_T) \in \mathbb{R}_+^3 \mid \Delta H_E \hat{X}_E + \Delta H_C \hat{X}_C + \Delta H_T \hat{X}_T = \eta(t)\}$$

où $\eta(t) = \Delta H_E X_E(t) + \Delta H_C X_C(t) + \Delta H_T X_T(t)$.

Conclusion

Nous n'avons pas de preuve de stabilité dans ce cas mais l'on observe en simulation une échelle de temps de convergence comparable au temps de séjour. Pour l'estimateur en temps réel qui en résulte, ses estimations sont en bon accord avec les mesures du laboratoire, quand il n'y pas de problème de biais sur les débits. On a une garantie de convergence partielle, puisque l'on sait que les variables convergent vers un sous ensemble borné de \mathbb{R}^3 .

Après coup, nous avons réalisé que l'un des intérêts pratiques de cet estimateur asymptotique construit à partir d'un modèle dynamique est le suivant. Il permet de calculer numériquement et de façon très efficace les régimes stationnaires. Si l'on voulait, avec notre modèle cinétique et en faisant l'hypothèse que l'on est en régime stabilisé, remonter algébriquement aux compositions stationnaires en chacun des réactifs, on aurait trois équations algébriques à résoudre à chaque instant. Les compositions correspondent alors aux racines d'un système de trois polynômes, chacun de trois variables et de degré trois. Résoudre ce système algébrique n'est une chose si facile à faire. On pourrait utiliser des techniques de calculs formels et d'élimination pour se ramener à une équation polynomiale unique en une seule variable. Mais alors son degré sera bien supérieur à 3. Des techniques numériques efficaces existent pour trouver les racines d'un polynôme. Cependant il faudrait alors, trier parmi ses racines, celles qui donnent des fractions massiques entre 0 et 1.

3.1.4 Stabilité avec la recycle

Nous n'avons pas de preuve de stabilité avec la recycle pour le modèle du réacteur entier développé au chapitre 2. Nous avons néanmoins des indications fortes de stabilité. On observe en simulation qu'avec recycle une convergence rapide vers des valeurs compatibles avec les mesures laboratoire de composition. Dans cette section nous donnons une indication de stabilité dans un cas simplifié en nous appuyant sur une séparation d'échelle de temps.

On sait que le temps de recycle est plus long que le temps de séjour. On se place ici dans le cas où l'on produit un copolymère. Le bilan est fait en régime quasi-statique pour le réacteur que l'on suppose en régime stationnaire. On découpe donc le problème en échelles de temps : on considère que le temps de recycle est long devant l'établissement de l'équilibre dans le réacteur. On note $X_C(t)$ la fraction massique en comonomère à l'instant t supposée uniforme. Soit F_C le débit en comonomère pur dans le réacteur, F_{out} le débit de sortie du réacteur, et $F_{E,ext}$ le débit d'éthylène frais injecté. En régime statique le débit de polymère en sortie est $F_{E,ext} + F_C$ et le débit de recycle est $F_{out} - (F_{E,ext} + F_C)$. La création d'enthalpie est imposée par la régulation de température, et les enthalpies massiques de l'éthylène et du comonomère sont du même ordre. Nous allons supposer le taux de conversion r constant. Il s'écrit par définition $r = (F_{E,ext} + F_C)/F_{out}$. Soit $\varphi : [0, 1] \mapsto [0, 1]$ la fonction qui donne le taux de comonomère incorporé dans le polymère pour une composition en comonomère $X_C(t)$ donnée dans le réacteur. Faisons un bilan statique de comonomère dans le réacteur :

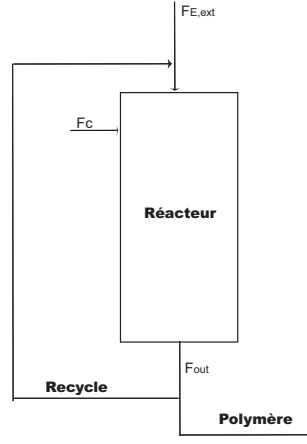


FIG. 3.1 – Schéma simplifié du réacteur avec recycle

- Comonomère en entrée : $F_C + F_{out}X_C(t - \Delta)$ avec Δ temps de de recycle. Le premier terme est l'injection de comonomère frais, le second la quantité de comonomère qui recycle.
- Comonomère en sortie : $F_{out}X_C(t) + (F_{E,ext} + F_C)\varphi\left(\frac{F_C + F_{out}X_C(t - \Delta)}{F_{out}}\right)$. Le premier terme est la quantité de comonomère dans le gaz, le second l'acrylate dans le polymère.

En notant $X_0 = F_C/F_{out}$ on obtient en tout divisant par F_{out}

$$X_0 + X_C(t - \Delta) = X_C(t) + r\varphi(X_0 + X_C(t - \Delta))$$

Pour montrer que le système simplifié avec recycle est stable, il nous faut donc déterminer que la fonction $f : X \mapsto (X_0 + X) - r\varphi(X_0 + X)$ est contractante et admet un point fixe. On ne connaît pas la fonction φ car elle est définie de manière implicite grâce au modèle cinétique développé au paragraphe 2.2. Ce modèle nous donne $\frac{dX_C}{dX_E}$ en fonction de X_E et de X_C . On a supposé qu'il y a trois espèces dans le réacteur, et que la fraction massique de polymère est r , cela nous donne la relation $X_E + X_C + r = 1$. Elle nous permet de tracer numériquement $\frac{dX_C}{dX_E}$ en fonction de X_C . On trouve une fonction quasiment linéaire, ce qui laisse à penser que φ , et donc f , est faiblement non linéaire ici. Par conséquent si l'on a $f(0) > 0$ et $f(1 - r) < 1 - r$ (sachant que $1 - r$ est la quantité maximale d'acrylate puisque $X_E + X_C + r = 1$ et $X_E > 0$) on peut penser sans preuve pour autant que f (qui est proche d'une fonction linéaire) est contractante. On a vu à la section 3.1.2 que la structure du système est tel que $\dot{X}_C > 0$ pour $\{X_C = 0, X_E > 0\}$. Cela entraîne que $f(0) > 0$. Puisque l'on a également $\dot{X}_E > 0$ pour $\{X_E = 0, X_C > 0\}$ on a $f(1 - r) < 1 - r$. En effet la fonction f donne la composition en comonomère à l'instant $t + \Delta$. Si à l'instant t il n'y a pas de comonomère dans le réacteur, à taux de conversion constant, le comonomère va moins réagir, et sa composition dans le mélange réactionnel va augmenter. Les deux inégalités ci-dessus et l'allure de f tracée numériquement laissent à penser que la pente de f est partout plus petite que 1.

3.1.5 Récapitulatif des mesures et des grandeurs estimées

Les mesures relatives à la zone i sont récapitulées dans le tableau 2.1, chapitre 2 à la page 11. Le tableau 3.1 suivant donne le récapitulatif des grandeurs estimées toutes déduites de la connaissance des entrées et de l'estimation des compositions internes à l'aide des formules précédentes.

Zone i		
<i>Description</i>	<i>Formule</i>	<i>unités</i>
Fraction massique Ethylène dans la zone	X_E^i	
Fraction massique Copolymère dans la zone	X_C^i	
Fraction massique Terpolymère dans la zone	X_T^i	
Fraction massique Polymère dans la zone	$1 - X_E^i - X_C^i - X_T^i$	
Masse éthylène dans le polymère	$F_E^i = F_{out}^i (X_{m,E,e}^i - X_E^i)$	t/h
Masse copolymère dans le polymère	$F_C^i = F_{out}^i (X_{m,C,e}^i - X_C^i)$	t/h
Masse terpolymère dans le polymère	$F_T^i = F_{out}^i (X_{m,T,e}^i - X_T^i)$	t/h
Production par zone	$Prod^i = F_E + F_C + F_T$	t/h
Tauxconversion	$Prod^i / F_{out}^i$	%
Copolymère incorporé dans le polymère	$F_C^i / Prod^i$	%
Terpolymère incorporé dans le polymère	$F_T^i / Prod^i$	%
Temps séjour	V^i / F_{out}^i	s
Conso spec cata	$F_{cata}^i / Prod^i$	(kg/t)

TAB. 3.1 – grandeurs estimées relatives à la zone i

3.2 Résultats

On ne peut valider le modèle et l'observateur qu'indirectement avec les mesures de production de polymère, et la mesure de la composition du polymère final. L'estimation du taux de conversion conduit à une estimation de la production, qui a été validée en comparant à la production réelle. La composition du polymère final est mesurée par le laboratoire avec une fréquence de l'ordre de la minute : un échantillon est acheminé au laboratoire, puis le polymère est fondu et transformé en film plastique, et on analyse sa composition par infrarouge. Ces mesures sont disponibles en salle de contrôle au bout d'un temps significatif après production du polymère. Ce retard de mesure est bien supérieur aux échelles de temps de la dynamique du réacteur, en particulier au temps de séjour. Nous allons comparer les estimations de la composition du polymère en temps réel fournies par notre observateur asymptotique avec les mesures du laboratoire dans les cas copolymère et terpolymère (le cas homopolymère est ici sans objet).

Cas copolymère

La figure 3.2 montre les résultats de l'estimation en ligne et la mesure en laboratoire du pourcentage massique d'acrylate (comonomère) incorporé au polymère lors d'une transition entre deux qualités (copolymérisation Ethylène-Acrylate de Méthyle à des teneurs différentes). Lors d'une transition, l'équilibre est quasiment réalisé à chaque instant au cours de la transition. Le décalage horizontal des courbes est lié au retard de la mesure. Le décalage vertical des courbes est lié aux biais sur les mesures de débits Haute-Pression. Ces divers biais ne peuvent pas être estimés indépendamment, ils sont "quasiment" inobservables, sauf à rajouter un biais agrégé qui n'aurait pas beaucoup de sens et que l'on ne pourrait estimer qu'au bout d'un temps significatif. On voit que le modèle fournit correctement l'évolution relative des grandeurs estimées, avec un biais globalement additif facile à estimer (avec un décalage temporel significatif) et qui peut être fait à vue sans problème par l'opérateur. Noter le décrochage brutal à $2/3$ (de l'axe horizontal) de l'estimateur : il est dû à une dérive brutale d'une des mesures de débit Haute-Pression. NB : les étiquettes des axes ont été retirées.

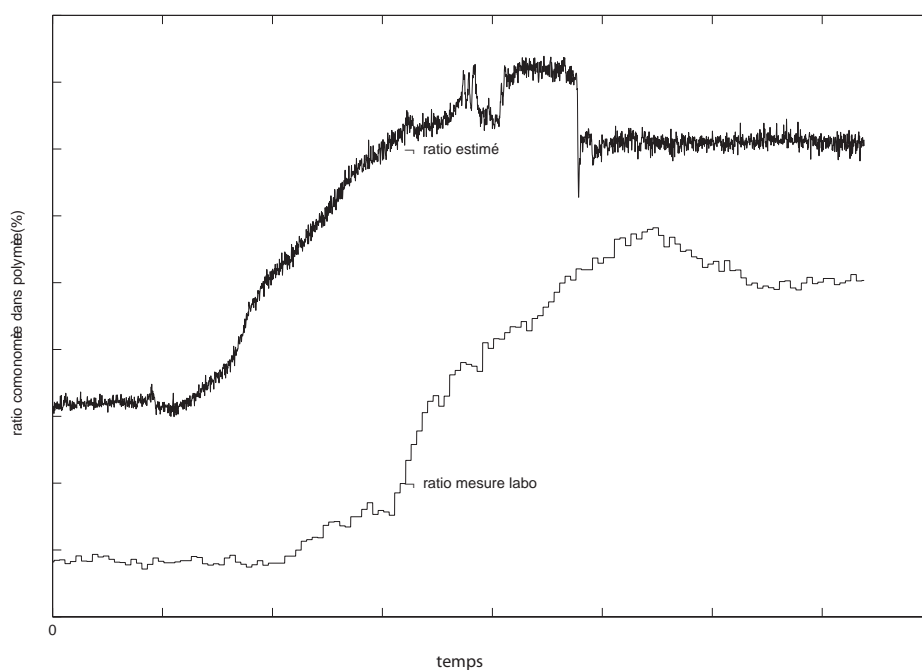


FIG. 3.2 – Quantité d'acrylate incorporé dans le polymère produit (% massique) pour une transition en copolymérisation Ethylène-Acrylate de Méthyle le 22-05-2005. L'estimation est faite en temps réel, le résultat de la mesure laboratoire a un retard.

Cas terpolymère

Les figures 3.3 et 3.4 montrent les résultats de l'estimation en ligne et la mesure en laboratoire du pourcentage massique d'acrylate (comonomère) incorporé au polymère et du pourcentage massique d'anhydride maléique (termonomère) incorporé au polymère lors d'une transition entre deux qualités (terpolymérisation Ethylène-Acrylate de Buthyle-Anhydride Maléique à des teneurs différentes). Lors d'une transition, l'équilibre est quasiment réalisé à chaque instant au cours de la transition. Le décalage horizontal des courbes est lié au retard de la mesure. Le décalage vertical des courbes est lié aux biais sur les mesures de débits Haute-Pression. Ces divers biais ne peuvent pas être estimés indépendamment, ils sont "quasiment" inobservables, sauf à rajouter un biais agrégé qui n'aurait pas beaucoup de sens et que l'on ne pourrait estimer qu'au bout d'un temps significatif. On voit que le modèle fournit correctement l'évolution relative des grandeurs estimées, avec un biais globalement additif facile à estimer (avec un décalage temporel significatif) et qui peut être fait à vue sans problème par l'opérateur. La variation brutale de résultat de la mesure labo qui intervient au même moment pour les deux graphiques est réellement un saut, et correspond à un recalage manuel des mesures labo, qui ont tendance à dériver (ré-étalonnage). L'estimateur est biaisé mais est donc ici plus précis dans son évolution relative que les mesures labo. NB : les étiquettes des axes ont été retirées.

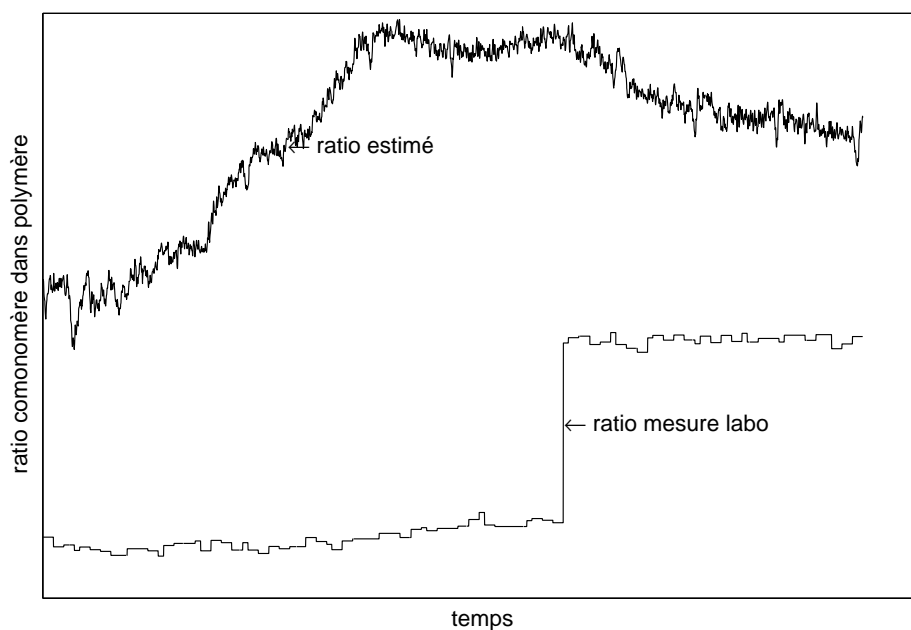


FIG. 3.3 – Quantité d'acrylate incorporé dans le polymère produit (% massique) pour une transition en terpolymérisation Ethylène-Acrylate de Butyle-Anhydride Maléique le 15-03-2007. L'estimation est faite en temps réel, le résultat de la mesure laboratoire a un retard. Le saut dans la mesure du laboratoire correspond à un ré-étalonnage manuel quand on effectue une transition de qualité produit.

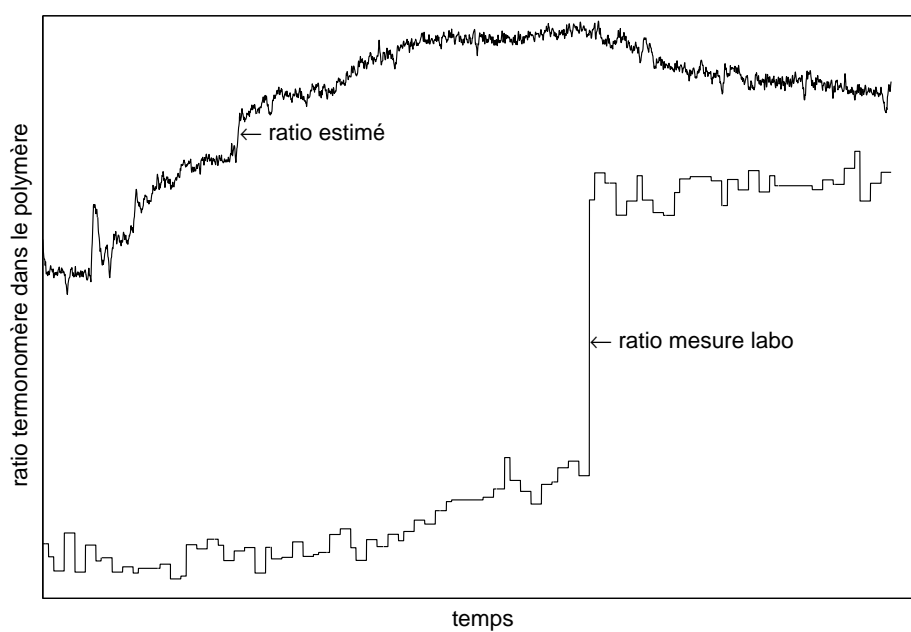


FIG. 3.4 – Quantité d’Anhydride Maléique incorporé dans le polymère produit (% massique) pour une transition en terpolymérisation Ethylène-Acrylate de Butyle-Anhydride Maléique le 15-03-2007. L’estimation est faite en temps réel, le résultat de la mesure laboratoire a un retard. Les quantités de termonomères incorporé sont de l’ordre du pourcent.

Chapitre 4

Conclusion

Nous avons établi un modèle du réacteur basé sur un bilan enthalpique, un bilan de matière et un modèle de cinétique chimique. Nous avons adopté des hypothèses simplificatrices pour que le modèle soit le plus simple possible, sachant que l'on fait déjà des erreurs non-négligeables quand on mesure les débits à des pressions de l'ordre de 2000 bars. Nous avons construit un observateur qui estime les concentrations en chacun des réactifs en nous appuyant sur une propriété forte des bilans matière re-écrits en utilisant l'hypothèse de quasi-stationnarité du bilan d'énergie (température constante car régulée par un régulateur PI supposé rapide et stable). Sous cette forme, on en déduit (après injection des mesures de débits et de températures) un système dynamique qui dépend du temps mais qui oublie sa condition initiale. Nous avons montré cette propriété dans le cas homopolymère et copolymère, et dans le cas terpolymère nous l'avons constatée numériquement. Il suffit donc de simuler le système pour avoir une estimation indépendante de la manière dont on initialise l'observateur. L'observateur permet d'avoir une estimation de la concentration des différentes espèces et donne ainsi une image interne de la réaction dans chaque zone du réacteur. Connaissant les débits, les concentrations d'entrée, ainsi que les concentrations dans chaque zone, on en déduit : la composition par zone et la composition finale du polymère, la production par zone, le taux de conversion, et les consommations spécifiques en catalyseur. Sur l'unité on ne mesure par ailleurs que la composition finale du polymère, et la production totale. On ne peut donc valider que partiellement l'observateur. Néanmoins, en régime statique, on a validé les résultats de l'observateur sur une seule zone, sur le réacteur pilote du GRL (voir annexe A). Les débits de comonomère sont biaisés (alors que ceux de termonomères sont mesurés avec une plus grande exactitude) ce qui donne souvent une sur-estimation de la quantité de comonomère dans le polymère. Mais l'erreur observée sur le pilote du GRL est en général inférieure à 2% de comonomère dans le polymère, et l'observateur permet de suivre les évolutions dynamiques de la composition du polymère.

Quand on a écrit les compositions, il a fallu choisir entre écrire des fractions molaires, ou des fractions massiques. Après avoir converti tous les débits en débits massiques, il paraissait naturel d'écrire les compositions en fractions massiques. Mais le modèle cinétique

est systématiquement donné dans la littérature ([50, 25]...) pour les fractions molaires. Il faut donc adapter la formule pour les fractions massiques. Par exemple pour le cas copolymère, comme on l'a vu dans le chapitre 2, le modèle cinétique s'écrit en molaire

$$\frac{R_{E,mol}^i}{R_{C,mol}^i} = \frac{[E]^i}{[C]^i} \frac{r_E[E]^i + [C]^i}{[E]^i + r_C[C]^i} \quad (4.1)$$

Ce qui donne pour les fractions massiques

$$\frac{R_{E,mass}^i}{R_{C,mass}^i} = \frac{\mathcal{M}_E X_E^i}{\mathcal{M}_C X_C^i} \frac{r_{E,mass} X_E^i + X_C^i}{X_E^i + r_{C,mass} X_C^i}$$

où $r_{E,mass} = \frac{\mathcal{M}_C}{\mathcal{M}_E} r_E$ et $r_{C,mass} = \frac{\mathcal{M}_E}{\mathcal{M}_C} r_C$. On était alors en droit de se demander si la vitesse de convergence de l'observateur, qui dépend uniquement des équations du modèle, était affectée par cette transformation, et alors quelles étaient les unités qui donnaient la convergence la plus rapide. Mais évidemment la mise à l'échelle $[E] \mapsto X_E = \mathcal{M}_E [E] V^i / M^i$ ne change pas la structure des équations et donc la convergence de l'observateur. Le modèle cinétique est donc compatible avec une mise à l'échelle, mais pour cela il faut adapter les paramètres cinétiques r_E et r_C . Si l'on considérait par exemple un modèle de cinétique chimique qui s'écrit

$$\frac{R_{E,mol}^i}{R_{C,mol}^i} = \frac{\log([E]^i)}{\log([C]^i)} \frac{r_E[E]^i + [C]^i}{[E]^i + r_C[C]^i}$$

une mise à l'échelle aurait modifié la structure des équations et donc la convergence de l'observateur. Plus généralement, tous les modèles cinétiques imaginables sont-ils compatibles avec les mises à l'échelle qui correspondent en chimie à des changements d'unités? Nous venons de voir que non. Selon la forme du modèle on peut changer la convergence de l'observateur. On peut se demander alors quand on ajoute aux équations un terme de correction non-linéaire (comme dans la formule générale de l'observateur de Luenberger) s'il conserve la structure des équations relativement à une mise à l'échelle? Faut-il changer certains paramètres dans le terme de correction pour conserver la structure? Ce sont ces réflexions qui motivent le développement théorique de la partie II. Cette partie sera introduite au chapitre 5 par l'exemple d'un réacteur chimique exothermique classique pour lequel on veut faire un observateur non-linéaire. Le modèle cinétique sera du premier ordre, donc plus simple que ce qu'on a vu jusqu'ici, mais on se concentrera sur les problèmes de mise à l'échelle des termes de correction.

Deuxième partie

Observateurs et symétries

Chapitre 5

Un réacteur exothermique

Ce chapitre a fait l'objet d'une publication [14]. On considère le réacteur chimique exothermique classique parfaitement agité de [5]. Avec des notations légèrement différentes la dynamique s'écrit :

$$\begin{aligned}\frac{d}{dt}X^{in} &= 0 \\ \frac{d}{dt}X &= D(t)(X^{in} - X) - k \exp\left(-\frac{E_A}{RT}\right) X \\ \frac{d}{dt}T &= D(t)(T^{in}(t) - T) + c \exp\left(-\frac{E_A}{RT}\right) X + v(t) \\ y &= T\end{aligned}\tag{5.1}$$

où (E_A, R, k, c) sont des paramètres positifs et connus, $D(t)$, $T^{in}(t)$ et $v(t)$ sont des fonctions connues du temps et le débit $D(t)$ vérifie $D(t) \geq 0$. Les mesures disponibles en temps réel sont la température à l'intérieur du réacteur T supposée uniforme. Le paramètre $X^{in} > 0$, qui est la composition d'entrée, supposée constante, n'est pas connu. La composition X n'est pas mesurée. On cherche donc à estimer X et X^{in} à partir de la seule mesure T . Nous allons montrer comment l'invariance de ces équations par rapport au choix des unités sur X nous a conduit à la synthèse d'un observateur asymptotique non linéaire invariant globalement convergent.

Ce modèle s'applique au réacteur considéré au chapitre 2 dans le cas homopolymère avec une quantité constante de catalyseur dans la zone : on retrouve les équations (2.10). On connaissait alors la variable X^{in} qui valait 1 et l'observateur de la composition était nettement plus simple.

5.1 Symétries

La première équation différentielle exprime que X^{in} est une constante. Les deux dernières équations différentielles consistent en un bilan massique et un bilan enthalpique. On peut remarquer que leur structure est indépendante des unités : ces équations s'écrivent exactement de la même manière que les unités soient massiques ou molaires. En effet cela se voit aisément en considérant la transformation de mise à l'échelle $X \mapsto gX$ et $X^{in} \mapsto gX^{in}$ avec $g > 0$ un réel positif. On choisit comme variables d'état $x = (X^{in}, X, T)$ et $u = (c, D(t), T^{in}(t), v(t))$ représente les entrées connues qu'elles soient constantes (paramètres) ou variables dans le temps (contrôles ou perturbations). L'opération de conversion des unités peut être vue comme une transformation sur l'état x et les entrées u :

$$\begin{pmatrix} X^{in} \\ X \\ T \end{pmatrix} \mapsto \varphi_g(x) = \begin{pmatrix} gX^{in} \\ gX \\ T \end{pmatrix}, \quad \begin{pmatrix} c \\ D \\ T^{in} \\ v \end{pmatrix} \mapsto \psi_g(u) = \begin{pmatrix} c/g \\ D \\ T^{in} \\ v \end{pmatrix}.$$

En fait ces transformations correspondent à l'action du groupe multiplicatif $G = \mathbb{R}_+^*$ sur l'espace $\mathcal{X} \times \mathcal{U}$ via les applications (linéaires) φ_g et ψ_g définies pour chaque élément du groupe g . On dit que la dynamique (5.1) est invariante sous l'action du groupe de symétrie G . On remarque qu'il est nécessaire que le groupe agisse également sur l'espace des entrées via ψ_g , car le paramètre c est une enthalpie molaire qui doit être convertie en enthalpie massique. On vérifie effectivement que l'on a bien pour tout $g > 0$ les mêmes équations pour les variables écrites dans les nouvelles unités :

$$\begin{aligned} \frac{d}{dt}(gX^{in}) &= 0 \\ \frac{d}{dt}(gX) &= D(t)((gX^{in}) - (gX)) - k \exp\left(-\frac{E_A}{RT}\right)(gX) \\ \frac{d}{dt}T &= D(t)(T^{in}(t) - T) + (c/g) \exp\left(-\frac{E_A}{RT}\right)(gX) + v(t) \\ y &= T \end{aligned}$$

Observateur et symétries

On désire faire un observateur (non-linéaire) pour l'état x , c'est à dire un filtre pour la température (éliminer les hautes-fréquences du signal de mesure), et un estimateur pour les compositions X^{in} et X . Ici on ne peut pas simplement copier les équations sur X et X^{in} en remplaçant T par sa mesure : par exemple la dynamique sur X^{in} n'est pas asymptotiquement stable (comme on l'a fait pour le cas de réacteur de polyéthylène haute pression dans la partie I). Il faut maintenant construire un observateur complet de type

Lunenberger qui a la structure suivante :

$$\begin{cases} \frac{d}{dt} \hat{X}^{in} = L_1(\hat{T} - T) \\ \frac{d}{dt} \hat{X} = D(t)(\hat{X}^{in} - \hat{X}) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + L_2(\hat{T} - T) \\ \frac{d}{dt} \hat{T} = D(t)(T^{in}(t) - \hat{T}) + c \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + v(t) + L_3(\hat{T} - T) \end{cases} \quad (5.2)$$

Si l'on effectue un chagement d'unité comme précédemment, et que l'on ré-écrit les équations de l'observateur on obtient :

$$\begin{cases} \frac{d}{dt}(g\hat{X}^{in}) = L_1(\hat{T} - T) \\ \frac{d}{dt}(g\hat{X}) = D(t)((g\hat{X}^{in}) - (g\hat{X})) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) (g\hat{X}) + L_2(\hat{T} - T) \\ \frac{d}{dt}(\hat{T}) = D(t)(T^{in}(t) - \hat{T}) + (c/g) \exp\left(-\frac{E_A}{R\hat{T}}\right) (g\hat{X}) + v(t) + L_3(\hat{T} - T) \end{cases}$$

ou encore

$$\begin{cases} \frac{d}{dt} \hat{X}^{in} = (L_1/g)(\hat{T} - T) \\ \frac{d}{dt} \hat{X} = D(t)(\hat{X}^{in} - \hat{X}) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + (L_2/g)(\hat{T} - T) \\ \frac{d}{dt} \hat{T} = D(t)(T^{in}(t) - \hat{T}) + c \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + v(t) + L_3(\hat{T} - T) \end{cases}$$

C'est-à-dire que les gains L_1 et L_2 ont été divisés par g . Par conséquent les gains de l'observateur (5.2) et donc la convergence dépendent des unités choisies et donc la structure des équations ne respecte pas les symétries. Cet observateur respecte néanmoins une forme d'invariance. En effet si l'on rajoute dans les entrées u les gains L_i avec $i \in \{1, 2\}$ et on étend ψ_g de manière à ce que les L_i soient transformés en L_i/g alors on obtient un système encore invariant. Mais ce que l'on cherche ici c'est à avoir des gains L_i qui soit eux mêmes des invariants : i.e. les paramètres de réglage de l'observateur qui soient indépendants du système d'unité choisi. C'est pour cela qu'on considère qu'un tel observateur n'est pas invariant. Dans la suite, nous allons nous efforcer de construire un observateur non-linéaire indépendant des unités, donc qui respecte les symétries. Nous dirons indifféremment que "l'observateur est invariant", ou "respecte les symétries", ou pour cet exemple "est indépendant des unités".

Construction de l'observateur

La construction des observateurs qui respectent les symétries est détaillée dans le cas général au chapitre suivant. Certaines affirmations de cette section trouvent leur justifi-

cation au chapitre suivant. La construction est basée sur les notions d'erreur de sortie invariante, et de champs invariants.

Erreur de sortie invariante

Une erreur de sortie invariante pour notre exemple est une fonction à valeurs réelles $E(\hat{x}, u, y)$ qui est invariante par changement d'unités : $E(\varphi_g(\hat{x}), \psi_g(u), y) = E(\hat{x}, u, y)$ et qui est nulle lorsque la sortie estimée égale la sortie mesurée : $E(x, u, y) = 0$. Par exemple une erreur de sortie invariante est $E_0(\hat{x}, u, y) = \hat{T} - y = \hat{T} - T$. L'erreur de sortie invariante la plus générale s'écrit comme une fonction de E_0 et de tous les invariants fondamentaux $I(\hat{x}, u)$ de l'action du groupe sur l'espace d'état et des entrées de l'observateurs. Les invariants fondamentaux de l'action du groupe consistent ici en toutes les variables que l'on peut former avec \hat{x} et u et invariants par l'action du groupe, donc par une conversion de *mol* en *kg* :

$$I(\hat{x}, u) = (\hat{X}/\hat{X}^{in}, \hat{T}, c\hat{X}^{in}, D, T^{in}, v).$$

et alors l'erreur de sortie invariante la plus générale s'écrit :

$$\mathcal{L}\left(I(\hat{x}, u), \hat{T} - T\right)$$

où \mathcal{L} est n'importe quelle fonction régulière de ses arguments qui vérifie $\mathcal{L}(I, 0) \equiv 0$.

Repère invariants

Les champs de vecteurs invariants sont des champs de vecteurs sur l'espace d'état $w(x)$ tels que la dynamique définie par $\frac{d}{dt}x = w(x)$ est invariante sous l'action du groupe. Un repère invariant est un ensemble de champs invariants définis sur \mathcal{X} qui constituent en tout point une base de l'espace tangent. Ici un repère invariant par exemple est l'ensemble des trois champs de vecteurs

$$w_1 = X^{in} \frac{\partial}{\partial X^{in}}, \quad w_2 = X \frac{\partial}{\partial X}, \quad w_3 = \frac{\partial}{\partial T}$$

L'observateur

L'observateur non-linéaire invariant le plus général sous l'action du groupe, autrement dit l'observateur indépendant du choix des unités, est

$$\left\{ \begin{array}{l} \frac{d}{dt} \hat{X}^{in} = \mathcal{L}_1\left(I(\hat{x}, u), \hat{T} - T\right) \hat{X}^{in} \\ \frac{d}{dt} \hat{X} = D(t)(\hat{X}^{in} - \hat{X}) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + \mathcal{L}_2\left(I(\hat{x}, u), \hat{T} - T\right) \hat{X} \\ \frac{d}{dt} \hat{T} = D(t)(T^{in}(t) - \hat{T}) + c \exp\left(-\frac{E_A}{R\hat{T}}\right) \hat{X} + v(t) + \mathcal{L}_3\left(I(\hat{x}, u), \hat{T} - T\right) \end{array} \right. \quad (5.3)$$

où les \mathcal{L}_i sont des fonctions scalaires régulières qui vérifient $\mathcal{L}_i(I, 0) \equiv 0$. On remarquera que tout observateur de cette forme préserve notamment automatiquement le fait que \hat{X} et \hat{X}^{in} sont des quantités positives. En effet le domaine $\{(\hat{X}^{in}, \hat{X}, T) \in \mathbb{R}^3 \mid \hat{X}^{in} > 0, \hat{X} > 0\}$ est positivement invariant pour (5.3), quel que soient les choix que l'on fait pour \mathcal{L}_1 , \mathcal{L}_2 et \mathcal{L}_3 (car on a $D(t) \geq 0$).

Convergence autour d'un équilibre

Il semble intéressant, avant même d'étudier une éventuelle convergence globale, de se pencher sur les possibilités de design des gains autour d'un point d'équilibre pour un observateur invariant. On remarque en fait qu'autour d'un équilibre, on ne perd rien en termes de propriétés de convergence locale à choisir un observateur qui respecte les symétries. Supposons qu'autour de l'équilibre $(\bar{X}^{in}, \bar{X}, \bar{T})$ de (5.1), l'on ait construit trois gains constants L_1 , L_2 , et L_3 , tels que l'observateur

$$\begin{aligned}\frac{d}{dt}\hat{X}^{in} &= L_1(\hat{T} - T) \\ \frac{d}{dt}\hat{X} &= D(t)(\hat{X}^{in} - \hat{X}) - k \exp\left(-\frac{E_A}{R\hat{T}}\right)\hat{X} + L_2(\hat{T} - T) \\ \frac{d}{dt}\hat{T} &= D(t)(T^{in}(t) - \hat{T}) + c \exp\left(-\frac{E_A}{R\hat{T}}\right)\hat{X} + v(t) + L_3(\hat{T} - T)\end{aligned}$$

est localement convergent autour de $(\bar{X}^{in}, \bar{X}, \bar{T})$. Cet observateur de type Luenberger a été considéré plus haut (5.2) et on a montré sa dépendance au choix des unités. On peut construire l'observateur invariant suivant

$$\begin{aligned}\frac{d}{dt}\hat{X}^{in} &= L_1(\hat{T} - T)\frac{\hat{X}^{in}}{\bar{X}^{in}} \\ \frac{d}{dt}\hat{X} &= D(t)(\hat{X}^{in} - \hat{X}) - k \exp\left(-\frac{E_A}{R\hat{T}}\right)\hat{X} + L_2(\hat{T} - T)\frac{\hat{X}}{\bar{X}} \\ \frac{d}{dt}\hat{T} &= D(t)(T^{in}(t) - \hat{T}) + c \exp\left(-\frac{E_A}{R\hat{T}}\right)\hat{X} + v(t) + L_3(\hat{T} - T)\end{aligned}$$

qui a la même approximation linéaire autour de l'équilibre. De plus il permet d'avoir X et X^{in} automatiquement positifs, et ses performances sont indépendantes du choix des unités.

5.2 Convergence globale

Pour étudier la convergence globale et donc faire le design des gains de l'observateur, il s'avère utile de choisir un jeu de coordonnées liées à l'action du groupe. En effet considérons

sur le domaine physique $\{(X^{in}, X, T) \in \mathbb{R}^3 \mid X^{in} > 0, X > 0\}$ le changement de variable suivant :

$$\begin{pmatrix} X \\ X^{in} \\ T \end{pmatrix} \mapsto \begin{pmatrix} Z = \log(X) \\ \xi = \log(X/X^{in}) \\ T \end{pmatrix}$$

Ces coordonnées s'appellent les coordonnées de base et de fibre. Si l'on considère l'action du groupe G sur l'espace \mathcal{X} , elle engendre des orbites qui sont (par définition) les ensembles $\{\varphi_g(x), g \in \mathbb{R}_+^*\}$. Les coordonnées de base numérotent les orbites et les coordonnées de fibre paramètrent les orbites. Une définition plus précise sera donnée au chapitre suivant. En effet ici ξ, T sont les coordonnées de base, puisqu'elles sont constantes sur les orbites car invariantes par changement d'unité, et Z est la coordonnée de fibre qui permet de distinguer deux points d'une même orbite. La dynamique (5.1) s'écrit dans ces coordonnées :

$$\begin{aligned} \frac{d}{dt}Z &= D(\exp(-\xi) - 1) - k \exp\left(-\frac{E_A}{RT}\right) \\ \frac{d}{dt}\xi &= D(\exp(-\xi) - 1) - k \exp\left(-\frac{E_A}{RT}\right) \\ \frac{d}{dt}T &= D(T^{in} - T) + c \exp Z \exp\left(-\frac{E_A}{RT}\right) + v(t) \end{aligned}$$

et l'observateur invariant (5.3) s'écrit plus simplement :

$$\begin{aligned} \frac{d}{dt}\hat{Z} &= D(\exp(-\hat{\xi}) - 1) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) + \mathcal{L}_2 \\ \frac{d}{dt}\hat{\xi} &= D(\exp(-\hat{\xi}) - 1) - k \exp\left(-\frac{E_A}{R\hat{T}}\right) + \mathcal{L}_2 - \mathcal{L}_1 \\ \frac{d}{dt}\hat{T} &= D(T^{in} - \hat{T}) + c \exp \hat{Z} \exp\left(-\frac{E_A}{R\hat{T}}\right) + v(t) + \mathcal{L}_3 \end{aligned}$$

Prenons le design des gains suivants ($\beta > 0$ et $\kappa > 0$ sont deux paramètres arbitraires)

$$\begin{aligned} \mathcal{L}_2 &= -\beta c \exp \hat{Z} \exp\left(-\frac{E_A}{RT}\right) (\hat{T} - T) + k \exp\left(-\frac{E_A}{R\hat{T}}\right) - k \exp\left(-\frac{E_A}{RT}\right) \\ \mathcal{L}_2 - \mathcal{L}_1 &= k \exp\left(-\frac{E_A}{R\hat{T}}\right) - k \exp\left(-\frac{E_A}{RT}\right) \\ \mathcal{L}_3 &= (-\kappa c \exp\left(-\frac{E_A}{RT}\right) (\hat{T} - T) - c \exp\left(-\frac{E_A}{R\hat{T}}\right) \\ &\quad + c \exp\left(-\frac{E_A}{RT}\right)) \exp \hat{Z} + D(\hat{T} - T) \end{aligned}$$

Nous allons montrer qu'un tel choix conduit à la stabilité asymptotique globale du moment que les mesures de concentration d'entrée, de débit et de température sont bornées dans le sens où il existe M et $\alpha > 0$ tels que $t \geq 0$, $M \geq X^{in}, D(t), T(t) \geq \alpha$. Cela implique d'après l'équation sur X du système (5.1) qu'il existe $\sigma > 0$ tel que $t \geq 0$, $M \geq X^{in}, X(t), D(t), T(t) \geq \sigma$. Le design des gains, bien que spécifique à cet exemple, repose sur la notion d'erreur d'état invariante (voir 6.2.6) présentée au chapitre suivant. En effet on considère comme erreur d'état le terme invariant par changement d'unités (invariant sous l'action du groupe) $\eta = (\tilde{Z}, \tilde{\xi}, \tilde{T})$ où

$$\begin{aligned}\tilde{Z} &= \hat{Z} - Z = \log(\hat{X}/X) \\ \tilde{\xi} &= \hat{\xi} - \xi = \log(\hat{X}/\hat{X}^{in}) - \log(X/X^{in}) \\ \tilde{T} &= \hat{T} - T\end{aligned}$$

Le terme η est appelée erreur d'état invariante car il est nul uniquement quand l'état estimé est égal au vrai état du système, et de plus il est invariant sous l'action du groupe. Ici cette erreur coïncide avec l'erreur d'état linéaire usuelle à cause du choix particulier des coordonnées (Z, ξ, T) , mais en général l'erreur d'état invariante est une fonction non-linéaire de l'état. Sa dynamique est la suivante :

$$\begin{aligned}\frac{d}{dt}\tilde{Z} &= D(\exp(-\hat{\xi}) - \exp(-\xi)) - \beta c \exp\left(-\frac{E_A}{RT} + Z\right) \exp \tilde{Z} \tilde{T} \\ \frac{d}{dt}\tilde{\xi} &= D \exp(-\xi)(\exp(-\tilde{\xi}) - 1) \\ \frac{d}{dt}\tilde{T} &= c \exp\left(-\frac{E_A}{RT} + Z\right) (\exp \tilde{Z} - 1) - \kappa c \exp\left(-\frac{E_A}{RT(t)} + Z\right) \exp \tilde{Z} \tilde{T}\end{aligned}$$

Puisque $M \geq X^{in}, D(t), X(t), T(t) \geq \sigma > 0$, $\xi(t)$ reste borné pour tout $t \geq 0$ et $\lim_{t \rightarrow +\infty} (\hat{\xi}(t) - \xi(t)) = 0$, ce qui veut dire que la dynamique sur la coordonnée de base ξ est stable. On considère maintenant le système réduit en (\tilde{Z}, \tilde{T}) :

$$\begin{aligned}\frac{d}{dt}\tilde{Z} &= -\beta c X \exp\left(-\frac{E_A}{RT} + Z\right) \exp \tilde{Z} \tilde{T} + \epsilon_1(t) \\ \frac{d}{dt}\tilde{T} &= c X \exp\left(-\frac{E_A}{RT} + Z\right) (\exp \tilde{Z} - 1) - \kappa c X \exp\left(-\frac{E_A}{RT(t)} + Z\right) \exp \tilde{Z} \tilde{T}.\end{aligned}$$

où l'on a posé $\epsilon_1(t) = D(\exp(-\hat{\xi}(t)) - \exp(-\xi(t)))$ et l'on sait que $\lim_{t \rightarrow +\infty} \epsilon_1(t) = 0$ et $\int |\epsilon_1(t)| dt < \infty$. Faisons un changement d'échelle de temps $\tau = \int_0^t c X(s) \exp\left(-\frac{E_A}{RT(s)} + Z(s)\right) ds$. Alors :

$$\begin{aligned}\frac{d\tilde{Z}}{d\tau} &= -\beta \exp \tilde{Z} \tilde{T} + \epsilon(t) \\ \frac{d\tilde{T}}{d\tau} &= (\exp \tilde{Z} - 1) - \kappa \exp \tilde{Z} \tilde{T}\end{aligned}\tag{5.4}$$

où $\epsilon(t) = cX \exp\left(\frac{E_A}{RT} - Z\right) \epsilon_1(t)$. On considère $V = \tilde{Z} + \exp(-\tilde{Z}) - 1 + \frac{\beta}{2}\tilde{T}^2$ comme fonction de Lyapunov. V est une fonction positive qui atteint son minimum unique pour $\tilde{Z} = \tilde{T} = 0$. On a

$$\frac{d}{dt}V = -\beta\kappa\tilde{T}^2 \exp \tilde{Z} + \epsilon(t)(1 - \exp(-\tilde{Z}))$$

Cela implique que $\frac{d}{dt}V \leq |\epsilon(t)(1 - \exp(-\tilde{Z}))|$. Comme pour tout $x \in \mathbb{R}$ on a $2(\exp(-x) + x) - 1 \geq |1 - \exp(-x)|$, cela implique

$$\frac{d}{dt}V \leq |\epsilon(t)|(1 + 2V)$$

Par conséquent V est majorée pour $t > 0$ par la fonction

$$t \mapsto [V(0) + \int_0^t |\epsilon(\tau)| e^{(-2 \int_0^t |\epsilon(s)| ds)} d\tau] \exp\left(\int_0^t |\epsilon(\tau)| d\tau\right)$$

qui est une fonction bornée. Cela implique que V est bornée et donc que les trajectoires $(\tilde{Z}(t), \tilde{T}(t))$ sont bornées dans \mathbb{R}^2 . Soit une trajectoire particulière du système $(\bar{Z}(t), \bar{T}(t))$. On considère maintenant pour cette trajectoire la fonction

$$U(t) = V(\bar{Z}, \bar{T}, t) - \int_t^\infty \epsilon(\tau)(1 - \exp(-\bar{Z}(\tau))) d\tau$$

Elle est bornée inférieurement. Sa dérivée est $\frac{d}{dt}U = -\beta\kappa\bar{T}^2 \exp \bar{Z} < 0$ et c'est une fonction uniformément continue en temps. D'après le lemme de Barbalat elle tend vers 0 quand t tend vers $+\infty$. Comme \bar{Z} est bornée cela veut dire que \bar{T} tend vers 0. Revenons au système initial (5.4). On voit que \bar{T} est bornée. Donc $\frac{d}{dt}T$ est uniformément continue. Encore d'après le lemme de Barbalat cela veut dire que $\frac{d}{dt}T$ tend vers 0 quand $t \rightarrow +\infty$. Cela implique donc que $\exp \bar{Z}$ tend vers 1 et donc $\lim_{t \rightarrow +\infty} \bar{Z} = 0$. Nous avons montré que $(0, 0)$ est globalement asymptotiquement stable.

Guidés par des considérations sur l'invariance sous l'action d'un groupe, donc la prise en compte des symétries du système, on a montré le

Théorème 1. *Soit le système (5.1). On suppose qu'il existe M et $\alpha > 0$ tels que pour tout $t \geq 0$, $M \geq X^{in}, D(t), T(t) \geq \alpha$. Alors pour tout choix de gains invariants $\beta, \kappa > 0$ l'observateur non-linéaire suivant :*

$$\begin{aligned} \frac{d}{dt} \hat{X}^{in} &= -\beta \exp\left(-\frac{E_A}{RT(t)}\right) (\hat{T} - T(t)) c\hat{X} \hat{X}^{in} \\ \frac{d}{dt} \hat{X} &= D(t)(\hat{X}^{in} - \hat{X}) - \exp\left(-\frac{E_A}{RT(t)}\right) \left(k + \beta(\hat{T} - T(t))c\hat{X}\right) \hat{X} \\ \frac{d}{dt} \hat{T} &= \exp\left(-\frac{E_A}{RT(t)}\right) \left(1 - \kappa(\hat{T} - T(t))\right) c\hat{X} + D(t)(T^{in}(t) - T(t)) + v(t) \end{aligned}$$

est globalement convergent.

Chapitre 6

Théorie des observateurs invariants

Ce chapitre reprend une partie du rapport disponible sous arxiv [12] et dont une partie est acceptée pour publication dans IEEE Automatic Control.

6.1 Introduction

Les symétries ont été utilisées en théorie du contrôle pour la synthèse de bouclage et également en contrôle optimal, voir par exemple [22, 24, 46, 47, 40, 49] mais beaucoup moins pour le design d'observateurs [3, 2, 38, 37, 26]. Dans ce chapitre, on développe une théorie des observateurs invariants. La théorie est motivée par des exemples liés à des problématiques de l'ingénieur : un réacteur chimique exothermique présenté au chapitre précédent, une voiture non-holonome, et un système de navigation inertielle assistée par mesure de vitesse. Dans chacun des cas les symétries ont une interprétation physique simple. Pour le premier exemple, nous avons exhibé au chapitre précédent un observateur globalement convergent (théorème 1, page 48). Pour le second nous proposons un observateur non-linéaire globalement convergent pour n'importe quelle condition initiale à l'exception d'une (théorème 5, page 76). En ce qui concerne le troisième, l'observateur est localement convergent autour de n'importe quelle trajectoire. De plus, le comportement global de l'observateur est indépendant de la trajectoire du système. Cette théorie peut s'appliquer à beaucoup d'autres systèmes comme ceux traités dans [21, 37, 26] où l'invariance est relative au choix de l'orientation et de l'origine du repère 3D, et est exploitée dans le design de l'observateur et dans l'analyse de la convergence.

La contribution théorique de ce chapitre et de la thèse en général est la suivante : pour le système régulier d'état x , d'entrée u et de sortie y , l'invariance sous l'action d'un groupe de Lie G est définie quand G agit séparément sur l'espace d'état, l'espace des entrées et l'espace de sortie. L'invariance signifie que la dynamique $\frac{d}{dt}x = f(x, u)$ ainsi que l'application de sortie $y = h(x)$ sont inchangées par une famille de changements de coordonnées particuliers sur les variables d'état, d'entrée, et de sortie, et qui sont l'action de G sur ces variables.

On définit l'invariance pour un observateur asymptotique de Luenberger sous l'action du groupe G de manière similaire, en copiant l'action de groupe sur l'espace des états estimés ainsi que sur l'espace des sorties estimées.

Quand la dimension du groupe est plus petite que la dimension de l'espace d'état on propose (théorème 2, page 57) une méthode de construction des observateurs invariants candidats ou pré-observateurs (on appelle un pré-observateur un système dynamique qui a la structure d'un observateur mais dont on n'a pas montré qu'il converge). La construction repose sur les notions de repère invariant, et sur la notion d'erreur de sortie invariante. De telles erreurs de sortie invariantes (definition 8, page 57) avaient déjà été introduites par les auteurs de [2] et peuvent être calculées grâce à la méthode du repère mobile d'Elie Cartan (théorème 3, page 58).

Quant aux problèmes de convergence, nous n'avons pas de méthode pour obtenir la convergence systématique des observateurs non-linéaires quand il y a des symétries. En revanche, nous introduisons la notion d'erreur d'état invariante. Les trois exemples montrent que cette notion joue un rôle clé dans l'analyse de la convergence. Quand la dimension du groupe coïncide avec la dimension de l'état, i.e, quand on peut identifier l'espace d'état au groupe, et son action à une multiplication à gauche sur le groupe, on a des formules explicites pour les pré-observateurs invariants, les erreurs de sorties et les erreurs d'état invariantes, l'équation d'erreur et son approximation linéaire.

Dans le chapitre 8 nous verrons comment l'on peut transformer n'importe quel observateur de Luenberger autour d'un point d'équilibre en un observateur invariant ayant la même approximation linéaire autour du point, ce qui prouve qu'il n'est pas contraignant d'imposer que l'observateur respecte les symétries autour d'un point d'équilibre. De plus on introduira la notion de trajectoire permanente qui étend la notion de point d'équilibre pour les systèmes qui présentent des symétries. En particulier autour de ces trajectoires l'équation d'erreur à la propriété remarquable qu'elle est une équation différentielle non-linéaire autonome.

Ce chapitre est organisé comme suit : dans la section 6.2, on définit les systèmes invariants et les observateurs invariants. La forme générale des observateurs invariants est donnée par le théorème 2 : elle repose sur les erreurs de sorties invariantes et les champs de vecteurs invariants. La construction explicite repose que la méthode du repère mobile de Cartan [45], qui est résumée dans la sous-section 6.2.3. Afin d'étudier la convergence, on définit la notion d'erreur d'état invariante. C'est une manière de définir une équation d'erreur qui préserve les symétries, puisque l'erreur linéaire usuelle $\hat{x} - x$ ne respecte pas les symétries en général. L'erreur d'état invariante est solution d'un système différentiel où seulement la partie invariante de la trajectoire apparaît (lemme 4). Cette propriété permet de réduire la dimension du problème de la convergence et joue un rôle crucial dans les exemples.

Dans la section 6.3, on considère le cas particulier où la dimension de G est égale à la dimension de l'espace d'état. On identifie alors G et l'espace d'état, et on voit l'action de G comme une multiplication à gauche. Alors les formules qui donnent la forme des

pré-observateurs, les erreurs de sortie invariantes, et les erreurs d'état invariantes sont plus simples, et peuvent être explicitées. On peut alors raffiner les résultats des sections précédentes.

Des résultats préliminaires peuvent être trouvés dans [2, 1, 14, 11, 10].

6.2 Systèmes invariants, observateurs et erreurs invariants

6.2.1 Système invariant et sortie compatible

Définition 1. Soit G un groupe de Lie dont on note e l'élément neutre et Σ un ouvert de \mathbb{R}^n (ou plus généralement une variété de dimension n). Un groupe de transformation $(\phi_g)_{g \in G}$ sur Σ est une application régulière

$$(g, \xi) \in G \times \Sigma \mapsto \phi_g(\xi) \in \Sigma$$

telle que :

- $\phi_e(\xi) = \xi$ pour tout ξ
- $\phi_{g_2}(\phi_{g_1}(\xi)) = \phi_{g_2 g_1}(\xi)$ pour tout g_1, g_2, ξ .

On peut remarquer que ϕ_g est par construction un difféomorphisme sur Σ pour tout g . Le groupe de transformation est dit *local* si l'action $\phi_g(\xi)$ est définie pour g suffisamment proche de e . La loi de transformation n'est alors valide que quand la formule $\phi_{g_2}(\phi_{g_1}(\xi)) = \phi_{g_2 g_1}(\xi)$ a un sens. Tous les résultats de ce chapitre sont des résultats locaux, puisqu'ils sont basés sur des hypothèses de rang constant, on ne considère donc que des groupes de transformations locaux agissant sur des ouverts. Quand on dit "pour tout g " cela veut dire "pour g suffisamment proche de l'élément neutre e de G "; de manière similaire "pour tout ξ " veut dire "pour tout ξ générique dans Σ ". On utilisera systématiquement ces raccourcis stylistiques afin de rendre la lecture plus aisée.

Considérons le système régulier

$$\frac{d}{dt}x = f(x, u) \tag{6.1}$$

$$y = h(x, u) \tag{6.2}$$

où x appartient à un ouvert $\mathcal{X} \subset \mathbb{R}^n$, u à un ouvert $\mathcal{U} \subset \mathbb{R}^m$ et y à un ouvert $\mathcal{Y} \subset \mathbb{R}^p$, $p \leq n$.

On suppose que les signaux $u(t), y(t)$ sont connus (y mesuré, et u est une entrée connue ou mesurée - contrôle, perturbation mesurée, paramètre constant).

Considérons le groupe de transformations local sur $\mathcal{X} \times \mathcal{U}$ défini par

$$(X, U) = (\varphi_g(x), \psi_g(u)), \tag{6.3}$$

où φ_g et ψ_g sont des difféomorphismes locaux. On remarque que φ_g agit sur \mathcal{X} et ψ_g agit sur \mathcal{U} . u peut aussi désigner le temps t mais dans ce cas ψ_g est la fonction identité. Les deux définitions suivantes s'inspirent de [40].

Définition 2. Le système $\frac{d}{dt}x = f(x, u)$ est dit G -invariant si $f(\varphi_g(x), \psi_g(u)) = D\varphi_g(x) \cdot f(x, u)$ pour tout g, x, u .

La propriété s'écrit également $\frac{d}{dt}X = f(X, U)$, i.e., le système est inchangé par la transformation (6.3).

Définition 3. La sortie $y = h(x, u)$ est G -compatible s'il existe un groupe de transformations $(\varrho_g)_{g \in G}$ sur \mathcal{Y} tel que $h(\varphi_g(x), \psi_g(u)) = \varrho_g(h(x, u))$ pour tout g, x, u .

On peut aussi utiliser la terminologie standard G -équivariante au lieu de G -compatible. Si l'on note $(X, U) = (\varphi_g(x), \psi_g(u))$ et $Y = \varrho_g(y)$, la définition veut dire que $Y = h(X, U)$. Les deux définitions précédentes peuvent être illustrées par le diagramme commutatif :

$$\begin{array}{ccc}
 T\mathcal{X} & \xrightarrow{D\varphi_g} & T\mathcal{X} \\
 f \uparrow & & f \uparrow \\
 \mathcal{X} \times \mathcal{U} & \xrightarrow{\varphi_g \times \psi_g} & \mathcal{X} \times \mathcal{U} \\
 h \downarrow & & h \downarrow \\
 \mathcal{Y} & \xrightarrow{\varrho_g} & \mathcal{Y}
 \end{array}$$

6.2.2 Hypothèses de base

A partir de maintenant on considère un système G -invariant $\frac{d}{dt}x = f(x, u)$ avec une sortie G -équivariante $y = h(x, u)$. Soit $r \leq n$ la dimension du groupe G . On suppose que pour tout x , l'application $g \mapsto \varphi_g(x)$ est de rang plein.

6.2.3 Méthode du repère mobile, champs de vecteurs invariants, coordonnées de base et de fibre

Méthode du repère mobile

On suit ici la présentation de [45, théorème 8.25]. Soit un groupe de transformation de dimension r , $(\phi_g)_{g \in G}$ qui agit sur $\Sigma \subset \mathbb{R}^s$ tel que $r \leq s$. Nous supposons que $\partial_g \phi_g$ est de rang plein $r := \dim G$ au point $(e, \xi^0) \in G \times \Sigma$. On peut alors décomposer ϕ_g en (ϕ_g^a, ϕ_g^b) dont les composantes sont de dimensions respectives r et $s - r$, et de sorte que ϕ_g^a est inversible par rapport à g autour de (e, ξ^0) . Les équations de normalisation s'obtiennent en fixant

$$\phi_g^a(\xi) = c,$$

où c est une constante appartenant à l'image de ϕ^a . Le théorème des fonctions implicites assure l'existence d'une solution locale $g = \gamma(\xi)$ (l'application $\gamma : \Sigma \rightarrow G$ s'appelle le *repère mobile*). On peut aussi dire que $\{\phi_e^a(\xi) = c\}$ définit une section transverse aux orbites, et que $g = \gamma(\xi)$ est l'unique élément qui envoie ξ sur la section transverse. Finalement, on obtient un ensemble complet J de $s - r$ invariants fonctionnellement indépendants en substituant $g = \gamma(\xi)$ dans les autres composantes de la transformation,

$$J(\xi) := \phi_{\gamma(\xi)}^b(\xi).$$

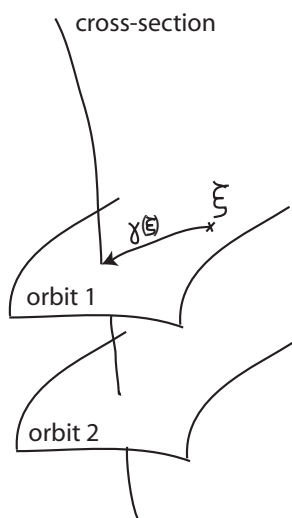
La propriété d'invariance veut dire que $J(\phi_g(\xi)) = J(\xi)$ pour tout g, ξ . Pour le prouver soit $\zeta = \phi_g(\xi)$. On a $\phi_{\gamma(\phi_g(\xi))}^a(\phi_g(\xi)) = \phi_{\gamma(\zeta)}^a(\zeta) = c$. Mais la loi de composition de groupe implique $\phi_{\gamma(\phi_g(\xi))}^a(\phi_g(\xi)) = \phi_{\gamma(\phi_g(\xi))g}^a(\xi)$. Donc $\phi_{\gamma(\phi_g(\xi))g}^a(\xi) = c$ ce qui prouve par unicité de $\gamma(\xi) \in G$

$$\gamma(\phi_g(\xi))g = \gamma(\xi) \tag{6.4}$$

qui la propriété principale (équivariance) du repère mobile et qui prouve en effet

$$J(\phi_g(\xi)) = \phi_{\gamma(\phi_g(\xi))}^b(\phi_g(\xi)) = \phi_{\gamma(\phi_g(\xi))g}^b(\xi) = \phi_{\gamma(\xi)}^b(\xi) = J(\xi)$$

De plus, tout invariant local J' , i.e, n'importe quelle fonction à valeurs réelles J' qui vérifie $J'(\phi_g(\xi)) = J'(\xi)$ pour tout g, ξ , peut s'écrire comme une fonction de l'ensemble complet (d'où la terminologie) d'invariants : $J' = \mathcal{H}(J)$.

FIG. 6.1 – Méthode du repère mobile et définition de la fonction γ

Champs de vecteurs invariants et repère invariant

La méthode du repère mobile permet de construire des repères invariants, qui jouent un rôle dans la construction des observateurs invariants.

Définition 4. *Un champ de vecteurs w sur \mathcal{X} est dit G -invariant si le système $\frac{d}{dt}x = w(x)$ est lui-même invariant. Cela veut dire que $w(\varphi_g(x)) = D\varphi_g(x) \cdot w(x)$ pour tout g, x .*

Définition 5. *Un repère invariant (w_1, \dots, w_n) sur \mathcal{X} est un ensemble de n champs de vecteurs linéairement indépendants et G -invariants, i.e, $(w_1(x), \dots, w_n(x))$ est une base de l'espace tangent à \mathcal{X} en x .*

Nous allons maintenant montrer comment l'on peut construire un repère invariant. On suit [44], théorème 2.84 et on applique la méthode du repère mobile à $\Sigma := \mathcal{X}$ et les équations de normalisation $g = \gamma(x)$ sont données par $\varphi_g^a(x) = c$.

Lemme 1. *Les champs de vecteurs définis par :*

$$w_i(x) := (D\varphi_{\gamma(x)}(x))^{-1} \cdot \frac{\partial}{\partial x_i}, \quad i = 1, \dots, n, \quad (6.5)$$

où $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ est la base canonique de \mathcal{X} , forment un repère invariant.

Démonstration. En effet, ils sont clairement linéairement indépendants. Chaque w_i est invariant car pour tout élément b du groupe on a

$$- w_i(\varphi_b(x)) = (D\varphi_{\gamma(\varphi_b(x))}(\varphi_b(x)))^{-1} \frac{\partial}{\partial x_i} \text{ et donc}$$

$$(D\varphi_b(x))^{-1} w_i(\varphi_b(x)) = [D\varphi_{\gamma(\varphi_b(x))}(\varphi_b(x)) D\varphi_b(x)]^{-1} \frac{\partial}{\partial x_i}$$

- la structure de groupe implique le fait que tous éléments du groupe c, d , on a $\varphi_c(\varphi_d)(x) = \varphi_{cd}(x)$; donc

$$D\varphi_c(\varphi_d(x)) D\varphi_d(x) = D\varphi_{cd}(x)$$

Par conséquent, avec $c = \gamma(\varphi_b(x))$ et $d = b$, on a

$$D\varphi_{\gamma(\varphi_b(x))}(\varphi_b(x)) D\varphi_b(x) = D\varphi_{\gamma(\varphi_b(x))b}(x)$$

- puisque $\gamma(\varphi_b(x))b \equiv \gamma(x)$, on a (cela correspond à la définition 4)

$$(D\varphi_b(x))^{-1} w_i(\varphi_b(x)) = (D\varphi_{\gamma(x)}(x))^{-1} \frac{\partial}{\partial x_i} = w_i(x).$$

□

Coordonnées de base et de fibre

On introduit les coordonnées de base et de fibre qui sont des coordonnées locales utiles pour exprimer le système G-invariant de la définition 2. On suppose désormais que G est le groupe de dimension r ($r \leq n$) qui agit sur $\mathcal{X} \times \mathcal{U}$ (voir (6.3)) et que pour tout x l'application $g \mapsto \varphi_g(x)$ est de rang plein. La méthode du repère mobile permet de trouver un ensemble complet d'invariants locaux $z_b \in \mathbb{R}^{n-r}$ de la réduction de l'action du groupe à son action sur l'espace \mathcal{X} . On complète cet ensemble avec $z_a \in \mathbb{R}^r$ de manière à ce que (z_a, z_b) soient des coordonnées de \mathcal{X} . Ces coordonnées sont appelées les coordonnées de fibre z_a et de base z_b (voir [45]). $\{z_a = c\}$ définit une section transverse aux orbites (voir fig 6.1). On peut dire que z_b numérote les orbites et z_a les paramètres. De plus, on peut toujours choisir z_a tel que pour tout $g \in G$ l'action du groupe de transformation s'écrive

$$\varphi_g(z_a, z_b) = (\varpi_g(z_a), z_b)$$

avec $g \mapsto \varpi_g(x)$ inversible pour tout $x \in \mathbb{R}^r$ (ce qui veut dire que z_a est une paramétrisation des éléments du groupe). Soit γ le repère mobile qui envoie z sur la section transverse aux orbites $\{z_a = c\}$. La dynamique invariante (2) s'écrit localement dans les nouvelles coordonnées :

$$\begin{aligned} \frac{d}{dt} z_a &= D\varpi_{\gamma(z)^{-1}} f_a(c, z_b, \psi_{\gamma(z)}(u)) \\ \frac{d}{dt} z_b &= f_b(c, z_b, \psi_{\gamma(z)}(u)) \end{aligned} \quad (6.6)$$

puisque le système est invariant. L'exemple du chapitre 5 illustre l'intérêt de telles coordonnées.

6.2.4 Caractérisation des pré-observateurs invariants

Définition 6 (pré-observateur). *Le système $\frac{d}{dt} \hat{x} = F(\hat{x}, u, y)$ est un pré-observateur de (6.1)-(6.2) si pour tout x, u $F(x, u, h(x, u)) = f(x, u)$.*

La définition ne traite pas de la convergence. Si l'on a de plus $\hat{x}(t) \rightarrow x(t)$ quand $t \rightarrow +\infty$ pour toute condition initiale (proche), le pré-observateur est un *observateur asymptotique*. On aurait également pu donner au pré-observateur le nom d'"observateur candidat", par analogie aux fonctions de Lyapunov candidates.

Définition 7. *Le pré-observateur $\frac{d}{dt} \hat{x} = F(\hat{x}, u, y)$ est G-invariant si pour tout g, \hat{x}, u, y ,*

$$F(\varphi_g(\hat{x}), \psi_g(u), \varrho_g(y)) = D\varphi_g(\hat{x}) \cdot F(\hat{x}, u, y).$$

La propriété s'écrit également $\frac{d}{dt} \hat{X} = F(\hat{X}, U, Y)$, avec $X = \varphi_g(x)$, $U = \psi_g(u)$ et $Y = \varrho_g(y)$. Cela veut dire que le pré-observateur est inchangé sous l'action de G sur

chacun des trois espaces \mathcal{X} , \mathcal{U} , et \mathcal{Y} via (resp.) φ_g , ψ_g et ϱ_g . On appelle évidemment *observateur invariant* un observateur asymptotique G -invariant.

En général l'erreur de sortie usuelle $\hat{y} - y = h(\hat{x}, u) - y$ ne préserve pas la géométrie du système, donc son utilisation dans la formule de l'observateur ne permet pas d'obtenir un observateur invariant. L'idée "clé" pour construire des (pré-)observateurs invariants consiste à utiliser, comme dans [2], une erreur de sortie invariante plutôt qu'une erreur de sortie linéaire.

Définition 8. *L'application régulière $(\hat{x}, u, y) \mapsto E(\hat{x}, u, y) \in \mathbb{R}^p$ est une erreur de sortie invariante si*

- *l'application $y \mapsto E(\hat{x}, u, y)$ est inversible pour tout \hat{x}, u*
- *$E(\hat{x}, u, h(\hat{x}, u)) = 0$ pour tout \hat{x}, u*
- *$E(\varphi_g(\hat{x}), \psi_g(u), \varrho_g(y)) = E(\hat{x}, u, y)$ pour tout \hat{x}, u, y*

La première et la seconde propriétés signifient que E est une "erreur de sortie", i.e. vaut zéro si, et seulement si, $h(\hat{x}, u) = y$; la troisième propriété s'écrit également $E(\hat{X}, U, Y) = E(\hat{x}, u, y)$, et correspond à l'invariance.

Théorème 2. *Considérons un système G -invariant $\frac{d}{dt}x = f(x, u)$ avec une sortie G -compatible $y = h(x, u)$.*

$\frac{d}{dt}\hat{x} = F(\hat{x}, u, y)$ est un observateur G -invariant si, et seulement si,

$$F(\hat{x}, u, y) = f(\hat{x}, u) + \sum_{i=1}^n \mathcal{L}_i(I(\hat{x}, u), E(\hat{x}, u, y))w_i(\hat{x}),$$

avec E erreur de sortie invariante, les \mathcal{L}_i sont des fonctions régulières telles que pour tout \hat{x}, u , $\mathcal{L}_i(I(\hat{x}, u), 0) = 0$, les (w_1, \dots, w_n) forment un repère invariant, et I est un ensemble complet d'invariants des variables (\hat{x}, u) .

Puisque tout \mathcal{L}_i est régulier et satisfait $\mathcal{L}_i(I, 0) = 0$, le terme constant du développement de Taylor de \mathcal{L}_i par rapport à sa deuxième variable est nul, et on peut écrire en regroupant les autres termes $\mathcal{L}_i(I, E) = \bar{\mathcal{L}}_i(I, E) \cdot E$ où $\bar{\mathcal{L}}_i(I, E)$ est une matrice $p \times 1$ dont les entrées dépendent de (I, E) . Par conséquent,

$$\begin{aligned} \sum_{i=1}^n \mathcal{L}_i(I, E)w_i &= \sum_{i=1}^n w_i(\bar{\mathcal{L}}_i(I, E) \cdot E) \\ &= (w_1 \quad \dots \quad w_n) \begin{pmatrix} \bar{\mathcal{L}}_1(I, E) \\ \vdots \\ \bar{\mathcal{L}}_n(I, E) \end{pmatrix} E \end{aligned}$$

L'observateur peut s'écrire

$$F(\hat{x}, u, y) = f(\hat{x}, u) + W(\hat{x})\bar{\mathcal{L}}(I(\hat{x}, u), E(\hat{x}, u, y))E(\hat{x}, u, y) \quad (6.7)$$

où $W(\hat{x}) = (w_1(\hat{x}), \dots, w_n(\hat{x}))$ et $\bar{\mathcal{L}}$ est une matrice $n \times p$ dont les entrées peuvent dépendre de (I, E) . L'observateur peut être vu comme un observateur avec une matrice $n \times p$ de gain $W \cdot \bar{\mathcal{L}}$ multipliée par l'erreur non linéaire E .

Afin de prouver le théorème 2 il nous faut démontrer le théorème suivant qui assure l'existence d'une erreur de sortie invariante (locale). La preuve est constructive et s'appuie sur la méthode du repère mobile de Cartan. (voir section 6.2.3).

Théorème 3. *Nous avons les résultats suivants*

- il existe une erreur de sortie invariante $(\hat{x}, u, y) \mapsto E(\hat{x}, u, y)$
- il existe une fonction invariante de rang plein $(\hat{x}, u) \mapsto I(\hat{x}, u) \in \mathbb{R}^{n+m-r}$ (un ensemble complet de $n+m-r$ invariants scalaires indépendants)
- toute autre erreur de sortie s'écrit

$$\tilde{E}(\hat{x}, u, y) = \mathcal{L}(I(\hat{x}, u), E(\hat{x}, u, y)).$$

où \mathcal{L} est n'importe quelle fonction régulière telle que $\mathcal{L}(I, 0) = 0$ et $E \mapsto \mathcal{L}(I, E)$ est inversible.

Démonstration. On applique la méthode du repère mobile (vue à la section 6.2.3) au cas suivant : $\Sigma = \mathcal{X} \times \mathcal{U} \times \mathcal{Y}$, et ϕ_g est la transformation composite

$$\phi_g(\hat{x}, u, y) := (\varphi_g(\hat{x}), \psi_g(u), \varrho_g(y)).$$

Puisque l'action de G sur \mathcal{X} est de rang plein on peut décomposer $\hat{x} \mapsto \varphi_g(\hat{x})$ en $\varphi_g^a(\hat{x}) \in \mathbb{R}^r$, qui est inversible par rapport à g , et en $\varphi_g^b(\hat{x}) \in \mathbb{R}^{n-r}$. Les r équations de normalisation

$$\varphi_g^a(\hat{x}) = c \tag{6.8}$$

peuvent alors être résolues et s'écrivent $g = \gamma(\hat{x})$. En substituant dans les équations restantes on obtient un ensemble complet de $n + m + p - r$ invariants indépendants fonctionnellement indépendants

$$I(\hat{x}, u) := (\varphi_{\gamma(\hat{x})}^b(\hat{x}), \psi_{\gamma(\hat{x})}(u)) \tag{6.9}$$

$$J_h(\hat{x}, y) := \bar{\varrho}_{\gamma(\hat{x})}(y). \tag{6.10}$$

Une erreur de sortie invariante est alors donnée par

$$E(\hat{x}, u, y) := J_h(\hat{x}, h(\hat{x}, u)) - J_h(\hat{x}, y) \tag{6.11}$$

Comme c'est en fait une fonction invariante des variables \hat{x}, u et y , tout erreur de sortie invariante \tilde{E} a la forme

$$\begin{aligned} \tilde{E}(\hat{x}, u, y) &= \mathcal{F}(I(\hat{x}, u), J_h(\hat{x}, y)) \\ &= \mathcal{F}(I(\hat{x}, u), J_h(\hat{x}, h(\hat{x}, u)) - E(\hat{x}, u, y)) \\ &= \mathcal{L}(I(\hat{x}, u), E(\hat{x}, u, y)) \end{aligned}$$

Nous avons utilisé le fait que $J_h(\hat{x}, h(\hat{x}, u))$, invariant par construction, doit être une fonction de $I(\hat{x}, u)$ (ensemble complet d'invariants de \hat{x} et de u). \square

Nous sommes maintenant en mesure de prouver le théorème 2.

Démonstration. du théorème 2 : Le champ de vecteurs F du théorème est clairement invariant. En effet,

$$\begin{aligned} F(x, u, h(x)) &= f(x, u) + \sum_{i=1}^n \mathcal{L}_i \left(I(\hat{x}), E(x, u, h(x)) \right) w_i(\hat{x}) \\ &= f(x, u) + \sum_{i=1}^n \mathcal{L}_i (I(\hat{x}), 0) w_i(\hat{x}) \\ &= f(x, u) \end{aligned}$$

Il est donc invariant par construction.

Réciproquement, supposons que $\frac{d}{dt}\hat{x} = F(\hat{x}, u, y)$ est un observateur G-invariant. Il peut alors être décomposé selon les champs de vecteurs linéairement indépendants w_i puisque

$$F(\hat{x}, u, y) = \sum_{i=1}^n F_i(\hat{x}, u, y) w_i(\hat{x}),$$

où les F_i sont des fonctions régulières. Puisque c'est un pré-observateur,

$$f(x, u) = F(x, u, h(x, u)) = \sum_{i=1}^n F_i(x, u, h(x, u)) w_i(x).$$

Puisqu'il est G-invariant

$$\sum_{i=1}^n F_i(\varphi_g(\hat{x}), \psi_g(u), \varrho_g(y)) w_i(\varphi_g(\hat{x})) = D\varphi_g(\hat{x}) \cdot \sum_{i=1}^n F_i(\hat{x}, u, y) w_i(\hat{x})$$

mais les w_i vérifient $D\varphi_g(\hat{x}) \cdot w_i(\hat{x}) = w_i(\varphi_g(x))$, et donc

$$F_i(\varphi_g(\hat{x}), \psi_g(u), \varrho_g(y)) = F_i(\hat{x}, u, y), \quad i = 1, \dots, n.$$

Par conséquent,

$$\begin{aligned} F(\hat{x}, u, y) &= f(\hat{x}, u) + [F(\hat{x}, u, y) - f(\hat{x}, u)] \\ &= f(\hat{x}, u) + \sum_{i=1}^n [F_i(\hat{x}, u, y) - F_i(\hat{x}, u, h(\hat{x}, u))] w_i(\hat{x}). \end{aligned}$$

Les fonctions $F_i(\hat{x}, u, y) - F_i(\hat{x}, u, h(\hat{x}, u))$ sont clairement invariantes; donc d'après le théorème 3, $F_i(\hat{x}, u, y) - F_i(\hat{x}, u, h(\hat{x}, u)) = \mathcal{L}_i(I(\hat{x}, u), E(\hat{x}, u, y))$. \square

Remarque additionnelle

On a supposé que la sortie était systématiquement compatible (c'est à dire que l'application de sortie est une application G -équivariante). En effet cette hypothèse est motivée par le résultat de [14] : si un pré-observateur $\frac{d}{dt}\hat{x} = F(\hat{x}, u, y) = F(\hat{x}, u, h(x, u))$ (voir la définition 6) est invariant, i.e.,

$$F(\varphi_g(\hat{x}), \psi_g(u), h(\varphi_g(x), \psi_g(u))) = D\varphi_g(\hat{x}) \cdot F(\hat{x}, u, h(x, u)).$$

et si le rang de F versus y est égal à $\dim(y)$, alors, l'application de sortie y est G -compatible au sens de la définition 3. La preuve est la suivante

Démonstration. On a

$$F(\hat{x}, y, t) = f(\hat{x}, t) + \sum_{i=1}^n E_i(\hat{x}, y, t)w_i(\hat{x})$$

où les E_i sont des erreurs invariantes et w_i des champs de vecteurs invariants. Comme le rang de $\frac{\partial F}{\partial y}$ est égal à $p = \dim(y)$, il existe p erreurs invariantes, $E = (E_{i_1}, \dots, E_{i_p})$, $1 \leq i_1 < i_2 \dots < i_p \leq n$ telles que le jacobien $\frac{\partial E}{\partial y}$ est inversible. On note G l'application inverse versus y (définie localement par le théorème des fonctions implicites), et on a

$$G(\hat{x}, E(\hat{x}, y, t), t) \equiv y.$$

De plus on a, pour tout $\hat{x}, x \in \mathcal{X}$ et $g \in G$,

$$E(\varphi_g(\hat{x}), h(\varphi_g(x)), t) = E(\hat{x}, h(x), t).$$

Donc

$$h(\varphi_g(x), t) = G(\varphi_g(\hat{x}), E(\hat{x}, h(x), t), t).$$

Fixons \hat{x} et t à des valeurs nominales, qu'on note \bar{x} et \bar{t} . L'identité ci-dessus veut dire que pour tout $g \in G$, il existe ρ_g (qui dépend aussi de \bar{x} and \bar{t} , mais ces dépendances ne sont pas exprimées ici) une application de \mathcal{Y} dans \mathcal{Y} telle que, pour tout $x \in \mathcal{X}$,

$$h(\varphi_g(x), t) = \rho_g(h(x)).$$

Donc l'application de sortie est G -compatible et ρ_g est une transformation sur \mathcal{Y} d'inverse ρ_g^{-1} . \square

6.2.5 Pre-observateurs invariants : une méthode constructive

Nous appliquerons cette méthode à tous les exemples considérés. Le système doit être vérifié invariant (i.e, inchangé par la transformation (6.3)) avec sortie équivariante (définition 3). On sait maintenant construire tous les observateurs candidats qui préservent les symétries grâce à la section précédente : *a)* Résoudre les équations de normalisation (6.8). Construire une erreur de sortie invariante avec (6.11), et un ensemble complet d'invariants I grâce à (6.9). *b)* Construire un repère invariant grâce à (6.5). *c)* La forme générale des observateurs invariants est donnée par le théorème 2. On donne une forme alternative matricielle (6.7). L'analyse de la convergence (et donc comment choisir des gains \mathcal{L}_i de l'observateur du théorème 2) est souvent simplifiée par l'utilisation d'une erreur d'état invariante.

6.2.6 Erreur d'état invariante et convergence

Pour n'importe quel système non-linéaire qui présente des symétries, nous n'avons pas de méthode générale systématique pour choisir des gains \mathcal{L}_i 's du théorème 2 qui assurent la convergence asymptotique de \hat{x} vers x . Néanmoins, la manière dont on définit l'erreur d'état joue un rôle fondamental dans l'analyse de la convergence (voir section 6.3). Au lieu de considérer l'erreur d'état linéaire $\hat{x} - x$, on considérera plutôt l'erreur d'état invariante

$$\eta(x, \hat{x}) = \varphi_{\gamma(x)}(\hat{x}) - \varphi_{\gamma(x)}(x)$$

où γ est défini par (6.8). On remarquera qu'il est équivalent de choisir \hat{x} comme variable de normalisation et alors l'erreur s'écrit

$$\eta(x, \hat{x}) = \varphi_{\gamma(\hat{x})}(x) - \varphi_{\gamma(\hat{x})}(\hat{x})$$

Un résultat remarquable est que la dynamique de l'erreur ne dépend de la trajectoire que via les invariants I :

Théorème 4. *La dynamique de l'erreur d'état invariante $\eta(\hat{x}, x) = \varphi_{\gamma(x)}(\hat{x}) - \varphi_{\gamma(x)}(x)$ ne dépend que de η et des invariants scalaires de x et de u :*

$$\frac{d}{dt}\eta = \Upsilon(\eta, I(x, u))$$

avec Υ fonction régulière de ses arguments et $I(x, u)$ ensemble complet d'invariants.

Démonstration. η est un invariant : pour tout $g \in G$ on a $\eta(\varphi_g(x), \varphi_g(\hat{x})) = \eta(x, \hat{x})$. Donc $\frac{d}{dt}\eta(\varphi_g(x), \varphi_g(\hat{x})) = \frac{d}{dt}\eta(x, \hat{x})$, i.e,

$$\begin{aligned} & \partial_1 \eta D\varphi_g(x) f(x, u) + \partial_2 \eta D\varphi_g(\hat{x}) F(\hat{x}, u, h(x, u)) \\ &= \partial_1 \eta f(x, u) + \partial_2 \eta F(\hat{x}, u, h(x, u)) \end{aligned} \tag{6.12}$$

où ∂_i est la dérivée partielle par rapport à la i -ème variable. Soit $\sigma(\hat{x}, x, u) = \frac{d}{dt}\eta(x, \hat{x}) = \partial_1\eta f(x, u) + \partial_2\eta F(\hat{x}, u, h(x, u))$. L'égalité (6.12) exprime que $\sigma(\varphi_g(\hat{x}), \varphi_g(x), \psi_g(u)) = \sigma(\hat{x}, x, u)$. Puisque $\hat{x} = \varphi_{\gamma(x)^{-1}}(\eta + \varphi_{\gamma(x)}(x))$, σ est une fonction invariante des variables (η, x, u) . Puisque η est un invariant, toute fonction invariante de (η, x, u) (en particulier $\frac{d}{dt}\eta$) est une fonction de η , et d'un ensemble complet d'invariants de x et de $u : I(x, u)$. \square

De telles coordonnées η ne sont pas uniques. Plus généralement, (et sous nos hypothèses de base) on voit que n'importe quelle erreur d'état invariante peut se mettre sous la forme

$$\mathcal{F}(I(x, u), \eta(x, \hat{x}))$$

où $I(x, u)$ forment un ensemble complet d'invariants pour l'action de G sur $\mathcal{X} \times \mathcal{U}$. En effet $\varphi_{\gamma(x)}(x)$ est invariant et peut donc être exprimé comme une fonction de I , et $(I(x, u), \varphi_{\gamma(x)}(\hat{x}))$ est un ensemble complet de $2n + m - r$ invariants de x, \hat{x} et u . Tous les exemples illustrent l'intérêt de ces coordonnées pour l'analyse de la convergence.

6.3 Observateurs invariants sur un groupe de Lie

Quand le groupe G , qui agit librement (hypothèse de rang plein) sur l'espace d'état \mathcal{X} , a la même dimension que \mathcal{X} ($r = n$), il n'y a qu'une seule orbite (action transitive). On peut alors identifier (à un sous-groupe discret près) G et \mathcal{X} , et la multiplication sur le groupe peut être définie de manière à ce que l'action de G sur $\mathcal{X} = G$ soit vue comme une multiplication à gauche L_g^1 :

$$\varphi_g(x) = gx = L_g(x).$$

Dans ce cas le système vit sur un groupe de Lie G et est invariant à gauche. Alors tous les calculs de la section 6.2 ainsi que la structure de l'observateur invariants deviennent plus simples.

Observateur invariant et équation d'erreur

Soit une dynamique invariante $x \in G$ au sens de la définition 2 : $\frac{d}{dt}x = f(x, u)$. Pour tout $g \in G$,

$$DL_g \cdot f(x, u) = f(gx, \psi_g(u))$$

La construction des erreurs invariantes et la formulation de l'équation d'erreur sont bien plus simples.

¹Un groupe de Lie peut également agir sur lui-même via les multiplications à droite $\varphi_g(x) = R_{g^{-1}}(x) = xg^{-1}$.

Erreurs invariantes

Afin de construire des erreurs invariantes on applique la méthode développée dans la preuve du théorème 3. L'équation de normalisation qui correspond à (6.8) est l'équation naturelle $gx = e$. Donc $\gamma(x, u) = x^{-1}$. Une erreur de sortie invariante E présentée au théorème 3 s'écrit

$$E(\hat{x}, u, y) = \varrho_{\hat{x}^{-1}}(h(\hat{x}, u)) - \varrho_{\hat{x}^{-1}}(y) \quad (6.13)$$

Et un ensemble complet d'invariants $I(\hat{x}, u)$ est

$$I(\hat{x}, u) = \psi_{\hat{x}^{-1}}(u). \quad (6.14)$$

Repère invariant

La méthode de construction développée à la section 6.2.3 revient à la démarche suivante. Soient n vecteurs linéairement indépendants (W_1, \dots, W_n) dans l'algèbre de Lie $TM_e = \mathcal{G}$ du groupe G . Les valeurs des champs de vecteurs sur l'espace tangent TM_x en x sont données par les relations suivantes qui expriment l'invariance

$$w_i(x) = DL_x \cdot W_i \in TM_x, i = 1 \dots n$$

Pré-observateurs invariants

Le théorème 3 implique que tout observateur invariant s'écrit

$$\frac{d}{dt}\hat{x} = f(\hat{x}, u) + DL_{\hat{x}} \cdot \left(\sum_{i=1}^n \mathcal{L}_i(\psi_{\hat{x}^{-1}}(u), h(e, \psi_{\hat{x}^{-1}}(u)) - \varrho_{\hat{x}^{-1}}(y)) W_i \right) \quad (6.15)$$

où les \mathcal{L}_i sont des fonctions régulières de leur argument qui s'annulent quand leur second argument est nul.

Formulation alternative

On remarque qu'on peut aussi définir l'observateur de la façon totalement intrinsèque suivante

$$\frac{d}{dt}\hat{x} = f(\hat{x}, u) + DL_{\hat{x}} \cdot \left(\sum_{i=1}^n \mathcal{L}_i(\psi_{\hat{x}^{-1}}(u), \varrho_{\hat{x}^{-1}}(y)) W_i \right)$$

où l'on impose $\mathcal{L}_i(I, h(e, \psi_{\hat{x}^{-1}}(u))) = 0$. On a supposé que la sortie h était à valeur dans un ouvert de \mathbb{R}^p . Cette formule est adaptée au cas d'une sortie à valeur dans G .

Dynamique de l'erreur d'état invariante

L'erreur d'état invariante définie à la section 6.2.6 s'écrit $\eta = x^{-1}\hat{x}$ et elle vaut e (au lieu de 0) quand $\hat{x} = x$. Elle est bien définie quand l'espace d'état \mathcal{X} est un ouvert de \mathbb{R}^n comme dans la section 6.2. Une erreur d'état petite correspond donc à η proche de e . La dérivée temporelle de η peut être calculée explicitement. On rappelle que R_g est la multiplication à droite sur G . Comme on a

- pour tout $g_1, g_2 \in G$, $DL_{g_1}DL_{g_2} = DL_{g_1g_2}$, $DR_{g_1}DR_{g_2} = DR_{g_2g_1}$ and $DL_{g_1}DR_{g_2} = DR_{g_2}DL_{g_1}$
- $I(\hat{x}, u) = \psi_{(x\eta)^{-1}}(u)$
- $\varrho_{\hat{x}^{-1}}(h(x, u)) = h(\hat{x}^{-1}x, \psi_{\hat{x}^{-1}}(u))$ writes $\varrho_{\hat{x}^{-1}}(y) = h(\eta^{-1}, \psi_{(x\eta)^{-1}}(u))$
- $\frac{d}{dt}(x^{-1}\hat{x}) = DL_{x^{-1}}\frac{d}{dt}\hat{x} - DR_{\hat{x}}\frac{d}{dt}x^{-1}$ with $\frac{d}{dt}x^{-1} = -DL_{x^{-1}}DR_{x^{-1}}\frac{d}{dt}x$

la dynamique de l'erreur s'écrit

$$\begin{aligned} \frac{d}{dt}\eta &= DL_{\eta} \cdot f(e, \psi_{(x\eta)^{-1}}(u)) - DR_{\eta} \cdot f(e, \psi_{x^{-1}}(u)) \\ &\quad + DL_{\eta} \cdot \left(\sum_{i=1}^n \mathcal{L}_i(\psi_{(x\eta)^{-1}}(u), h(e, \psi_{(x\eta)^{-1}}(u)) - h(\eta^{-1}, \psi_{(x\eta)^{-1}}(u))) W_i \right). \end{aligned} \quad (6.16)$$

Cette équation peut être transportée dans l'algèbre de Lie et ré-écrite sous une forme alternative

$$\begin{aligned} (DL_{\eta})^{-1} \frac{d}{dt}\eta &= f(e, \psi_{(x\eta)^{-1}}(u)) - DL_{\eta^{-1}}DR_{\eta} \cdot f(e, \psi_{x^{-1}}(u)) \\ &\quad + \left(\sum_{i=1}^n \mathcal{L}_i(\psi_{(x\eta)^{-1}}(u), h(e, \psi_{(x\eta)^{-1}}(u)) - h(\eta^{-1}, \psi_{(x\eta)^{-1}}(u))) W_i \right). \end{aligned}$$

Comme on l'a vu au lemme 4, l'erreur invariante η est solution d'une équation différentielle uniquement couplée à la trajectoire $t \mapsto (x(t), u(t))$ via les invariants $I(x, u) = \psi_{x^{-1}}(u)$. Lorsqu'on a $\psi_g(u) \equiv u$ (comme dans l'exemple de la voiture développé à la section 7.1) la dynamique de l'erreur est indépendante de la trajectoire $x(t)$.

Approximation linéaire invariante

Puisque G est une variété différentiable on sait que l'on peut identifier un morceau de l'algèbre de Lie de G à un voisinage U de e grâce à l'application exponentielle. On suppose maintenant que η est proche de e , dans le sens où $\eta \in U$ et il existe un élément ξ de l'algèbre de Lie \mathfrak{g} tel que l'on puisse écrire $r = \exp \epsilon \xi$ où $\epsilon \in \mathbb{R}$ est un petit paramètre. Jusqu'aux termes d'ordre 2 en ϵ , on a l'équation d'erreur invariante linéarisée dans l'espace

tangent à e

$$\begin{aligned} \frac{d}{dt}\xi = & [\xi, f(e, \psi_{x^{-1}}(u))] - \frac{\partial f}{\partial u}(e, \psi_{x^{-1}}(u)) \frac{\partial \psi}{\partial g}(e, \psi_{x^{-1}}(u))\xi \\ & - \sum_{i=1}^n \left(\frac{\partial \mathcal{L}_i}{\partial E}(\psi_{x^{-1}}(u), 0) \frac{\partial h}{\partial x}(e, \psi_{x^{-1}}(u))\xi \right) W_i \end{aligned} \quad (6.17)$$

où $[\cdot, \cdot]$ est le crochet de Lie de \mathcal{G} . Les gains $\frac{\partial \mathcal{L}_i}{\partial E}(u, 0)$ peuvent être réglés à l'aide des techniques usuelles de réglage pour les systèmes linéaires afin d'avoir une convergence locale.

6.4 Observateurs invariants sur un groupe de Lie dans un cas particulier

Cette section a fait l'objet d'une communication au CIFA 2006 [10]. Dans l'exemple qui concerne la navigation inertielle et sera vu au chapitre 7, le mouvement correspond à une dynamique invariante sur un groupe de Lie, et on s'aperçoit que l'équation d'erreur d'état est autonome : alors que la dynamique dépend du temps, la dérivée temporelle de l'erreur d'état invariante ne dépend que de l'erreur elle-même, ce qui rappelle le cas linéaire stationnaire. Cette propriété est inattendue dans l'étude d'un système non-linéaire, et son intérêt est de permettre un réglage des gains plus facile. On peut alors faire un observateur non-linéaire localement exponentiellement convergent autour de toute trajectoire, de manière à ce que le comportement global de l'erreur soit indépendant de la trajectoire du système. Dans cette section on cherche à déterminer la forme générale de la dynamique qui permet d'avoir cette propriété. Quand on regarde l'équation d'erreur linéarisée (6.17), on comprend qu'une condition nécessaire pour que l'équation (6.16) soit autonome, est que l'action de G sur l'espace des entrées \mathcal{U} soit telle que le terme $\frac{\partial f}{\partial u}(e, \psi_{x^{-1}}(u)) \frac{\partial \psi}{\partial g}(e, \psi_{x^{-1}}(u))\xi$ soit égal au crochet de Lie $[\xi, f(e, \psi_{x^{-1}}(u))]$. C'est en particulier le cas quand la dynamique est invariante à gauche sur un groupe de Lie et la sortie est compatible à droite. Dans ce cas, on peut en effet considérer l'action du groupe sur lui-même par multiplication à droite, faire un observateur invariant à droite, et montrer que l'équation d'erreur est autonome. La possibilité de définir les actions à gauche et à droite de manière à ce que la dynamique soit invariante à gauche, mais également invariante à droite (à condition de pouvoir considérer la dépendance en temps de la dynamique comme une entrée comme nous allons le voir) est typique du mouvement d'un solide dans l'espace tournant autour d'un point fixe. Car dans ce cas les invariances galiléennes impliquent que le mouvement est invariant par changement de repère attaché à l'espace, et également par changement de repère attaché au solide. Nous allons considérer dans cette section le cas du mouvement d'un "solide généralisé", et nous donnons ainsi une explication à la possibilité d'obtenir une équation d'erreur autonome, comme au chapitre 7.

Dynamique invariante à gauche et sortie compatible à droite

On considère une dynamique invariante à gauche sur un groupe de Lie

$$\frac{d}{dt}x(t) = f(x, t) \quad (6.18)$$

avec :

$$\forall g, x \in G \quad f(L_g(x), t) = DL_g f(x, t)$$

C'est un cas particulier de celui de la section 6.3, où l'espace des entrées est réduit à $\mathcal{U} = \mathbb{R}_+$ puisque $u(t) \equiv t$ représente le temps. Comme dans [6] on appelle

$$\omega_s = DL_{x^{-1}}\dot{x} \in \mathfrak{g}$$

ω_s est un élément de l'algèbre de Lie \mathfrak{g} de G indépendant de x . On peut donc voir toute dynamique invariante à gauche sur un groupe de Lie G comme le mouvement d'un "solide généralisé" d'espace de configuration G . Et $\omega_s(t) = f(e, t)$ est "la vitesse angulaire par rapport au solide", avec e élément neutre de G . Nous écrirons systématiquement toute dynamique invariante à gauche (6.18) sous la forme

$$\frac{d}{dt}x(t) = DL_x \omega_s(t) \quad (6.19)$$

Supposons que la sortie G -compatible $h : G \mapsto Y$ l'est en fait à droite (propriété de droite-gauche équivariance), i.e, pour tout $g \in G$, il existe une application régulière $\varrho_g : Y \mapsto Y$, telle que $h(x \cdot g) = \varrho_g(h(x))$ i.e

$$h(R_g(x)) = \varrho_g(h(x))$$

où R_g est la multiplication à droite sur G .

Dans l'annexe C, on donnera un exemple simple d'application de cette section. En effet les résultats s'appliquent alors car la dynamique est bien invariante à gauche puisque l'on a $\frac{d}{dt}q(t) = q \cdot \vec{\omega} = DL_q \omega(t)$ et la sortie $h(q) = q^{-1} \cdot \vec{B} \cdot q$ est invariante à droite

$$\begin{aligned} h(R_p(q)) &= ((q \cdot p)^{-1} \cdot \vec{G} \cdot (q \cdot p), (q \cdot p)^{-1} \cdot \vec{B} \cdot (q \cdot p)) \\ &= (p^{-1} \cdot y_G \cdot p, p^{-1} \cdot y_B \cdot p) = \varrho_p(h(q)) \end{aligned}$$

Observabilité

Lorsque la dimension de l'espace de sortie est strictement plus petite que la dimension de l'espace d'état ($\dim y < \dim x$), et que la sortie est compatible à droite, le système est nécessairement non-observable. Dans ce cas on peut trouver deux éléments distincts g_1 et g_2 de G tels que $h(g_1) = h(g_2)$ (puisque h est régulière et la dimension de l'espace

d'arrivée est strictement inférieure à celle de l'espace de départ). Si $x(t)$ est une trajectoire du système, on a

$$\frac{d}{dt}x(t) = DL_x\omega_s(t)$$

et à cause de l'invariance à gauche, $g_1x(t)$ et $g_2x(t)$ sont également deux trajectoires du système :

$$\frac{d}{dt}(g_1 \cdot x(t)) = DL_{g_1x}\omega_s(t), \quad \frac{d}{dt}(g_2 \cdot x(t)) = DL_{g_2x}\omega_s(t).$$

Mais puisque h est compatible à droite :

$$h(g_1 \cdot x(t)) = \varrho_{x(t)}h(g_1) = \varrho_{x(t)}h(g_2) = h(g_2x(t)).$$

Les trajectoires $g_1 \cdot x(t)$ et $g_2 \cdot x(t)$ sont distinctes pour tout t et pourtant donnent la même sortie. Le système est inobservable.

La dynamique est invariante à droite

On considère encore la dynamique (6.18). Elle peut être vue comme une dynamique invariante à droite sur G ! En effet considérons $\omega_s(t)$ comme une entrée : $u(t) = \omega_s(t) \in \mathcal{U}$. On peut définir l'action de G via ψ_g sur l'espace des entrées \mathcal{U} de la manière suivante pour tout g :

$$\psi_g = DL_{g^{-1}}DR_g$$

Alors ψ_g est l'automorphisme intérieur de G . Et la dynamique (6.18) s'écrit

$$\frac{d}{dt}x = F(x, u) = DL_xu$$

et peut être vue comme une dynamique invariante à droite. Pour tout x, g on a bien :

$$\frac{d}{dt}R_g(x) = DR_gDL_x\omega_s(t) = DL_xDL_gDL_{g^{-1}}DR_g\omega_s(t) = DL_{R_g(x)}\psi_g(\omega_s(t)) = F(R_g(x), \psi_g(u))$$

Construction de l'observateur

Puisque la dynamique est invariante à droite et la sortie est également invariante à droite, on peut appliquer la théorie développée à la section 6.3 en remplaçant partout multiplication à gauche par multiplication à droite. Soit (W_1, \dots, W_n) un repère invariant à droite. Les champs de vecteurs qui le composent vérifient

$$W_i(g) = DR_gW_i(e)$$

De plus, grâce à la méthode du repère mobile de Cartan on sait construire des erreurs de sortie invariantes à droite, i.e, des fonctions scalaires régulières E_i de \hat{x} et de $h(x)$, qui vérifient :

$$\begin{aligned} \forall g, x \quad E_i(R_g(\hat{x}), h(R_g(x))) &= E_i(\hat{x}, h(x)) \\ E_i(x, h(x)) &= 0 \end{aligned}$$

On considère les observateurs de la forme

$$\begin{aligned} \frac{d}{dt}\hat{x} &= DL_{\hat{x}}\omega_s(t) + \sum_{i=1}^n E_i(\hat{x}, y)W_i(\hat{x}) \\ &= DL_{\hat{x}}\omega_s(t) + DR_{\hat{x}}\left(\sum_{i=1}^n E_i(\hat{x}, y)W_i(e)\right) \end{aligned} \quad (6.20)$$

où les E_i sont des erreurs de sortie invariantes à droite, (W_1, \dots, W_n) est un repère invariant à droite et $y = h(x)$ est la sortie. Ces observateurs sont invariants à droite quand ω_s est vu comme une entrée sur laquelle agit ψ_g .

Equation d'erreur

On définit l'erreur d'état invariante à droite $G \ni r = (\hat{x}x^{-1}) = L_{\hat{x}}(x^{-1})$. La dynamique d'erreur vérifie

$$\frac{d}{dt}r = DR_r\left(\sum_{i=1}^n E_i(e, h(r^{-1}))W_i(e)\right) \quad (6.21)$$

En effet

$$\frac{d}{dt}r = DL_{\hat{x}}(x^{-1}) + D_x L_{\hat{x}}(x^{-1})\dot{\hat{x}}$$

et on a

$$\begin{aligned} D_x L_{\hat{x}}(x^{-1})\dot{\hat{x}} &= DR_{x^{-1}}(\dot{\hat{x}}) \\ &= DR_{x^{-1}}(f(\hat{x}, t) + \sum_{i=1}^n E_i(\hat{x}, y)W_i(\hat{x})) \\ &= DR_{x^{-1}}DL_{\hat{x}} f(e, t) + DR_{x^{-1}}DR_{\hat{x}}\sum_{i=1}^n E_i(\hat{x}, y)W_i(e) \\ &= DR_{x^{-1}}DL_{\hat{x}} \omega_s + DR_r\sum_{i=1}^n E_i(\hat{x}, y)W_i(e) \end{aligned}$$

Mais l'invariance implique que

$$E_i(\hat{x}, y) = E_i(\hat{x}, H(x)) = E_i(e, H(r^{-1}))$$

On a également

$$\begin{aligned} DL_{\hat{x}}(\dot{x}^{-1}) &= -DL_{\hat{x}}DR_{x^{-1}}DL_{x^{-1}}\dot{x} = -DL_{\hat{x}}DR_{x^{-1}}\omega_s \\ &= -DR_{x^{-1}}DL_{\hat{x}}\omega_s \end{aligned}$$

L'équation d'erreur est donc autonome et indépendante de la trajectoire $t \mapsto x(t)$ (et du temps t) :

$$\frac{d}{dt}r = DR_r\left(\sum_{i=1}^n E_i(e, H(r^{-1}))W_i(e)\right).$$

Approximation linéaire

Puisque G est une variété différentiable on sait que l'on peut identifier un morceau de l'algèbre de Lie de G à un voisinage U de e grâce à l'application exponentielle. On suppose maintenant que r est proche de e , dans le sens où r est dans U et il existe un petit élément ξ de l'algèbre de Lie \mathfrak{g} tel que l'on puisse écrire $r = \exp \xi$. Soit $p = \dim y$ la dimension de la sortie, on a

$$\frac{d}{dt}\xi = -D_y E|_{e, h(e)} Dh|_e(\xi)$$

où l'on appelle $E = (E_1, \dots, E_n)$.

Définissons le produit scalaire sur l'espace tangent \mathfrak{g} en e , et soit $(Dh|_e)^T$ l'opérateur adjoint de $Dh|_e$ au sens de la métrique associée au produit scalaire. On prend pour tout $\eta \in \mathfrak{g}$ proche de zéro

$$E(e, \exp \eta) = K(Dh|_e)^T \eta.$$

Grâce à l'invariance de E il est possible de définir E sans ambiguïté pour tout (\hat{x}, x) avec $\hat{x}x^{-1}$ proche de e .

L'approximation au premier ordre s'écrit

$$\dot{\xi} = -K Dh^T Dh \xi \tag{6.22}$$

et pour $K > 0$, admet comme fonction de Lyapunov $\|\xi\|^2$ qui est la longueur de ξ au sens du produit scalaire.

Une classe d'observateurs non-linéaires convergents au premier ordre

Soient les observateurs suivants :

$$\frac{d}{dt}\hat{x} = DL_{\hat{x}}\omega_s(t) + DR_{\hat{x}}\left[\sum_{i=1}^n [E_i(\varrho_{\hat{x}}^{-1}(h(x)) - E_i(h(e)))W_i(e)]\right]$$

où les E_i sont des fonctions scalaires régulières. Prenons $E = (E_1, \dots, E_n)$ de manière à ce que la partie symétrique (au sens du produit scalaire sur TG_e) de l'application linéaire

$$\xi \mapsto \frac{\partial E}{\partial y} \Big|_{h(e)} \frac{\partial h}{\partial x} \Big|_e \xi$$

soit négative. Quand elle est définie négative, nous obtenons un observateur non-linéaire convergent localement exponentiellement autour de toute trajectoire.

6.5 Conclusion

Nous avons proposé une méthode systématique de design des (pré-)observateurs qui sont des observateurs candidats avec des termes de correction préservant les symétries du système originel. Les hypothèses nécessaires sont assez modérées sur la régularité de l'action du groupe sur l'espace d'état. Nous n'avons pas de méthode systématique pour attaquer la convergence de l'observateur non-linéaire et les singularités de l'action du groupe sur l'espace d'état. Néanmoins, l'exemple du réacteur chimique (chapitre 5), de la voiture non-holonome, et de la navigation inertielle (chapitre 7) illustrent l'intérêt des erreurs de sortie et des erreurs d'état invariante dans la construction d'observateurs non-linéaire asymptotiquement convergents.

Dans le cas où l'espace d'état peut être identifié à un groupe de Lie, les calculs sont plus simples et l'équation d'erreur est explicite.

Dans un cas particulier (dynamique invariante à gauche sur un groupe de Lie et sortie compatible à droite) nous avons proposé une classe d'observateurs non-linéaires localement convergents autour de toute trajectoire, et tels que le comportement global est indépendant de la trajectoire du système.

Chapitre 7

Exemples

Ce chapitre reprend le rapport disponible sous arxiv [12] et l'ensemble de ce chapitre est contenu dans la partie qui est acceptée pour publication dans IEEE Automatic Control.

On présente ici deux exemples qui illustrent la théorie. Tous deux montrent que la théorie développée au chapitre 6 fournit une méthode précise pour trouver tous les observateurs invariants candidats (pré-observateurs) ainsi qu'un guide d'étude de la convergence grâce à la notion d'erreur d'état invariante. Le premier exemple est une voiture non-holonome équipée d'un GPS. C'est un exemple théorique et pédagogique. Nous faisons un observateur qui est presque globalement convergent (théorème 5). Le deuxième exemple est emprunté au domaine de l'ingénierie. Il traite d'un problème de navigation inertielle. Nous construisons un observateur invariant qui est localement exponentiellement convergent autour de toute trajectoire. De plus le comportement global est indépendant de la trajectoire du système (théorème 6). Les propriétés remarquables des observateurs proviennent du fait que les deux dynamiques que l'on considère peuvent être vues comme des dynamiques invariantes à gauche sur un groupe de Lie.

Ceci est la raison pour laquelle nous nous sommes intéressés à des exemples de la mécanique, et pas seulement aux réacteurs chimiques, car ils possèdent moins de symétries. En mécanique, on bénéficie des invariances galiléennes, et donc nous avons des chances de trouver le même nombre de symétries engendrées par des actions locale que la dimension de l'espace d'état. Dans ce cas on a vu que l'erreur d'état peut être vue comme un élément d'un groupe de Lie, et est donc définie intrinsèquement. Cette propriété se révèle intéressante au vu des exemples suivants.

7.1 La voiture non-holonome

Considérons une voiture non-holonome (représentée sur la figure 7.1) dont la dynamique est la suivante :

$$\frac{d}{dt}x = u \cos \theta, \quad \frac{d}{dt}y = u \sin \theta, \quad \frac{d}{dt}\theta = uv \quad h(x, y, \theta) = (x, y) \quad (7.1)$$

où u est la vitesse et v une fonction de l'angle de braquage ϕ : ($v = \tan(\phi)/L$). La mesure est la position $(x, y, \theta) = (x, y)$ (i.e, on dispose d'un GPS dans la voiture).

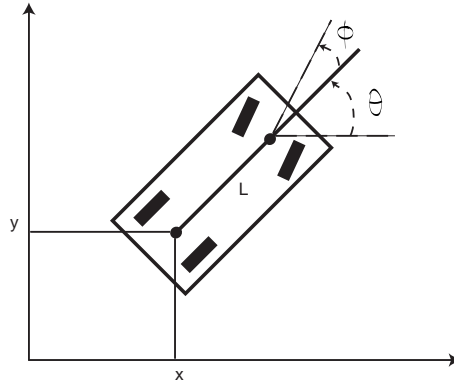


FIG. 7.1 –

La dynamique est indépendante de l'origine et de l'orientation du repère choisi, i.e, elle est invariante sous l'action du groupe $SE(2)$, le groupe des rotations et translations du plan. De plus l'espace d'état $\mathcal{X} = \mathbb{R}^2 \times \mathbb{S}^1$ coïncide avec le groupe $G = SE(2)$ et la dynamique est invariante par multiplication à gauche (voir la section 6.3). Pour tout $(x_0, y_0, \theta_0) \in G$ l'application $\varphi_{(x_0, y_0, \theta_0)}$ correspond à l'action de G sur lui-même par multiplication à gauche :

$$L_{(x_0, y_0, \theta_0)}(x, y, \theta) = \varphi_{\begin{pmatrix} x_0 \\ y_0 \\ \theta_0 \end{pmatrix}}(x, y, \theta) = \begin{pmatrix} x_0 \\ y_0 \\ \theta_0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} = \begin{pmatrix} x \cos \theta_0 - y \sin \theta_0 + x_0 \\ x \sin \theta_0 + y \cos \theta_0 + y_0 \\ \theta + \theta_0 \end{pmatrix}$$

$$\text{and } \psi_{\begin{pmatrix} x_0 \\ y_0 \\ \theta_0 \end{pmatrix}}(u, v) = \begin{pmatrix} u \\ v \end{pmatrix}$$

La dynamique est en effet invariante au sens de la définition 2. Soit $(x_0, y_0, \theta_0) \in G$ et $(x, y, \theta) \in G$ et $(u, v) \in \mathcal{U} = \mathbb{R}^2$. On pose $\varphi_{(x_0, y_0, \theta_0)}(x, y, \theta) = (X, Y, \Theta)$ et $\psi_{(x_0, y_0, \theta_0)}(u, v) = (U, V)$ (voir fig 7.2). La dynamique s'écrit avec les nouvelles variables de la même manière

$$\frac{d}{dt}X = U \cos \Theta, \quad \frac{d}{dt}Y = V \sin \Theta, \quad \frac{d}{dt}\Theta = UV$$

La fonction de sortie est compatible au sens de la définition 3 : pour tout $x_0, y_0, \theta_0, x, y \in G$ on a $q_{(x_0, y_0, \theta_0)}(x, y) = \begin{pmatrix} x \cos \theta_0 - y \sin \theta_0 + x_0 \\ x \sin \theta_0 + y \cos \theta_0 + y_0 \end{pmatrix}$
 On applique maintenant la méthode de construction des observateurs invariants développée

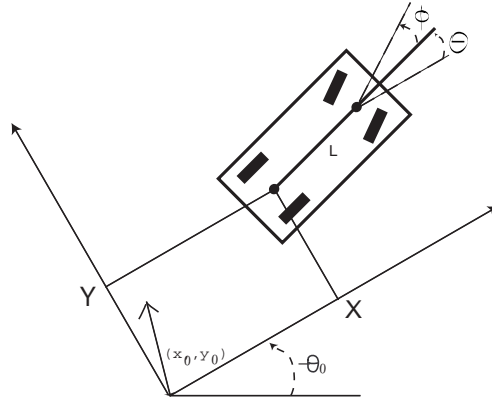


FIG. 7.2 –

au chapitre précédent.

Erreur de sortie invariante

Les équations de normalisation (6.8) s'écrivent avec $c = 0$:

$$\begin{aligned} x \cos \theta_0 - y \sin \theta_0 + x_0 &= 0 \\ x \sin \theta_0 + y \cos \theta_0 + y_0 &= 0 \\ \theta + \theta_0 &= 0 \end{aligned}$$

Par conséquent

$$\begin{pmatrix} x_0 \\ y_0 \\ \theta_0 \end{pmatrix} = \begin{pmatrix} -x \cos \theta - y \sin \theta \\ x \sin \theta - y \cos \theta \\ -\theta \end{pmatrix} = \gamma \begin{pmatrix} x \\ y \\ \theta \end{pmatrix}$$

Un ensemble complet d'invariants fondamentaux est donné par (voir 6.9) : $I(x, y, \theta, u, v) =$

$\psi_{\gamma(x, y, \theta)}(u, v) = \begin{pmatrix} u \\ v \end{pmatrix}$. Soit $\begin{pmatrix} x_0 \\ y_0 \\ \theta_0 \end{pmatrix} = \gamma \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix}$. Une erreur de sortie invariante s'écrit

(voir (6.11)) :

$$\begin{aligned}
E &= \varrho_{(x_0, y_0, \theta_0)}(\hat{x}, \hat{y}) - \varrho_{(x_0, y_0, \theta_0)}(x, y) \\
&= \begin{pmatrix} \cos \theta_0 & -\sin \theta_0 \\ \sin \theta_0 & \cos \theta_0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \cos \theta_0 & -\sin \theta_0 \\ \sin \theta_0 & \cos \theta_0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \\
&= \begin{pmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \end{pmatrix}
\end{aligned}$$

Repère Invariant

Un repère invariant (w_1, w_2, w_3) est donné par l'image de la base canonique de $\mathbb{R}^2 \times \mathbb{S}^1$ (l'algèbre de Lie de $G = SE(2)$) par l'application $DL_{(x, y, \theta)}$, i.e, les colonnes de la matrice

$$DL_{(x, y, \theta)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = (w_1 \ w_2 \ w_3)$$

On remarque que cela correspond au repère de Frenet.

Pré-observateur invariant

Tout pré-observateur invariant s'écrit(voir (6.7))

$$\frac{d}{dt} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix} = \begin{pmatrix} u \cos \hat{\theta} \\ u \sin \hat{\theta} \\ uv \end{pmatrix} + \begin{pmatrix} \cos \hat{\theta} & -\sin \hat{\theta} & 0 \\ \sin \hat{\theta} & \cos \hat{\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\mathcal{L}} \begin{pmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \end{pmatrix} \quad (7.2)$$

où $\bar{\mathcal{L}}$ est une matrice 3×2 de gain régulière dont les entrées dépendent de l'erreur de sortie invariante E mais aussi des invariants $I(\hat{x}, \hat{y}, \hat{\theta}, u, v)$.

Equation d'erreur

La variable que l'on choisit pour normaliser est $(\hat{x}, \hat{y}, \hat{\theta})^T$. L'erreur d'état invariante s'écrit alors (voir (6.2.6)) :

$$\begin{aligned}
\eta &= \gamma(\hat{x}, \hat{y}, \hat{\theta}) \cdot \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix} - \gamma(\hat{x}, \hat{y}, \hat{\theta}) \cdot \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \\
&= - \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix}^{-1} \cdot \begin{pmatrix} x \\ y \\ \theta \end{pmatrix} = \begin{pmatrix} (\hat{x} - x) \cos \hat{\theta} + (\hat{y} - y) \sin \hat{\theta} \\ -(\hat{x} - x) \sin \hat{\theta} + (\hat{y} - y) \cos \hat{\theta} \\ (\hat{\theta} - \theta) \end{pmatrix}
\end{aligned} \quad (7.3)$$

et l'on note $\eta = (\eta_x, \eta_y, \eta_\theta)^T$ ses coordonnées dans $\mathbb{R}^2 \times \mathbb{S}^1$. On remarque que les deux premières coordonnées de l'erreur d'état coïncident avec l'erreur de sortie $(\eta_x, \eta_y) = (E_x, E_y)$.

Des calculs simples à partir des relations suivantes

$$\begin{aligned} - \begin{pmatrix} \eta_x \\ \eta_y \end{pmatrix} &= \begin{pmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \end{pmatrix} \\ - \frac{d}{dt} \begin{pmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} &= (uv + \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y) \begin{pmatrix} \cos(\hat{\theta} + \pi/2) & \sin(\hat{\theta} + \pi/2) \\ -\sin(\hat{\theta} + \pi/2) & \cos(\hat{\theta} + \pi/2) \end{pmatrix} \\ - \frac{d}{dt} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \\ \hat{\theta} - \theta \end{pmatrix} &= \begin{pmatrix} u(\cos \hat{\theta} - \cos \theta) \\ u(\sin \hat{\theta} - \sin \theta) \\ 0 \end{pmatrix} + \begin{pmatrix} \cos \hat{\theta} & -\sin \hat{\theta} & 0 \\ \sin \hat{\theta} & \cos \hat{\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\mathcal{L}} \begin{pmatrix} \eta_x \\ \eta_y \end{pmatrix} \end{aligned}$$

aboutissent à l'équation d'erreur suivante :

$$\frac{d}{dt} \begin{pmatrix} \eta_x \\ \eta_y \\ \eta_\theta \end{pmatrix} = \begin{pmatrix} u(1 - \cos \eta_\theta) + (uv + \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y)\eta_y \\ u \sin \eta_\theta - (uv + \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y)\eta_x \\ 0 \end{pmatrix} + \bar{\mathcal{L}} \begin{pmatrix} \eta_x \\ \eta_y \end{pmatrix}$$

En effet l'erreur d'état invariante est indépendante de la trajectoire du système et ne dépend que des quantités relatives η_x , η_y et η_θ comme le prévoit le lemme 4 puisqu'ici les invariants fondamentaux sont u, v .

Convergence de l'erreur

On peut régler les gains de manière à ce que le système d'erreur soit presque globalement convergent. L'équation d'erreur s'écrit :

$$\begin{aligned} \frac{d}{dt}\eta_x &= u(1 - \cos \eta_\theta) + (uv + \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y)\eta_y + \bar{\mathcal{L}}_{11}\eta_x + \bar{\mathcal{L}}_{12}\eta_y \\ \frac{d}{dt}\eta_y &= u \sin \eta_\theta - (uv + \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y)\eta_x + \bar{\mathcal{L}}_{21}\eta_x + \bar{\mathcal{L}}_{22}\eta_y \\ \frac{d}{dt}\eta_\theta &= \bar{\mathcal{L}}_{31}\eta_x + \bar{\mathcal{L}}_{32}\eta_y \end{aligned}$$

Prenons

$$\bar{\mathcal{L}} = \begin{pmatrix} -|u|a & 0 \\ 0 & -|u|c \\ 0 & -ub \end{pmatrix} + \begin{pmatrix} 0 & ubE_y - uv \\ uv - ubE_y & 0 \\ 0 & 0 \end{pmatrix} \quad (7.4)$$

où a, b, c sont des scalaires positifs, l'équation d'erreur s'écrit

$$\begin{aligned} \frac{d}{dt}\eta_x &= u(1 - \cos \eta_\theta) - |u|a\eta_x \\ \frac{d}{dt}\eta_y &= u \sin \eta_\theta - |u|c\eta_y \\ \frac{d}{dt}\eta_\theta &= -ub\eta_y \end{aligned} \quad (7.5)$$

Supposons que la vitesse de la voiture soit telle que $\int_{t_0}^{\infty} |u(t)| dt = +\infty$ pour tout $t_0 > 0$. On pourra remarquer que quand la voiture est à l'arrêt, le système est inobservable. Faisons un changement régulier d'échelle de temps $t : ds = |u| dt$, on a ($\epsilon_1 = \pm 1$ est le signe de u)

$$\begin{aligned}\frac{d}{ds}\eta_x &= \epsilon_1(1 - \cos \eta_\theta) - a\eta_x \\ \frac{d}{ds}\eta_y &= \epsilon_1 \sin \eta_\theta - c\eta_y \\ \frac{d}{ds}\eta_\theta &= -\epsilon_1 b\eta_y\end{aligned}$$

avec la structure triangulaire suivante :

$$\begin{aligned}\frac{d^2}{ds^2}\eta_\theta &= -c\frac{d}{ds}\eta_\theta - b \sin \eta_\theta \\ \frac{d}{ds}\eta_x &= \epsilon_1(1 - \cos \eta_\theta) - a\eta_x\end{aligned}$$

La première équation est la dynamique d'un pendule non-linéaire amorti. L'équilibre $\eta_\theta = 0$ est "presque stable" (toutes les trajectoires convergent vers cet équilibre à part une qui s'obtient en initialisant η à π). La seconde équation est un système stable au premier ordre avec pour terme source $\epsilon_1(1 - \cos \eta_\theta)$. Grâce à la notion d'erreur d'état invariante définie par (7.3) nous avons prouvé

Théorème 5. *Considérons le système (7.1). Supposons que $\int_{t_0}^{\infty} |u(t)| dt = +\infty$ pour tout $t_0 > 0$. L'observateur non-linéaire*

$$\frac{d}{dt} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix} = \begin{pmatrix} u \cos \hat{\theta} \\ u \sin \hat{\theta} \\ uv \end{pmatrix} + \begin{pmatrix} \cos \hat{\theta} & -\sin \hat{\theta} & 0 \\ \sin \hat{\theta} & \cos \hat{\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\mathcal{L}} \begin{pmatrix} \cos \hat{\theta} & \sin \hat{\theta} \\ -\sin \hat{\theta} & \cos \hat{\theta} \end{pmatrix} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \end{pmatrix}$$

avec

$$\bar{\mathcal{L}} = \begin{pmatrix} -|u|a & 0 \\ 0 & -|u|c \\ 0 & -ub \end{pmatrix} + \begin{pmatrix} 0 & ubE_y - uv \\ uv - ubE_y & 0 \\ 0 & 0 \end{pmatrix}$$

est presque globalement convergent.

7.2 Navigation inertielle avec mesure de vitesse

Dans les systèmes de navigation inertielle à bas coût, les gyroscopes et accéléromètres sont "assistés" par des mesures de vitesse (qui proviennent d'un "air-data system" ou d'un radar Doppler) et des magnétomètres. Les différentes mesures sont ensuite fusionnées en

accord avec les équations (terre plate) du mouvement de l'avion, usuellement par un observateur à gains programmés ou par un filtre de Kalman étendu. L'analyse de la convergence est loin d'être triviale. En utilisant notre théorie, on obtient une équation d'erreur simple et indépendante de la trajectoire du système. Le réglage des gains de manière à obtenir la convergence locale autour de toute trajectoire est immédiat.

Les simulations illustrent le bon comportement de l'observateur en présence de biais et de bruit sur les capteurs. Elles indiquent même que le domaine de convergence est très large (bien que nous n'ayons pas exploré le comportement global).

Les formules de l'observateur et son implémentation sont fortement simplifiées quand l'orientation du solide est décrite par un quaternion de norme 1 (plutôt que par les angles d'Euler). Notamment car il est plus simple de maintenir la norme d'un quaternion égale à 1 numériquement qu'une matrice dans $SO(3)$, et les calculateurs embarqués peuvent être de capacité de calcul limitée. Dans l'annexe B on trouve un récapitulatif sur les quaternions. On écrit tout quaternion $q = q^0 + q^1 e_1 + q^2 e_2 + q^3 e_3$ où e_1, e_2, e_3 sont les nombres imaginaires usuels i, j, k et $*$ la multiplication des quaternions.

Equations de la dynamique

Le mouvement d'un objet volant (quand on fait l'approximation terre plate) s'écrit

$$\begin{aligned} \frac{d}{dt}q &= \frac{1}{2}q * \omega \\ \frac{d}{dt}v &= v \times \omega + q^{-1} * \mathbf{A}_{grav} * q + a \\ y &= (y_v, y_b) = (v, q^{-1} * \mathbf{B} * q) \end{aligned} \tag{7.6}$$

où (ω, a) sont les entrées et

- q est le quaternion de norme 1 qui représente l'orientation du repère attaché au solide dans le repère attaché à la terre. On remarque que la norme de q est préservée par la première équation car ω est un vecteur de \mathbb{R}^3 (i.e, un quaternion dont la première coordonnée est 0).
- $\omega = \omega^1 e_1 + \omega^2 e_2 + \omega^3 e_3$ est le vecteur rotation instantané exprimé dans le repère du solide.
- $v = v^1 e_1 + v^2 e_2 + v^3 e_3$ est le vecteur vitesse du centre de masse exprimé dans le repère du solide.
- $\mathbf{A}_{grav} = \mathbf{A}_{grav}^1 e_1 + \mathbf{A}_{grav}^2 e_2 + \mathbf{A}_{grav}^3 e_3$ est le vecteur gravité exprimé dans le repère terrestre et qu' on suppose constant dans l'aire de vol.
- $a = a^1 e_1 + a^2 e_2 + a^3 e_3$ est vecteur accélération spécifique, i.e, l'accélération mesurée par les accéléromètres, et représente donc les forces aérodynamiques (toutes les forces sauf la gravité) divisées par la masse du solide.
- $\mathbf{B} = \mathbf{B}^1 e_1 + \mathbf{B}^2 e_2 + \mathbf{B}^3 e_3$ est le champ magnétique terrestre exprimé dans le repère terrestre et supposé constant sur l'aire de vol.

La première équation décrit la cinématique du solide, et la seconde la loi de Newton. Les mesures sont $\omega(t)$, $a(t)$, $v(t)$ and $q^{-1}(t)*\mathbf{B}*q(t)$ (respectivement mesurés par les gyroscopes, les accéléromètres, l'air-data system). Leurs coordonnées sont connues dans le repère du solide. On veut estimer q et v . L'observateur est un filtre pour v et permet d'estimer q . On applique maintenant la méthode de construction des observateurs invariants développée au chapitre précédent.

Invariance des équations

Les symétries sont issues de considérations liées à la physique du problème (indépendance du choix du repère) et sont associées au groupe $SE(3)$ qui consistent en les rotations et translations en dimension 3. Nous pouvons considérer les rotations et translations dans le repère attaché à la terre, ou le repère attaché au solide. D'après la dernière section du chapitre 6, en vue d'une équation d'erreur simple, il est plus intéressant de considérer les rotations-translations dans le solide (donc des multiplications à droite). On identifie (au groupe multiplicatif près $\{-1, +1\}$) $SE(3)$ et l'espace d'état $\mathcal{X} = \mathbb{H}_1 \times \mathbb{R}^3$, où \mathbb{H}_1 désigne le sous-groupe des quaternions de norme 1, (\mathbb{H}_1 est isomorphe à $SU(2)$). Pour tout $(q_0, v_0) \in G$, l'application $\varphi_{(q_0, v_0)}$ correspond à l'action de G sur lui-même par multiplication à gauche :

$$L_{(q_0, v_0)}(q, v) = \varphi_{(q_0, v_0)}(q, v) = \begin{pmatrix} q_0 \\ v_0 \end{pmatrix} \cdot \begin{pmatrix} q \\ v \end{pmatrix} = \begin{pmatrix} q * q_0 \\ q_0^{-1} * v * q_0 + v_0 \end{pmatrix}$$

Ici $u = \begin{pmatrix} a \\ \omega \end{pmatrix}$. Pour tout $(q_0, v_0) \in G$, l'application $\psi_{(q_0, v_0)}$ est donnée par

$$\psi_{(q_0, v_0)}(a, \omega) = \begin{pmatrix} q_0^{-1} * a * q_0 - v_0 \times (q_0^{-1} * \omega * q_0) \\ q_0^{-1} * \omega * q_0 \end{pmatrix}$$

Vérifions que la dynamique est bien invariante au sens de la définition 2. Prenons $(q_0, v_0) \in G$ et $(q, v) \in G$ et $(a, \omega) \in \mathcal{U}$. On pose

$$\varphi_{(q_0, v_0)}(q, v) = (Q, V), \quad \psi_{(q_0, v_0)}(a, \omega) = (A, \Omega).$$

On a d'une part

$$\frac{d}{dt}Q = \frac{1}{2}q * \omega * q_0 = \frac{1}{2}q * q_0 * q_0^{-1} * \omega * q_0 = \frac{1}{2}Q * \Omega$$

et d'autre part

$$\begin{aligned} \frac{d}{dt}V &= q_0^{-1}(v \times \omega + q^{-1} * \mathbf{A}_{grav} * q + a) * q_0 \\ &= (V - v_0) \times \Omega + Q^{-1} * \mathbf{A}_{grav} * Q + A + v_0 \times (q_0^{-1} * \omega * q_0) \\ &= V \times \Omega + Q^{-1} * \mathbf{A}_{grav} * Q + A \end{aligned}$$

donc la dynamique dans les nouvelles variables s'écrit de la même manière

$$\begin{aligned}\frac{d}{dt}Q &= \frac{1}{2}Q * \Omega \\ \frac{d}{dt}V &= V \times \Omega + Q^{-1} * \mathbf{A}_{grav} * Q + A\end{aligned}$$

et la dynamique est invariante au sens de la définition 2. La fonction de sortie

$$y = (y_v, y_b) = (v, q^{-1} * \mathbf{B} * q)$$

est G -compatible au sens de la définition 3 avec

$$(Y_v, Y_b) = \varrho_{(q_0, v_0)}(y_v, y_b) = (q_0^{-1} * y_v * q_0 + v_0, q_0^{-1} * y_b * q_0).$$

7.2.1 Un pré-observateur invariant

On suit la méthode du chapitre 6.

Erreur de sortie invariante , ensemble complet d'invariants fondamentaux et repère invariant .

Les équations de normalisation (6.8) s'écrivent

$$\begin{aligned}q * q_0 &= 1 \\ q_0^{-1} * v * q_0 + v_0 &= 0\end{aligned}$$

(où 1 est le quaternion unité : $1 + 0e_1 + 0e_2 + 0e_3$), et par conséquent

$$\begin{pmatrix} q_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} q \\ v \end{pmatrix}^{-1} = \begin{pmatrix} q^{-1} \\ -q * v * q^{-1} \end{pmatrix} = \gamma \begin{pmatrix} q \\ v \end{pmatrix}$$

un ensemble complet d'invariants fondamentaux est

$$I(q, v, a, \omega) = \psi_{\gamma(q, v)} \begin{pmatrix} \omega \\ a \end{pmatrix} = \begin{pmatrix} q * \omega * q^{-1} \\ q * (a + v \times \omega) * q^{-1} \end{pmatrix}$$

Soit $(q_0, v_0) = \gamma(\hat{q}, \hat{v})$. Une erreur de sortie invariante est :

$$E = \varrho_{(q_0, v_0)}(\hat{y}) - \varrho_{(q_0, v_0)}(y) = \begin{pmatrix} \hat{q} * (\hat{v} - v) * \hat{q}^{-1} \\ \mathbf{B} - \hat{q} * y_b * \hat{q}^{-1} \end{pmatrix}$$

Afin de construire un repère invariant on doit choisir une base de l'espace tangent à l'identité aux sous groupe des quaternions de norme 1. Il est constitué de l'ensemble des quaternions dont la première coordonnée est nulle. C'est un espace de dimension 3. e_1, e_2, e_3 est la base canonique de cet espace, et peut être identifiée à la base canonique de \mathbb{R}^3 . Un repère invariant est alors donné par l'ensemble des 6 champs de vecteurs dont les valeurs en (q, v) sont les suivantes :

$$DL_{(q,v)} \begin{pmatrix} e_i \\ 0 \end{pmatrix}_{1 \leq i \leq 3} = \begin{pmatrix} e_i * q \\ 0 \end{pmatrix}_{1 \leq i \leq 3}, \quad DL_{(q,v)} \begin{pmatrix} 0 \\ e_i \end{pmatrix}_{1 \leq i \leq 3} = \begin{pmatrix} 0 \\ q^{-1} * e_i * q \end{pmatrix}_{1 \leq i \leq 3}$$

Pré-observateur invariant

Tout observateur invariant s'écrit

$$\begin{aligned} \frac{d}{dt} \hat{q} &= \frac{1}{2} \hat{q} * \omega + \sum_{i=1}^3 \mathcal{L}_i^q(I, E) e_i * \hat{q} \\ \frac{d}{dt} \hat{v} &= \hat{v} \times \omega + \hat{q}^{-1} * \mathbf{A}_{grav} * \hat{q} + a + \sum_{i=1}^3 \mathcal{L}_i^v(I, E) \hat{q}^{-1} * e_i * \hat{q}, \end{aligned}$$

où les $\mathcal{L}_i^q, \mathcal{L}_i^v$ sont des fonctions régulières de I et E et telles que $\mathcal{L}_i^q(I, 0) = 0$ et $\mathcal{L}_i^v(I, 0) = 0$. Afin de les écrire dans leur forme alternative (6.7) on décompose E en $(E_v, E_b) = (\hat{q} * (\hat{v} - v) * \hat{q}^{-1}, \mathbf{B} - \hat{q} * y_b * \hat{q}^{-1})$ et l'on écrit

$$\begin{aligned} \mathcal{L}_i^q(I, E_v, E_b) &= \bar{\mathcal{L}}_{v,i}^q(I, E_v, E_b) \cdot E_v + \bar{\mathcal{L}}_{b,i}^q(I, E_v, E_b) \cdot E_b \\ \mathcal{L}_i^v(I, E_v, E_b) &= \bar{\mathcal{L}}_{v,i}^v(I, E_v, E_b) \cdot E_v + \bar{\mathcal{L}}_{b,i}^v(I, E_v, E_b) \cdot E_b, \end{aligned}$$

où les $\bar{\mathcal{L}}_{v,i}^q, \bar{\mathcal{L}}_{b,i}^q, \bar{\mathcal{L}}_{v,i}^v, \bar{\mathcal{L}}_{b,i}^v$ sont des matrices 3×1 dont les entrées dépendent de (I, E_v, E_b) . Ainsi,

$$\begin{aligned} \sum_{i=1}^3 \mathcal{L}_i^q(I, E) e_i * \hat{q} &= \left(\sum_{i=1}^3 e_i (\bar{\mathcal{L}}_{v,i}^q(I, E_v, E_b) \cdot E_v + \bar{\mathcal{L}}_{b,i}^q(I, E_v, E_b) \cdot E_b) \right) * \hat{q} \\ &= \left(\sum_{i=1}^3 e_i (\bar{\mathcal{L}}_{v,i}^q \cdot E_v + \bar{\mathcal{L}}_{b,i}^q \cdot E_b) \right) * \hat{q} \\ &= \left(\begin{pmatrix} \bar{\mathcal{L}}_{v,1}^q \\ \bar{\mathcal{L}}_{v,2}^q \\ \bar{\mathcal{L}}_{v,3}^q \end{pmatrix} E_v + \begin{pmatrix} \bar{\mathcal{L}}_{b,1}^q \\ \bar{\mathcal{L}}_{b,2}^q \\ \bar{\mathcal{L}}_{b,3}^q \end{pmatrix} E_b \right) * \hat{q} \end{aligned}$$

En procédant de la même manière avec les autres termes de correction $\sum_{i=1}^3 \mathcal{L}_i^v(I, E) \hat{q}^{-1} * e_i * \hat{q}$, l'observateur invariant général s'écrit

$$\frac{d}{dt} \hat{q} = \frac{1}{2} \hat{q} * \omega + (\bar{\mathcal{L}}_v^q E_v + \bar{\mathcal{L}}_b^q E_b) * \hat{q} \quad (7.7)$$

$$\frac{d}{dt} \hat{v} = \hat{v} \times \omega + \hat{q}^{-1} * \mathbf{A}_{grav} * \hat{q} + a + \hat{q}^{-1} * (\bar{\mathcal{L}}_v^v E_v + \bar{\mathcal{L}}_b^v E_b) * \hat{q} \quad (7.8)$$

où $\bar{\mathcal{L}}_v^q$, $\bar{\mathcal{L}}_b^q$, $\bar{\mathcal{L}}_v^v$ et $\bar{\mathcal{L}}_b^v$, sont des matrices 3×3 de gain dont les entrées dépendent de E_v et E_b et aussi de l'invariant $I(\hat{q}, \hat{v}, a, \omega)$.

Une propriété intéressante de l'observateur lié à sa structure géométrique qui respecte la physique des équations, est qu'on a automatiquement la conservation de la norme de \hat{q} par (7.7), car ω et $\bar{\mathcal{L}}_v^q E_v + \bar{\mathcal{L}}_b^q E_b$ sont des vecteurs de \mathbb{R}^3 (i.e. des quaternions de première coordonnée nulle).

Equation d'erreur

L'erreur d'état invariante définie comme à la section 6.3 s'écrit $\eta = (\eta_q, \eta_v)$ avec

$$\eta_q = \hat{q} * q^{-1}, \quad \eta_v = q * (\hat{v} - v) * q^{-1}.$$

On remarque que η est une erreur au sens de la loi de composition du groupe. Donc une erreur petite correspond à (η_q, η_v) proche de l'élément neutre du groupe $(1, 0)$. Sa dérivée temporelle vérifie (analogue de (6.16)) :

$$\begin{aligned} \frac{d}{dt} \eta_q &= \left(\frac{1}{2} \hat{q} * \omega + (\bar{\mathcal{L}}_v^q E_v + \bar{\mathcal{L}}_b^q E_b) * \hat{q} \right) * q^{-1} - \hat{q} * (q^{-1} * \frac{1}{2} q * \omega * q^{-1}) \\ &= 0 + (\bar{\mathcal{L}}_v^q E_v + \bar{\mathcal{L}}_b^q E_b) * \eta_q \\ \frac{d}{dt} \eta_v &= q * ((\hat{v} - v) \times \omega + \hat{q}^{-1} * \mathbf{A}_{grav} * \hat{q} - q^{-1} * \mathbf{A}_{grav} * q \\ &\quad + \hat{q}^{-1} * (\bar{\mathcal{L}}_v^v E_v + \bar{\mathcal{L}}_b^v E_b) * \hat{q}) * q^{-1} + \frac{1}{2} q * \omega * (\hat{v} - v) * q^{-1} - q * (\hat{v} - v) * \omega * q^{-1} \\ &= q * (\hat{v} - v) \times \omega * q^{-1} + q * \omega \times (\hat{v} - v) * q^{-1} + \eta_q^{-1} * \mathbf{A}_{grav} * \eta_q \\ &\quad - \mathbf{A}_{grav} + \eta_q^{-1} * (\bar{\mathcal{L}}_v^v E_v + \bar{\mathcal{L}}_b^v E_b) * \eta_q \\ &= \eta_q^{-1} * \mathbf{A}_{grav} * \eta_q - \mathbf{A}_{grav} + \eta_q^{-1} * (\bar{\mathcal{L}}_v^v E_v + \bar{\mathcal{L}}_b^v E_b) * \eta_q \end{aligned}$$

où $E_v = \eta_q * \eta_v * \eta_q^{-1}$ et $E_b = \mathbf{B} - \eta_q * \mathbf{B} * \eta_q^{-1}$. Le système d'erreur est donc autonome :

$$\begin{aligned} \frac{d}{dt} \eta_q &= [\bar{\mathcal{L}}_v^q(\eta_q * \eta_v * \eta_q^{-1}) + \bar{\mathcal{L}}_b^q(\mathbf{B} - \eta_q * \mathbf{B} * \eta_q^{-1})] * \eta_q \quad (7.9) \\ \frac{d}{dt} \eta_v &= \eta_q^{-1} * [\mathbf{A}_{grav} + \bar{\mathcal{L}}_v^v(\eta_q * \eta_v * \eta_q^{-1}) + \bar{\mathcal{L}}_b^v(\mathbf{B} - \eta_q * \mathbf{B} * \eta_q^{-1})] * \eta_q - \mathbf{A}_{grav} \end{aligned}$$

Il ne dépend pas de la trajectoire, ni des entrées $\omega(t)$ ni de $a(t)$! Dans le cas général (voir le lemme 4 ou (6.16) pour les groupes de Lie) l'équation d'erreur ne dépend de la trajectoire que via les invariants I . De plus, ici la dynamique est telle que les deux premiers membres de (6.16) se compensent exactement.

Convergence de l'équation d'erreur linéarisée

Supposons que \hat{q} et \hat{v} sont proches de q et v . L'approximation linéaire s'écrit $\delta E_v = \delta \eta_v$ et $\delta E_b = -\delta \eta_q * \mathbf{B} + \mathbf{B} * \delta \eta_q = 2\mathbf{B} \times \delta \eta_q$. Donc l'équation d'erreur linéarisée est

$$\begin{aligned} \frac{d}{dt} \delta \eta_q &= \bar{\mathcal{L}}_v^q \delta \eta_v + 2\bar{\mathcal{L}}_b^q (\mathbf{B} \times \delta \eta_q) \\ \frac{d}{dt} \delta \eta_v &= 2\mathbf{A}_{grav} \times \delta \eta_q + \bar{\mathcal{L}}_v^v \delta \eta_v + 2\bar{\mathcal{L}}_b^v (\mathbf{B} \times \delta \eta_q) \end{aligned}$$

Choisissons

$$\begin{aligned} \bar{\mathcal{L}}_v^q &= \begin{pmatrix} 0 & -M_{12} & 0 \\ M_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \bar{\mathcal{L}}_v^v &= - \begin{pmatrix} N_{11} & 0 & 0 \\ 0 & N_{22} & 0 \\ 0 & 0 & N_{33} \end{pmatrix} \\ \bar{\mathcal{L}}_b^q &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\lambda \mathbf{B}^2 & \lambda \mathbf{B}^1 & 0 \end{pmatrix} & \bar{\mathcal{L}}_b^v &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Dans les coordonnées terrestres,

$$\delta \eta_q := \begin{pmatrix} 0 \\ \delta \eta_q^1 \\ \delta \eta_q^2 \\ \delta \eta_q^3 \end{pmatrix}, \quad \delta \eta_v := \begin{pmatrix} \delta \eta_v^1 \\ \delta \eta_v^2 \\ \delta \eta_v^3 \end{pmatrix} \quad \text{and} \quad \mathbf{A}_{grav} = \begin{pmatrix} 0 \\ 0 \\ \mathbf{a}_{grav} \end{pmatrix},$$

Le système d'erreur se décompose en quatre sous-systèmes découplés :

- le sous-système longitudinal

$$\begin{pmatrix} \delta \frac{d}{dt} \eta_q^2 \\ \delta \frac{d}{dt} \eta_v^1 \end{pmatrix} = \begin{pmatrix} 0 & M_{21} \\ -2\mathbf{a}_{grav} & -N_{11} \end{pmatrix} \begin{pmatrix} \delta \eta_q^2 \\ \delta \eta_v^1 \end{pmatrix} \quad (7.10)$$

- le sous-système latéral

$$\begin{pmatrix} \delta \frac{d}{dt} \eta_q^1 \\ \delta \frac{d}{dt} \eta_v^2 \end{pmatrix} = \begin{pmatrix} 0 & -M_{12} \\ 2\mathbf{a}_{grav} & -N_{22} \end{pmatrix} \begin{pmatrix} \delta \eta_q^1 \\ \delta \eta_v^2 \end{pmatrix} \quad (7.11)$$

– le sous-système vertical

$$\delta \frac{d}{dt} \eta_v^3 = -N_{33} \delta \eta_v^3 \quad (7.12)$$

– le sous-système de cap

$$\delta \frac{d}{dt} \eta_q^3 = \lambda \mathbf{B}^3 (\mathbf{B}^1 \delta \eta_q^1 - \mathbf{B}^2 \delta \eta_q^2) - \lambda ((\mathbf{B}^1)^2 + (\mathbf{B}^2)^2) \delta \eta_q^3 \quad (7.13)$$

Les matrices ont été choisies de manière à ce que le système d'erreur se décompose en quatre sous-systèmes. On peut placer les pôles sur chacun des sous-systèmes. Guidés par des considérations sur l'invariance nous avons obtenu le résultat suivant non-trivial :

Théorème 6. *Soit la dynamique (7.6). L'observateur non-linéaire*

$$\begin{aligned} \frac{d}{dt} \hat{q} &= \frac{1}{2} \hat{q} * \omega(t) + (\bar{\mathcal{L}}_v^q E_v + \bar{\mathcal{L}}_b^q E_b) * \hat{q} \\ \frac{d}{dt} \hat{v} &= \hat{v} \times \omega(t) + \hat{q}^{-1} * \mathbf{A}_{grav} * \hat{q} + a(t) + \hat{q}^{-1} * (\bar{\mathcal{L}}_v^v E_v + \bar{\mathcal{L}}_b^v E_b) * \hat{q} \end{aligned} \quad (7.14)$$

avec

$$E_v = \hat{q} * (\hat{v} - y_v(t)) * \hat{q}^{-1}, \quad E_b = \mathbf{B} - \hat{q} * y_b(t) * \hat{q}^{-1}$$

et avec les matrices de gain constantes $\bar{\mathcal{L}}_v^q$, $\bar{\mathcal{L}}_b^q$, $\bar{\mathcal{L}}_v^v$ et $\bar{\mathcal{L}}_b^v$ choisies telles que les systèmes linéaires (7.10), (7.11), (7.12) et (7.13), soient asymptotiquement stables, converge localement et exponentiellement autour de toute trajectoire du système. De plus la convergence et les exposants de Lyapunov sont complètement indépendants de la trajectoire du système puisque l'équation d'erreur (7.9) est autonome.

Les simulations ci-dessous montrent que la convergence est loin d'être locale. Nous suspectons une plus grande stabilité. Nous proposons comme conjecture que l'observateur non-linéaire est en fait presque globalement convergent. Il ne peut pas être globalement convergent à cause de l'effet "spin" : si $(\eta_q = 1, \eta_v = 0)$ est un équilibre globalement stable pour l'erreur d'équation invariante (7.9), $(\eta_q = -1, \eta_v = 0)$ est également un équilibre globalement stable. D'un point de vue physique ce phénomène n'a aucune importance car \hat{q} et $-\hat{q}$ correspondent à la même rotation $R_{\hat{q}}$ dans $SO(3)$.

7.2.2 Simulations

Afin d'obtenir des valeurs réalistes pour ω , v , \mathbf{a} et $q^{-1} * \mathbf{B} * q$ exprimées dans le repère attaché au solide, on a généré une trajectoire d'un drone du type VTOL simplifié. On modélise le vol de la manière suivante : initialement q est le quaternion unité. Soit k l'axe vertical dirigé vers le bas (quaternion e_3). Soit P la position du centre de masse du solide. On suppose que le mouvement est tel que k est toujours colinéaire à $\ddot{P} - \mathbf{A}_{grav}$. On suppose

que q correspond à la rotation qui envoie \mathbf{A}_{grav} sur k et dont l'axe de rotation est colinéaire à $\mathbf{A}_{grav} \times k$. Donc avec ce modèle la connaissance de la trajectoire de $P(t)$ est suffisante pour calculer $q(t)$, $v(t)$, $\omega(t)$, $\mathbf{a}(t)$ et $q^{-1}(t) * \mathbf{B} * q(t)$.

On suppose qu'initialement $P(0) = \frac{d}{dt}P(0) = \ddot{P}(0) = 0$. $P(t)$ décrit un cercle de rayon 5 mètres, paramétré par l'angle $\theta(t)$. La fonction $t \mapsto P(t)$ est C^3 avec

- Pour $0 \leq t \leq t_1$ on a $0 \leq \theta \leq \pi/2$ et $\ddot{\theta}(t) = c(1 - \cos(2\pi t/t_1))$ où $c = \frac{1}{t_1^2} \frac{2\pi^3}{2\pi^2+1}$ est choisi pour que $t_1 = 2$ s.
- Pour $t_1 \leq t \leq t_2$ on a $\pi/2 \leq \theta \leq 3\pi/2$ et $\ddot{\theta}(t) = 0$ avec $t_2 = 4.15$ s.
- Pour $t_2 \leq t \leq t_3$ on a $3\pi/2 \leq \theta \leq 2\pi$ et $\ddot{\theta}(t) = -c(1 - \cos(2\pi(t - t_2)/t_1))$ avec $t_3 = 6.15$ s.

Le drone s'arrête finalement après avoir décrit un cercle. L'accélération horizontale maximale est d'environ 10 ms^{-2} . Des modèles cinématiques inverses de ce type permettent d'obtenir des valeurs réalistes pour $\mathbf{a}(t)$, $\omega(t)$, $v(t)$ et $q^{-1}(t) * \mathbf{B} * q(t)$. On prend $\mathbf{B} = [\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]$.

Pour les deux simulations illustrées par les figures (7.3,7.4,7.5) et par les figures (7.6,7.7,7.8), les conditions initiales sont identiques :

	Vrai système	Observateur (7.14)
q_0	1	$\cos(\pi/3)$
q_1	0	$\sin(\pi/3)/\sqrt{3}$
q_2	0	$-\sin(\pi/3)/\sqrt{3}$
q_3	0	$\sin(\pi/3)/\sqrt{3}$
v_1	0	10
v_2	0	-10
v_3	0	5

Cela veut dire que la rotation initiale diffère d'un angle $2\pi/3$ de la vraie rotation. Les gains sont les suivants : $M_{12} = M_{21} = 0.4$, $N_{11} = N_{22} = 4$, $N_{33} = 2$ et $\lambda = 4$. Avec $\mathbf{a}_{grav} = 10 \text{ ms}^{-2}$, les pôles des sous-systèmes longitudinaux et latéraux sont : $-2(1 \pm i) \text{ s}^{-1}$ et ceux des sous-systèmes verticaux et de cap -2 s^{-1} .

La seule différence entre la première (figures (7.3,7.4,7.5)) et la seconde (figures (7.6,7.7,7.8)) simulation est que les signaux mesurés de la seconde sont bruités et biaisés. En effet on a ajouté des hautes fréquences et des biais aux signaux $\mathbf{a}(t)$, $\omega(t)$, $v(t)$ et $q^{-1} * \mathbf{B} * q(t)$ de manière à rendre compte des imperfections des capteurs. Les signaux bruités et biaisés sont définis par : $\mathbf{a}_m(t) = \mathbf{a}(t) + 0.5 [1, -1, 1] + \sigma_1$, et $\omega_m(t) = \omega(t) + 4\pi/360 [1, -1, 1] + 0.25 \sigma_2$, et $v_m(t) = v(t) + 0.5 [1, -1, 1] + \sigma_3$, et $y_{bm}(t) = y_b(t) + 0.05 [1, -1, 1] + 0.1 \sigma_4$, où les σ_i sont des vecteurs aléatoires indépendants distribués normalement avec moyenne 0 et variance 1, et $\mathbf{a}(t)$, $\omega(t)$, $v(t)$, $q^{-1} * \mathbf{B} * q(t)$ sont les signaux parfaits et réguliers calculés avec le modèle de drone de type VTOL (figure 7.3). Les simulations montrent que le bassin d'attraction de l'observateur asymptotique (7.14) est large, et qu'il est robuste aux biais et aux bruits de mesure.

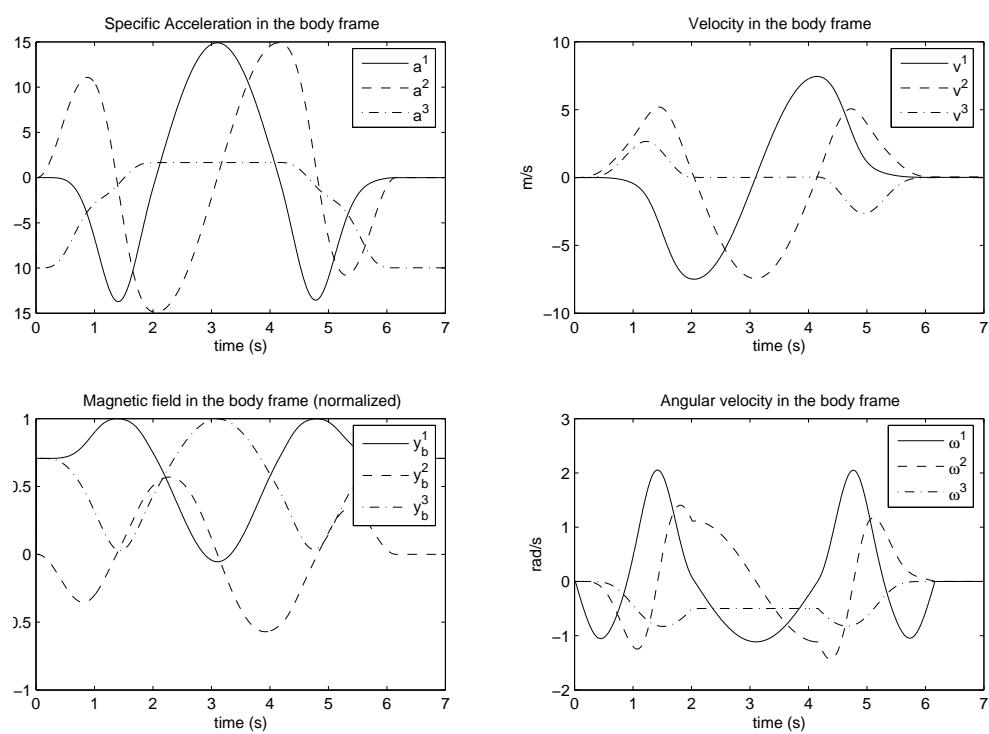


FIG. 7.3 – Signaux mesurés (pas de bruit ni de biais) : accélération spécifique a , vitesse $y_v = v$, champ magnétique normalisé y_b , et vitesse angulaire ω dans le repère du solide.

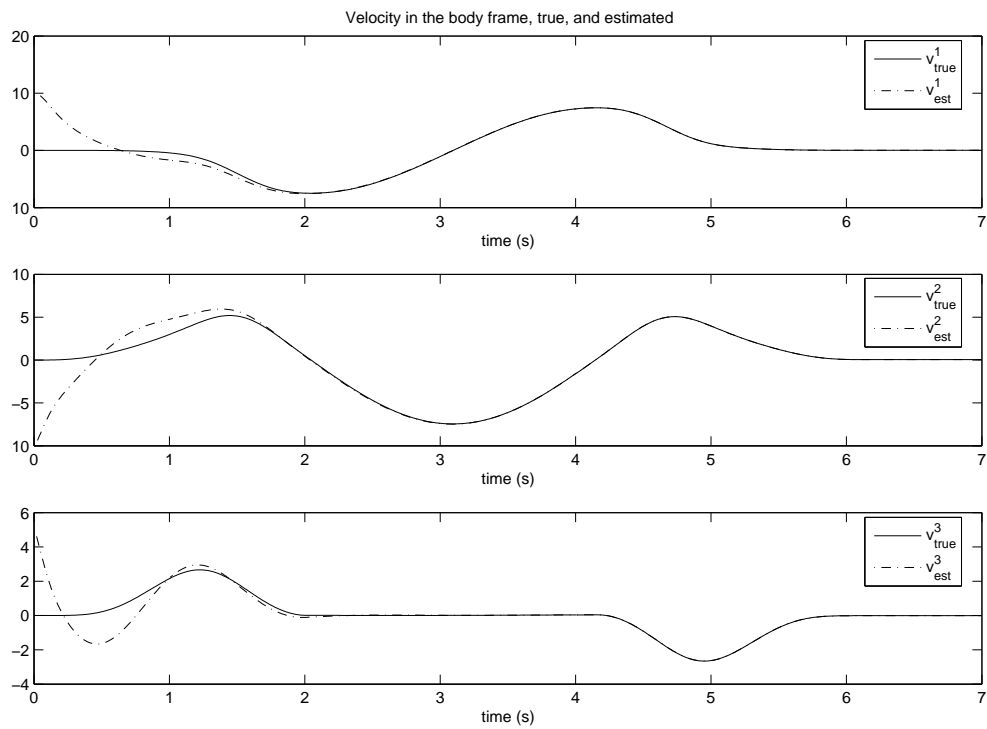


FIG. 7.4 – Vitesse v (trait plein) et vitesse estimée \hat{v} (tirets) via l'observateur invariant (7.14) (pas de bruit ni de biais).

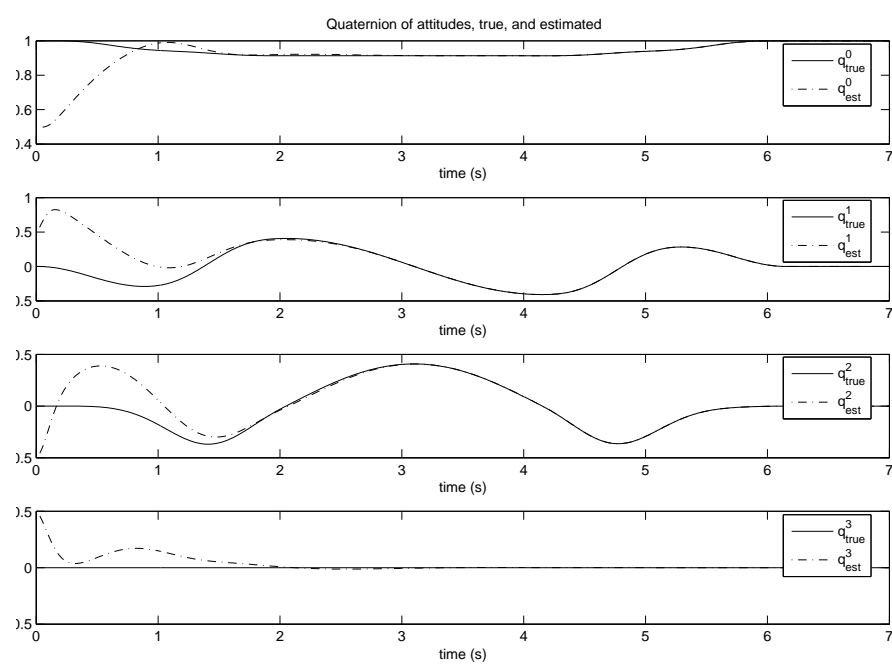


FIG. 7.5 – Orientation q (trait plein) et estimation de l'orientation \hat{q} (tirets) via l'observateur invariant (7.14) (pas de bruit ni de biais).

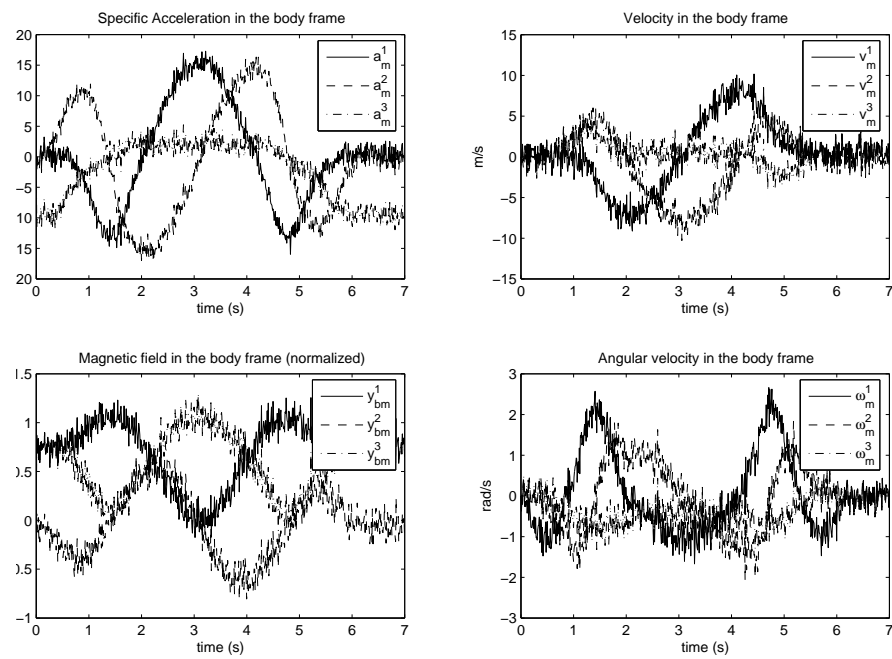


FIG. 7.6 – Signaux mesurés (avec bruit et biais) : accélération spécifique a , vitesse $y_v = v$, champ magnétique normalisé y_b , et vitesse angulaire ω dans le repère du solide.

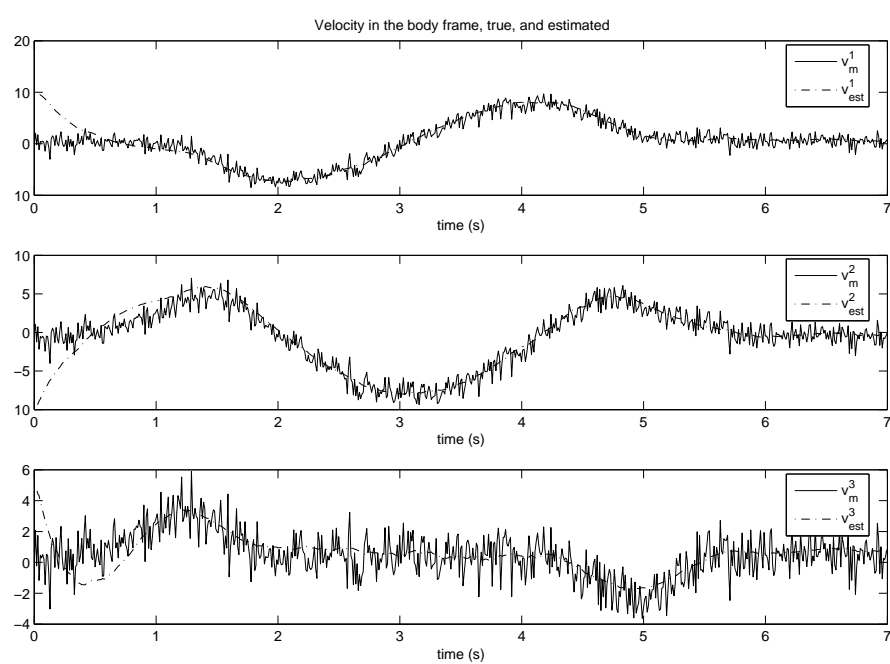


FIG. 7.7 – Vitesse v (trait plein) et vitesse estimée \hat{v} (tirets) via l'observateur invariant (7.14) (avec bruit et biais).

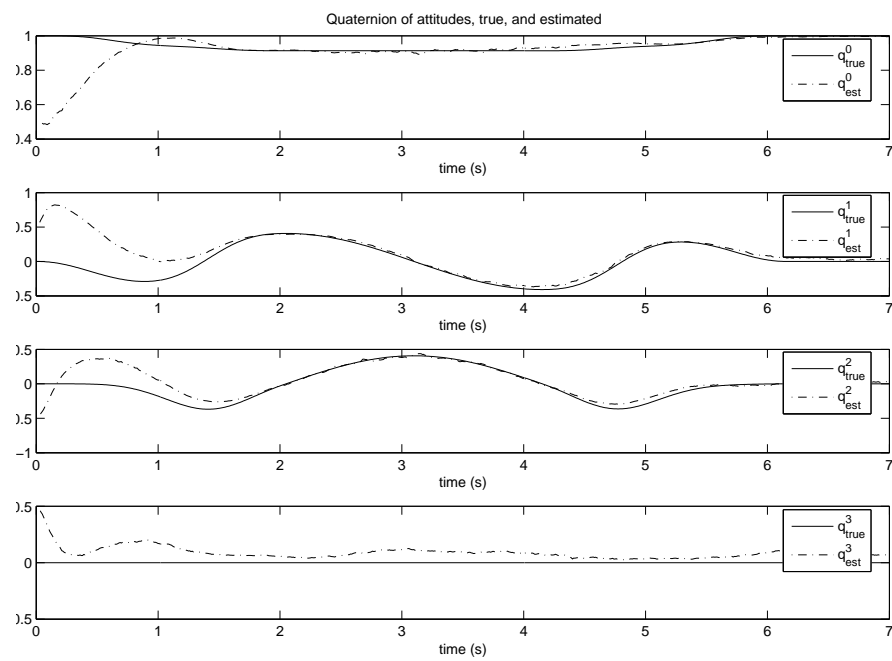


FIG. 7.8 – Orientation q (trait plein) orientation estimée \hat{q} (tirets) via l'observateur invariant (7.14) (avec bruit et biais).

Chapitre 8

Luenberger, Kalman et invariance

Dans ce chapitre nous explorons les liens entre les observateurs de Luenberger, le filtre de Kalman étendu, et les observateurs invariants. Dans la première section on compare un observateur de Luenberger et un observateur invariant autour d'un point d'équilibre. On montre qu'autour de tout point d'équilibre on peut construire un observateur invariant qui admet la même approximation linéaire que tout observateur de Luenberger autour de ce point. Par conséquent si le système linéarisé autour de l'équilibre est observable on peut placer les pôles de l'observateur invariant. Comme nous l'avons vu au chapitre 7 pour la navigation inertielle, cette propriété est intéressante. En effet on peut faire le design des gains pour l'équation d'erreur au premier ordre. Quand l'erreur est plus grande le design est basé sur l'invariance (il y a tout de même une infinité de manière de prolonger les gains du tangent de manière invariante). L'invariance est donc une manière d'"étendre" loin de l'équilibre les gains dont on a fait le design pour le système au premier ordre.

Dans la seconde section, on ira un peu plus loin et l'on montrera qu'il existe une classe de trajectoires non-réduite aux équilibres autour desquelles construire un observateur invariant revient à construire un observateur de Luenberger autour d'un point d'équilibre. Dans le cas d'un mouvement sur un groupe de Lie, nous donnerons un sens géométrique à ces trajectoires qui sont alors des sous-groupes à un paramètre.

Dans la dernière section, nous explorons les liens entre invariance et filtre de Kalman. Par souci de simplicité, on considère une dynamique invariante à gauche sur un groupe de Lie. On suppose que les équations du système sont des équations différentielles stochastiques, i.e, il y a un bruit d'état et un bruit de mesure. Nous montrerons qu'on peut définir les bruits d'état et les bruits de mesure de façon à ce qu'ils respectent l'invariance à gauche des équations. Nous proposerons un filtre de Kalman étendu tel que l'erreur d'estimation invariante suit une équation différentielle indépendante de la trajectoire. La théorie est illustrée par un exemple simple d'estimation d'attitude (pour un drone par exemple...).

8.1 Point d'équilibre

Dans ce paragraphe, on montre qu'il est toujours possible de transformer un observateur asymptotique avec un design local des gains en un observateur invariant ayant le même comportement local (voir l'exemple du réacteur chimique du chapitre 5 section 1). En effet soit $(\bar{x}, \bar{u}, \bar{y})$ un équilibre caractérisé par $f(\bar{x}, \bar{u}) = 0$ et $\bar{y} = h(\bar{x}, \bar{u})$. Supposons que le linéaire tangent autour de cet équilibre est observable. Cela veut dire que la paire (A, C) est observable où

$$A = \frac{\partial f}{\partial x}(\bar{x}, \bar{u}), \quad B = \frac{\partial f}{\partial u}(\bar{x}, \bar{u}), \quad C = \frac{\partial h}{\partial x}(\bar{x}, \bar{u}), \quad D = \frac{\partial h}{\partial u}(\bar{x}, \bar{u})$$

Soit l'observateur asymptotique

$$\frac{d}{dt}\hat{x} = f(\hat{x}, u) + L(\hat{y} - y) \quad (8.1)$$

où l'on a choisi la matrice de gain L telle que $A + LC$ est une matrice stable. En général un tel observateur n'est pas invariant. Sous les hypothèses de base du chapitre 6, on peut construire un observateur invariant avec la même approximation linéaire, i.e, un observateur asymptotique de la forme (6.7)

$$\frac{d}{dt}\hat{x} = F(\hat{x}, u, y) = f(\hat{x}, u) + W(\hat{x})\bar{\mathcal{L}}(I(\hat{x}, u), E(\hat{x}, u, y))E(\hat{x}, u, y)$$

avec

$$\frac{\partial F}{\partial \hat{x}}(\bar{x}, \bar{u}, \bar{y}) = A + LC, \quad \frac{\partial F}{\partial u}(\bar{x}, \bar{u}, \bar{y}) = B + LD, \quad \frac{\partial F}{\partial y}(\bar{x}, \bar{u}, \bar{y}) = -L \quad (8.2)$$

Nous allons proposer un choix possible pour $\bar{\mathcal{L}}$ qui satisfasse les conditions sur $\frac{\partial F}{\partial \hat{x}}$, $\frac{\partial F}{\partial x}$ et $\frac{\partial F}{\partial y}$ à l'équilibre. Puisque $E(\hat{x}, u, \hat{y}) \equiv 0$, par différenciation versus \hat{x} , u et y , on a à l'équilibre

$$\frac{\partial E}{\partial \hat{x}} = -\frac{\partial E}{\partial y}C, \quad \frac{\partial E}{\partial u} = -\frac{\partial E}{\partial y}D$$

Soit V la matrice $p \times p$ inversible $V = \frac{\partial E}{\partial y}(\bar{x}, \bar{u}, \bar{y})$. Soit

$$\bar{\mathcal{L}} = -W(\bar{x})^{-1}LV^{-1}$$

Le choix de $\bar{\mathcal{L}}$ est tel que les conditions ci-dessus (8.2) sont remplies. Nous avons fait un observateur invariant avec même approximation linéaire que (8.1).

8.2 Trajectoires permanentes

Cas général

Le but de ce paragraphe est d'étendre les résultats autour d'un point d'équilibre vus à la section 8.1 à une certaine classe de trajectoires appelées trajectoires permanentes.

Supposons que la trajectoire $(x(t), u(t))$ du système (6.1) est telle que $I(x(t), u(t)) = \bar{I}$ est constant au cours du temps t . Selon le lemme 4 l'erreur d'estimation invariante $\eta(\hat{x}(t), x(t))$ vérifie une équation d'erreur autonome

$$\frac{d}{dt}\eta = \Upsilon(\eta, \bar{I}).$$

Si les gains \mathcal{L}_i sont tels que la matrice $\frac{\partial \Upsilon}{\partial \eta}(0, \bar{I})$ est stable, il y a convergence locale de l'observateur G -invariant du théorème 2 autour de la trajectoire $t \mapsto (x(t), u(t))$ telle que $I(x(t), u(t)) = \bar{I}$. Cette propriété motive fortement la définition suivante :

Définition 9. *Une trajectoire (6.1) est dite permanente si $I(x(t), u(t)) = \bar{I}$ est indépendant de t .*

Afin de fixer les idées donnons une caractérisation locale de ces trajectoires. La dynamique écrite dans les variables locales de base et de fibre (z_a, z_b) de la section 6.2.3 est :

$$\begin{aligned} \frac{d}{dt}z_a &= D\varphi_{\gamma(z)}^a f_a(c, z_b, \psi_{\gamma(z)}(u)) \\ \frac{d}{dt}z_b &= f_b(c, z_b, \psi_{\gamma(z)}(u)) \end{aligned}$$

Puisque $I(z, u) = (z_b, \psi_{\gamma(z)}(u))$ la trajectoire $(z(t), u(t))$ est permanente si et seulement si pour tout t

$$z_b(t) = \bar{z}_b, \quad \psi_{\gamma(z)}(u) = \bar{u}$$

où \bar{z}_b et \bar{u} sont indépendants de t . Cela implique que $f_b(c, \bar{z}_b, \bar{u}) = 0$. Donc toute trajectoire permanente s'écrit comme le produit d'un point d'équilibre sur la base avec une trajectoire particulière sur la fibre qui n'est en général pas un équilibre. La trajectoire sur la fibre aura un sens géométrique à la prochaine section. L'approximation linéaire de la dynamique (dans le sens de l'erreur invariante) autour une trajectoire permanente est indépendante de t , ce qui rappelle le cas linéaire stationnaire. En effet on a

Lemme 2. *Soit $(x_r(t), u_r(t))$ une trajectoire permanente. Soit $\eta_x(x, x_r) = \varphi_{\gamma(x)}(x) - \varphi_{\gamma(x)}(x_r)$ et $\eta_u = \psi_{\gamma(x)}(u) - \psi_{\gamma(x)}(u_r)$. On a*

$$\frac{d}{dt}\eta_x = \Upsilon(\eta_x, \eta_u, \bar{I})$$

pour une fonction régulière Υ telle que $\Upsilon(0, 0, \bar{I}) = 0$ ou $\bar{I} = I(x_r(t), u_r(t))$.

Démonstration. La preuve est analogue à la preuve du lemme 4. $\frac{d}{dt}\eta_x$ est une fonction invariante de (x, x_r, u, u_r) . C'est donc une fonction invariante de $(\eta_x, x_r, \eta_u, u_r)$. Mais un ensemble complet d'invariants de x_r et de u_r consiste en les composants de $I(x_r, u_r)$. \square

Soit un observateur invariant donné par le théorème 2. Si $(x_r(t), u_r(t))$ est une trajectoire permanente, on a selon le lemme 2 que l'approximation linéaire de l'équation d'erreur d'état $\eta_x(\hat{x}, x_r)$ est indépendante du temps t . Donc obtenir la convergence autour d'une trajectoire permanente revient à l'obtenir autour d'un point d'équilibre. De plus, si un observateur invariant converge autour d'une trajectoire permanente particulière $(x_r(t), u_r(t))$ il converge autour de toute trajectoire image de cette trajectoire par l'action du groupe, i.e tout trajectoire permanente de la forme $(\varphi_g(x_r(t)), \psi_g(u_r(t)))$ pour tout g puisque la dérivée temporelle de l'équation d'erreur est la même.

Trajectoires permanentes sur des groupes de Lie

Le but de cette section est d'appliquer les résultats de la section précédente au cas $\mathcal{X} = G$, et de montrer qu'il est dans ce cas possible de faire un observateur convergent autour de toute trajectoire permanente. Toute trajectoire du système vérifie

$$\frac{d}{dt}x(t) = DL_{x(t)}f(e, \psi_{x^{-1}}(u(t))) \quad (8.3)$$

grâce à l'invariance de la dynamique. Soit $(x(t), u(t))$ une trajectoire. Elle est dite permanente (voir la définition 9) si $\psi_{x^{-1}(t)}(u(t)) = \bar{u}$ est indépendant de t . La trajectoire permanente $x(t)$ est alors donnée par $x(0) \exp(t\bar{w})$ où \bar{w} est le vecteur $f(e, \bar{u})$. Donc $x(t)$ correspond, à une multiplication à gauche près définie par la condition initiale, à un groupe à un paramètre.

Grâce au lemme 2 l'erreur invariante suit une équation différentielle non linéaire autonome. Construisons un observateur autour de toute trajectoire permanente. On note $(x_r(t), u_r(t))$ une trajectoire permanente associée à $\bar{u} = \psi_{\bar{x}_r^{-1}(t)}u_r(t)$. Supposons que l'on a fait un observateur invariant de la forme (6.15). Alors l'équation d'erreur (6.16) s'écrit

$$\begin{aligned} \frac{d}{dt}\eta &= DL_\eta \cdot f(e, \psi_{\eta^{-1}}(\bar{u})) - DR_\eta \cdot f(e, \bar{u}) \\ &+ DL_\eta \cdot \left(\sum_{i=1}^n \mathcal{L}_i(\psi_{\eta^{-1}}(\bar{u}), h(e, \psi_{\eta^{-1}}(\bar{u})) - h(\eta^{-1}, \psi_{\eta^{-1}}(\bar{u}))) W_i \right). \end{aligned} \quad (8.4)$$

puisque $\psi_{(\bar{x}_\eta)^{-1}}(u) = \psi_{\eta^{-1}}(\psi_{\bar{x}^{-1}}(u)) = \psi_{\eta^{-1}}(\bar{u})$. L'approximation linéaire (6.17) est maintenant indépendante du temps :

$$\frac{d}{dt}\xi = [\xi, f(e, \bar{u})] - \frac{\partial f}{\partial u}(e, \bar{u}) \frac{\partial \psi}{\partial g}(e, \bar{u})\xi - \sum_{i=1}^n \left(\frac{\partial \mathcal{L}_i}{\partial E}(\bar{u}, 0) \frac{\partial h}{\partial x}(e, \bar{u})\xi \right) W_i$$

Ecrivons ξ et $f(e, u)$ dans le repère défini par les W_i : $\xi = \sum_{k=1}^n \xi^k W_k$ et $f(e, \bar{u}) = \sum_{k=1}^n \bar{f}^k W_k$. Soient C_{ij}^k les constantes de structure associées à l'algèbre de Lie G : $[W_i, W_j] =$

$\sum_{k=1}^n C_{ij}^k W_k$. Le système ci-dessus s'écrit :

$$\frac{d}{dt} \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} = (A + \bar{\mathcal{L}}C) \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix} \quad (8.5)$$

où

$$A = \left(\left[\frac{\partial f}{\partial u}(e, \bar{u}) \frac{\partial \psi}{\partial g}(e, \bar{u}) \right]_{i,j} - \sum_{k=1}^n C_{jk}^i \bar{f}^k \right)_{1 \leq i, j \leq n},$$

$$\bar{\mathcal{L}} = \left(-\frac{\partial \mathcal{L}_i}{\partial E_k}(\bar{u}, 0) \right)_{\substack{1 \leq i \leq n \\ 1 \leq k \leq p}}, \quad C = \left(\frac{\partial h_k}{\partial x_j}(\bar{u}, 0) \right)_{\substack{1 \leq k \leq p \\ 1 \leq j \leq n}}$$

et (x_1, \dots, x_n) sont les coordonnées locales autour de e définies par l'application exponentielle : $x = \exp(\sum_{i=1}^n x_i W_i)$. Si l'on suppose la paire (A, C) observable on peut choisir les pôles de $A + \bar{\mathcal{L}}C$ de manière à obtenir un observateur invariant et localement convergent autour de toute trajectoire permanente associée à \bar{u} . Il suffit de prendre :

$$\frac{d}{dt} \hat{x} = f(\hat{x}, u(t)) + W(\hat{x}) \bar{\mathcal{L}} E(\hat{x}, u(t), y(t)) \quad (8.6)$$

où $E(\hat{x}, u(t), y(t))$ est le vecteur $p \times 1$ qui correspond aux composantes de l'erreur de sortie invariante : $h(e, \psi_{\hat{x}^{-1}}(u(t))) - \varrho_{\hat{x}^{-1}}(y(t))$ et $W(x) = [W_1(x), \dots, W_n(x)]$.

Trajectoires permanentes : Exemples

Les trajectoires permanentes ne coïncident par forcément avec les équilibres : par exemple pour l'exemple de la voiture de la section 7.1, les trajectoires permanentes sont les cercles et les droites décrites à vitesse constante (angle de braquage et vitesse constants). En effet puisque $\psi_{x^{-1}}(u) = (u, uv)$, une trajectoire est permanente si u et v sont indépendants de t . Il y a donc deux sortes de trajectoires permanentes au sens de la définition 9 : cercles et droites avec vitesse constante. Il est aisé de construire un observateur de Luenberger autour de la droite horizontale par exemple : $v = 0, \theta = 0$. On prend

$$\frac{d}{dt} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\theta} \end{pmatrix} = \begin{pmatrix} u \cos \hat{\theta} \\ u \sin \hat{\theta} \\ uv \end{pmatrix} + \begin{pmatrix} -|u|a & 0 \\ 0 & -|u|c \\ 0 & -ub \end{pmatrix} \begin{pmatrix} \hat{x} - x \\ \hat{y} - y \end{pmatrix}$$

et on trouve une approximation de (5) autour de cette trajectoire particulière.

On considère maintenant l'exemple de navigation inertielle vu à la section 7.2. Puisqu'on a $\psi_{x^{-1}}(u) = \begin{pmatrix} q * \omega * q^{-1} \\ q * (a + v \times \omega) * q^{-1} \end{pmatrix}$, une trajectoire de (7.6) est permanente si $q * \omega * q^{-1}$

et $q*(a+v \times \omega)*q^{-1}$ sont indépendantes de t . On montre que toute trajectoire permanente s'écrit :

$$q(t) = \exp\left(\frac{\Omega}{2}t\right) * q_0$$

$$v(t) = q_0^{-1} * \left((\lambda\Omega t + \Upsilon + \exp\left(-\frac{\Omega}{2}t\right) * \Gamma * \exp\left(\frac{\Omega}{2}t\right) \right) * q_0$$

où Ω , Υ et Γ sont des vecteurs constants de \mathbb{R}^3 , λ est un scalaire constant et q_0 quaternion de norme 1. Ces constantes peuvent être choisies arbitrairement. La trajectoire permanente générale s'écrit, à une transformation galiléenne près, comme un mouvement hélicoïdal uniformément accéléré le long de l'axe de rotation quand $\lambda \neq 0$; et quand λ tend vers l'infini et Ω vers 0, nous retrouvons le cas dégénéré d'un mouvement rectiligne accéléré. Quand $\lambda = 0$ et $\Gamma = 0$ on retrouve le virage stabilisé.

8.3 Observateur invariant et filtre de Kalman étendu

Cette section fait l'objet d'une communication soumise au CDC 2007 [9].

8.3.1 Introduction

Dans cette section on cherche à définir les processus stochastiques de bruit et à modifier les gains du filtre de Kalman étendu de manière à ce qu'il respecte les symétries. Afin de relier la théorie des observateurs au filtre de Kalman, on doit donner une formulation stochastique du problème de l'estimation de l'état x . Cela peut être fait de la manière suivante. Considérons donc le problème qui consiste à estimer le processus $x(t)$ à partir de l'observation d'un processus $y(t)$. Cela peut être formulé en termes stochastiques comme un problème de filtrage non-linéaire décrit par les équations différentielles stochastiques :

$$\dot{x} = f(x, t) + M(x)w \quad (8.7)$$

$$y = h(x) + N(x)v \quad (8.8)$$

$$x(t_0) = x^0 \quad (8.9)$$

Le processus d'état et de sortie vivent sur les variétés de dimensions respectives n et p . x^0 est gaussien, w et v sont des bruits blancs gaussiens indépendants de l'état et de sortie.

Le problème du filtrage non-linéaire consiste à calculer en temps réel l'espérance conditionnelle de l'état $x(t)$ sachant les observations passées $\{y(s), t_0 < s < t\}$. Le filtre de Kalman étendu (EKF) fournit une approximation de l'espérance conditionnelle décrite par un nombre fini de statistiques. Elles évoluent selon des équations différentielles ordinaires (à la place d'équations stochastiques aux dérivées partielles) et peuvent donc être calculées en temps réel. En effet l'EKF est un observateur, et c'est le filtre non-linéaire

le plus utilisé. Soit $A(t) = \frac{\partial f}{\partial x}(\hat{x}(t))$ et $C(t) = \frac{\partial h}{\partial x}(\hat{x}(t))$ (et MM^T , NN^T les matrices de covariance du bruit d'état et du bruit d'observation) les équation de l'EKF s'écrivent

$$\frac{d}{dt}\hat{x} = f(\hat{x}, t) + K(t)(y - h(\hat{x})) \quad (8.10)$$

$$K(t) = P(t)C^T(t)(NN^T)^{-1} \quad (8.11)$$

$$\begin{aligned} \frac{d}{dt}P(t) &= A(t)P(t) + P(t)A^T(t) + MM^T \\ &\quad - P(t)C^T(t)(NN^T)^{-1}C(t)P(t) \end{aligned} \quad (8.12)$$

Elles reviennent à construire un filtre de Kalman pour l'approximation linéaire de (8.7)-(8.8)-(8.9) et à l'implémenter sur le modèle non-linéaire. Comme l'observateur de Luenberger elles ne prennent pas en compte la géométrie spécifique du problème (8.7). Quand le modèle admet des symétries, les équations de l'EKF ne respectent pas ces symétries. On voudrait que ce soit le cas, donc on veut transformer l'EKF (qui est un observateur) en un observateur invariant. Dans cette section, par souci de simplicité, on considérera le cas où la dynamique f admet n symétries. Sous quelques hypothèses sur la régularité de l'action de groupe (voir chapitre 6 section sur les groupes de Lie) cela revient au cas où $\mathcal{X} = G$ et $\dot{x} = f(x, t)$ est une dynamique invariante à gauche. On verra dans ce cas comment l'on peut définir les bruits $M(x)dw$ et $N(x)dv$ de manière à ce que (i) ils préservent les symétries (ii) L'EKF dans son ensemble prenne en compte la géométrie du système et préserver les symétries. Notre observateur est donné par (8.17)-(8.18)-(8.19)-(8.20). Cette formulation invariante à gauche, donc intrinsèque, du filtre de Kalman étendu est nouvelle à notre connaissance.

L'idée du filtre de Kalman invariant est la suivante. Les bruits dépendent des variables d'état de telle manière que les deux méthodes suivantes sont équivalentes :

1. Soit l'observateur non-linéaire (8.17)-(8.18)-(8.19)-(8.20) ; on considère une erreur d'état entre l'estimée et le vrai système sans les bruits bien posée (i.e, formulée de manière intrinsèque), on calcule sa dérivée temporelle ; on la linéarise ; on ajoute les bruits d'état et d'observation additifs au système linéarisé, on fait un filtre de Kalman pour ce système linéaire ; et on en déduit les gains K de l'observateur.
2. On considère la dynamique (8.13)-(8.14)-(8.15) où le bruit d'état est non-linéaire et invariant à gauche ; on calcule l'équation d'erreur pour l'erreur bien posée ; on fait un filtre de Kalman pour le système linéarisé, et on en déduit le gain K .

Les filtres de Kalman étendu, filtre de Kalman unscented, et filtre de Kalman d'ensemble sont des tentatives de prendre en compte les non-linéarités. Il y a eu plusieurs autres tentatives d'introduire de la géométrie dans le problème de filtrage non-linéaire ([15, 39, 31]). Ici on aborde le problème du filtrage par le biais des groupes et on fait un filtre de Kalman pour une dynamique non-linéaire invariante à gauche sur un groupe de Lie. On ne montre aucun résultat sur la convergence. On ne s'intéresse qu'à la structure de

l'observateur. Des résultats partiels sur la convergence de l'EKF peuvent être trouvés, e.g., dans [48].

Dans la section 8.3.2 on montre comment on peut définir des bruits qui respectent l'invariance. Dans la section 8.3.3 on construit un filtre de Kalman étendu sur un groupe de Lie. Dans la section 8.3.4 on applique la théorie à un exemple de [37]. Des simulations sont faites à la section 8.3.5.

8.3.2 Bruits et symétries

A partir de maintenant on considère la dynamique

$$\frac{d}{dt}x = f(x, t) + M(x)w \quad (8.13)$$

$$y = h(x) + N(h(x))v \quad (8.14)$$

$$x(t_0) = x^0 \quad (8.15)$$

où x est un élément d'un groupe de Lie réel de dimension n noté G , $y \in \mathbb{R}^p$, x^0 est gaussien, w et v sont des bruits blancs gaussiens indépendants d'observation et de mesure, $M(x)$ est un morphisme qui dépend de la variable x et qui envoie l'espace tangent à l'élément neutre du groupe e sur l'espace tangent en x . $N(h(x))$ est un opérateur de \mathbb{R}^p . De plus on suppose que le modèle est tel que f est une dynamique invariante à gauche :

$$\text{pour tout } g, x \in G \quad f(L_g x, t) = DL_g f(x, t) \quad (8.16)$$

où $L_g(x) = gx$ est la multiplication à gauche sur G et DL_g est l'application tangente induite. On rappelle que D correspond à la différentiation. Un bon choix pour les matrices de gain quand on s'intéresse à des applications pratiques n'est pas forcément évident. Une façon logique (plus respectueuse de la physique du système) de définir les bruits est d'utiliser des considérations géométriques relatives à la physique du système.

Bruit d'état invariant

Considérons le modèle dynamique déterministe, c'est-à-dire sans bruit $\frac{d}{dt}z = f(z, t)$. L'invariance à gauche de f (8.16) veut dire que pour tout $g \in G$ $Z = L_g(z)$ implique que $\frac{d}{dt}Z = f(Z, t)$. On verra w comme un vecteur aléatoire de l'algèbre de Lie \mathcal{G} de G . De manière à ce que les bruits préservent les symétries on supposera que $M(x)$ est défini par :

$$M(x) = DL_x M(e)$$

où $M(e)$ est n'importe quel endomorphisme de \mathcal{G} . Cela implique que pour tout $g \in G$, $DL_g M(x) = M(L_g(x))$ puisque $DL_g DL_x = DL_{gx}$. Donc la dynamique (8.13) affectée par

le bruit blanc est également invariante. En effet si l'on note $X = L_g(x)$ on a

$$\begin{aligned} \frac{d}{dt}X &= DL_g \frac{d}{dt}x = DL_g(f(x, t) + M(x)w) \\ &= f(L_g(x), t) + DL_g M(x)w \\ &= f(X, t) + M(X)w \end{aligned}$$

ce qui prouve l'invariance de la dynamique (8.13). $M(e)$ représente l'amplitude du bruit quand x est proche de e .

Bruit d'observation invariant

On suppose que l'application de sortie $h : G \mapsto \mathbb{R}^p$ préserve les symétries au sens où elle est G -compatible au sens de la définition 3. Comme dans les problèmes de filtrage non-linéaire usuels (équation (8.8)), on considère que la sortie y est la mesure de $h(x)$ affectée par un bruit blanc gaussien (dû aux imperfections des capteurs). Si la dynamique est invariante par une certaine transformation (changement de coordonnées etc.) on veut que le bruit de mesure préserve cette invariance au sens suivant : $N(h(x))$ est une matrice qui dépend de $h(x)$ et qui vérifie pour tout $x, g \in G$

$$N(\rho_g(h(x))) = D\rho_g N(h(x))$$

où $D\rho_g$ est l'application tangente ρ_g . Donc la sortie Y en $X = L_g(x)$ est une fonction de $g, h(x)$ et du bruit $N(h(x))v$

$$Y = h(X) + N(h(X))v = \rho_g(h(x)) + D\rho_g N(h(x))v$$

De plus, on suppose que les difféomorphismes ρ_g sont linéaires comme dans l'exemple suivant de la section 8.3.4, ainsi que dans les exemples que l'on a traités au chapitre 5 et 7. Donc $\rho_g(h(x) + N(h(x))v) = \rho_g(h(x)) + D\rho_g N(h(x))v$ et

$$Y = \rho_g(y)$$

où l'on rappelle que Y est la sortie en $X = L_g(x)$. Cela veut dire que le bruit de mesure est tel que l'application de sortie ne détruit pas les symétries, au sens où même affectée d'un bruit elle reste compatible avec l'action du groupe via ρ_g .

8.3.3 Filtre de Kalman étendu invariant à gauche

On considère la dynamique (8.13)-(8.14)-(8.15). Soit

$$\begin{aligned} A(t) : \mathcal{G} \ni \xi &\mapsto [\xi, f(e, t)] \in \mathcal{G} & C &= Dh(e) \\ N &= N(h(e)) & M &= M(e) \end{aligned}$$

où $[\cdot, \cdot]$ est le crochet de Lie associé à l'algèbre de Lie \mathcal{G} . Soit l'observateur

$$\frac{d}{dt}\hat{x} = f(\hat{x}, t) + DL_{\hat{x}}K(t)(\rho_{\hat{x}^{-1}}(y) - \rho_{\hat{x}^{-1}}(h(\hat{x}))) \quad (8.17)$$

$$K(t) = P(t)C^T(t)(NN^T)^{-1} \quad (8.18)$$

$$\frac{d}{dt}P(t) = A(t)P(t) + P(t)A^T(t) + MM^T \quad (8.19)$$

$$- P(t)C^T(t)(NN^T)^{-1}C(t)P(t) \\ P(0) = P^0 \quad (8.20)$$

C'est une généralisation de filtre de Kalman étendu pour (8.13)-(8.14)-(8.15), quand les bruits d'état et de mesure sont choisis comme on l'a expliqué à la section 8.3.2. De plus il présente les mêmes symétries que le système (8.13)-(8.14)-(8.15), i.e, il est invariant à gauche. Contrairement à l'EKF usuel $K(t)$ dépend de t mais ne dépend pas de l'estimation $\hat{x}(t)$! Cela rappelle le filtre de Kalman classique pour une système linéaire instationnaire. La structure de l'observateur est celle de l'observateur invariant sur un groupe de Lie du chapitre 6.

L'observateur préserve les symétries

Nous allons prouver que l'observateur est invariant par multiplication à gauche comme on l'a fait au chapitre 6. En effet on considère maintenant (8.17). Pour tout $g \in G$ si l'on prend $L_g(x) = gx = X$ et $Y = \rho_g(y)$ on a

$$\begin{aligned} \frac{d}{dt}\hat{X} &= DL_g f(\hat{x}, t) + DL_g DL_{\hat{x}}K(t)(\rho_{\hat{x}^{-1}}(y) - \rho_{\hat{x}^{-1}}(h(\hat{x}))) \\ &= f(g\hat{x}, t) + DL_{g\hat{x}}K(t)(\rho_{(g\hat{x})^{-1}}(\rho_g(y)) - \rho_{(g\hat{x})^{-1}}(h(g\hat{x}))) \\ &= f(\hat{X}, t) + DL_{\hat{X}}K(t)(\rho_{\hat{X}^{-1}}(Y) - \rho_{\hat{X}^{-1}}(h(\hat{X}))) \end{aligned}$$

puisque

$$\begin{aligned} - DL_{g_1}f(g_2, t) &= f(g_1g_2, t) \\ - \rho_{(g\hat{x})^{-1}} &= \rho_{\hat{x}^{-1}}\rho_{g^{-1}} \\ - h(g\hat{x}) &= \rho_g h(\hat{x}) \end{aligned}$$

L'observateur est un filtre de Kalman étendu invariant à gauche

L'idée du filtre de Kalman invariant est de linéariser les équations (8.13)-(8.14) autour de la trajectoire $\hat{x}(t)$, puis de construire un filtre de Kalman pour le modèle linéarisé et de l'implémenter sur le modèle non-linéaire. On peut alors dire que P est une approximation de $\mathbb{E}((\hat{x} - x)(\hat{x} - x)^T)$.

L'idée de l'EKF invariant à gauche pour une dynamique invariante à gauche est de calculer la dérivée temporelle de l'erreur d'état invariante, de linéariser cette équation et de construire un filtre de Kalman pour l'équation d'erreur linéaire, et de l'implémenter sur le modèle non-linéaire. Quand l'EKF consiste à implémenter les gains linéaires sur le modèle non-linéaire, l'EKF invariant donne une manière géométrique d'adapter ces gains linéaires au modèle non-linéaire.

Erreur d'état invariante

Considérons l'erreur d'état invariante $\eta = x^{-1}\hat{x}$. Quand l'observateur est défini par (8.17)-(8.18)-(8.19)-(8.20), grâce à (6.16) on a l'équation d'erreur suivante

$$\begin{aligned} \frac{d}{dt}\eta &= DL_\eta f(e, t) - DR_\eta(f(e, t) + M(e)w) \\ &\quad + DL_\eta K(h(\eta^{-1}) + N(h(\eta^{-1}))v - h(e)) \end{aligned} \quad (8.21)$$

Un fait remarquable est que l'équation d'erreur non-linéaire ne dépend que des bruits w et v , du temps t et de η . Elle est donc indépendante de la trajectoire et ne dépend que des positions relatives de \hat{x} et x . On bénéficie en effet directement des résultats du chapitre 6.

Equation d'erreur linéarisée

On veut faire un filtre de Kalman pour l'erreur linéarisée. Pour une erreur petite (η proche de e), on peut poser $\eta = \exp(\epsilon\xi)$ où ξ est un élément de l'algèbre de Lie \mathcal{G} et $\epsilon \in \mathbb{R}$ est petit. Jusqu'aux termes d'ordre 2 en ϵ , on a l'équation d'erreur linéarisée suivante sur \mathcal{G}

$$\begin{aligned} \frac{d}{dt}\xi &= [\xi, f(e, t)] - (Id + r(\xi))M(e)w \\ &\quad + (Id + l(\xi))K((Id - d(\xi))N(h(e))v - Dh(e)\xi) \end{aligned} \quad (8.22)$$

où Id est l'opérateur identité, l , r et d sont les applications tangentes induites par DL , DR et $D\rho_g$: $l(\xi) = \frac{d}{ds}DL_{exp(s\xi)}$, $r(\xi) = \frac{d}{ds}DR_{exp(s\xi)}$ et $d(\xi) = \frac{d}{ds}D\rho_{exp(s\xi)}$. Une telle équation différentielle stochastique s'appelle une équation multiplicative. Un traitement heuristique des équations différentielles stochastiques multiplicatives peut être trouvé dans [53]. Nous nous en inspirons fortement pour justifier les assertions suivantes. Le détail des calculs est donné au paragraphe suivant. Soit α un paramètre qui représente l'amplitude des bruits. Approximer l'équation d'erreur par l'équation suivante (8.23) revient à faire une erreur d'ordre α^2 quand on calcule la dynamique de l'espérance de l'erreur ξ et également la dynamique de la covariance de ξ . Par conséquent : si l'observateur pour (8.23) est robuste, l'espérance de ξ (qui vérifie (8.22)) tendra aussi vers 0 quand $t \rightarrow \infty$. L'erreur sur la matrice de covariance quand on fait cette approximation est d'ordre α^2 . Donc la solution

de (8.23) fournit une approximation de (8.22). On approxime l'équation d'erreur linéarisée maintenant par :

$$\frac{d}{dt}\xi = [\xi, f(e, t)] - M(e)w + K(N(h(e))v - Dh(e)\xi) \quad (8.23)$$

avec A, C, M et N définies au début de la section, les approximations s'écrivent

$$\frac{d}{dt}\xi = (A - KC)\xi - Mw + KNv \quad (8.24)$$

et on retrouve l'équation d'erreur usuelle pour un filtre de Kalman linéaire. Un filtre de Kalman pour le modèle linéaire est donné par les équations (8.24) et (8.18)-(8.19)-(8.20) où P est la matrice de covariance intrinsèque $P(t) = \mathbb{E}(\xi\xi^T)$ de l'erreur linéarisée. Donc (8.17)-(8.18)-(8.19)-(8.20) peut être vu comme un filtre de Kalman étendu invariant à gauche.

Justifications des approximations précédentes

On considère l'équation différentielle stochastique (EDS) (8.22) qui peut être ré-écrite

$$\frac{d}{dt}\xi = a_0(t)\xi + \alpha L(t) + \alpha a_1(t)\xi \quad (8.25)$$

où $a_0(t)$ est une fonction non stochastique du temps t , $L(t)$ et $a_1(t)$ sont des fonctions linéaires des bruits blancs $w(t)$ et $v(t)$, et α est un paramètre qui mesure l'amplitude des bruits. En s'inspirant de [52], on prouve de manière semi-rigoureuse en employant des arguments heuristiques que (8.25) peut être approximée à l'ordre α^2 par l'EDS (8.23) qui s'écrit aussi

$$\frac{d}{dt}\xi = a_0(t)\xi + \alpha L(t) \quad (8.26)$$

dans le sens où l'espérance et la covariance des solutions vérifient les mêmes équations différentielles à des termes près du second ordre en α . Supposons que ξ est de dimension 1. La même méthode peut être appliquée quand la dimension de ξ est plus de 1 en utilisant une matrice $A_0(t)$ à la place de $a_0(t)$ et le "propagateur" déterminé par l'équation matricielle $\dot{Y}(t) = A_0(t)Y(t)$, $Y(0) = 1$. Puisque a_1 et L sont des fonctions linéaires des bruits blancs on suppose qu'on a

$$\mathbb{E}(a_1(t)) = \mathbb{E}(L(t)) = 0$$

et

$$\mathbb{E}(a_1(t)a_1(t')) \simeq \mathbb{E}(L(t)L(t')) \simeq k(t, t')$$

avec $k(t, t') = \delta(t - t')$. Pour ne pas sous-estimer la corrélation entre a_1 et L on suppose qu'elle est maximale $\mathbb{E}(a_1(t)L(t')) \simeq k(t, t')$. L'équation (8.25) peut être résolue explicitement : Prenons $\xi(t) = \exp(\int_0^t a_0(\tau)d\tau)v(t)$. On a

$$\frac{d}{dt}v = \alpha a_1(t)v + \alpha L'(t)$$

avec $L'(t) = \exp(-\int_0^t a_0)L(t)$. On a

$$v(t) = v_0 \exp(\int_0^t \alpha a_1) + \alpha \int_0^t [\exp(\int_\tau^t \alpha a_1(s)ds)]L'(\tau)d\tau$$

Mais à l'ordre α^2 , $\exp(\int_0^t \alpha a_1)$ est égal à

$$1 + \alpha \int_0^t dt_1 a_1(t_1) + \alpha^2 \int_0^t dt_1 \int_0^{t_1} dt_2 a_1(t_1)a_1(t_2)$$

dont l'espérance est $1 + 0 + \alpha^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \mathbb{E}(a_1(t_1)a_1(t_2))$, i.e. $\mathbb{E}(\exp(\int_0^t \alpha a_1)) = 1 + \int_0^t \alpha^2$. Donc $\mathbb{E}(v(t))$ peut être approximée à l'ordre α^2 par

$$\mathbb{E}(v_0)(1 + \int_0^t \alpha^2) + \alpha^2 \int_0^t \exp(-\int_0^\tau a_0)d\tau$$

Mais, cependant, ce terme est la solution à l'ordre α^2 de l'équation déterministe

$$\frac{d}{dt}\mathbb{E}(v(t)) = \alpha^2(\mathbb{E}(v(t)) + \exp(-\int_0^t a_0))$$

pour $t \ll \alpha^{-1}$. Par conséquent

$$\frac{d}{dt}\mathbb{E}(\xi) = (a_0(t) + \alpha^2)\mathbb{E}(\xi) + \alpha^2$$

Les calculs s'appliquaient à un intervalle $\Delta t \ll \alpha^{-1}$. Mais la même méthode s'applique à l'intervalle suivant Δt . Cette équation est en fait correcte pour tout temps t (voir [52]). Soit ζ une solution (8.26). On a

$$\frac{d}{dt}\mathbb{E}(\zeta) = a_0(t)\mathbb{E}(\zeta)$$

$\mathbb{E}(\xi)$ et $\mathbb{E}(\zeta)$ sont des solutions de la même équation différentielle stochastique à des termes près d'ordre α^2 . Des calculs analogues montrent que la covariance $\mathbb{E}(\xi^2)$ de la solution de (8.25) et la covariance $\mathbb{E}(\zeta^2)$ de la solution de (8.26) vérifient des équations différentielles (qui sont les équations suivantes (8.27) et (8.28)) dont les coefficients diffèrent par des

termes d'ordre α^2 . Si le bruit est typiquement de 10% les coefficients des équations diffèrent de 1%. Mais l'on rappelle que l'équation (8.25) est d'ores et déjà une approximation puisque que c'est l'équation d'erreur linéarisée où l'on a négligé les termes en $\|\xi\|^2$.

NB : La méthode employée dans cette section est une méthode heuristique dérivée de la physique. L'équation (8.25) est l'équation de Langevin lorsqu'on rajoute une stochasticité interne. Il est possible d'aboutir aux mêmes conclusions en utilisant le formalisme rigoureux du calcul stochastique. La formalisation des calculs précédents provient d'une discussion avec Tony Lelièvre du CERMICS. Calculons rigoureusement les équations vérifiées par les covariances par exemple, et montrons qu'elles ne diffèrent que par des termes d'ordre α^2 . Prenons $L(t) = a_1(t) = dW_t$ avec W_t processus de Wiener ou mouvement Brownien et donc dW_t bruit blanc. L'équation (8.25) s'écrit alors avec le formalisme usuel

$$dX = a_0 X dt + \alpha(1 + X)dW_t$$

La formule d'Itô dit que $d(X^2) = 2X(a_0 X)dt + 2\alpha X(1 + X)dW_t + \alpha^2(1 + X)^2 dt$. Il faut intégrer cette dernière équation en temps. Le terme central du second membre est d'espérance nulle car l'espérance d'une intégrale stochastique d'Ito est nulle car c'est une martingale. Et donc

$$\frac{d}{dt} \mathbb{E}(X^2) = 2a_0 \mathbb{E}(X^2) + \alpha^2 [1 + 2\mathbb{E}(X) + \mathbb{E}(X^2)] \quad (8.27)$$

L'équation (8.26) est $dY = a_0 Y dt + \alpha dW_t$. Et on a donc $\frac{d}{dt} \mathbb{E}(Y) = a_0 \mathbb{E}(Y)$. On a aussi $d(Y^2) = 2a_0 Y^2 dt + 2\alpha Y dW_t + \alpha^2 dt$. Et donc

$$\frac{d}{dt} \mathbb{E}(Y^2) = 2a_0 \mathbb{E}(Y^2) + \alpha^2 \quad (8.28)$$

8.3.4 Exemple

Afin de donner une illustration on va considérer un exemple pédagogique emprunté au domaine de la navigation inertielle qui est une version simplifiée du problème vu au chapitre 7. Nous allons montrer comment définir les processus de bruit, et construire un EKF qui préserve la géométrie du problème.

Un problème simple d'estimation d'attitude

Les équations de la cinématique pour un solide s'écrivent à l'aide des quaternions \mathbb{H} (voir annexe B)

$$\frac{d}{dt} q = \frac{1}{2} q * \omega + q * Mw \quad (8.29)$$

où

- $q \in \mathbb{H}$ est le quaternion de norme 1 qui représente la rotation qui envoie le repère terrestre sur le repère du solide.
- $\omega(t)$ est la vitesse de rotation instantanée mesurée par les gyroscopes.
- M est une matrice 3×3 constante et w est un bruit gaussien de \mathbb{R}^3 .
- $*$ est la multiplication quaternionique.

Les mesures sont la vitesse angulaire instantanée $\omega(t)$, ainsi que le champ magnétique et l'accélération spécifique. On suppose que les deux dernières mesures (issues des magnétomètres et accéléromètres) sont combinées de manière à fournir une estimation algébrique de l'attitude q . Cette approximation est valide quand l'accélération du solide est petite comparée à l'attraction terrestre (10 m.s^{-2}), comme c'est expliqué dans [37] qui considère la même dynamique (8.29) avec la même sortie mais sans bruit ($M = N = 0$). Par conséquent, contrairement à l'exemple traité au chapitre 7 le modèle qu'on considère n'est pas correct quand l'accélération du solide est nulle. On considère que le bruit d'état $q * Mw$ correspond à l'erreur que l'on fait avec cette approximation. La sortie est

$$q_y = q + q * Nv$$

où v est un bruit gaussien standard \mathbb{R}^3 qui correspond aux imperfections de capteurs, et N est une matrice 3×3 réelle.

Symétries

Les équations de la cinématique (8.29) sont indépendantes du choix du repère terrestre (notamment). Il n'y a pas de raison que le bruit d'état et le bruit de mesure dépendent de ces coordonnées non plus. Ils ont été choisis de manière à respecter ces invariances. Sinon, le bruit d'état en particulier n'a aucun sens physique, et dépend même du choix d'écrire les rotations avec des quaternions plutôt qu'avec des matrices. On veut faire un filtre de Kalman étendu invariant à gauche pour ce problème, i.e, qui ne dépende pas du choix du repère terrestre et donc préserve les symétries, ce qui est une propriété intrinsèque.

Vérifions que la dynamique est une dynamique invariante à gauche sur un groupe de Lie. Le sous-groupe des quaternions de norme 1 est un groupe de Lie. Soit r n'importe quel quaternion de norme 1. Soit $Q = r * q$. Les équations pour Q sont inchangées

$$\begin{aligned} \frac{d}{dt}Q &= \frac{1}{2}r * q * \omega + r * q * Mw \\ &= \frac{1}{2}Q * \omega + Q * Mw \end{aligned}$$

De plus la sortie est compatible puisque

$$Q_y = r * q_y = r * (q + q * Nv) = Q + Q * Nv$$

La dynamique et la sortie correspondent bien aux équations (8.13)-(8.14) quand on choisit les bruits comme à la section 8.3.2.

Interprétation physique

Le bruit d'état préserve le fait que q est un quaternion de norme 1 et donc représente bien une rotation. L'équation stochastique usuelle appliquée à l'exemple s'écrirait $\frac{d}{dt}q = \frac{1}{2}q * \omega + Mw$ et ne préserve pas le fait que q est une rotation, et n'a donc pas de sens.

De toute façon le bruit d'état est censé modéliser les incertitudes du modèle mais ici le modèle est une équation cinématique, on lui fait donc confiance. Il est bien plus sensé de ne pas faire confiance à ω .

Un filtre de Kalman étendu invariant à gauche

Au lieu de considérer l'erreur d'état linéaire $\hat{q} - q$ on considère l'erreur d'état invariante $\eta = q^{-1}\hat{q}$ (voir section 8.3.3). η correspond à la rotation qui envoie le repère attaché au vrai solide sur celui attaché au solide estimé, et est donc totalement indépendant du choix du repère terrestre. $P(0)$ est une matrice définie positive qui est la covariance initiale de l'erreur $\mathbb{E}(\eta\eta^T)$. Le commutateur $p \mapsto \frac{1}{2}(p * \omega(t) - \omega(t) * p)$ est une fonction linéaire du quaternion p noté $A(t)$. Soit l'observateur suivant (correspondant à (8.17))-(8.18)-(8.19)) :

$$\begin{aligned} \frac{d}{dt}\hat{q} &= \frac{1}{2}\hat{q} * \omega + \hat{q} * K(t)(\hat{q}^{-1} * q_y - 1) \\ K(t) &= P(t)(NN^T)^{-1} \\ \frac{d}{dt}P(t) &= A(t)P(t) + P(t)A^T(t) + MM^T \\ &\quad - P(t)(NN^T)^{-1}P(t) \end{aligned} \tag{8.30}$$

Le gain $P(t)$ est calculé de sorte que l'approximation linéaire de l'erreur vérifie les équations linéaires de Kalman. En effet l'erreur d'estimation suit l'équation différentielle suivante (correspondant à (8.21)) :

$$\begin{aligned} \frac{d}{dt}\eta &= \frac{1}{2}(\eta * \omega - \omega * \eta) - Mw * \eta \\ &\quad + \eta * K(\eta^{-1} - 1 + \eta^{-1} * Nv) \end{aligned}$$

puisque $\frac{d}{dt}q^{-1} = -q^{-1} * \frac{d}{dt}q * q^{-1} = -\omega * q^{-1}$ et $\hat{q}^{-1} * q = \eta^{-1}$. Supposons que \hat{q} est proche de q et soit $\eta = 1 + \delta\eta$ avec $\delta\eta$ un petit quaternion dont la première coordonnée vaut 0, i.e $\delta\eta$ est un vecteur tangent en l'unité au groupe des quaternions de norme 1. On a jusqu'au termes du second ordre en $\delta\eta$

$$\begin{aligned} \frac{d}{dt}\delta\eta &= A(t)\delta\eta - Mw - Mw * \delta\eta \\ &\quad - \delta\eta * KNv + KNv - K\delta\eta \end{aligned}$$

Et si l'on suppose que les bruits Mw et Nv sont suffisamment petits pour que $Mw * \delta\eta$ et $\delta\eta * KNv$ soient considérés comme des termes du "second ordre". Si les bruits sont trop grands, il n'y a aucune de raison que l'erreur η reste petite au cours du temps, et la méthode du filtre de Kalman étendu n'a plus de sens. L'approximation linéaire de de l'erreur devient (correspondant à (8.23)) :

$$\frac{d}{dt}\delta\eta = A(t)\delta\eta - K\delta\eta - Mw + KNv$$

Le filtre (8.30) est un observateur invariant à gauche tel que les gains correspondent à un filtre de Kalman pour l'approximation linéaire de l'erreur. Il est intéressant de remarquer que $\hat{q}(t)$ (voir chapitre 6) et donc la convergence (et le réglage des gains, donc des matrices M et N) ne dépendent plus de la trajectoire bien que la dynamique soit non-linéaire.

8.3.5 Simulations

Pour avoir des valeurs réalistes de ω et q_y on reprend les simulations de la section 7.2. Le centre de masse du drone décrit un cercle de rayon 5 mètres et s'arrête. Des signaux haute fréquence et des biais sont ajoutés aux signaux correspondant aux trajectoires régulières du drone $\omega(t)$ et $q(t)$. On prend $q_y(t) = q(t) * (1 + .2 \sigma_1)$ et $\omega(t) + .5 \sigma_2$, où les σ_i sont des vecteurs aléatoires de \mathbb{R}^3 d'espérance 0 et de variance 1. L'amplitude du bruit est 20% de la valeur maximale des variables.

Les conditions initiales sont telles que l'orientation initiale estimée diffère de la vraie orientation d'un angle $\pi/3$. Cela s'écrit grâce aux quaternions :

$$\begin{aligned} q(0) &= [1, 0, 0, 0] \\ \hat{q}(0) &= [\cos(\pi/6), \sin(\pi/6)/\sqrt{3}, \\ &\quad -\sin(\pi/6)/\sqrt{3}, \sin(\pi/6)/\sqrt{3}] \end{aligned}$$

On prend les matrices de bruit : $M = 0.5 * I_3$ et $N = 0.2 * I_3$ et $P(0) = 0.1 * I_3$ (par exemple) où I_3 est la matrice identité 3×3 . Les mesures et le comportement de \hat{q} et de q sont représentés par les figures 8.1 and 8.2.

8.3.6 Conclusion

L'EKF invariant a deux avantages pratiques : une manière systématique de propager les matrices de bruit qu'on ne choisit qu'une fois autour d'un point, et donc de faire le réglage des gains ; et l'équation d'erreur (donc le placement des pôles) ne dépend pas de la trajectoire $x(t)$ mais seulement des positions relatives de x et \hat{x} et du temps t via l'entrée $\omega(t)$, bien que la dynamique soit non-linéaire. Ainsi, si les gains sont bien réglés autour d'une trajectoire particulière correspondant à l'entrée $\{\omega(t), t \geq 0\}$, ils sont réglés

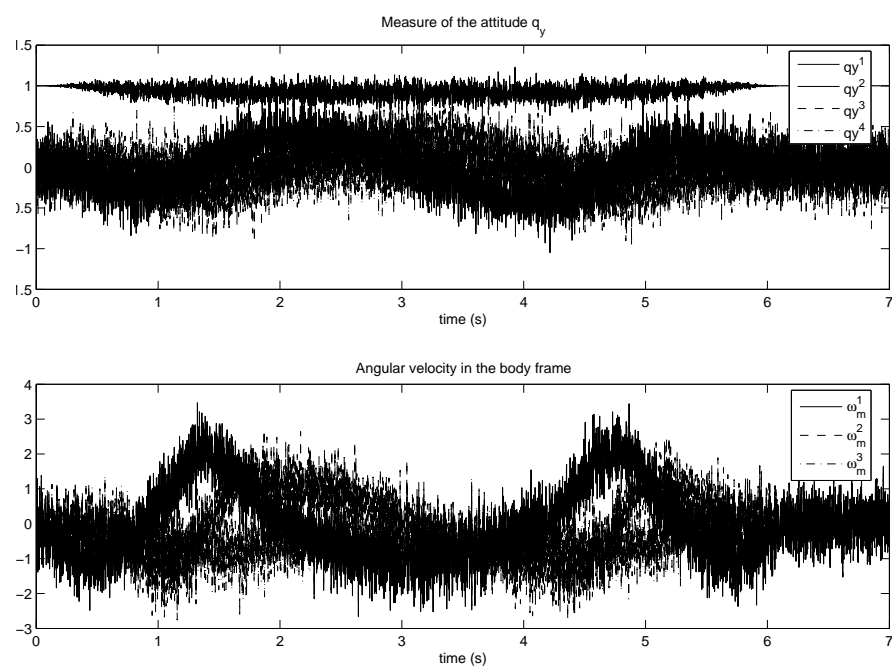


FIG. 8.1 – Signaux mesurés : attitude q_y et vitesse angulaire ω dans le repère attaché au solide .

autour de toute trajectoire correspondant à cette entrée, quelque soit la condition initiale (puisque l'équation d'erreur est identique). On peut également traiter, avec la méthode que l'on vient de développer, le cas général (voir 6.2) d'une dynamique sur une variété de dimension n sous l'action d'un groupe de Lie de dimension r quand $r < n$. Dans ce cas on définit l'erreur d'état comme à la section 6.2.6 et l'on fait un filtre de Kalman étendu invariant comme on l'a fait dans ce paragraphe.

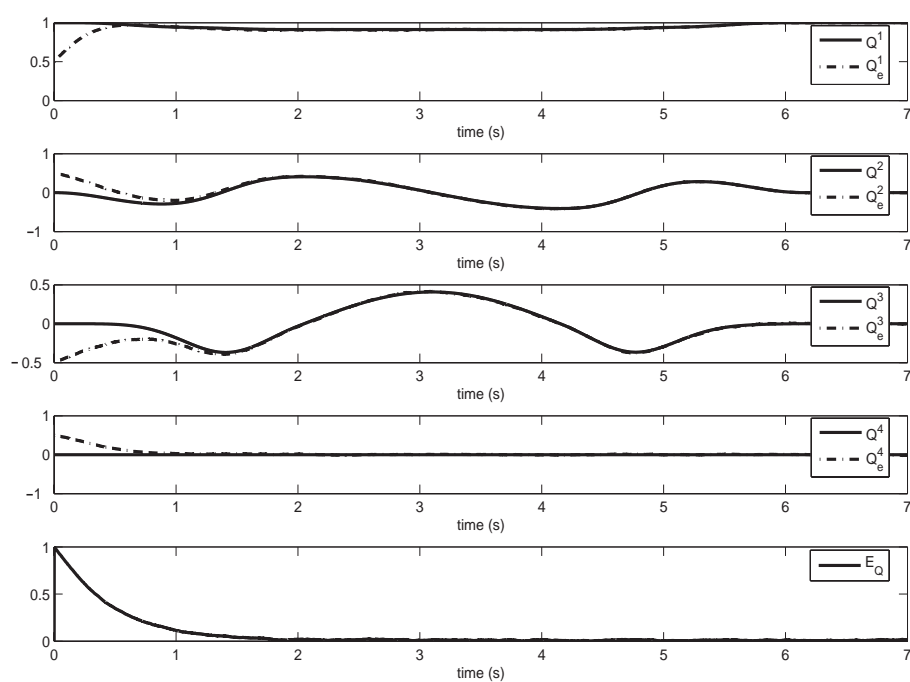


FIG. 8.2 – attitude (trait plein) attitude estimée \hat{q} (tirets); erreur d'estimation $E_Q = \|q^{-1}\hat{q} - 1\|$ avec l' EKF invariant (8.30).

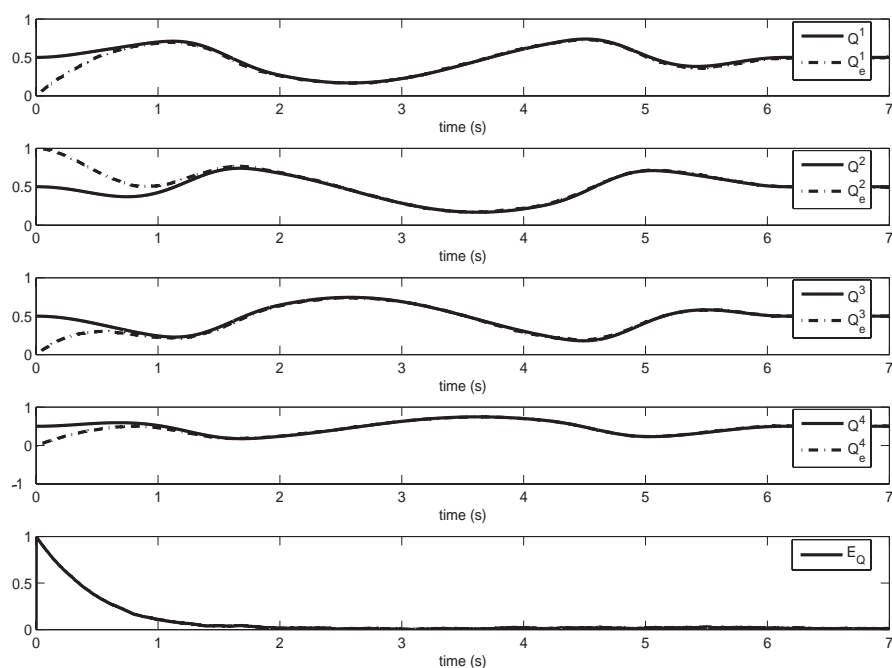


FIG. 8.3 – attitude (trait plein) attitude estimée \hat{q} (tirets); erreur d'estimation $E_Q = \|q^{-1}\hat{q} - 1\|$ avec l' EKF invariant (8.30). La seule différence avec la figure 8.2 est que la condition initiale est différente $q(0) = [0.5, 0.5, 0.5, 0.5]$, par contre $\eta(0)$ est identique. On voit que l'erreur E_Q a le même comportement malgré une trajectoire différente.

Chapitre 9

Observateurs invariants en temps discret

Dans ce chapitre on considère le cas des systèmes en temps discret de la forme (k est un indice entier)

$$x_{k+1} = f(x_k, u_k), \quad y_k = h(x_k, u_k)$$

et l'on reprend les notions de base de la théorie développée au chapitre 6 pour ce type de système. Les systèmes obtenus par discrétisation exacte en temps des exemples considérés au chapitre 7 font partie de ce type de système. En effet, pour fixer les idées, on peut penser à la version discrétisée de l'exemple de la voiture non-holonyme

$$x_{k+1} = x_k + s_k \cos \theta_k, \quad y_{k+1} = y_k + s_k \sin \theta_k, \quad \theta_{k+1} = \theta_k + \frac{s_k}{l} \tan \phi_k$$

avec comme contrôle $s_k = u_k \delta t$ et ϕ_k . Les sorties sont (x_k, y_k) . Mais nous ne traiterons pas l'exemple. Ce chapitre consiste juste en une reprise du chapitre 6 pour les systèmes discrets. Les définitions de l'invariance pour les systèmes et les observateurs sont analogues. Les résultats sur l'équation d'erreur sont les mêmes. Par contre la construction et la caractérisation des observateurs invariants présentées à la section 9.2 prennent une forme différente.

9.1 Systèmes invariants, sorties compatibles et observateurs invariants

Soit le système discret

$$x_{k+1} = f(x_k, u_k) \tag{9.1}$$

$$y_k = h(x_k, u_k) \tag{9.2}$$

où x appartient à l'ouvert $\mathcal{X} \subset \mathbb{R}^n$, u à un ouvert $\mathcal{U} \subset \mathbb{R}^m$ et y à un ouvert $\mathcal{Y} \subset \mathbb{R}^p$, $p \leq n$. On suppose les signaux u_k, y_k connus (u_k est l'entrée du canal et y_k la sortie par exemple en traitement du signal). On considère le groupe de transformations défini par

$$(X, U) = (\varphi_g(x), \psi_g(u)), \quad (9.3)$$

avec φ_g et ψ_g difféomorphismes locaux paramétrés par g élément de G groupe de Lie de dimension r . φ_g agit sur \mathcal{X} et ψ_g sur \mathcal{U} .

Définition 10. Le système $x_{k+1} = f(x_k, u_k)$ est G -invariant si $f(\varphi_g(x_k), \psi_g(u_k)) = \varphi_g(f(x_k, u_k))$ pour tous g, x_k, u_k .

On peut également écrire $X_{k+1} = f(X_k, U_k)$, i.e., le système demeure inchangé par la transformation (9.3).

Définition 11. La sortie $y_k = h(x_k, u_k)$ est G -compatible s'il existe un groupe de transformation $(\varrho_g)_{g \in G}$ sur \mathcal{Y} tel que $h(\varphi_g(x_k), \psi_g(u_k)) = \varrho_g(h(x_k, u_k))$ pour tout g, x_k, u_k .

Avec $(X_k, U_k) = (\varphi_g(x_k), \psi_g(u_k))$ et $Y_k = \varrho_g(y_k)$, la sortie s'écrit $Y_k = h(X_k, U_k)$.

Définition 12 (pre-observateur). Le système $\hat{x}_{k+1} = F(\hat{x}_k, u_k, y_k)$ est un pré-observateur de (9.1)-(9.2) si pour tout x_k, u_k on a $F(x_k, u_k, h(x_k, u_k)) = f(x_k, u_k)$.

Définition 13. Le pré-observateur $\hat{x}_{k+1} = F(\hat{x}_k, u_k, y_k)$ est G -invariant si pour tout g, \hat{x}_k, u_k, y_k ,

$$F(\varphi_g(\hat{x}_k), \psi_g(u_k), \varrho_g(y_k)) = \varphi_g(F(\hat{x}_k, u_k, y_k)).$$

Cela s'écrit également $\hat{X}_{k+1} = F(\hat{X}_k, U_k, Y_k)$, avec $X_{k+1} = \varphi_g(x_{k+1})$, $X_k = \varphi_g(x_k)$, $U_k = \psi_g(u_k)$ et $Y_k = \varrho_g(y_k)$. On dit que c'est un observateur invariant si l'on a de plus $\hat{x}_k \rightarrow x_k$ quand $k \rightarrow \infty$. On aurait pu également appeler un "pré-observateur" un "candidat pour un observateur invariant".

9.2 Construction et caractérisation des observateurs invariants discrets

La construction des observateurs invariants repose sur la méthode du repère mobile développée au paragraphe 6.2.3.

Définition 14. L'application régulière $(\hat{x}_k, u_k, y_k) \mapsto E(\hat{x}_k, u_k, y_k) \in \mathbb{R}^p$ est une erreur de sortie invariante si

- L'application $y \mapsto E(\hat{x}_k, u_k, y)$ est inversible pour tout \hat{x}_k, u_k
- $E(\hat{x}_k, u_k, h(\hat{x}_k, u_k)) = 0$ pour tout \hat{x}_k, u_k
- $E(\varphi_g(\hat{x}_k), \psi_g(u_k), \varrho_g(y_k)) = E(\hat{x}_k, u_k, y_k)$ pour tout g, \hat{x}_k, u_k, y_k

Cette définition coïncide avec la définition 8 et l'existence d'erreur de sortie invariante est assurée par le théorème 3. Le théorème suivant est une caractérisation des observateurs invariants discrets et indique comment on les construit avec la méthode du repère mobile.

Théorème 7. *Soit un système discret G -invariant $x_{k+1} = f(x_k, u_k) \in \mathcal{X}$ avec une sortie G -compatible $y_k = h(x_k, u_k)$. Soit $r \leq n$ la dimension du groupe G . Supposons que pour tout x , l'application $g \mapsto \varphi_g(x)$ est de rang plein. Alors $\hat{x}_{k+1} = F(\hat{x}_k, u_k, y_k)$ est un pré-observateur G -invariant si et seulement si*

$$F(\hat{x}_k, u_k, y_k) = \varphi_{\gamma(\hat{x}_k)}^{-1}(\mathcal{L}(E(\hat{x}_k, u_k, y_k), I(\hat{x}_k, u_k)))$$

où E est une erreur de sortie invariante, I un ensemble complet d'invariants fondamentaux des variables \hat{x}_k et u_k et \mathcal{L} une fonction régulière qui vérifie pour tout \hat{x}_k, u_k ,

$$\mathcal{L}(0, I(\hat{x}_k, u_k)) = \varphi_{\gamma(\hat{x}_k)}(f(\hat{x}_k, u_k))$$

Démonstration. Si l'application γ est le repère mobile, la fonction $\varphi_{\gamma(\hat{x}_k)}(F(\hat{x}_k, u_k, y_k))$ est un invariant des variables (\hat{x}_k, u_k, y_k) et s'écrit donc $\mathcal{L}(E(\hat{x}_k, u_k, y_k), I(\hat{x}_k, u_k))$ où E est une erreur de sortie invariante, I un ensemble complet d'invariants fondamentaux des variables \hat{x}_k et u_k et \mathcal{L} une fonction régulière. \square

9.3 Erreur d'état invariante et convergence

On définit l'erreur de sortie invariante comme au chapitre 6

$$\eta(x_k, \hat{x}_k) = \varphi_{\gamma(x_k)}(\hat{x}_k) - \varphi_{\gamma(x_k)}(x_k)$$

avec γ défini par (6.8). Sous les hypothèse du théorème 7 on voit que les erreurs de sortie invariantes sont toutes de la forme :

$$\mathcal{F}(I(x_k, u_k), \eta(x_k, \hat{x}_k))$$

avec $I(x_k, u_k)$ ensemble complet d'invariants fondamentaux de l'action du groupe G sur $\mathcal{X} \times \mathcal{U}$. En effet $\varphi_{\gamma(x_k)}(x_k)$ est invariant et peut donc s'écrire comme fonction de I , et $(I(x_k, u_k), \varphi_{\gamma(x_k)}(\hat{x}_k))$ est un ensemble complet de $2n + m - r$ invariants de \hat{x}_k, x_k et u_k . Un résultat remarquable est que la dynamique d'erreur ne dépend de la trajectoire qu'à travers les invariants I :

Lemme 3. *Sous les hypothèses du théorème 7, la dynamique de l'erreur invariante $\eta(\hat{x}, x) = \varphi_{\gamma(x)}(\hat{x}) - \varphi_{\gamma(x)}(x)$ dépend seulement de η et des invariants scalaires de x et u :*

$$\eta_{k+1} = \eta(\hat{x}_{k+1}, x_{k+1}) = \Upsilon(\eta(\hat{x}_k, x_k), I(x_k, u_k)) = \Upsilon(\eta_k, I(x_k, u_k))$$

où Υ est une fonction régulière de ses arguments.

Démonstration. Soit σ la fonction définie par

$$\eta(\hat{x}_{k+1}, x_{k+1}) = \eta(F(\hat{x}_k, u_k, y_k), f(x_k, u_k)) = \sigma(\hat{x}_k, x_k, u_k)$$

Elle est invariante : $\sigma(\varphi_g(\hat{x}_k), \varphi_g(x_k), \psi_g(u_k)) = \sigma(\hat{x}_k, x_k, u_k)$. C'est donc une fonction invariante de $(\eta(\hat{x}_k, x_k), x_k, u_k)$ et donc c'est une fonction Υ de $(\eta(\hat{x}_k, x_k), I(x_k, u_k))$. \square

L'analyse de la dynamique de l'erreur invariante peut être un guide utile pour le design des gains de l'observateurs, comme cela a pu être le cas en temps continu au chapitre 5 et 7. Par exemple pour l'exemple de la voiture en formulation discrète mentionné en introduction de ce chapitre, le lemme 3 nous dit que la dynamique d'erreur est indépendante de la trajectoire x_k , et ne dépend que des u_k, v_k et de l'erreur elle-même.

Troisième partie

Extensions

La théorie a été développée dans le cadre où l'on a un système avec un espace d'état de dimension n , un espace des entrées de dimension m , et un groupe de symétries engendrées par des actions locales, avec $r \leq n$ paramètres. Le caractère local n'a pas été un problème pour traiter les exemples car les formules de la physique que nous avons manipulées étaient analytiques, et l'action du groupe globalement définie en fait. De nombreux problèmes de la physique présentent des symétries mais ne rentrent pas dans ce cadre. Nous pensons que la théorie précédente peut être un guide utile pour aborder d'autres types de problèmes. Dans cette partie, nous allons envisager quatre exemples empruntés à quatre domaines divers, pour lesquels notamment (et dans le désordre) la dimension du groupe est plus grande que n , le groupe est de dimension infinie, le groupe est de dimension finie mais l'état de dimension infinie.

Tout d'abord, on remarque que quand l'état est de dimension n , les paramètres de dimension m , et le groupe de dimension r avec $n < r \leq n + m$, on peut appliquer sans difficulté la théorie développée à la partie précédente en utilisant u dans la normalisation. On a écrit comme équation de normalisation $\varphi_g(x) = c$ où c est une constante arbitraire de dimension n dans l'image de la famille des φ_g . La solution est le repère mobile $g = \gamma(x)$. Dans le cas où $n < r \leq n + m$ on peut écrire comme équation de normalisation

$$(\varphi_g(x), \psi_g(u)) = c'$$

où c' est une constante de dimension $n + m$. Le repère mobile est alors une application $g = \gamma(x, u)$ qui dépend également des entrées u . On applique ensuite la théorie. Les champs invariants $w(x, u) \in TM_x$ s'écrivent alors comme des fonctions de x et de u . L'erreur d'état invariante $\eta(\hat{x}, x, u)$ dépend également de u . L'identification des paramètres d'un système quantique à deux états traité dans le chapitre suivant est une illustration de ce cas : l'état est un vecteur sur la sphère de Bloch \mathbb{S}^2 et les entrées u regroupent l'amplitude complexe d'un laser en quasi-résonance avec la fréquence de transition et les matrices de Pauli qui sont alors vues comme des paramètres connus. Le groupe de transformation est $SU(2)$: il est donc de dimension 3, alors que l'état (pour être précis la réduction de l'espace d'état à sa composante selon laquelle l'action du groupe est non triviale) est de dimension 2 et u de dimension $2 + 3 = 5$ (2 pour l'amplitude du laser et 3 pour celle des matrices de Pauli qui sont connues à une transformation de $SU(2)$ près). Ceci dit, l'hypothèse basique que nous avons utilisé dans la théorie comme quoi les difféomorphismes $(\varphi_g)_{g \in G}$ correspondant à l'action sur l'espace d'état sont de rang plein ne peut être vérifiée. Ici on procédera un peu différemment bien que $n < n + m$ et on définira un observateur globalement défini en suivant la méthode de construction usuelle, sans bénéficier directement des résultats liés à l'hypothèse de rang plein. Le problème peut être ramené à trouver les champs invariants et les invariants scalaires de l'action transitive de $SO(3)$ sur \mathbb{S}^2 .

Au chapitre 11, nous allons nous intéresser à des équations aux dérivées partielles. La construction d'observateurs pour des EDP est également appelée méthode du "nudging", et nous allons nous intéresser plus particulièrement aux équations de Saint-Venant dans

le cadre du modèle Shallow Water. La version de [29] du modèle des eaux peu profondes (Shallow Water) est très utilisée en océanographie pour l'assimilation de données altimétriques de la surface océanique et provenant de satellites pour l'estimation des courants marins en profondeur. Il s'agit d'équations aux dérivées partielles qui ne dépendent pas du choix que l'on fait du repère pour les écrire (origine et orientation). Du coup la formulation des propriétés d'invariances prend une forme différente, puisque l'état est constitué de fonctions des coordonnées d'espace, et est donc de dimension infinie, et l'invariance consiste en des opérations sur les coordonnées d'espace. Nous allons exprimer l'invariance, et donner la forme des observateurs invariants, en recherchant tous les termes de correction qui préservent l'invariance par rotation-translation. On établit la convergence de l'observateur pour le cas simple des équations de Saint-Venant linéarisées.

Au chapitre 12, on reprend l'exemple de navigation inertielle du chapitre 7 lorsqu'on utilise des caméras CCD pour estimer la vitesse. On ne considère pas le problème de la navigation inertielle dans sa globalité mais on s'interroge sur la possibilité d'écrire un observateur qui donnerait une estimation de la vitesse d'un objet volant à partir de vidéos de l'environnement considéré comme fixe, et filmé depuis l'objet volant. On écrit un aussi un observateur invariant par rotation pour recalibrer les biais des gyroscopes quand l'environnement est lointain.

Finalement, au chapitre 13 on reprendra le problème posé dans [3]. Les auteurs ont construit un observateur intrinsèque pour estimer la vitesse des systèmes mécaniques lagrangiens dont on mesure toutes les positions. On cherche à faire un observateur intrinsèque différent de celui introduit par les auteurs pour ce problème, qui est du type observateur réduit. Nous montrons alors que, contrairement au cas de l'observateur intrinsèque complet proposé dans [3], le fait d'avoir une courbure négative renforce la stabilité de cet observateur réduit.

Chapitre 10

Identification de paramètres pour un système quantique

Ce chapitre a été complété et a fait l'objet d'une soumission pour publication [13]. Dans ce chapitre on s'intéresse à un système quantique dont on cherche à identifier certains paramètres. Schématiquement, on éclaire le système avec un laser sonde dont la fréquence est proche de la fréquence de transition entre l'état fondamental et un état excité très instable. Cela entraîne pour une partie de la population ainsi éclairée une transition vers l'état excité suivie rapidement d'une descente vers l'état fondamental avec émission spontanée de photons. On mesure le nombre de photons ainsi émis qui est directement proportionnel à la population de l'état fondamental. Par ailleurs on dispose d'un autre laser dont la fréquence est proche sans être pour autant exactement identique à celle d'une transition vers un autre état excité mais de durée de vie beaucoup plus longue. Ainsi la dynamique entre l'état fondamental et ce second état excité obéit à une équation de Schrödinger (en première approximation car le dispositif de mesure décrit ci-dessus entraîne une intrication avec l'environnement). On suppose à ce niveau que certains paramètres caractéristiques sont mal connus : Δ la différence entre la fréquence du second laser et celle de cette transition atomique ; μ le moment di-polaire qui caractérise le couplage entre ce système à deux niveaux et le second laser. Le but est d'identifier Δ et μ à partir de la mesure de la population de l'état fondamental, mesure obtenue grâce au premier laser sonde et au photo-détecteur. D'un point de vue pratique on peut voir le système physique considéré comme un ensemble de systèmes indentiquement préparés, comme un gaz monoatomique dilué, qui présente l'avantage de n'être constitué que de particules identiques et inter-agissant très faiblement. Schématiquement, l'expérience de base sous-jacente ne prend pas beaucoup plus de place en surface qu'un bureau. Le gaz est enfermé dans un tube transparent et éclairé par les deux lasers. On utilise pour la mesure un photo-détecteur ou plutôt une photo-diode dans l'alignement du laser sonde qui mesure l'absorption (proportionnelle au nombre d'atomes dans l'état fondamental) par le gaz de la lumière issue du laser.

On propose dans ce chapitre un observateur asymptotique qui estime en même temps la

fonction d'onde (la matrice densité en fait) et les deux paramètres Δ et μ . Nous proposons ici une extension de l'observateur décrit dans [43] où seul μ était estimé et nous montrons le rôle important joué par les symétries (invariance des équations par rapport au changement de repère décrit par un élément de $U(2)$, en particulier de $SU(2)$). On trouvera dans [20] un excellent cours de base de mécanique quantique où les notions physiques utiles à ce chapitre sont exposées (voir aussi [28] pour un cours plus avancé).

10.1 Dynamique et invariances

10.1.1 Dynamique

L'équation de Schrödinger pour le système s'écrit :

$$i\frac{d}{dt}\Psi = \left(\frac{\Delta}{2}\sigma_z + \frac{u\mu}{2}\sigma_x\right)\Psi, \quad \Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} \in \mathbb{C}^2$$

où l'on note

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

les matrices de Pauli, où Δ est la différence entre la fréquence de transition atomique de l'état fondamental à l'état excité et la fréquence du laser, $\mu > 0$ est le moment dipolaire et $u \in \mathbb{R}$ est l'amplitude lentement variable du laser. On a les formules utiles $\sigma_x^2 = 1$; $\sigma_x\sigma_y = i\sigma_z$ (avec permutations circulaires). La sortie est la différence de population

$$y = \langle \sigma_z \Psi, \Psi \rangle = |\Psi_1|^2 - |\Psi_2|^2 = 2|\Psi_1|^2 - 1$$

car $|\Psi_1|^2 + |\Psi_2|^2 = 1$ (conservation de la probabilité). Pour une justification (dans le cadre des perturbations singulières) de ce modèle avec la loi de sortie (hypothèse de mesure faible) nous renvoyons à [42].

Il est pratique d'écrire le système dans le langage de la matrices densité. Si l'on note la matrice complexe $\rho = \Psi\Psi^\dagger$ associée au projecteur sur l'état Ψ , on a notamment les propriétés $Tr(\rho) = \Psi_1^\dagger\Psi_1 + \Psi_2^\dagger\Psi_2 = 1$ et $\rho^2 = \Psi\Psi^\dagger\Psi\Psi^\dagger = \rho$. On a

$$\begin{aligned} \frac{d}{dt}\rho &= -i \left[\frac{\Delta}{2}\sigma_z + \frac{u\mu}{2}\sigma_x, \rho \right] \\ y &= Tr(\sigma_z\rho) \end{aligned} \tag{10.1}$$

où $[,]$ est le commutateur. On suppose ici que l'amplitude du laser varie lentement devant la pulsation de Rabi $|u\mu|$: $|\dot{u}| \ll |u\mu||u|$. On suppose aussi que les fréquences de transition atomique et du laser sont proches, c'est-à-dire à une distance inférieure à la largeur atomique (sinon le système ne réagit pas et l'approximation résonnante sur laquelle repose la dynamique précédente n'est plus valable), on a donc $|\Delta| \ll |u\mu|$.

10.1.2 Invariances

Le système est invariant sous l'action de $SU(2)$ (qui est isomorphe au groupe des quaternions de norme 1 que l'on a noté \mathbb{H}_1 au chapitre 7). Soit $U \in SU(2)$, l'action de groupe s'écrit sur l'état $\varpi = U\rho U^\dagger$, et sur les paramètres $\zeta_x = U\sigma_x U^\dagger$, $\zeta_y = U\sigma_y U^\dagger$, $\zeta_z = U\sigma_z U^\dagger$. On remarque que cette invariance correspond à l'invariance par changement de base $\Psi \mapsto U\Psi$ sur la fonction d'onde. La dynamique s'écrit de la même façon à l'aide des nouvelles variables

$$\begin{aligned} \frac{d}{dt}\varpi &= -i\left[\frac{\Delta}{2}\zeta_z + \frac{u\mu}{2}\zeta_x, \varpi\right] \\ y &= \text{Tr}(\zeta_z\varpi) \end{aligned}$$

et $\zeta_x, \zeta_y, \zeta_z$ sont des matrices de Pauli car elles obéissent aux mêmes relations de commutations. Le système est donc invariant sous l'action de $SU(2)$ selon la définition 2. Mais ici l'état est représenté par ρ et les entrées sont associées à l'amplitude du laser u et aussi les matrices de Pauli qui sont des paramètres constants.

10.2 Un observateur invariant

On choisit comme observateur pour le système étendu (avec les paramètres)

$$\frac{d}{dt}\hat{\rho} = -i\left[\frac{\hat{\Delta}}{2}\sigma_z + \frac{u\hat{\mu}}{2}\sigma_x, \hat{\rho}\right] \quad (10.2)$$

$$- K_\rho(\text{Tr}(\sigma_z\hat{\rho}) - y) (\sigma_z\hat{\rho} + \hat{\rho}\sigma_z - 2\text{Tr}(\sigma_z\hat{\rho})\hat{\rho}) \quad (10.3)$$

$$\frac{d}{dt}\hat{\mu} = -uK_\mu\text{Tr}(\sigma_y\hat{\rho}) (\text{Tr}(\sigma_z\hat{\rho}) - y) \quad (10.4)$$

$$\frac{d}{dt}\hat{\Delta} = -uK_\Delta\text{Tr}(\sigma_x\hat{\rho}) (\text{Tr}(\sigma_z\hat{\rho}) - y) \quad (10.5)$$

où K_ρ, K_μ et K_Δ sont des gains positifs. C'est bien un observateur invariant de la forme présentée au théorème 2 page 57 car la première ligne peut se ré-écrire sous la forme (on montre cela à la section suivante)

$$\begin{aligned} \frac{d}{dt}\hat{\rho} &= -i\left[\frac{\hat{\Delta}}{2}\sigma_z + \frac{u\hat{\mu}}{2}\sigma_x, \hat{\rho}\right] \\ &\quad - K_\rho(\text{Tr}(\sigma_z\hat{\rho}) - y) (\text{Tr}(\sigma_y\hat{\rho})[\sigma_x, \hat{\rho}] - \text{Tr}(\sigma_x\hat{\rho})[\sigma_y, \hat{\rho}]) \end{aligned}$$

La forme de l'observateur où l'on met Δ et $\hat{\Delta}$ à 0 est directement issue de [43].

On suppose que u est constant, que $\hat{\Delta} = \epsilon\hat{d}$ et en vue d'appliquer les techniques perturbatives habituelles pour ce type de système physique on fait le choix de gains suivant

$$\hat{K}_\rho = 4k_\rho\epsilon|u|\mu, \quad K_\mu = 2k_\mu\epsilon^2\mu^2, \quad K_\Delta = 2k_\Delta\epsilon^2|u|\mu^2$$

avec $\epsilon > 0$ petit ($\epsilon \ll 1$), et $k_\rho, k_\mu, k_\Delta \sim 1$.

10.2.1 Interprétation géométrique avec la sphère de Bloch

La sphère de Bloch est une représentation géométrique d'un système quantique pur à deux états. Une propriété importante de la matrice densité ρ est qu'elle peut se décomposer dans la forme

$$\rho = \frac{1 + X\sigma_x + Y\sigma_y + Z\sigma_z}{2}, \quad \text{avec } \zeta = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \in \mathbb{S}^2$$

où 1 la matrice identité 2×2 . On a $\text{Tr}(\sigma_x \rho) = X$, $\text{Tr}(\sigma_y \rho) = Y$ et $\text{Tr}(\sigma_z \rho) = Z$. La coordonnée Z dans cette section est la sortie y . L'opération de commutation $-i[\sigma_x, \rho]$ correspond au produit vectoriel $(1, 0, 0)^T$ avec $(X, Y, Z)^T$ (les permutations circulaires permettent de compléter les correspondances). La dynamique de ζ (voir eq (10.1)) est

$$\frac{d}{dt}\zeta = \left(\frac{u\mu}{2}, 0, \frac{\Delta}{2}\right)^T \wedge \zeta \quad (10.6)$$

Les symétries associées à l'invariance par changement de base correspondent dans représentation de la sphère de Bloch à l'invariance des équations sous l'action de $\text{SO}(3)$, i.e indépendantes du choix du repère orthonormé de la sphère dans lequel on écrit les équations. D'un point de vue physique faire un observateur invariant consiste à préserver le fait que le système est pur, et donc à obliger ζ à évoluer sur \mathbb{S}^2 .

Le terme de correction (10.3) s'écrit $-K_\rho(\hat{Z} - Z)[- \hat{X}\hat{Z}\sigma_x - \hat{Y}\hat{Z}\sigma_y + (1 - \hat{Z}^2)\sigma_z]$ et correspond dans la sphère de Bloch à $-K_\rho(\hat{Z} - Z)(\hat{Y}, -\hat{X}, 0)^T \wedge \hat{\zeta}$, avec $\hat{\zeta} = (\hat{X}, \hat{Y}, \hat{Z})^T$. Cela veut dire que (10.2)-(10.3) peut être écrit

$$\begin{aligned} \frac{d}{dt}\hat{\rho} = -i & \left[\frac{\hat{\Delta}}{2}\sigma_z + \frac{u\hat{\mu}}{2}\sigma_x, \hat{\rho} \right] \\ & + iK_\rho(\text{Tr}(\sigma_z \hat{\rho}) - y) (\text{Tr}(\sigma_y \hat{\rho})[\sigma_x, \hat{\rho}] - \text{Tr}(\sigma_x \hat{\rho})[\sigma_y, \hat{\rho}]) \end{aligned}$$

et alors l'observateur (10.2)-(10.3) pour le système réduit (10.1) est bien un observateur invariant.

Heuristique du choix des termes de correction invariants

Ici on se concentre sur l'interprétation géométrique des gains. Des preuves plus rigoureuses sont données dans la suite. Soit $e_z = (0, 0, 1)^T$. On a prouvé que (10.2)-(10.3) s'écrit dans la sphère de Bloch :

$$\frac{d}{dt}\hat{\zeta} = \frac{1}{2} \begin{pmatrix} u\hat{\mu} \\ 0 \\ \hat{\Delta} \end{pmatrix} \times \hat{\zeta} - K_\rho(\hat{Z} - Z)(\hat{\zeta} \times e_z) \times \hat{\zeta} \quad (10.7)$$

Donc le terme de correction (10.3) est simplement "reflété" dans la sphère de Bloch par la projection du gradient de $-\frac{K_\rho}{2}(\hat{Z} - Z)^2$ sur l'espace tangent $TS^2|_{\hat{\zeta}}$. C'est un vecteur qui pointe toujours vers le nord (i.e, $Z = 1$) si $\hat{Z} < Z$ et vers le sud (i.e, $Z = -1$) sinon.

K_ρ est bien plus grand que les gains K_μ et K_Δ (le ratio est d'ordre ϵ). Supposons que $\hat{\zeta} - \zeta \approx 0$. Utilisant (10.6) et (10.7), on a $\frac{d}{dt}(\hat{Z} - Z) \approx \frac{u}{2}(\hat{\mu} - \mu)\hat{Y}$. Donc si $\hat{\mu} > \mu$, la différence $\hat{Z} - Z$ (proche de zéro) croît ssi $\hat{Y} > 0$. Cela explique le choix de (8.18) qui s'écrit dans la représentation de la sphère de Bloch $\frac{d}{dt}\hat{\mu} = -uK_\mu\hat{Y}(\hat{Z} - Z)$. Le choix de (10.5) qui s'écrit $\frac{d}{dt}\hat{\Delta} = -uK_\Delta\hat{X}(\hat{Z} - Z)$ s'explique de la même manière en supposant que $\hat{\mu} - \mu \approx 0$ et en remarquant qu'alors $\frac{d^2}{dt^2}(\hat{Z} - Z) \approx \frac{u}{4}\mu(\hat{\Delta} - \Delta)\hat{X}$.

10.2.2 Diverses dimension et normalisation

La dimension de la composante de l'état sur laquelle le group agit non-trivialement est $n = 2$, puisque ρ vit sur la sphère de Bloch. L'action sur les paramètres μ et Δ est triviale (c'est-à-dire laisse les paramètres inchangés). La dimension du groupe est $r = 3$. On a donc $r > n$. L'hypothèse que l'action du groupe sur l'espace d'état est de rang plein faite au chapitre 6 ne peut être vérifiée. Cependant, le nombre m d'entrées est suffisant pour qu'une normalisation (6.8) soit possible au travers de l'action du groupe sur l'espace d'état et celui des entrées (les matrices de Pauli). La sortie est compatible au sens de la définition 3 du chapitre 6 puisque elle est inchangée par la transformation. Nous n'allons pas utiliser les équations de normalisation et nous allons construire un observateur invariant (globalement défini) pour ce système, en suivant les étapes de la méthode de la section 6 comme un guide.

Erreur de sortie invariante

La sortie $y = Tr(\sigma_z\rho)$ est un invariant scalaire. L'erreur invariante la plus simple est donc donnée $\hat{y} - y = Tr(\sigma_z(\hat{\rho} - \rho))$. On peut en trouver naturellement d'autres.

Champs invariants

L'espace d'état sur lequel l'action est non-triviale est de dimension 2, et comme le groupe est de dimension 3, il suffit localement de deux champs invariants indépendants pour faire un repère invariant. Ceci dit l'utilisation de 3 champs de vecteurs invariants permet une paramétrisation globale par des champs de vecteurs invariants. Nous prendrons donc comme champs de vecteurs invariants en nous inspirant directement de la forme de la dynamique $[\sigma_x, \cdot]$, $[\sigma_y, \cdot]$ et $[\sigma_z, \cdot]$.

Invariants scalaires

On cherche un ensemble d'invariants scalaires. Le groupe $SU(2)$ est de dimension 3 et agit sur $\hat{\rho}$, σ_x , σ_y et σ_z . On retiendra donc les trois invariants simples $Tr(\sigma_x\hat{\rho})$, $Tr(\sigma_y\hat{\rho})$, $Tr(\sigma_z\hat{\rho})$,

dont chacun s'exprime en fonction des deux autres, à une incertitude sur le signe près (ce signe est constant localement mais varie globalement). En effet on a $Tr^2(\sigma_x \hat{\rho}) + Tr^2(\sigma_y \hat{\rho}) + Tr^2(\sigma_z \hat{\rho}) = 1$.

10.3 Convergence

Dans cette section nous allons établir qu'il existe une fonction de Lyapunov pour le système simplifié par moyennisation des perturbations (périodiques). En vue d'appliquer l'approximation des champs tournants on se place dans le repère d'interaction, c'est-à-dire qu'on fait le changement de variable qui dépend du temps :

$$\rho = e^{-i\frac{u\mu t\sigma_x}{2}} \xi e^{i\frac{u\mu t\sigma_x}{2}}, \quad \hat{\rho} = e^{-i\frac{u\mu t\sigma_x}{2}} \hat{\xi} e^{i\frac{u\mu t\sigma_x}{2}}.$$

On a alors

$$\frac{d}{dt}\xi = \left[i\frac{u\mu\sigma_x}{2}, \xi \right] + e^{i\frac{u\mu t\sigma_x}{2}} \left(\frac{d}{dt}\rho \right) e^{-i\frac{u\mu t\sigma_x}{2}}$$

et la formule analogue pour $\frac{d}{dt}\hat{\xi}$ ce qui donne

$$\frac{d}{dt}\xi = -i \left[\frac{\Delta}{2} e^{iu\mu t\sigma_x} \sigma_z, \xi \right]$$

On remarque que pour chaque k entier $\sigma_z \sigma_x^{2k} = \sigma_x^{2k} \sigma_z$ et $\sigma_z \sigma_x^{2k+1} = -\sigma_x^{2k+1} \sigma_z$ et cela entraîne par continuité de la multiplication matricielle

$$\sigma_z e^{i\frac{u\mu t\sigma_x}{2}} = e^{-i\frac{u\mu t\sigma_x}{2}} \sigma_z$$

Ce raisonnement est valable aussi en remplaçant σ_z par σ_y , et l'observateur s'écrit dans le repère d'interaction :

$$\begin{aligned} \frac{d}{dt}\hat{\xi} &= -i \left[\frac{\hat{\Delta}}{2} e^{iu\mu t\sigma_x} \sigma_z + \frac{u(\hat{\mu} - \mu)}{2} \sigma_x, \hat{\xi} \right] \\ &\quad - K_\rho \text{Tr} \left(e^{iu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right) \left(e^{iu\mu t\sigma_x} \sigma_z \hat{\xi} + \hat{\xi} e^{iu\mu t\sigma_x} \sigma_z - 2\text{Tr} \left(e^{iu\mu t\sigma_x} \sigma_z \hat{\xi} \right) \hat{\xi} \right) \\ \frac{d}{dt}\hat{\mu} &= -uK_\mu \text{Tr} \left(e^{iu\mu t\sigma_x} \sigma_y \hat{\xi} \right) \text{Tr} \left(e^{iu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right) \\ \frac{d}{dt}\hat{\Delta} &= -uK_\Delta \text{Tr} \left(\sigma_x \hat{\xi} \right) \text{Tr} \left(e^{iu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right) \end{aligned}$$

L'approximation séculaire au premier ordre

Egalement appelée approximation quasi-résonnante, elle consiste à supposer que puisque $u\mu$ est une fréquence élevée, par moyennisation des perturbations, on peut négliger les

termes qui "tournent" dans le plan complexe à la fréquence $u\mu$ et $2u\mu$ sur les échelles de temps caractéristiques de l'évolution du vecteur d'état. Le mouvement s'écrit comme un déplacement le long d'une trajectoire non-perturbée accompagnée de petites oscillations de moyenne nulle et de période $2\pi/(u\mu)$ autour de cette trajectoire. On s'intéresse à la trajectoire non-perturbée. On utilise la formule pour $a > 0$

$$e^{ia\sigma_x}\sigma_z = \sigma_z + a\sigma_y - \frac{a^2}{2}\sigma_z - \frac{a^3}{6}\sigma_y + \dots = \cos(a)\sigma_z + \sin(a)\sigma_y \quad (10.8)$$

et son analogue

$$e^{ia\sigma_x}\sigma_y = \cos(a)\sigma_y - \sin(a)\sigma_z \quad (10.9)$$

Ces deux formules dérivent directement de la formule

$$e^{ia\sigma_x} = \cos(a) + i\sin(a)\sigma_x$$

très naturelle car $(i\sigma_x)^2 = -1 \in \mathbb{C}^{2 \times 2}$. En écrivant les cosinus et sinus comme combinaisons d'exponentielles et en utilisant alors des formules d'interférences du type :

$$\begin{aligned} & \text{Tr} \left(e^{iu\mu t\sigma_x}\sigma_z(\hat{\xi} - \xi) \right) e^{iu\mu t\sigma_x}\sigma_z\hat{\xi} \\ &= \frac{1}{4} \left(2\text{Tr} \left(\sigma_z(\hat{\xi} - \xi) \right) \sigma_z\hat{\xi} + 0 + 2\text{Tr} \left(\sigma_y(\hat{\xi} - \xi) \right) \sigma_y\hat{\xi} + 0 \right) + \text{termes oscillants} \\ & \quad \text{de moyenne nulle} \end{aligned}$$

on obtient pour le système, et pour l'observateur, après moyennisation, le système autonome suivant :

$$\begin{aligned} \frac{d}{dt}\xi &= 0 \\ \frac{d}{dt}\hat{\xi} &= -i \left[\frac{u(\hat{\mu} - \mu)}{2}\sigma_x, \hat{\xi} \right] \\ & \quad - \frac{K_\rho}{2} \text{Tr} \left(\sigma_y(\hat{\xi} - \xi) \right) \left(\sigma_y\hat{\xi} + \hat{\xi}\sigma_y - 2\text{Tr} \left(\sigma_y\hat{\xi} \right) \hat{\xi} \right) \\ & \quad - \frac{K_\rho}{2} \text{Tr} \left(\sigma_z(\hat{\xi} - \xi) \right) \left(\sigma_z\hat{\xi} + \hat{\xi}\sigma_z - 2\text{Tr} \left(\sigma_z\hat{\xi} \right) \hat{\xi} \right) \\ \frac{d}{dt}\hat{\mu} &= -\frac{uK_\mu}{2} \left(\text{Tr} \left(\sigma_y\hat{\xi} \right) \text{Tr} \left(\sigma_z(\hat{\xi} - \xi) \right) - \text{Tr} \left(\sigma_z\hat{\xi} \right) \text{Tr} \left(\sigma_y(\hat{\xi} - \xi) \right) \right) \\ \frac{d}{dt}\hat{\Delta} &= 0 \end{aligned}$$

On en déduit la convergence locale de $\hat{\xi}$ et $\hat{\mu}$ pour tous $K_\rho, K_\mu > 0$ en choisissant comme fonction de Lyapounov (c.f. [43])

$$V = \frac{1}{2} \text{Tr}^2(\sigma_y(\hat{\xi} - \xi)) + \frac{1}{2} \text{Tr}^2(\sigma_z(\hat{\xi} - \xi)) + \frac{1}{K_\mu} (\hat{\mu} - \mu)^2$$

Afin de calculer $\frac{d}{dt}V$ on remarque d'abord que $\text{Tr}\left(\frac{d}{dt}\sigma_y(\hat{\xi} - \xi)\right) = -\frac{u(\hat{\mu}-\mu)}{2}\text{Tr}\left(2\sigma_z\hat{\xi}\right) + \dots$ et $\text{Tr}\left(\frac{d}{dt}\sigma_z(\hat{\xi} - \xi)\right) = \frac{u(\hat{\mu}-\mu)}{2}\text{Tr}\left(2\sigma_y\hat{\xi}\right) + \dots$ et donc les termes dont u est en facteur disparaissent et il ne reste que des termes facteurs de K_ρ . On note $\delta\xi = \hat{\xi} - \xi$ et on calcule alors la quantité

$$\text{Tr}(\sigma_y\delta\xi)\text{Tr}(\sigma_y\dot{\delta\xi}) = 2(\text{Tr}^2(\sigma_y\delta\xi)(1 - \text{Tr}^2(\sigma_y\hat{\xi})) - \text{Tr}(\sigma_y\delta\xi)\text{Tr}(\sigma_z\delta\xi)\text{Tr}(\sigma_y\hat{\xi})\text{Tr}(\sigma_z\hat{\xi}))$$

et l'on a également la même formule si l'on intervertit σ_y et σ_z . $\frac{d}{dt}V$ est la somme de ces deux termes multipliée par $-K_\rho/2$. On note $\text{Tr}(\sigma_x\hat{\xi}) = \hat{x}$ et $\text{Tr}(\sigma_x\xi) = x$ et similairement avec y et z . On a la relation immédiate sur la sphère de Bloch $\hat{x}^2 + \hat{y}^2 + \hat{z}^2 = 1$ grâce au fait que l'observateur invariant préserve les propriétés géométriques de ρ et donc de ξ . On peut écrire alors

$$\begin{aligned} -\frac{1}{K_\rho}\frac{d}{dt}V &= (\hat{y} - y)^2(1 - \hat{y}^2) + (\hat{z} - z)^2(1 - \hat{z}^2) - 2\hat{y}\hat{z}(\hat{y} - y)(\hat{z} - z) \\ &= (\hat{y} - y)^2(\hat{x}^2 + \hat{z}^2) + (\hat{z} - z)^2(\hat{x}^2 + \hat{y}^2) - 2\hat{y}\hat{z}(\hat{y} - y)(\hat{z} - z) \\ &\geq ((\hat{y} - y)\hat{z} - (\hat{z} - z)\hat{y})^2 \geq 0. \end{aligned}$$

On renvoie à [43] pour la fin de la preuve de la convergence locale via l'invariance de Lasalle.

Approximation séculaire au second ordre

Puisque les termes séculaires au premier ordre disparaissent dans le calcul de la dynamique de ξ et $\hat{\Delta}$, on ne peut pas se contenter de ce développement au premier ordre pour montrer la convergence (dans cette approximation $\hat{\Delta}$ est constant et ne peut converger). Pour pousser le développement aux termes d'ordre supérieurs, il faut faire un calcul à l'ordre deux pour ξ et $\hat{\Delta}$ uniquement. On reprend la méthode de Kapitsa décrite dans [33] pour l'étude du mouvement d'une particule dans un champ à oscillations rapides. On a supposé que $u\mu = \nu$ est grand. ξ vérifie une équation différentielle avec une excitation périodique haute fréquence du type $\frac{d}{dt}\xi = f(\xi, \nu t)$. On a vu qu'au premier ordre et en moyenne ξ est constant. L'intégration des termes oscillants haute fréquence conduit à des termes oscillants petits : on cherche une solution de la forme

$$\xi = \zeta + \frac{g_1(\zeta, t)}{\nu} + \frac{g_2(\zeta, t)}{\nu^2} \dots$$

où les g_i sont de l'ordre de ζ et où ζ est la moyenne de ξ sur une période, et l'on a vu que $\dot{\zeta} = 0 + O(1/\nu)$. Nous allons pousser au premier ordre en $\frac{1}{\nu}$ et donc négliger les effets d'ordre 2. On a d'une part en s'arrêtant aux termes d'ordre 0 :

$$\frac{d}{dt}\xi = 0 + \frac{g'_1(\zeta, t)}{\nu}$$

où g'_1 désigne la dérivée de g_1 par rapport à sa seconde variable. D'autre part en utilisant la dynamique de ξ on a en s'arrêtant à l'ordre 1 :

$$\frac{d}{dt}\xi = -\imath \left[\frac{\Delta}{2} e^{\imath\nu t\sigma_x} \sigma_z, \zeta + \frac{g_1(\zeta, t)}{\nu} \right]$$

et on obtient la relation $g'_1(\zeta, t) = -\nu \imath \left[\frac{\Delta}{2} e^{\imath\nu t\sigma_x} \sigma_z, \zeta \right] + \dots$ où l'on n'écrit pas les autres termes qui sont petits devant ν . Intégrée par rapport au temps t la dernière équation donne :

$$g_1(\zeta, t) = \imath \left[\frac{\Delta}{2} e^{\imath\nu t\sigma_x} \sigma_y, \xi \right] + \frac{1}{\nu} (\dots)$$

A partir de cette relation on peut ré-écrire la dynamique de ξ en lui injectant cette dernière relation.

$$\dot{\xi} = -\imath \left[\frac{\Delta}{2} e^{\imath\nu t\sigma_x} \sigma_z, \zeta + g_1(\zeta, t)/\nu \right] = -\imath \left[\frac{\Delta}{2} e^{\imath\nu t\sigma_x} \sigma_z, \zeta \right] + \frac{\Delta^2}{4\nu} [e^{\imath\nu t\sigma_x} \sigma_z, [e^{\imath\nu t\sigma_x} \sigma_y, \xi]] + \dots$$

Calculons la moyenne temporelle (sur une période) de cette relation et ne gardons que les termes séculaires. On utilise les formules (10.8) et (10.9) et l'identité de Jacobi pour calculer le double crochet. On développe les cosinus et sinus en exponentielles complexes. On remarque que les termes en $(\cos(\nu t) \sin(\nu t))$ oscillent à la fréquence 2ν et sont donc de moyenne nulle. On obtient alors comme terme séculaire

$$\frac{d}{dt}\xi = -\imath \frac{\Delta^2}{2u\mu} [\sigma_x, \xi] + \dots$$

et on retrouve ainsi le terme $\frac{\Delta^2}{2u\mu}$ qui correspond classiquement à la correction dite de Bloch-Siegert (Bloch-Siegert shift).

On est également obligé de développer au premier ordre en $\frac{1}{\nu}$ la variable $\hat{\Delta}$ qui, comme ξ , est constante en moyenne. Sa dynamique s'écrit

$$\frac{d}{dt}\hat{\Delta} = -uK_{\Delta} \text{Tr} \left(\sigma_x \hat{\xi} \right) \text{Tr} \left(e^{\nu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right)$$

ξ et $\hat{\xi}$ sont solutions d'équations différentielles dont la dépendance en temps est liée à des termes source oscillants à la fréquence ν . Comme on ne va que jusqu'au second ordre on va écrire la dynamique comme une somme de deux termes. 1. $\text{Tr} \left(\sigma_x \hat{\xi} \right)$ est supposé non-oscillant et on calcule les termes séculaires du produit $\text{Tr} \left(e^{\nu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right)$. 2. on suppose $\sigma_z (\hat{\xi} - \xi)$ non oscillant et on calcule les termes séculaires de $\text{Tr} \left(\sigma_x \hat{\xi} \right) \text{Tr} \left(e^{\nu\mu t\sigma_x} \right)$. Seuls les termes de fréquence ν peuvent avoir un effet séculaire, ceux de fréquence 2ν

ne rencontreront pas de terme pouvant annuler leur oscillation (par d'effet séculaire). On considère alors la dynamique de $\hat{\xi}$

$$\begin{aligned} \frac{d}{dt}\hat{\xi} = -i \left[\frac{\hat{\Delta}}{2} e^{i\nu\mu t\sigma_x} \sigma_z + \frac{u(\hat{\mu} - \mu)}{2} \sigma_x, \hat{\xi} \right] - K_\rho \text{Tr} \left(e^{i\nu\mu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right) \\ \dots \left(e^{i\nu\mu t\sigma_x} \sigma_z \hat{\xi} + \hat{\xi} e^{i\nu\mu t\sigma_x} \sigma_z - 2\text{Tr} \left(e^{i\nu\mu t\sigma_x} \sigma_z \hat{\xi} \right) \hat{\xi} \right) \end{aligned}$$

On va appliquer la même méthode que pour le calcul de ξ , mais l'on ne cherche que la partie de $\hat{\xi}$ oscillant à une fréquence ν . Le terme de correction en facteur de K_ρ , une fois qu'on a développé les deux termes sources comme dans le calcul au premier ordre, ne donne que des termes non-oscillants et des termes de fréquence 2ν . La partie de $\hat{\xi}$ qui oscille à la fréquence ν ne peut provenir que de l'intégration du terme $-i \left[\frac{\hat{\Delta}}{2} e^{i\nu\mu t\sigma_x} \sigma_z + \frac{u(\hat{\mu} - \mu)}{2} \sigma_x, \hat{\xi} \right]$ et le calcul est exactement le même que pour ξ . On a donc

$$\hat{\xi} = \hat{\zeta} + \frac{i\hat{\Delta}}{2u\mu} \left[e^{i\nu\mu t\sigma_x} \sigma_y, \hat{\xi} \right] + \dots$$

où $\hat{\zeta}$ est solution de l'équation moyennisée à l'ordre 0 vérifiée par $\hat{\xi}$, et les termes que l'on a pas écrits sont des termes de fréquence 2ν et les termes d'ordre supérieur en $\frac{1}{\nu}$. Dans le calcul de la dynamique de $\hat{\Delta}$ il suffit donc de remplacer alternativement $\hat{\xi}$ puis $\hat{\xi} - \xi$ par leurs parties oscillant à la fréquence ν et de ne retenir que les termes séculaires. On a d'une part

$$\begin{aligned} i\frac{\hat{\Delta}}{2\nu} e^{i\nu t\sigma_x} \sigma_z [e^{i\nu t\sigma_x} \sigma_y, \hat{\xi}] = i\frac{\hat{\Delta}}{2\nu} (\sigma_z [\sigma_y, \hat{\xi}] - i\sigma_y [\sigma_y, \hat{\xi}]) + \text{termes oscillants} \\ \text{de moyenne nulle} \\ \text{donc } \text{Tr} \left(e^{i\nu t\sigma_x} \sigma_z (\hat{\xi} - \xi) \right) = \frac{1}{\nu} (\hat{\Delta} \text{Tr} (\sigma_x \hat{\xi}) - \Delta \text{Tr} (\sigma_x \xi)) + \text{termes oscillants} \\ \text{de moyenne nulle} \end{aligned}$$

Cela permet le calcul du premier membre du résultat final (10.10). D'autre part pour remplacer $\text{Tr} (\sigma_x \hat{\xi})$ par sa partie qui oscille à la fréquence ν on écrit que

$$\begin{aligned} \text{Tr} \left(\sigma_x [e^{i\nu t\sigma_x} \sigma_y, \hat{\xi}] \right) &= \text{Tr} \left(\cos \nu t \sigma_x [\sigma_y, \hat{\xi}] - \sin \nu t \sigma_x [\sigma_z, \hat{\xi}] \right) \\ &= 2i \text{Tr} \left(\cos \nu t \sigma_z \hat{\xi} + \sin \nu t \sigma_y \hat{\xi} \right) \end{aligned}$$

Puis après multiplication et élimination des termes oscillants on trouve le second membre du résultat final (10.10). On a finalement

$$\begin{aligned} \frac{d}{dt} \hat{\Delta} = -\frac{K_\Delta}{\mu} \left(\text{Tr} (\sigma_x \hat{\xi})^2 \hat{\Delta} - \text{Tr} (\sigma_x \hat{\xi}) \text{Tr} (\sigma_x \xi) \Delta \right) \\ + \frac{K_\Delta \hat{\Delta}}{2\mu} \left(\text{Tr} (\sigma_y \hat{\xi}) \text{Tr} (\sigma_y (\hat{\xi} - \xi)) - \text{Tr} (\sigma_z \hat{\xi}) \text{Tr} (\sigma_z (\hat{\xi} - \xi)) \right) \end{aligned} \quad (10.10)$$

Et on obtient pour le système et l'observateur le système triangulaire suivant qui est localement convergent :

$$\begin{aligned}
\frac{d}{dt}\xi &\stackrel{\text{ordre } 1}{=} -i\frac{\Delta^2}{2u\mu}[\sigma_x, \xi] \\
\frac{d}{dt}\hat{\xi} &\stackrel{\text{ordre } 0}{=} -i\left[\frac{u(\hat{\mu} - \mu)}{2}\sigma_x, \hat{\xi}\right] \\
&\quad - \frac{K_\rho}{2}\text{Tr}\left(\sigma_y(\hat{\xi} - \xi)\right)\left(\sigma_y\hat{\xi} + \hat{\xi}\sigma_y - 2\text{Tr}\left(\sigma_y\hat{\xi}\right)\hat{\xi}\right) \\
&\quad - \frac{K_\rho}{2}\text{Tr}\left(\sigma_z(\hat{\xi} - \xi)\right)\left(\sigma_z\hat{\xi} + \hat{\xi}\sigma_z - 2\text{Tr}\left(\sigma_z\hat{\xi}\right)\hat{\xi}\right) \\
\frac{d}{dt}\hat{\mu} &\stackrel{\text{ordre } 0}{=} -\frac{uK_\mu}{2}\left(\text{Tr}\left(\sigma_y\hat{\xi}\right)\text{Tr}\left(\sigma_z(\hat{\xi} - \xi)\right) - \text{Tr}\left(\sigma_z\hat{\xi}\right)\text{Tr}\left(\sigma_y(\hat{\xi} - \xi)\right)\right) \\
\frac{d}{dt}\hat{\Delta} &\stackrel{\text{ordre } 1}{=} -\frac{K_\Delta}{\mu}\left(\text{Tr}\left(\sigma_x\hat{\xi}\right)^2\hat{\Delta} - \text{Tr}\left(\sigma_x\hat{\xi}\right)\text{Tr}\left(\sigma_x\xi\right)\Delta\right) \\
&\quad + \frac{K_\Delta\hat{\Delta}}{2\mu}\left(\text{Tr}\left(\sigma_y\hat{\xi}\right)\text{Tr}\left(\sigma_y(\hat{\xi} - \xi)\right) - \text{Tr}\left(\sigma_z\hat{\xi}\right)\text{Tr}\left(\sigma_z(\hat{\xi} - \xi)\right)\right).
\end{aligned}$$

Il converge en deux temps. D'abord $\hat{\xi}$ et $\hat{\mu}$ d'après la convergence exponentielle de l'approximation séculaire au premier ordre. Il est ensuite aisé de voir que $\hat{\Delta}$ converge.

Réglage des gains sur le système linéarisé

On considère les variables d'erreur linéarisée

$$\hat{\xi} - \xi = \frac{1 + \tilde{x}\sigma_x + \tilde{y}\sigma_y + \tilde{z}\sigma_z}{2}, \quad \tilde{\mu} = \hat{\mu} - \mu, \quad \tilde{\Delta} = \hat{\Delta} - \Delta$$

où l'on note $\text{Tr}(\sigma_x\xi) = x$, $\text{Tr}(\sigma_x\hat{\xi}) = \hat{x}$ et similairement sur y, z . On se place autour de $\xi = \frac{1-\sigma_z}{2}$ c'est à dire $z = -1$ et on va exprimer les dynamiques linéarisées. Pour calculer la dynamique de $\hat{\xi} - \xi$ on remplace les termes du type $\text{Tr}(\sigma_z(\hat{\xi} - \xi))$ par \tilde{z} et les $\hat{\xi}$ en facteurs de termes d'erreur linéarisée (petits) par $\frac{1-\sigma_z}{2}$. Puis on identifie les coefficients devant les matrices de Pauli. Il ne reste alors que des termes en facteurs de σ_y . Et on a au premier ordre

$$\frac{d}{dt}\tilde{y} = -u\tilde{\mu} - K_\rho\tilde{y}$$

Il est aisé de voir que

$$\frac{d}{dt}\tilde{\mu} = -uK_\mu(0 - \tilde{y})/2 = uK_\mu\tilde{y}/2$$

et autour de $\rho = \frac{1-\sigma_x}{2}$ (i.e, $x = -1$)

$$\frac{d}{dt}\tilde{\Delta} = -\frac{K_\Delta}{\mu}\tilde{\Delta}.$$

De manière à respecter les échelles de temps on choisit $0 < \epsilon \ll 1$ et on choisit terme de correction dans un rapport d'ordre ϵ à la haute fréquence $u\mu$ pour recalculer en moyenne $\hat{\xi}$. Avec le choix de gains

$$K_\rho = 2k_\rho\epsilon|u|\mu, \quad K_\mu = 2\epsilon^2\mu^2$$

on a

$$\frac{d^2}{dt^2}\tilde{y} + 2(\epsilon u\mu)\tilde{y} + (\epsilon u\mu)^2\tilde{y} = 0$$

Le système oscille à la fréquence $u\mu$ et l'observateur pour la partie estimation d'état converge avec un temps caractéristique d'ordre $1/\epsilon$ oscillations. Puis on choisit un recalage lent des paramètres constants Δ et μ par rapport au temps de convergence de $\hat{\xi}$

$$K_\Delta = k_\Delta\epsilon^2|u|\mu^2$$

où $k_\rho, k_\Delta > 0$ sont à choisir mais doivent être du même ordre de grandeur que 1.

On a donc procédé au réglage des gains sur le système linéarisé. On pense que la structure non-linéaire invariante de l'observateur basée sur les symétries et donc très proche de la structure du système, permet d'assurer un bon comportement global, comme le montrent les simulations.

10.4 Simulations

Simulation avec mesures parfaites

On prend pour conditions initiales :

$$\rho_0 = \frac{1 + \cos\left(\frac{\pi}{5}\right)\sigma_x + \sin\left(\frac{\pi}{5}\right)\cos\left(\frac{\pi}{1.4}\right)\sigma_y + \sin\left(\frac{\pi}{5}\right)\sin\left(\frac{\pi}{1.4}\right)\sigma_z}{2}, \quad \mu = 1, \quad \Delta = \frac{1}{5}, \quad \hat{\rho}_0 = \sigma_x\rho_0\sigma_x$$

On choisit comme contrôle u et comme gains : $u = 1$, $K_\rho = 2\epsilon|u|\mu$, $K_\mu = 2\epsilon^2\mu^2$ et $K_\Delta = 2\epsilon^2|u|\mu^2$ et $\epsilon = \frac{1}{5}$.

Simulation avec mesures bruitées

On ajoute aux signaux un bruit blanc gaussien d'amplitude 20% et un bruit d'amplitude 10% au contrôle u . On choisit comme conditions initiales :

$$\rho_0 = \frac{1 + \cos\left(\frac{\pi}{5}\right)\sigma_x + \sin\left(\frac{\pi}{5}\right)\cos\left(\frac{\pi}{1.4}\right)\sigma_y + \sin\left(\frac{\pi}{5}\right)\sin\left(\frac{\pi}{1.4}\right)\sigma_z}{2}, \quad \mu = 1, \quad \Delta = \frac{1}{5}, \quad \hat{\rho}_0 = \sigma_x\rho_0\sigma_x$$

On choisit comme contrôle u et comme gains : $u = 1$, $K_\rho = 2\epsilon|u|\mu$, $K_\mu = 2\epsilon^2\mu^2$ et $K_\Delta = 2\epsilon^2|u|\mu^2$ avec $\epsilon = \frac{1}{5}$.

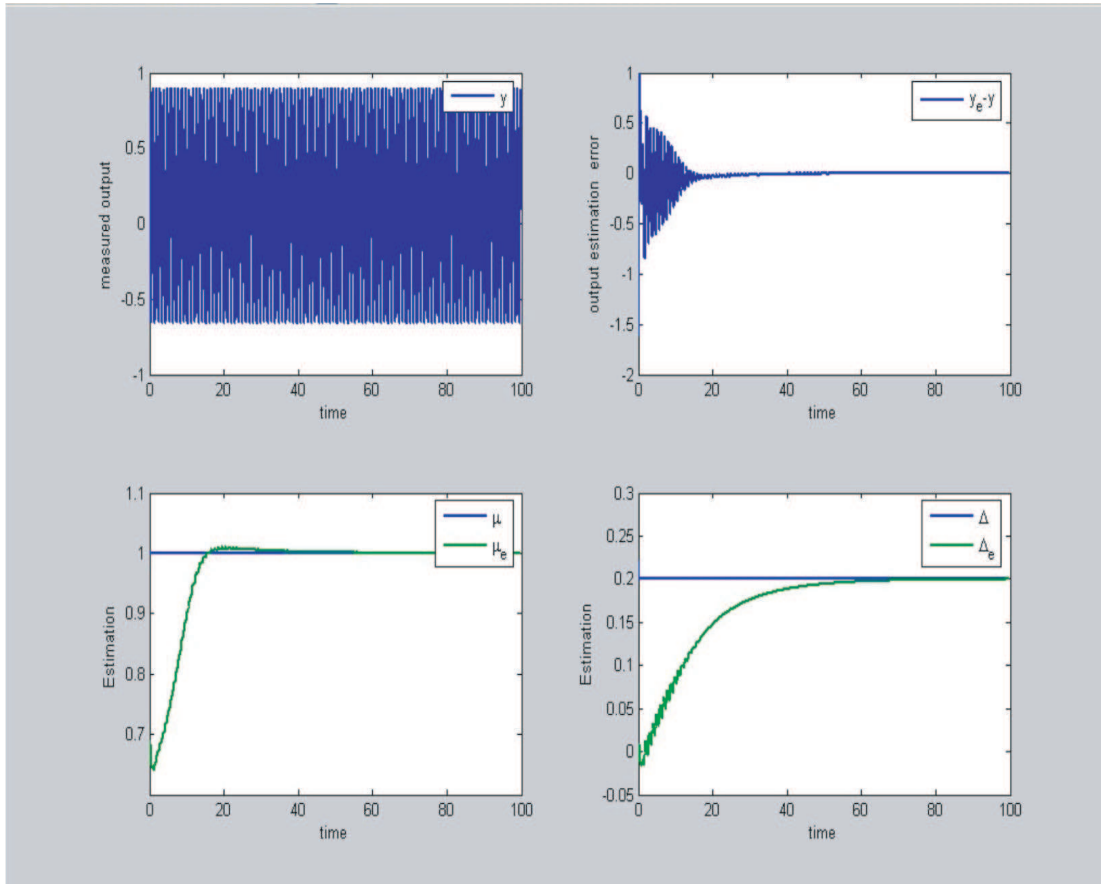


FIG. 10.1 – Sortie mesurée, erreur de sortie, estimations des paramètres μ et Δ

10.5 Conclusion

On pourrait se passer de l'utilisation des matrices et directement écrire les équations sur la sphère de Bloch avec les coordonnées (x, y, z) . Cependant, notre méthode doit pouvoir s'étendre à des systèmes quantiques plus généraux que les systèmes à deux états. Comme le montre [13] ou [43] lorsque $\Delta = 0$, des extensions à des systèmes ayant plus que deux états sont possibles, systèmes pour lesquels l'interprétation avec la sphère de Bloch de la matrice densité n'est plus possible.

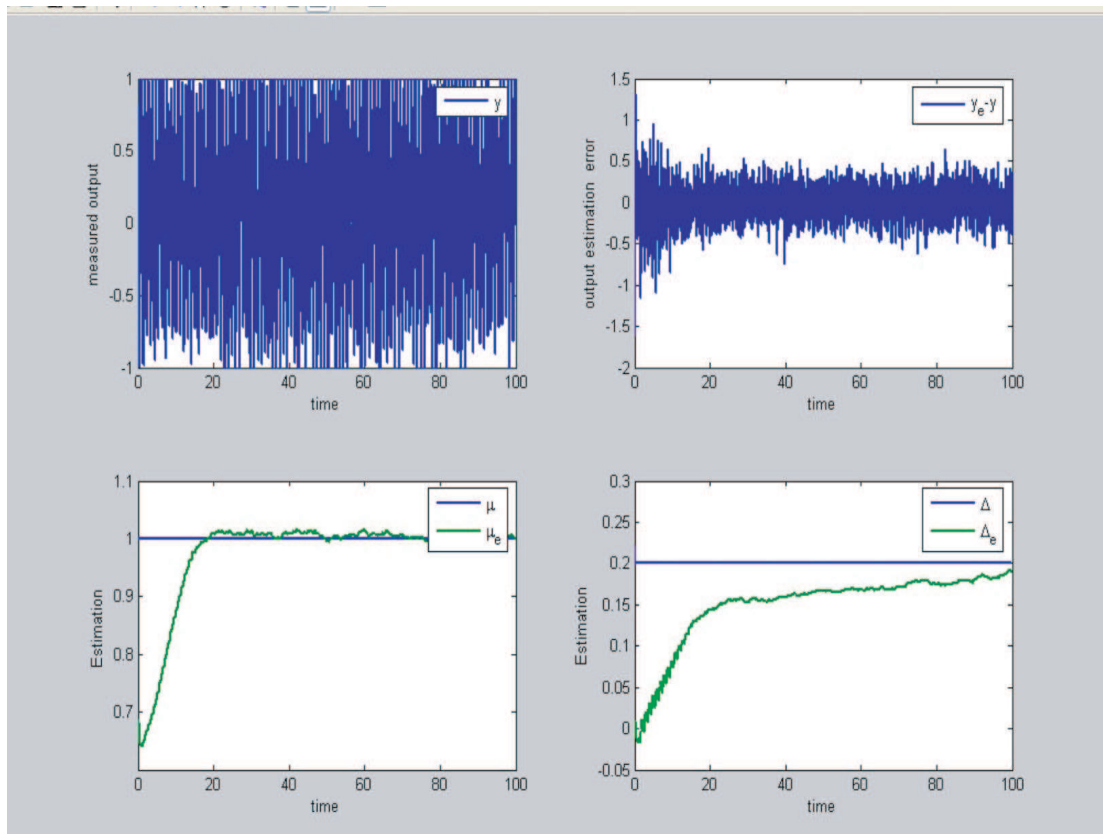


FIG. 10.2 – Sortie mesurée bruitée, amplitude du laser bruitée, erreur de sortie, estimations des paramètres μ et Δ

Chapitre 11

Nudging invariant

Dans ce chapitre, nous présentons le modèle des eaux profondes (Shallow Water) très utilisé en océanographie pour l'assimilation de données altimétriques provenant de satellites pour l'estimation des courants marins. Une des techniques pour remonter aux courants à partir des données altimétriques consiste à construire des observateurs pour ce type d'équations, c'est ce qu'on appelle le Nudging. Nous allons nous interroger sur la possibilité de définir des termes de correction qui respectent les symétries du modèle, à savoir invariance par rotation et translation du plan. Ici le groupe de symétries est donc fini, mais l'état est de dimension infinie, et le groupe n'agit pas sur l'état mais sur les coordonnées d'espace qui "indiquent" l'état, et les équations de la dynamique sont des équations aux dérivées partielles (EDP). Nous allons faire un observateur invariant pour un système décrit par les équations de Saint-Venant, qui sont une simplification du modèle Shallow Water. Nous ne présenterons pas de résultats rigoureux, mais une manière d'aborder le nudging par des considérations sur les symétries.

Notre intérêt pour la fusion de données en océanographie et pour le Nudging vient de discussions que nous avons eues dans le cadre d'un programme d'action LEFE intitulé "le back and forth nudging : études théoriques, numériques et applications" financé par l'Institut Nationales des Sciences de l'Univers, projet piloté par Jacques Blum de l'université de Nice. Pour des résultats numériques récents sur le nudging, nous renvoyons un lecteur intéressé à la thèse de Didier Auroux [7].

11.1 Modèle Shallow Water

Sur les conseils de Louis Prieur du Laboratoire d'Océanographie de Villefranche nous avons considéré le modèle standard de Jiang et al [29] de type "Shallow Water" (eaux peu profondes¹). On considère que le fluide est composé d'une seule couche active, de densité

¹Pour les eaux plus profondes ces modèles peuvent être adaptés en modèle multi-couches, dont chaque couche est considérée comme vérifiant un modèle en eaux peu profondes avec des termes de forçage liés à

constante ρ et d'épaisseur variable $h(x, y, t)$, qui recouvre une couche plus profonde et immobile de densité $\rho + \Delta\rho$. Le domaine consiste en une base rectangulaire $0 < x < L$ et $0 < y < L$ où x et y sont les coordonnées cartésiennes correspondant aux directions Est et Nord. On note ∇ l'opérateur

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$$

Les équations de la dynamique s'écrivent :

$$\frac{\partial(hv)}{\partial t} + (\nabla \cdot (hv)) + (hv) \cdot \nabla v = -g'h\nabla h - \mathbf{k} \times f(hv) + (\alpha_A A \nabla^2 - R)(hv) + \alpha_{tau} \tilde{\tau} \mathbf{i} / \rho \quad (11.1)$$

$$\frac{\partial h}{\partial t} = -\nabla \cdot (hv) \quad (11.2)$$

où $hv = h(v_x \mathbf{i} + v_y \mathbf{j})$ est le transport horizontal, avec \mathbf{i} et \mathbf{j} pointant vers l'est et le nord respectivement, $f = f_0 + \beta(y - 0.5D)$ est le paramètre de Coriolis, \mathbf{k} pointe vers le haut, g' est la gravité réduite. Le terme $\tilde{\tau} \mathbf{i}$ est un terme de forçage par les vents d'intensité variable $\tilde{\tau}$ dans l'espace et connue. R et A sont la friction du fond et la viscosité latérale.

On suppose qu'on mesure à l'aide de satellites la hauteur de l'eau $h(x, y)$ dans tout le domaine. La sortie est donc $h(x, y, t)$. On veut estimer grâce à un observateur la vitesse $v(x, y)$ en tout point (x, y) du domaine. Les variables d'état sont $(v(x, y), h(x, y))$. L'espace d'état est de dimension infinie. Les autres variables sont considérées comme des entrées connues.

11.2 Invariance par permutation d'indices

Cette section est pédagogique et peut être sautée par le lecteur. Elle est juste destinée à appréhender que (contrairement à ce qui précède) l'espace des phases est ici de dimension infinie car le système est paramétré par une infinité de variables et l'invariance par SE(2) du système que nous allons introduire peut être vue comme une permutation des variables d'espace. Les variables d'espace peuvent être vues comme des indices continus. Si l'on discrétise le système en espace (de la manière suivante par exemple), on obtient un système de dimension finie n^2 indicé par les variables d'espace discrètes $(x_i, y_j) = (i \frac{L}{n}, j \frac{L}{n})$. $h(x_i, y_j)$ est une fonction réelle, et $v(x_i, y_j)$ une fonction vectorielle à valeurs dans \mathbb{R}^2 dont les composantes sont v_x et v_y . Pour fixer les idées nous n'allons considérer que la discrétisée de la seconde équation (11.2)

$$\frac{dh(x_i, y_j)}{dt} = -((hv_x)(x_{i+1}, y_j) - (hv_x)(x_i, y_j)) \frac{n}{L} - ((hv_y)(x_i, y_{j+1}) - (hv_y)(x_i, y_j)) \frac{n}{L}$$

la friction avec les autres couches.

Cette équation est invariante par certaines permutation d'indices. On peut translater (x_i, y_j) de n'importe quel vecteur de \mathbb{R}^2 de la forme $(k\frac{L}{n}, l\frac{L}{n})$, à condition d'adapter le domaine. Ici on peut aussi permuter naturellement à la fois x_i avec y_j et v_x avec v_y . Cela revient à faire une symétrie par rapport à la diagonale, ou si l'on adapte le domaine, c'est à dire $(x_i, y_j) \in [0, L] \times [-L, 0]$, cette opération peut être vue comme une rotation d'angle $\pi/2$. Le système est en fait invariant par toutes les rotations et translations qui envoient l'ensemble des (x_i, y_j) dans lui-même. Le système non-discrétisé est en particulier invariant par toutes les rotations et translations (les déplacements) de \mathbb{R}^2 . Si l'on veut simuler le système, et rajouter des termes de corrections qui dépendent de la mesure $h(x_i, y_j)$ (faire un observateur), on s'autorise naturellement dans les termes de correction à faire toutes les opérations d'addition et de soustraction des quantités $(\hat{h}(x_i, y_j), \hat{v}(x_i, y_j), h(x_i, y_j))$, compatibles avec les symétries du système évoquées plus haut.

11.3 Invariances des équations

Les vecteurs \mathbf{i} et \mathbf{j} ont été pris pointant vers l'Est et le Nord. Le choix de \mathbf{i} pointant vers le Nord est arbitraire et les équations de la physique du modèle doivent être indépendantes de la direction Nord, ainsi que de l'origine du repère choisi, donc par l'action du groupe $SE(2)$. On va montrer cette propriété des équations. Soit R_θ une rotation d'angle θ dans le plan horizontal. Soient X, Y les coordonnées associées au nouveau repère $R_\theta(\mathbf{i}, \mathbf{j}) + (x_0, y_0)$. On note avec des majuscules les variables dans ces nouvelles coordonnées de sorte que

$$\begin{aligned}(X, Y) &= R_\theta(x, y) + (x_0, y_0) \\ H(X, Y) &= h(x, y) \\ V(X, Y) &= R_\theta v(x, y)\end{aligned}$$

On a bien $\nabla h(x, y) = R_\theta \nabla H(X, Y)$ et les équations dans les nouvelles coordonnées sont inchangées avec $\mathbf{K} = \mathbf{k}$ et $\mathbf{I} = R_\theta \mathbf{i}$:

$$\begin{aligned}\frac{\partial(HV)}{\partial t} + (\nabla \cdot (HV) + (HV) \cdot \nabla)V &= -g'H\nabla H - \mathbf{K} \times f(HV) \\ &\quad + (\alpha_A A \nabla^2 - R)(HV) + \alpha_{\tau} \tilde{\tau} \mathbf{I} / \rho \\ \frac{\partial H}{\partial t} &= -\nabla \cdot (HV)\end{aligned}$$

car les opérateurs laplacien, et divergence sont invariants par rotation et translation.

11.4 Nudging invariant

On appelle "nudging" la construction d'observateurs pour ce type de système où les équations de la dynamique sont des EDP et où les mesures sont échantillonnées en espace

et en temps. On rappelle que les mesures sont $h(x, y, t)$. Un observateur s'écrit pour le système :

$$\begin{aligned} \frac{\partial(hv)}{\partial t} + (\nabla \cdot (hv) + (hv) \cdot \nabla)v &= -g'h\nabla h - \mathbf{k} \times f(hv) + (\alpha_A A \nabla^2 - R)(hv) \\ &\quad + \alpha_{\text{tau}} \tilde{\mathbf{i}}/\rho + F_v(h, \hat{v}, \hat{h}) \\ \frac{\partial h}{\partial t} &= -\nabla \cdot (hv) + F_h(h, \hat{v}, \hat{h}) \end{aligned}$$

avec

$$F_v(h, \hat{v}, h) = 0, \quad F_h(h, \hat{v}, h) = 0$$

et de la même manière que l'on s'autorise à utiliser toutes les composantes de l'état en dimension finie, et faire les opérations de soustraction et d'addition, ici on s'autorise à ce que F_v et F_h soient des fonctionnelles de leurs arguments, des opérateurs différentiels et intégraux par rapport aux deux variables d'espace (x, y) (pas par rapport au temps t , bien sûr). En effet ici, pour le cas où les indices sont dans $[0, L]^2$, les opérations de soustraction ou d'addition avec des points de l'espace infiniment voisins s'écrivent comme des opérations de différentiation ou des intégrales en espace. $F_v(h, \hat{v}, \hat{h})$ est un vecteur et $F_h(h, \hat{v}, \hat{h})$ un scalaire. On impose que les termes de corrections respectent l'invariance spatiale par SE(2), donc représentent une certaine quantité physique. On ne conservera donc que les opérations de différentiation qui sont invariantes par rotation et translation, et également des intégrales des fonctions h, \hat{v}, \hat{h} invariantes par rotation translation.

Caractérisation des termes de correction invariants

Dans cette section nous construisons toute une famille de termes de correction possibles F_v et F_h et qui sont bien invariants par rotation et translation des axes.

Pour la correction scalaire F_h , on s'appuie sur le résultat classique suivant [32] : tout opérateur différentiel scalaire invariant par tous les déplacements est de la forme $Q(\Delta)$ avec Q polynôme à une variable et Δ le laplacien . Plus généralement les opérateurs différentiels invariants que nous recherchons sont les polynômes que l'on peut former avec l'opérateur ∇ . Ces opérateurs différentiels peuvent être appliqués à des fonctions ponctuelles de (h, \hat{h}, \hat{v}) et invariantes par translation rotation, i.e., des fonctions scalaires de h, \hat{h} et \hat{v}^2 donc. Ainsi nous obtenons toutes une gamme de correction invariante scalaire F_h avec la formule suivante

$$F_h = Q_1(\Delta, h, \hat{v}^2, \hat{h} - h) + \nabla \left(Q_2((\Delta, h, \hat{v}^2, \hat{h} - h)) \right) \cdot \hat{v} + f_h$$

où Q_1 et Q_2 sont des sortes de "polynômes" scalaires en Δ et où f_h est un terme intégral en espace défini ci-dessous. Plus précisément pour $i = 1, 2$, on a

$$Q_i(\Delta, h, \hat{v}^2, \hat{h} - h) = \sum_{k=0}^N a_k^i(h, \hat{v}^2, \hat{h} - h) \Delta^k \left(b_k^i(h, \hat{v}^2, \hat{h} - h) \right)$$

où les a_k^i et b_k^i sont de fonctions scalaires régulières de leur arguments telles que

$$a_k^i(h, \hat{v}^2, 0) = b_k^i(h, \hat{v}^2, 0) = 0.$$

Pour les corrections vectorielles F_v , on obtient toute une famille invariante par les formules suivantes :

$$F_v = P_1(\Delta, h, \hat{v}^2, \hat{h} - h)\hat{v} + \nabla \left(P_2(\Delta, h, \hat{v}^2, \hat{h} - h) \right) + f_v$$

où P_1 et P_2 sont des "polynômes" en Δ analogues aux Q_i définis ci-dessus.

Nous nous intéressons maintenant à f_h et f_v qui sont des termes intégraux que l'on veut invariants par rotation et translation. On les construit à partir de ce qui précède par convolution avec des fonctions scalaires de (x, y) à support compact, ϕ_h et ϕ_v , qui ne dépendent que de $x^2 + y^2$ uniquement :

$$f_v(x, y, t) = \int \int \left[R_1(\Delta, h, \hat{v}^2, \hat{h} - h)\hat{v} + \nabla \left(R_2(\Delta, h, \hat{v}^2, \hat{h} - h) \right) \right]_{(x-\xi, y-\zeta, t)} \phi_v(\xi^2 + \zeta^2) d\xi d\zeta$$

$$f_h(x, y, t) = \int \int \left[S_1(\Delta, h, \hat{v}^2, \hat{h} - h) + \nabla \left(S_2(\Delta, h, \hat{v}^2, \hat{h} - h) \right) \cdot \hat{v} \right]_{(x-\xi, y-\zeta, t)} \phi_h(\xi^2 + \zeta^2) d\xi d\zeta$$

où les R_i et S_i sont des "polynômes" en Δ analogues aux Q_i définis ci-dessus.

Le support des fonctions ϕ_v et ψ_h est choisi égal au diamètre d'une zone que l'on considère être une "zone d'influence" de sorte que cela a un sens de "corriger" la valeur de l'estimation au centre de la zone avec l'ensemble des valeurs de h dans cette zone. L'invariance par translation découle directement de l'utilisation du produit de convolution. L'invariance par rotation découle de l'utilisation d'un terme radial $\xi^2 + \zeta^2$ uniquement et veut qu'on ne privilégie aucune direction (et donc un système de coordonnées) dans la définition des termes de correction.

11.5 Etude du système linéarisé

Nous reprenons les équations de l'observateur générales avec termes de corrections invariants par tous les déplacements. Nous prenons pour cette section un modèle de Shallow Water simplifié qui correspond aux équations de Saint-Venant :

$$\begin{aligned} \frac{\partial}{\partial t} h &= -\nabla(hv) \\ \frac{\partial}{\partial t} v &= -v\nabla v - g\nabla h \end{aligned} \tag{11.3}$$

On s'interdit de différencier en espace la mesure h ce qui amplifierait le bruit. En fait nous choisissons de ne retenir parmi les termes de correction, que les termes intégraux, car

ils ont un effet régularisant. Ainsi, on prend $Q_1 = Q_2 = P_1 = P_2 = 0$ et $R_1 = S_2 = 0$ avec $S_1 = R_2 = h - \hat{h}$. L'observateur invariant s'écrit alors (11.3) :

$$\begin{aligned}\frac{\partial}{\partial t} \hat{h} &= -\nabla(\hat{h}\hat{v}) + \int \int \phi_h(\xi^2 + \zeta^2)(h - \hat{h})_{(x-\xi, y-\zeta, t)} d\xi d\zeta \\ &= -\nabla(\hat{h}\hat{v}) + \phi_h * (h - \hat{h}) \\ \frac{\partial}{\partial t} \hat{v} &= -\hat{v}\nabla\hat{v} - g\nabla\hat{h} + \int \int \phi_v(\xi^2 + \zeta^2)\nabla(h - \hat{h})_{(x-\xi, y-\zeta, t)} d\xi d\zeta \\ &= -\hat{v}\nabla\hat{v} - g\nabla\hat{h} + \phi_v * \nabla(h - \hat{h})\end{aligned}\tag{11.4}$$

On retrouve dans le cas dégénéré où $\phi_h = K_h\delta_0$ et $\phi_v = K_v\delta_0$ (K_h et K_v sont des gains scalaires positifs) les termes de correction classiques en nudging

$$\begin{aligned}\frac{\partial}{\partial t} \hat{h} &= -\nabla(\hat{h}\hat{v}) + K_h(h - \hat{h}) \\ \frac{\partial}{\partial t} \hat{v} &= -\hat{v}\nabla\hat{v} - g\nabla\hat{h} + K_v\nabla(h - \hat{h})\end{aligned}$$

Nous ne ferons ici que l'étude du linéarisé. On se place autour de l'équilibre. Soit \bar{h} la hauteur d'équilibre de l'eau. On considère des petites vitesses $\delta v = v \ll \sqrt{g\bar{h}}$ et $\delta h = h - \bar{h} \ll \bar{h}$. Le système linéarisé s'écrit alors

$$\frac{\partial}{\partial t} \delta h = -\bar{h}\nabla\delta v, \quad \frac{\partial}{\partial t} \delta v = -g\nabla\delta h.$$

Les erreurs d'estimation, $\tilde{h} = \hat{h} - h$ et $\tilde{v} = \hat{v} - v$, obéissent alors aux équations linéarisées suivantes :

$$\frac{\partial}{\partial t} \tilde{h} = -\bar{h}\nabla\tilde{v} - \phi_h * \tilde{h}, \quad \frac{\partial}{\partial t} \tilde{v} = -g\nabla\tilde{h} - \phi_v * \nabla\tilde{h}.$$

Traitement heuristique de l'équation d'erreur linéarisée

Par élimination de \tilde{v} on obtient l'équation

$$\frac{\partial^2}{\partial t^2} \tilde{h} = g\bar{h}\Delta\tilde{h} + \phi_v * \Delta\tilde{h} - \phi_h * \frac{\partial}{\partial t} \tilde{h}.$$

Lorsque les noyaux sont positifs et centrés étroitement autour de zéro avec comme intégrales $K_h > 0$ et $K_v > 0$, ils sont quasiment proportionnels à des mesures de Dirac. Ainsi la dynamique de d'erreur \tilde{h} est proche de celle d'une équation des ondes

$$\frac{\partial^2}{\partial t^2} \tilde{h} = (g\bar{h} + K_v)\Delta\tilde{h} - K_h\frac{\partial}{\partial t} \tilde{h}$$

C'est l'équation qui décrit la propagation d'une onde amortie avec une vitesse de propagation $\sqrt{g\bar{h} + K_v}$ et coefficient d'amortissement K_h . Une analyse dimensionnelle rapide donne comme premier réglage des gains

$$K_h = \omega_0, \quad K_v = \max(0, (L_0\omega_0)^2 - g\bar{h})$$

où ω_0 est une pulsation caractéristique et L_0 est une longueur caractéristique.

Si dans (11.4) on avait rajouté un terme en $\phi_h^\Delta * \Delta(\hat{h} - h)$ au second membre de l'équation en $\frac{\partial}{\partial t}\hat{h}$ (ϕ_h^Δ est un noyau positif régularisant, $K_h^\Delta = \int \int \phi_h^\Delta > 0$), on aurait obtenu une équation linéarisée d'erreur en \tilde{h} de type onde amortie

$$\frac{\partial^2}{\partial t^2}\tilde{h} = (g\bar{h} + K_v)\Delta\tilde{h} - K_h\frac{\partial}{\partial t}\tilde{h} + K_h^\Delta\Delta\left(\frac{\partial}{\partial t}\tilde{h}\right)$$

qui aurait tendance à filtrer encore plus les hautes fréquences spatiales et temporelles (amortissement structurel interne). Un premier réglage conforme à ce qui précède de K_h^Δ est $L_0^2\omega_0$.

Chapitre 12

Fusion de données vidéo en navigation inertielle

Dans ce chapitre nous nous interrogeons sur la possibilité de remplacer dans l'exemple de navigation inertielle vu au chapitre 7 les mesures "air" de vitesse (Air-Data System) par du traitement d'image issu de caméras CCD. Nous reprenons donc les notations du chapitre 7, et l'on considère un objet muni de caméras qui se déplace dans un environnement fixe. Nous allons considérer un cas simplifié. On ne se préoccupera ici que de l'estimation de la vitesse de l'objet à partir d'images sans considérer le problème de la navigation inertielle dans sa globalité (estimation de l'orientation ...).

Dans un premier temps, nous procédons à la modélisation du problème. Puis l'on considèrera que l'environnement est lointain, et l'on montrera comment l'on peut faire un observateur pour recalibrer les biais des gyroscopes. L'invariance correspond à l'invariance des équations par changement de repère terrestre (ce qui est une petite différence avec l'exemple de la navigation inertielle). La principale contribution de ce chapitre est de donner une formulation totalement géométrique des équations du mouvement et de la dynamique de l'image pour une caméra qui se déplace dans un environnement fixe et de les relier au problème de la navigation inertielle.

12.1 Hypothèses

Ici nous établissons un modèle du problème qui repose sur des hypothèses simplificatrices. On considère un objet volant équipé d'une caméra sphérique, qui filme son environnement dans toutes les directions de l'espace. L'environnement est modélisé par une surface $\mathcal{S} \subset \mathbb{R}^3$ entourant l'objet et difféomorphe à la sphère \mathbb{S}^2 . On considère que la géométrie de l'environnement est telle qu'on peut voir tous les points de l'environnement depuis l'objet à tout moment, les éléments de l'environnement ne passent pas les uns derrière les autres. Cela veut dire que toute droite issue de l'objet ne rencontre la surface \mathcal{S} qu'en un point.

L'environnement est paramétré par une variable $s \in \mathbb{S}^2$ et à tout s correspond un point et un seul $M(s) \in \mathcal{S} \subset \mathbb{R}^3$. Le centre de gravité de l'objet est repéré par un point $C \in \mathbb{R}^3$. On reprend les notations du chapitre 7. L'orientation de l'objet est donc repérée par le quaternion $q \in \mathbb{H}_1$ qui envoie le repère fixe (lié à l'espace) sur le repère mobile (lié à l'objet). On rappelle que v est la vitesse dans le repère du solide. Ainsi $v(t) = q^{-1} * \frac{d}{dt}C * q$ car $\frac{d}{dt}C$ est la vitesse dans le repère fixe. On note $\eta \in \mathbb{S}^2$ le vecteur unitaire, exprimé dans le repère lié à l'objet, qui pointe dans une direction donnée de l'espace. On identifie (voir l'annexe sur les quaternions) les vecteurs de \mathbb{R}^3 aux quaternions imaginaires purs. La structure de l'environnement telle qu'on l'a décrite implique que pour tout temps t , à une direction de l'espace donnée η ne correspond qu'un point de l'environnement $M(s)$. Ainsi on a par définition

$$D(t, s) q * \eta * q^{-1} = \overrightarrow{C(t)M(s)} \quad \text{avec} \quad D(t, s) = \|\overrightarrow{C(t)M(s)}\|$$

et

$$s = \varphi(t, \eta)$$

où φ est une fonction bijective par rapport à η pour tout t . On notera la fonction réciproque $\eta = \psi(t, s)$. On suppose que tout point $M(s)$ de l'environnement rayonne avec une intensité lumineuse $I(s)$ constante au cours du temps. On fait l'hypothèse que le flux d'énergie lumineuse issue du point à travers une surface fermée est indépendant de la surface (analogue du théorème de Gauss pour le champ électrique). L'intensité lumineuse provenant de $M(s)$ perçue par l'objet situé au point C est alors $K \frac{I(s)}{D^2(t, s)}$ où $K > 0$ est un facteur de normalisation. La sortie est l'intensité perçue par la caméra à tout moment dans chaque direction :

$$y(t, \eta) = K \frac{I(s)}{D^2(t, s)} \quad \text{avec} \quad s = \varphi(t, \eta)$$

12.2 Modèle cinématique

Nous allons établir que les équations du mouvement sont données par les formules (12.1) avec comme entrée connue $\omega(t)$ et $a(t)$, comme paramètre connu \mathbf{A}_{grav} et comme sortie $y(t, \eta)$ l'intensité du pixel dans la direction $\eta \in \mathbb{S}^2$ de la caméra sphérique embarquée sur le solide. Comme au chapitre 7, on a $\frac{d}{dt}q = \frac{1}{2}q * \omega$ où $\omega(t)$ est le vecteur rotation instantanée (gyromètre).

Les fonctions que nous manipulons sont des fonctions de trois variables η, s, t mais qui ne dépendent en réalité que de deux variables indépendantes car on a $\eta = \psi(t, s)$. Nous allons dériver les relations par rapport à t à s fixé. Ainsi par η nous désignerons systématiquement la fonction de deux variables $\psi(t, s)$ que nous allons différencier par rapport à t . On note par $\frac{\partial}{\partial t}|_s$ la dérivation par rapport au temps t à s fixé. On note de même $\frac{\partial}{\partial t}|_\eta$ la dérivation par rapport à t à η fixé. On utilisera la variable vitesse réduite

notée ξ

$$\xi(t, s) = \frac{v(t)}{D(t, s)}$$

où $v = q^{-1} * \frac{d}{dt}C * q$. Nous allons maintenant différencier par rapport au temps la relation structurelle (on omet l'indice t dans les équations)

$$q * \eta * q^{-1} = \frac{\overrightarrow{C(t)M(s)}}{D(t, s)} = \frac{\overrightarrow{CM}}{\|\overrightarrow{CM}\|}$$

On a

$$\frac{\partial}{\partial t} \Big|_s \left(\overrightarrow{CM} \right) = -\frac{d}{dt}C = -q * v * q^{-1}$$

et

$$\frac{\partial}{\partial t} \Big|_s \|\overrightarrow{CM}(s)\| = -\frac{d}{dt}C \cdot \frac{\overrightarrow{CM}}{\|\overrightarrow{CM}\|} = -(q * v * q^{-1}) \cdot (q * \eta * q^{-1}) = -v \cdot \eta$$

où "." dénote le produit scalaire. Donc

$$\frac{\partial}{\partial t} \Big|_s \left(\frac{\overrightarrow{CM}}{\|\overrightarrow{CM}\|} \right) = q * (-\xi + (\eta \cdot \xi)\eta) * q^{-1}$$

Par ailleurs on a $\frac{\partial}{\partial t} \Big|_s (q * \eta * q^{-1}) = q * (\omega \times \eta) * q^{-1} + q * \frac{\partial \eta}{\partial t} \Big|_s * q^{-1}$. La première équation que nous obtenons est donc une relation sur $\eta = \psi(t, s)$ qui est

$$\frac{\partial \eta}{\partial t} \Big|_s = -\omega \times \eta + (\eta \cdot \xi)\eta - \xi$$

On a aussi

$$\frac{\partial y}{\partial t} \Big|_s = \frac{\partial}{\partial t} \Big|_s \left(K \frac{I(s)}{|C(t)M(s)|^2} \right) = -2K \frac{I(s)}{|CM|^2} (q * \xi * q^{-1} \cdot q * \eta * q^{-1}) = -2(\xi \cdot \eta)y.$$

La caméra nous donne en fait à chaque instant, le champ scalaire $\mathbb{S}^2 \ni \eta \mapsto y(t, \eta)$ qui est donc la sortie du système. Comme

$$\frac{\partial y}{\partial t} \Big|_s = \frac{\partial y}{\partial t} \Big|_\eta + \frac{\partial y}{\partial \eta} \Big|_t \frac{\partial \eta}{\partial t} \Big|_s$$

le champ scalaire de sortie vérifie l'équation cinématique suivante :

$$\frac{\partial y}{\partial t} \Big|_\eta + \frac{\partial y}{\partial \eta} \Big|_t (-\omega \times \eta + (\eta \cdot \xi)\eta - \xi) = -2(\xi \cdot \eta)y.$$

De $\frac{\partial}{\partial t}\Big|_s \|CM(s)\| = -\frac{d}{dt}C \cdot \frac{\overline{CM}}{\|CM\|} = -v \cdot \eta$ on déduit également ($D = \|CM\|$)

$$\frac{\partial D}{\partial t}\Big|_\eta + \frac{\partial D}{\partial \eta}\Big|_t (-\omega \times \eta + (\eta \cdot \xi)\eta - \xi) = -v \cdot \eta.$$

Comme $\xi = v(t)/D$ on a avec $\Lambda(t, \eta) = 1/D$, le modèle cinématique suivant (les dérivées partielles portent sur les variables indépendantes (t, η)) :

$$\begin{aligned} \frac{d}{dt}q &= \frac{1}{2}q * \omega \\ \frac{d}{dt}v &= v \times \omega + q^{-1} * \mathbf{A}_{grav} * q + a \\ \frac{\partial \Lambda}{\partial t} &= -\frac{\partial \Lambda}{\partial \eta}(\eta \times (\omega + \Lambda \eta \times v)) - \Lambda^2 v \cdot \eta \\ \frac{\partial y}{\partial t} &= -\frac{\partial y}{\partial \eta}(\eta \times (\omega + \Lambda \eta \times v)) - 2\Lambda(v \cdot \eta)y \end{aligned} \tag{12.1}$$

où

- $q(t) \in \mathbb{H}_1$ et $v(t) \in \mathbb{R}^3$ forment la partie non mesurée de l'état, partie de dimension finie.
- $\mathbb{S}^2 \ni \eta \mapsto \Lambda(t, \eta) \in \mathbb{R}_*^+$ est la partie non mesurée de l'état, partie de dimension infinie.
- $\mathbb{S}^2 \ni \eta \mapsto y(t, \eta) \in \mathbb{R}_*^+$ est la partie mesurée de l'état, partie de dimension infinie.
- les vecteurs variables $\omega(t), a(t) \in \mathbb{R}^3$ et le vecteur fixe $\mathbf{A}_{grav} \in \mathbb{R}^3/\{0\}$ sont connus.

On remarquera l'invariance de ce système par rapport au groupe $G = \mathbb{H}_1$ dont l'action (changement de repère dans le mobile) est définie par

$$(q, v, \eta, \omega, a, \mathbf{A}_{grav}) \mapsto (gq, g^{-1}vg, g^{-1}\eta g, g^{-1}\omega g, g^{-1}ag, g^{-1}\mathbf{A}_{grav}g)$$

pour tout $g \in G$. Les états de dimensions infinies Λ et y sont inchangés au sens où l'argument η est seulement remplacé par $g^{-1}\eta g$ (changement de paramétrage). Il est aussi invariant par rapport à l'autre action (changement de repère dans le fixe)

$$(q, \mathbf{A}_{grav}) \mapsto (gq, g\mathbf{A}_{grav}g^{-1})$$

où les autres quantités restent inchangées.

12.3 Environnement infiniment loin et calage des gyroscopes

Dans un premier temps, nous allons nous placer dans le cas simple où l'environnement est lointain (les étoiles, la terre vue depuis un avion..), de sorte qu'on peut faire

l'approximation $\Lambda = 1/\|CM\| \approx 0$ car $\|CM\|$ est très grand. Alors (12.1) devient

$$\begin{aligned}\frac{d}{dt}q &= \frac{1}{2}q * \omega \\ \frac{d}{dt}v &= v \times \omega + q^{-1} * \mathbf{A}_{grav} * q + a \\ \frac{\partial y}{\partial t} &= -\frac{\partial y}{\partial \eta}(\eta \times \omega)\end{aligned}$$

Ainsi, on voit que l'on ne peut pas estimer v et q à partir des données de la caméra. En revanche on peut estimer un biais constant $c \in \mathbb{R}^3$ au niveau de la mesure ω_m du gyroscope : $\omega = \omega_m + c$. La dynamique de l'application de sortie y s'écrit maintenant

$$\frac{\partial y}{\partial t} + \frac{\partial y}{\partial \eta}(\eta \times \omega) = 0$$

ou encore

$$\frac{\partial y}{\partial t} + \frac{\partial y}{\partial \eta}(\eta \times \omega_m) + \frac{\partial y}{\partial \eta}(\eta \times c) = 0$$

A ce niveau on peut imaginer plusieurs solutions pour estimer le paramètre $c \in \mathbb{R}^3$ à partir de la mesure $\mathbb{S}^2 \ni \eta \mapsto y(t, \eta)$. En voici une de type Lyapounov basée sur l'idée suivante. La moyenne de y sur l'orbite engendrée par l'action des rotations sur la sphère, $\int_{\mathbb{S}^2} y(t, \eta) d\sigma_\eta$ où $d\sigma$ est l'élément d'aire sur la sphère unité \mathbb{S}^2 , est un invariant scalaire, puisque l'action du groupe est une isométrie. Il est en plus indépendant du temps puisqu'il représente la lumière totale reçue issue de l'environnement lointain, qu'on a supposée constante.

On note par ∇y le gradient de la fonction scalaire $\mathbb{S}^2 \ni \eta \mapsto y(t, \eta)$. Ainsi ∇y peut être vu comme un vecteur de \mathbb{R}^3 , tangent à la sphère en η , et donc $\nabla y \cdot \eta = 0$. On a donc

$$\frac{\partial y}{\partial \eta}(\eta \times \omega_m) = \nabla y \cdot (\eta \times \omega_m) = (\nabla y \times \eta) \cdot \omega_m$$

et la dynamique s'écrit :

$$\frac{\partial y}{\partial t} = -(\nabla y \times \eta) \cdot \omega = -(\nabla y \times \eta) \cdot (\omega_m + c)$$

Soient $k_y, k_c > 0$, deux gains constants. On considère l'observateur asymptotique suivant

$$\frac{\partial \hat{y}}{\partial t} = -(\nabla \hat{y} \times \eta) \cdot (\omega_m + \hat{c}) - k_y(\hat{y} - y), \quad \frac{d}{dt}\hat{c} = k_c \int_{\mathbb{S}^2} (\hat{y} - y)(\nabla \hat{y} \times \eta) d\sigma$$

La convergence de l'observateur repose sur la fonction de Lyapounov suivante :

$$V = \frac{1}{2} \int_{\mathbb{S}^2} (\hat{y} - y)^2 d\sigma + \frac{1}{2k_c} \|\hat{c} - c\|^2.$$

Pour chaque q , $\eta \mapsto q^{-1} * \eta * q$ est une isométrie de \mathbb{S}^2 . Donc

$$\int_{\mathbb{S}^2} (\hat{y}(t, \eta) - y(t, \eta))^2 d\sigma_\eta = \int_{\mathbb{S}^2} (\hat{y}(t, q^{-1} * \eta * q) - y(t, q^{-1} * \eta * q))^2 d\sigma_\eta$$

car $d\sigma_\eta = d\sigma_{q^{-1} * \eta * q}$. Ici, à η fixé, $q^{-1} * \eta * q$ est ce dernier vecteur fixe mais exprimé dans le repère mobile, donc $y(t, q^{-1} * \eta * q)$ est la lumière issue d'un point fixe ("s=cste") de l'environnement. Cette dernière quantité est donc indépendante du temps. Ainsi on a

$$\frac{d}{dt} \left(\int_{\mathbb{S}^2} (\hat{y}(t, \eta) - y(t, \eta))^2 d\sigma_\eta \right) = \int_{\mathbb{S}^2} \frac{d}{dt} (\hat{y}(t, q^{-1} * \eta * q) - y(t, q^{-1} * \eta * q))^2 d\sigma_\eta.$$

Comme $\frac{d}{dt} q = \frac{1}{2} q * \omega$, et $\frac{\partial(\hat{y}-y)}{\partial \eta}(\eta \times \omega) = (\nabla(\hat{y} - y) \times \eta) \cdot \omega$ on a

$$\frac{d}{dt} \left(\int_{\mathbb{S}^2} (\hat{y} - y)^2 d\sigma \right) = \int_{\mathbb{S}^2} 2(\hat{y} - y) \left(\frac{\partial(\hat{y} - y)}{\partial t} + (\nabla(\hat{y} - y) \times \eta) \cdot \omega \right) d\sigma.$$

Comme $\frac{\partial(\hat{y}-y)}{\partial t} = -(\nabla(\hat{y} - y) \times \eta) \cdot \omega - k_y(\hat{y} - y) + (\nabla\hat{y} \times \eta) \cdot (c - \hat{c})$ et $\omega - c = \omega_m$, on a

$$\frac{d}{dt} \left(\int_{\mathbb{S}^2} (\hat{y} - y)^2 d\sigma \right) = 2 \int_{\mathbb{S}^2} (\hat{y} - y) (-k_y(\hat{y} - y) - (\nabla\hat{y} \times \eta) \cdot (\hat{c} - c)) d\sigma.$$

Ainsi

$$\frac{d}{dt} V = -2k_y \int_{\mathbb{S}^2} (\hat{y} - y)^2 d\sigma \leq 0.$$

Une étude plus fine de la convergence nécessite le principe d'invariance de Lasalle, principe dont l'application en dimension infinie est assez délicate. Intuitivement, on a asymptotiquement $\hat{y} = y$ et donc (ce qui n'est pas évident ici) $\nabla\hat{y} = \nabla y$, soit $(\nabla y \times \eta) \cdot (\hat{c} - c) = 0$. Ainsi si l'ensemble des vecteurs $\nabla y \times \eta$ engendre tout \mathbb{R}^3 quand η parcourt \mathbb{S}^2 , on doit avoir asymptotiquement $\hat{c} = c$. Le raisonnement précédent n'est pas rigoureusement justifié. Son seul mérite est de donner une condition plausible de convergence asymptotique : $\text{vect}(\nabla y \times \eta)_{\eta \in \mathbb{S}^2} = \mathbb{R}^3$, i.e., l'image doit avoir un contraste suffisant dans trois directions indépendantes.

Cas d'une caméra classique

Notre approche est compatible avec la modélisation classique des caméras qu'on appelle le "pinhole model" (voir fig 12.1) puisqu'elle lui correspond à un changement de variable près. La seule différence est l'utilisation de caméras sphériques qui permet de bien faire ressortir la géométrie du problème et son invariance par $SO(3)$ que l'on identifie ici à \mathbb{H}_1 . On peut éviter cette hypothèse en restreignant les différentes intégrales sur la sphère à une portion de sphère. Pour cela il faudrait définir une zone centrale pour l'image telle qu'on

puisse calculer le flux optique entre le moment où un pixel de l'environnement sort de la zone centrale, et le moment où il sort de l'image.

Pour cela soit $\phi : \mathbb{R}^3 \mapsto \mathbb{R}$, une fonction C^∞ uniquement de η , dont le support définit la fenêtre de la caméra. Elle vaut 0 en dehors de la fenêtre, et 1 dans l'intérieur de la fenêtre auquel on a enlevé un voisinage des bords. Sur la fenêtre on multiplie la sortie mesurée y par ϕ de manière à obtenir la nouvelle sortie $Y(t, \eta) = \phi(\eta) y(t, \eta)$ qui vérifie

$$\frac{\partial Y}{\partial t} + (\nabla Y \times \eta - y \nabla \phi \times \eta) \cdot \omega = 0$$

On modifie alors l'observateur comme suit

$$\begin{aligned} \frac{\partial \hat{Y}}{\partial t} &= -(\nabla \hat{Y} \times \eta - y \nabla \phi \times \eta) \cdot (\omega_m + \hat{c}) - k_y(\hat{Y} - Y), \\ \frac{d}{dt} \hat{c} &= k_c \int_{\mathbb{S}^2} (\hat{Y} - Y)(\nabla \hat{Y} \times \eta - y \nabla \phi \times \eta) d\sigma \end{aligned} \quad (12.2)$$

A nouveau on prend comme fonction de Lyapunov

$$V = \frac{1}{2} \int_{\mathbb{S}^2} (\hat{Y} - Y)^2 d\sigma + \frac{1}{2k_c} \|\hat{c} - c\|^2.$$

et on a maintenant

$$\frac{\partial(\hat{Y} - Y)}{\partial t} = -(\nabla(\hat{Y} - Y) \times \eta) \cdot \omega - k_y(\hat{Y} - Y) - (\nabla \hat{Y} \times \eta - y \nabla \phi \times \eta) \cdot (\hat{c} - c)$$

et le premier terme du membre de droite disparaît quand il est intégré sur toute la sphère pour les mêmes raisons que précédemment.

Formulation à l'aide des coordonnées cartésiennes

Dans le modèle de caméra classique, chaque pixel de l'image est repéré par ses coordonnées (x, y) . Un point M de l'environnement est repéré par ses coordonnées cartésiennes (x, y, z) (voir fig 12.1). On le projette sur la sphère et on obtient les nouvelles coordonnées $(x', y') = (\sin(\arctan x), \sin(\arctan y))$. On écrit $\eta = (0, x', y', \sqrt{1 - x'^2 - y'^2})^T$. En effet $\eta \in \mathbb{S}^2$ est un quaternion de norme 1 et de première coordonnée nulle. La sortie est maintenant une fonction $h(t, x, y)$ (on change y en h pour éviter les confusions avec les coordonnées cartésiennes) et l'observateur (12.2) s'écrit

$$\begin{aligned} \frac{\partial \hat{Y}}{\partial t} &= -(\nabla \hat{Y} \times \eta - h(t, x, y) \nabla \phi \times \eta) \cdot (\omega_m + \hat{c}) - k_y(\hat{Y} - Y), \\ \frac{d}{dt} \hat{c} &= k_c \int_{\substack{-L \leq x \leq L \\ -h \leq y \leq h}} (\hat{Y} - Y)(\nabla \hat{Y} \times \eta - h(t, x, y) \nabla \phi \times \eta) d\Sigma(x', y', z') \end{aligned}$$

où $d\Sigma$ est l'élément d'aire pour la sphère unité munie des coordonnées cartésiennes et $z' = \sqrt{1 - x'^2 - y'^2}$.

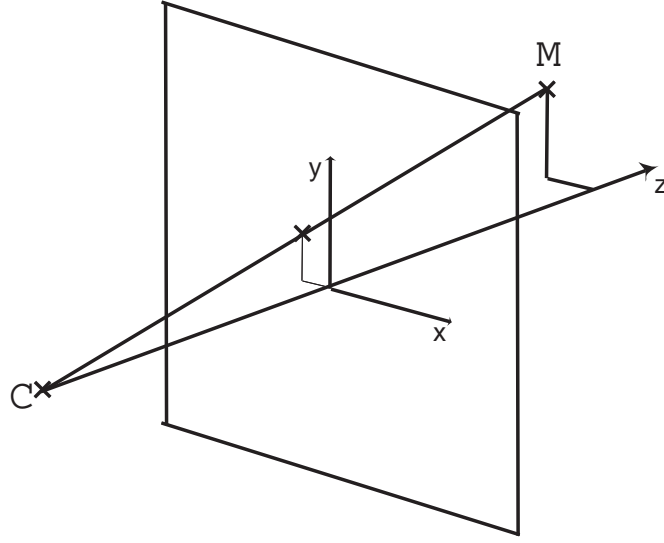


FIG. 12.1 – Coordonnées cartésiennes et modèle de caméra “pinhole model”. C est le centre optique, Cz est l’axe optique de la caméra, Cx et Cy sont parallèles au plan de l’image, et (Cxyz) forme un repère orthonormal. Le point M est repéré par ses coordonnées cartésiennes dans le repère caméra. La longueur focale est prise égale à 1.

12.4 Conclusion

Si on suppose Λ petit mais pas négligeable, alors on peut négliger dans (12.1) les termes d’ordre 2 en Λ et garder ceux d’ordre un, dérivées partielles en η incluses. On obtient alors le système approché suivant

$$\begin{aligned}\frac{d}{dt}q &= \frac{1}{2}q * \omega \\ \frac{d}{dt}v &= v \times \omega + q^{-1} * \mathbf{A}_{grav} * q + a \\ \frac{\partial \Lambda}{\partial t} &= -\frac{\partial \Lambda}{\partial \eta}(\eta \times \omega) \\ \frac{\partial y}{\partial t} &= -\frac{\partial y}{\partial \eta}(\eta \times (\omega + \Lambda \eta \times v)) - 2\Lambda(v \cdot \eta)y\end{aligned}$$

Il faudrait maintenant dans de futurs travaux considérer le modèle approché ci-dessus, et déterminer si l’on peut remonter à l’orientation q et la vitesse v (comme au chapitre 7).

Chapitre 13

Un observateur réduit intrinsèque

13.1 Introduction

Dans ce chapitre on reprend le problème posé dans [3] (voir aussi [38] où l'espace de configuration est un groupe de Lie). Dans cet article les auteurs considèrent les systèmes mécaniques lagrangiens pour lesquels on mesure les positions et proposent une méthode pour construire un observateur pour reconstituer l'état (positions et vitesses). L'observateur construit est indépendant du choix des variables de configuration, choix qui forme un groupe de dimension infinie. Sa construction repose sur une méthode originale sans lien direct avec la méthode de construction des observateurs invariants du chapitre 6. La méthode employée consiste à s'appuyer sur la structure Riemannienne de l'espace des configurations et la métrique définie par l'énergie cinétique. L'observateur construit est alors intrinsèque (invariant par changement de coordonnées) et converge localement.

On propose ici une méthode alternative, basée sur la notion d'observateur réduit, et ayant une interprétation géométrique simple. Nous montrons alors que, contrairement au cas de l'observateur intrinsèque complet proposé dans [3], le fait d'avoir une courbure négative renforce la stabilité de cet observateur réduit. On ne montre la convergence qu'en l'absence d'énergie potentielle et de forces extérieures mais il y a convergence sur un domaine qui peut être très large pour un espace de courbure négative, puisque ce domaine admet pour diamètre le rayon d'injectivité d'un point quelconque de la variété. Pour les variétés de courbure négative simplement connexes la convergence est globale. L'observateur proposé est défini intrinsèquement, et est invariant par les changements de coordonnées de configuration. Ce chapitre est une première tentative vers un observateur intrinsèque globalement convergent dans le cas d'une courbure négative, et différent de celui [3], destiné à estimer la vitesse pour un système lagrangien dont on mesure toutes les positions. Il faudrait dans une suite éventuelle de ce travail modifier la formule de l'observateur pour inclure les termes liés à l'énergie potentielle et aux forces extérieures.

Système Lagrangien et métrique Riemannienne

Afin de situer le problème, on reprend ici mot pour mot l'article [3]. "On considère un système mécanique Lagrangien avec une variété de configuration de dimension n munie d'une métrique Riemannienne. Les coordonnées locales de $q \in V$ sont notées $(q^i)_{i=1\dots n}$. Le Lagrangien est donné par :

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}g_{ij}(q)\dot{q}^i\dot{q}^j - U(q)$$

où la matrice définie positive $g(q) = (g_{ij}(q))_{i=1\dots n, j=1\dots n}$ définit la métrique (matrice d'inertie) et la fonction scalaire $U(q)$ l'énergie potentielle. Les équations d'Euler-Lagrange sont, dans les coordonnées locales,

$$\frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}^i} \mathcal{L} \right) = \frac{\partial}{\partial q^i} \mathcal{L} + u^i(q, t), \quad i = 1, \dots, n$$

où $u(q, t) = (u^i(q, t))_{i=1, \dots, n}$ est une fonction connue de t et q , qui correspond en général à des entrées connues. La formulation Riemannienne des équations d'Euler-Lagrange est

$$\nabla_{\dot{q}} \dot{q} = -\text{grad}_q U(q) + g(q)^{-1}u(q, t) \equiv S(q, t)$$

où ∇ , grad_q et $g(q)^{-1}$ sont respectivement, la connection de Levi-Civita, l'opérateur gradient associé à la structure Riemannienne et l'inverse de la matrice de la métrique $g(q)$. Comme la position q est mesurée, le terme source $S(q, t)$ est un champ de vecteur sur M dépendant du temps et connu. Dans les coordonnées locales, les équations d'Euler-Lagrange s'écrivent :

$$\ddot{q}^i = -\Gamma_{jk}^i(q)\dot{q}^j\dot{q}^k + S^i(q, t) \quad (13.1)$$

où les termes de connection Γ_{jk}^i (symboles de Christoffel) sont donnés par

$$\Gamma_{jk}^i = \frac{1}{2}g^{il} \left(\frac{\partial g_{lk}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^k} - \frac{\partial g_{jk}}{\partial q^l} \right)$$

avec g^{il} les composantes de la matrice g^{-1} . Nous utilisons ici la convention de sommation : quand un indice apparaît à la fois en exposant et en indice, cela signifie que la sommation selon cet indice doit être faite.

Rappelons que $\nabla_{\dot{q}} v$ est la dérivée covariante du champ de vecteur v le long de \dot{q} . Dans les coordonnées locales, on obtient

$$\{\nabla_{\dot{q}} v\}^i = \dot{v}^i + \Gamma_{jk}^i(q)v^j\dot{q}^k$$

où $\{ \}^i$ signifie coordonnée i ".

13.2 Un observateur réduit intrinsèque

Le problème revient donc à construire un observateur qui donne une estimation de \dot{q} quand q se déplace sur une variété M , et qu'on mesure sa position au cours du temps. Nous allons procéder de la manière suivante. Le problème est simple si q se déplace dans \mathbb{R}^n à une vitesse constante v . On peut en effet estimer sa vitesse v à l'aide d'un observateur réduit : on construit un point ξ qui le "suit" à une certaine distance. La trajectoire de ξ est définie par l'équation différentielle suivante

$$\dot{\xi} = -\frac{\xi - q}{\lambda} \quad (13.2)$$

On appelle \hat{v} la vitesse de ξ . On veut montrer que $\hat{v} \rightarrow v$. En dérivant cette équation on obtient $\dot{\hat{v}} = -(1/\lambda)(\hat{v} - v)$. On choisit $\lambda > 0$. On a $\hat{v} \rightarrow v$. L'interprétation géométrique est la suivante : puisqu'on a par définition $\xi = q + \lambda\hat{v}$, on voit que ξ "suit" le point q à une distance λv , à l'image d'un "chariot" attaché au point q par une sorte de "ressort". (13.2) est un observateur réduit dans \mathbb{R}^n .

Maintenant on aimerait utiliser cet outil pour estimer la vitesse d'un point q qui se déplace sur une variété Riemannienne M selon une géodésique. On définit de manière analogue ξ comme solution de l'équation

$$\dot{\xi} = -\frac{1}{2\lambda} \overrightarrow{\text{grad}}_{\xi} D^2(\xi, q) \quad \lambda > 0 \quad (13.3)$$

où $D(\xi, q)$ est la distance géodésique entre q et ξ . Cette dynamique est indépendante d'un choix de coordonnées. Dans \mathbb{R}^n , on retrouve bien l'équation (13.2). On veut montrer que $\dot{\xi}$ transporté parallèlement de ξ à q est une estimation de $v = \dot{q}$. On pose donc

$$\hat{v} = \mathcal{T}_{//\xi \rightarrow q} \dot{\xi} \quad (13.4)$$

où $\mathcal{T}_{//\xi \rightarrow q}$ est le transport parallèle de ξ à q sur la géodésique joignant ξ à q , et il est défini pour ξ et q suffisamment proches. Il faut donc montrer que ξ converge indépendamment de sa position initiale vers un point mobile qui est sur la géodésique parcourue par q et à une distance λv de q .

Dans le cas où la courbure riemannienne de l'espace est négative ou nulle, on a $\hat{v} \rightarrow v$ tant que ξ reste dans le rayon d'injectivité de q , qui peut dans ce cas être très grand.

Dans la cas où la courbure est positive et bornée, on a $\hat{v} \rightarrow v$ pour $1/\lambda$ pas trop grand et ξ initialisé proche de q . De manière imagée, cela veut dire que si $1/\lambda$ est trop grand alors le "chariot" est trop loin du point qu'il suit et on perd la convergence. On illustrera ce résultat par des simulations sur la sphère.

13.2.1 Lemme

Nous allons commencer par un lemme de géométrie.

Lemme 4. *Soit M une variété Riemannienne régulière. Soit P un point fixé de cette variété. Sur l'ensemble des éléments de M que l'on peut atteindre par une géodésique et une seule (rayon d'injectivité), on définit le système dynamique*

$$\frac{d}{dt}x = -\frac{1}{2\lambda} \overrightarrow{\text{grad}}_x D^2(P, x) \quad \lambda > 0 \quad (13.5)$$

Si la courbure K dans toutes les sections en chaque point est négative ou nulle, la dynamique est contractante au sens de [34], i.e, si δx est un déplacement virtuel à t fixé on a

$$\frac{d}{dt}\|\delta x\|^2 \leq -\frac{1}{\lambda}\|\delta x\|^2$$

où $\|\cdot\|$ désigne la norme au sens de la métrique Riemannienne de M .

Démonstration. Soit $x \in M$ appartenant au rayon d'injectivité de P . Le déplacement virtuel δx est défini dans [34] comme une forme linéaire différentielle tangente, et par dualité peut être considéré comme un vecteur de $TM|_x$. Nous allons définir une surface. Soit C la géodésique qui relie P à x . On considère un point infiniment voisin de x dans la direction δx . Il est relié à P par une géodésique. Les directions de cette géodésique et de C engendrent un 2-plan tangent en P . On considère toutes les géodésiques issues de P suivant les directions contenues dans ce plan. Elles forment une surface différentiable S . La surface construite est plongée dans M et hérite de la métrique reimannienne. Elle est invariante par le flot de (13.5), puisque le champ est tangent aux géodésiques issues de P . On a bien $x \in S$, et δx est tangent à S en x .

En suivant [51] (p 177) on choisit sur S un jeu de coordonnées spécial qui généralise les coordonnées polaires sur le plan. Un point est repéré par σ , qui est la longueur d'arc géodésique qui le relie à P , et θ , qui est l'angle (dans le plan tangent à S en P) entre C et la géodésique qui relie le point à P . Dans ces coordonnées polaires géodésiques l'élément de longueur au sens de la métrique est

$$ds^2 = d\sigma^2 + G(\sigma, \theta)d\theta^2$$

et G vérifie les relations initiales $\sqrt{G} = 0$ et $\frac{\partial\sqrt{G}}{\partial\sigma} = 1$ en $\sigma = 0$ ainsi que l'équation du second ordre $\frac{\partial^2\sqrt{G}}{\partial\sigma^2} = -K\sqrt{G}$ où $K(\sigma, \theta)$ est la courbure de Gauss au point de coordonnées (σ, θ) . On peut montrer que la courbure de Gauss de la surface S en un point est majorée par la courbure sectionnelle évaluée sur le plan tangent à la surface S en ce point, et est donc négative. On en déduit que l'on a le long de C

$$\frac{\partial^2 G}{\partial\sigma^2} = \frac{\partial}{\partial\sigma}(2\sqrt{G}\frac{\partial\sqrt{G}}{\partial\sigma}) \geq 0$$

et donc $\frac{\partial}{\partial \sigma}(\sigma \frac{\partial G}{\partial \sigma}) \geq \frac{\partial G}{\partial \sigma}$. Comme on a $G(\sigma, 0) = 0$, on a en intégrant l'inégalité $\sigma \frac{\partial G}{\partial \sigma} \geq G$. La dynamique s'écrit dans ces coordonnées

$$\dot{\sigma} = -\frac{1}{\lambda}\sigma; \quad \dot{\theta} = 0$$

En écrivant $\|\delta x\|^2 = ds^2$ on obtient immédiatement le long de C (repérée par $\theta = 0$)

$$\frac{d}{dt}\|\delta x\|^2 = -\frac{2}{\lambda}d\sigma^2 - \frac{\sigma}{\lambda} \frac{\partial G(\sigma, 0)}{\partial \sigma} d\theta^2 \leq -\frac{1}{\lambda}\|\delta x\|^2$$

Remarque : Pour une dynamique sur une surface de courbure de Gauss constante positive K , on a $\sqrt{G(\sigma, 0)} = \frac{1}{K} \sin(\sqrt{K}\sigma)$ et il y a contraction stricte pour $\sigma < \frac{\pi}{2K} - \epsilon$ avec $0 < \epsilon < 1$ fixé.

□

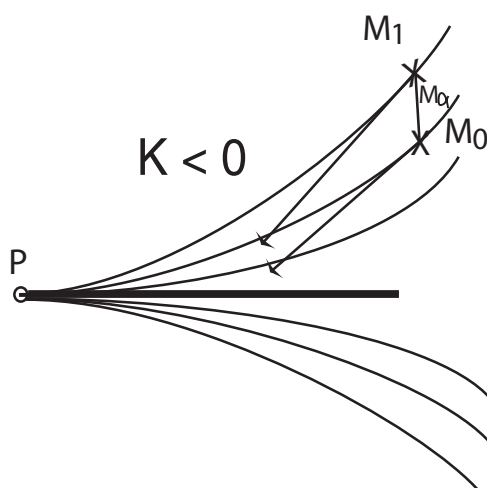


FIG. 13.1 – Variation des géodésiques sur une variété de courbure négative; et dynamique de deux points voisins

13.2.2 L'observateur réduit

On définit ξ (voir (13.3)) comme solution de l'équation

$$\dot{\xi} = -\frac{1}{2\lambda} \overrightarrow{\text{grad}}_{\xi} D^2(\xi, q) \quad \lambda > 0$$

On rappelle que q parcourt une géodésique à vitesse constante v . On transporte parallèlement le vecteur vitesse $\frac{1}{2\lambda} \overrightarrow{\text{grad}}_{\xi} D^2(\xi, q)$ de ξ à q le long de la géodésique pour avoir une estimation de v qui n'est pas plus bruitée que la mesure q .

On suppose que la courbure de la variété est négative. D'après le lemme 4, on sait que la dynamique ci-dessus est contractante au sens de [34] sur un domaine qui peut être très large. Toutes les trajectoires convergent donc entre elles et le système oublie sa condition initiale. Il suffit alors de trouver une trajectoire solution. Elle s'écrit comme suit. Soit ζ le point situé sur la géodésique parcourue par q à une distance $D(q, \zeta) = \lambda v$ dans le sens opposé à \dot{q} . La distance entre q et ζ s'écrit alors comme la longueur de l'arc géodésique qui les relie, et elle est constante :

$$\frac{d}{dt}D(\zeta, q) = v - \frac{1}{\lambda}D(\zeta, q) = 0$$

Dans le cas où la variété est simplement connexe de courbure sectionnelle négative (appelées alors variétés de Hadamard dans le cas courbure sectionnelle strictement négative), l'application exponentielle de pôle P est un difféomorphisme de l'espace tangent en P vers la variété, il est donc surjectif et injectif. La dynamique est définie quelque soit les conditions initiales et est strictement contractante. Il y a convergence globale.

Dans le cas où la courbure sectionnelle est positive et majorée par $a > 0$, la dynamique est contractante pour ξ suffisamment proche de q . (On peut même obtenir une borne inférieure pour le diamètre du domaine d'attraction dans le cas où la courbure de Gauss des surfaces construites dans la preuve du lemme est positive). Cela implique d'initialiser ξ proche de q et de prendre λ suffisamment grand. Dans ce cas ξ converge bien vers un point mobile de la géodésique parcourue par q et à une distance λv de q . L'exemple suivant illustre ce cas.

13.2.3 Cas où la courbure est positive : simulations sur la sphère

On considère la sphère de rayon 1 dans \mathbb{R}^3 muni d'un repère euclidien. Soit q un point de coordonnées q_x, q_y, q_z qui tourne autour de l'équateur à vitesse constante. L'équation de son mouvement est :

$$\begin{aligned}\dot{q}_x(t) &= \cos(t) \\ \dot{q}_y(t) &= \sin(t) \\ \dot{q}_z(t) &= 0\end{aligned}$$

On considère l'observateur réduit intrinsèque (13.3), et on note ξ_x, ξ_y, ξ_z les coordonnées du "chariot suiveur" ξ . La dynamique de l'observateur réduit s'écrit ici

$$\begin{aligned}\dot{\xi}_x &= (1/\lambda) \varphi \frac{((q \wedge \xi) \wedge \xi)_x}{\| (q \wedge \xi) \wedge \xi \|} \\ \dot{\xi}_y &= (1/\lambda) \varphi \frac{((q \wedge \xi) \wedge \xi)_y}{\| (q \wedge \xi) \wedge \xi \|} \\ \dot{\xi}_z &= (1/\lambda) \varphi \frac{((q \wedge \xi) \wedge \xi)_z}{\| (q \wedge \xi) \wedge \xi \|}\end{aligned}$$

où $\lambda < 0$ et φ est l'angle entre les vecteurs q et ξ et correspond donc à la distance géodésique. Pour les simulations les conditions initiales sont : $q(0) = [1, 0, 0]^T$ et $\xi(0) = \frac{1}{\sqrt{2}}[0, 1, 1]^T$. Un bruit blanc gaussien d'amplitude 20% de la valeur maximale des signaux a été ajouté pour simuler les imperfections des capteurs. On voit dans le cas convergent que ξ ne converge pas vers q mais converge bien vers un point mobile de l'équateur situé à une distance λ de q , et le transport parallèle de la vitesse de ξ à q fournit une estimation de la vitesse. Pour $\lambda < \pi/2$ l'observateur réduit converge, pour $\lambda > \pi/2$ on observe en simulation qu'il diverge. Cela est lié au comportement des géodésiques de la sphère (figure 13.4).

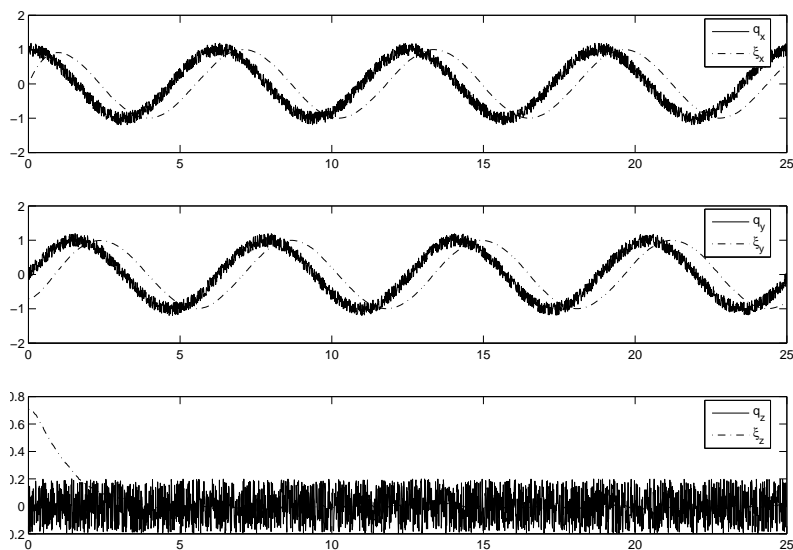


FIG. 13.2 – Résultat des simulations pour $\lambda = \pi/4$: q (ligne pleine) et ξ (pointillés)

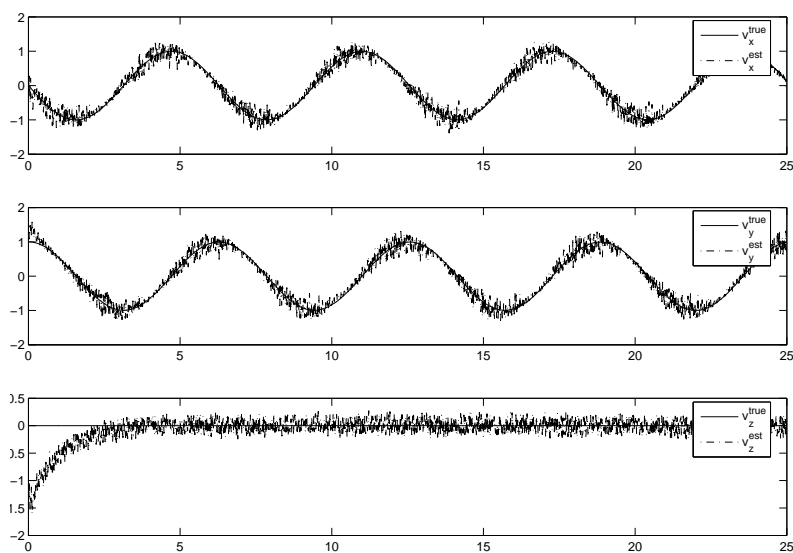


FIG. 13.3 – Résultat des simulations pour $\lambda = \pi/4$: vitesse v (ligne pleine) et estimation \hat{v} (pointillés)

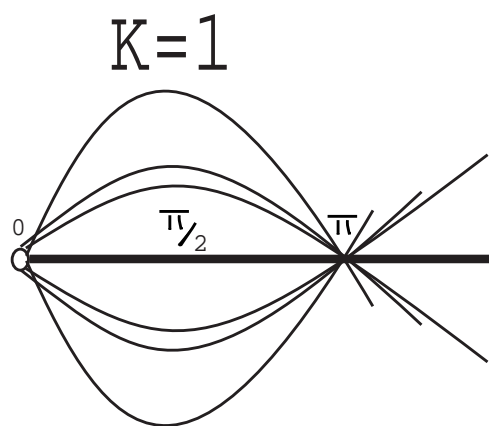


FIG. 13.4 – Variation des géodésiques sur la sphère

Remarque

Pour illustrer la théorie dans le cas où il n'y a pas de terme source et où la courbure est négative, et pour lequel il y a alors convergence sur un domaine de diamètre le rayon d'injectivité de la variété, on pourrait reprendre l'exemple du ball and beam traité dans [3] sans gravité et sans contrôle u puisque la courbure est négative.

Chapitre 14

Conclusion

Nous avons construit et implémenté un observateur pour un réacteur chimique exothermique industriel. Nous nous sommes interrogés ensuite sur la construction d'observateurs en général pour les systèmes chimiques de type réacteur exothermique, et l'on s'est interrogé sur la dépendance des termes de correction (de l'observateur) aux unités. On a donc cherché ensuite à construire de manière systématique des termes de correction qui respectent les invariances des équations de la dynamique dans le cas général (indépendance par rapport aux choix d'unités, indépendance par rapport aux changements de coordonnées ...). Les invariances sont définies relativement à l'action d'un groupe de Lie sur l'espace d'état et l'espace des entrées. La contribution théorique principale de la thèse est de donner une méthode systématique de construction de pré-observateurs invariants sous des hypothèses modérées de régularité de l'action du groupe. On a de plus la garantie théorique qu'on trouve ainsi la forme de tous les pré-observateurs invariants, autrement dit tous les observateurs candidats qui préservent les symétries. Nous n'avons jusqu'ici pas de méthode générale pour attaquer les singularités de l'action du groupe. Nous n'avons pas non plus de méthode générale pour "sélectionner" parmi les observateurs invariants candidats, ceux qui convergent. Mais il n'existe pas de méthode générale pour trouver des observateurs convergents pour des systèmes non-linéaires admettant des symétries. On ne peut qu'essayer de tirer parti de la structure apportée par les symétries dans la synthèse d'observateurs asymptotiques.

Pour cela nous avons remis en perspective le sens de la notion d'erreur d'état quand il y a des symétries. En effet, les symétries correspondent souvent à des invariances issues de la physique du système, et définir une erreur non-linéaire qui préserve les symétries, donc la structure du système, est donc a priori intéressant. Le cas le plus frappant est celui où l'on a une dynamique invariante à gauche sur un groupe de Lie. On peut alors donner à l'erreur un sens complètement intrinsèque, comme étant un élément du groupe. L'erreur linéaire usuelle dépend du choix des coordonnées et n'a alors pas de sens. Cette démarche qui consiste à définir une erreur qui respecte la structure du système est ici développée dans le cadre général mais a souvent un sens physique bien connu dans les exemples traités.

En effet la démarche de l'ingénieur consiste à manipuler des coefficients adimensionnels dans la mesure du possible, quand il y a plusieurs choix possible d'unités, comme dans les exemples issus de la chimie. Pour l'exemple de la voiture, ramener l'erreur dans le repère de Frénet pour qu'elle ne dépende pas du choix de l'orientation du repère (autrement dit d'où est le Nord) paraît très naturel également. Si on faisait un observateur-contrôleur pour la voiture, on serait surpris qu'elle fasse un créneau de manière différente selon l'orientation du trottoir.

La construction (de manière systématique) de pré-observateurs invariants, aboutit à des termes de correction non-linéaires invariants qui peuvent avoir une structure qui n'est pas entièrement prévisible a priori. Le calcul de l'équation d'erreur invariante pour ces observateurs permet d'obtenir des propriétés intéressantes pour l'étude de la convergence. On trouve que l'équation d'erreur ne dépend de la trajectoire que via les invariants fondamentaux de l'action du groupe. Dans les exemples, cette propriété rend l'étude de la convergence plus simple, et le réglage des gains plus facile et plus homogène. L'exemple de la navigation inertielle illustre le mieux cette propriété, car alors l'erreur obéit à une équation différentielle autonome, elle ne dépend donc ni de la trajectoire, ni des entrées (qui dépendent du temps). Dans des cas qui présentent moins de symétries la dynamique de l'erreur est liée à la décomposition en coordonnées base-fibre, et correspond à des changements de variables intéressants. Pour ces raisons l'utilisation des symétries va plus loin que la simple idée que quand le système vit dans un variété plongée, on peut rajouter des termes de corrections tangents à la variété qui forcent l'observateur à rester sur cette dernière.

Notre théorie ne s'applique directement qu'à une certaine classe de systèmes. Dans la partie III nous l'avons cependant utilisée comme un guide pour aborder d'autres systèmes, et notamment ceux gouvernés par des équations aux dérivées partielles. Ainsi nous avons proposé les premières pistes pour aborder la fusion de données en océanographie de façon invariante (nudging invariant) en tirant partie des symétries des modèles de type "Shallow Water". Les exemples qui portent sur la navigation inertielle, l'identification paramétrique d'un système quantique à deux états, et le nudging, correspondent à des problèmes non-linéaires non triviaux et pour lesquels on n'a obtenu qu'une convergence locale.

Au-delà du cadre de la théorie de la partie II se dégage en fait une méthode simple pour la construction d'observateurs sur un système non-linéaire avec symétries : on commence par chercher les observateurs, donc les termes de correction, qui préservent les symétries. On écrit ensuite l'équation d'erreur invariante. On peut espérer que ses propriétés permettent d'obtenir un domaine de convergence plus grand et un réglage des gains plus simple. Quand on n'arrive pas à montrer la convergence globale, on peut linéariser l'équation de la dynamique de l'erreur d'état invariante, et faire le design des gains sur le linéarisé tangent (autour d'un équilibre ou d'une trajectoire permanente par exemple). On a vu qu'au premier ordre le design des gains n'était pas soumis à des contraintes particulières et s'appuie sur les techniques linéaires usuelles. Quand l'état du système estimé et celui du vrai système sont loin l'un de l'autre, la structure de l'observateur invariant est une manière,

basée sur la physique du système, de prolonger les gains choisis localement. Il semble que, pour les exemples considérés dans cette thèse, cette façon de construire l'observateur permet d'avoir un bassin d'attraction plus large, et donc un meilleur comportement global.

On s'est concentré sur les estimateurs du type observateurs asymptotiques avec copie de la dynamique. D'autres types d'estimateurs existent comme par exemple ceux de Kazantzis-Kravaris/Luenberger où l'état de l'estimateur ne s'identifie plus nécessairement à l'état x du système (voir, e.g., [4] pour des résultats récents sur cette approche). Il existe aussi des méthodes algébriques alternatives radicalement différentes d'estimation d'état ou de paramètres comme par exemple les travaux conduits par Michel Fliess au sein du projet Alien de l'Inria-Futur ¹. Nous pensons que l'utilisation des symétries issues de la physique peut également jouer un rôle intéressant pour ces méthodes.

¹Voir le site web : <http://www.inria.fr/recherche/equipes/alien.fr.html>

Quatrième partie

Annexes

Annexe A

Cinétique de polymérisation

A.1 Justification théoriques des modèles de cinétique chimique

Nous présentons dans ce chapitre une brève synthèse de l'étude bibliographique que nous avons faite concernant les cinétiques de polymérisation, principalement [41, 50]. Cette étude nous a permis d'établir un modèle simplifié comme décrit au chapitre I pour la cinétique de polymérisation.

Cas Homopolymère

Pour établir le modèle cinétique, nous avons considéré que la cinétique est donnée essentiellement par la vitesse de propagation v_p :

$$v_p = k_p[M][M^\bullet]$$

où $[M]$, $[M^\bullet]$ représentent respectivement la concentration en monomère et en radicaux et k_p la constante cinétique de propagation.

Nous appliquons alors l'hypothèse d'état quasi-stationnaire pour déterminer le terme $[M^\bullet]$, et on suppose à cet effet, que les vitesses d'amorçage v_a et de terminaison v_t sont égales. Cette hypothèse repose sur le fait que les radicaux sont très réactifs et disparaissent quasiment instantanément après leur apparition. Cela conduit à

$$v_a = 2fk_d[Cata] = v_t = k_t[M^\bullet]^2$$

où f est un facteur d'efficacité, et (k_d, k_t) les constantes cinétiques d'amorçage et de terminaison.

Finalement, on obtient l'expression suivante pour la cinétique de polymérisation :

$$v_p = k_p[M]\sqrt{\frac{2fk_d[Cata]}{k_t}}$$

Si toutes les constantes cinétiques sont données par une loi du type Arrhenius (vérifiée expérimentalement [19, 18]) :

$$k = k_o \exp\left(-\frac{E_o}{RT}\right)$$

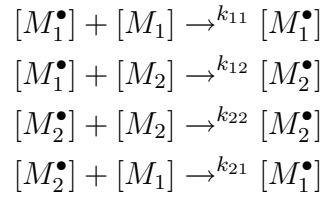
en négligeant l'effet de la pression qui est régulée dans le réacteur, on parvient au modèle de cinétique suivant que nous avons retenu au chapitre I :

$$v_p = k_o \exp\left(\frac{-E_o}{RT}\right) [M] \sqrt{[Cata]} \quad (\text{A.1})$$

en agglomérant les 3 constantes cinétiques (k_p, k_d, k_t) et le facteur d'efficacité f dans une seule constante.

Cas Copolymère

Dans le cas de la copolymérisation, l'action du catalyseur conduit à la formation de deux radicaux $[M_1^\bullet]$ et $[M_2^\bullet]$. Les réactions de transfert s'écrivent



On suppose comme dans le cas homopolymère (hypothèse de quasi stationnarité) que la concentration en chacun des deux radicaux $[M_1^\bullet]$ et $[M_2^\bullet]$ est constante. Cela veut dire que l'addition de $[M_1^\bullet]$ à $[M_2]$ est égale à l'addition de $[M_2^\bullet]$ à $[M_1]$. Donc

$$k_{12}[M_1^\bullet][M_2] = k_{21}[M_2^\bullet][M_1]$$

On pose alors

$$r_1 = \frac{k_{11}}{k_{12}} \text{ et } r_2 = \frac{k_{22}}{k_{21}}$$

et puisque

$$\begin{aligned} -\frac{d[M_1]}{dt} &= k_{11}[M_1^\bullet][M_1] + k_{21}[M_2^\bullet][M_1] \\ -\frac{d[M_2]}{dt} &= k_{12}[M_1^\bullet][M_2] + k_{22}[M_2^\bullet][M_2] \end{aligned}$$

on obtient

$$\frac{d[M_1]}{d[M_2]} = \frac{[M_1]}{[M_2]} \frac{r_1[M_1] + [M_2]}{[M_1] + r_2[M_2]}$$

avec les notations du chapitre I cela donne

$$\frac{R_{ethylene,mol}^i}{R_{A,mol}^i} = \frac{[Ethylene]^i r_e [Ethylene]^i + [A]^i}{[A]^i [Ethylene]^i + r_A [A]^i}$$

Cinétique de l'amorceur

Tous les articles que nous avons consultés attribuent la cinétique suivante à l'amorceur (décomposition du premier ordre pour le peroxyde) :

$$R_{Cata,mol} = V k_d \exp\left(\frac{-E_d}{RT}\right) [Cata]$$

Il est à noter qu'on a en plus une relation directe entre la quantité de peroxyde et la quantité de polymère dans le cas homopolymère puisque $-d[M]/dt = k_o \exp(\frac{-E_o}{RT})[M]\sqrt{[Cata]}$. On n'a pas de telle relation dans le cas copolymère ou terpolymère.

A.2 Justification expérimentale des paramètres des modèles

Ici on présente des résultats sur la composition du polymère, obtenus sur le réacteur pilote du Groupement de Recherche de Lacq du groupe ARKEMA pour différentes qualités de polymère, en utilisant l'observateur du chapitre 3. Les paramètres r_{ij} (ratios de réactivité) ont été pour certains mesurés par M. Buback de l'institut de physique chimique de Göttingen ([16, 17]), ou par G. Luft de l'institut de technologie chimique de Darmstadt ([36]). Nous avons dans le cas échéant repris ces valeurs. Les autres ont été estimés au GRL. Toutes les expériences ont été faites avec débits, pressions et température d'entrée, constantes. La figure A.1 montre les résultats (on a masqué l'échelle de l'axe horizontal) de la quantité d'acrylate incorporé dans le polymère en fonction du taux de conversion pour un grade Ethylène-Acrylate de Butyle, calculé par l'observateur d'une part, et mesuré par infra-rouge (IR) d'autre part. La figure A.2 montre la quantité de comonomère (acrylate) et de termonomère (MAH) incorporées dans le polymère pour divers grade de terpolymère, calculée par l'observateur d'une part, et mesuré par infra-rouge d'autre part.

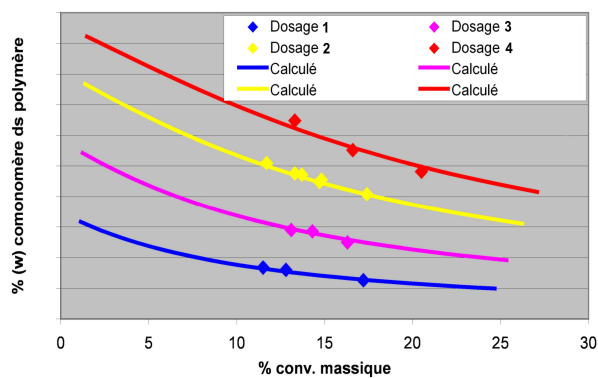


FIG. A.1 –

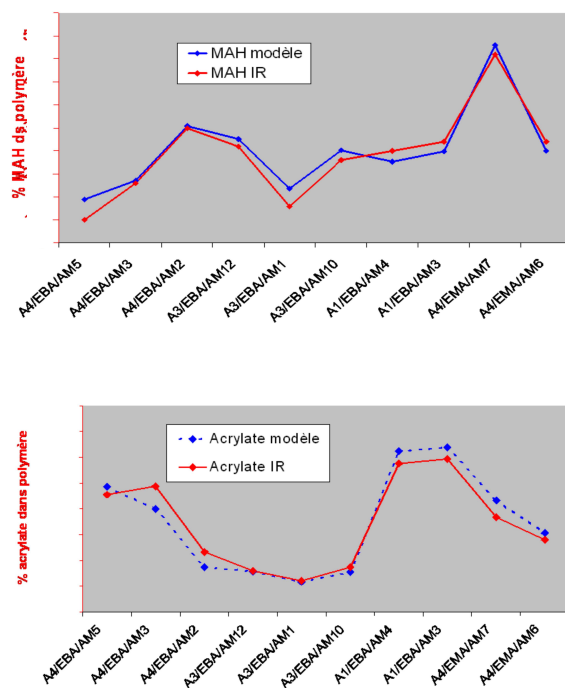


FIG. A.2 – Taux d'incorporations en terpolymère et en copolymère prédits par l'estimateur (modèle) et mesurés expérimentalement pour divers types de produits

Annexe B

Quaternions

Comme dans [21], on utilise la paramétrisation par les quaternions de $SO(3)$ afin de construire des observateurs pour estimer l'état. Les quaternions forment un groupe non-commutatif. Tout quaternion q peut s'écrire $q = q^0 + q^1 e_1 + q^2 e_2 + q^3 e_3$ avec $(q^0, q^1, q^2, q^3) \in \mathbb{R}^4$, la multiplication non-commutative $*$ est définie par

$$e_1 * e_1 = -1, \quad e_1 * e_2 = -e_2 * e_1 = e_3$$

et les relations qu'on déduit par permutation circulaire

$$e_2 * e_2 = e_3 * e_3 = -1, \quad e_2 * e_3 = -e_3 * e_2 = e_1 \quad e_3 * e_1 = -e_1 * e_3 = e_2$$

et la norme de q est $\sqrt{(q^0)^2 + (q^1)^2 + (q^2)^2 + (q^3)^2}$. Tout vecteur $\vec{p} \in \mathbb{R}^3$ peut être identifié avec le quaternion $p^1 e_1 + p^2 e_2 + p^3 e_3$. On fera cette identification dans toute la thèse pour tout ce qui touche à la navigation inertielle, où les vecteurs de \mathbb{R}^3 sont vus comme des quaternions.

On peut alors associer à tout quaternion de norme 1 une matrice de rotation $R_q \in SO(3)$ grâce à la relation relation :

$$q^{-1} * \vec{p} * q = R_q \vec{p}$$

pour tout \vec{p} . Réciproquement, à toute rotation R_q de $SO(3)$ sont associés deux quaternions $\pm q$ de longueur 1. Bien que l'espace d'état dans la section 7.2 soit $SO(3) \times \mathbb{R}^3$, on écrira les éléments de $SO(3)$ comme des quaternions dont la norme est 1 (on note \mathbb{H}_1 cet espace) et les vecteurs de \mathbb{R}^3 comme des quaternions dont la première coordonnée est 0. Numériquement, il est plus simple de conserver la norme d'un quaternion égale à 1 qu'une matrice 3×3 dans $SO(3)$. Les calculateurs embarqués sur les drones sont peu puissants. Le produit vectoriel $v \times \omega$ des vecteurs de \mathbb{R}^3 correspond pour les quaternions correspondants au commutateur $(v * \omega - \omega * v)/2$.

Annexe C

Observateur et équation d'erreur autonome : Un exemple

Navigation inertielle assistée par mesure de champ magnétique

Cette section est une reprise de l'exemple du chapitre 7, dans un cas plus simple, mais illustrant la théorie de la fin du chapitre 6.

Equations du mouvement

Les équations de la cinématique pour un objet volant s'écrivent \mathbb{H}

$$\frac{d}{dt}q = \frac{1}{2}q \cdot \omega$$

avec

- $q \in \mathbb{H}$ quaternion des attitudes,
- $\omega(t)$ vecteur rotation instantanée,
- $*$ multiplication des quaternions.

On prend pour état q . On fait une approximation supplémentaire qui est celle du vol quasi-stationnaire. C'est-à-dire que l'accélération du centre de masse est petite devant $\| \mathbf{A}_{grav} \|$. Cela veut dire que $\mathbf{a} \simeq q^{-1} * \mathbf{A}_{grav} * q$. La sortie est l'accélération spécifique $y_G = \mathbf{a}$ et le champ magnétique \vec{B} mesurés par les magnétomètres dans le repère avion $y_B = q^{-1} * \vec{B} * q$. La sortie mesurée y est

$$y = (y_G, y_B) = (q^{-1} * \mathbf{A}_{grav} * q, q^{-1} * \vec{B} * q)$$

Un observateur non-linéaire

On considère les observateurs invariants de la forme (voir fin du chapitre 6)

$$\frac{d}{dt}\hat{q} = \frac{1}{2}\hat{q} * \omega + \left(\sum_{i=1}^3 E_i \vec{e}_i \right) * \hat{q} \quad (\text{C.1})$$

où les E_i sont les erreurs de sortie invariantes $\hat{q} * y_B \hat{q}^{-1} - \vec{B}$ et $\hat{q} * y_G \hat{q}^{-1} - \vec{G}$. Les \vec{e}_i sont les quaternions associés à la base canonique de \mathbb{R}^3 (Voir annexe sur les quaternions).

Système d'erreur

On considère l'erreur d'état invariante à droite :

$$r := \hat{q} * q^{-1}$$

Soit

$$E(r) = \sum_{i=1}^3 E_i \left(r * \vec{B} * r^{-1} - \vec{B}, r * \vec{G} * r^{-1} - \vec{G} \right) \vec{e}_i$$

Elle est bien invariante par multiplication à droite $q \mapsto q * h \in \mathbb{H}$ pour $h \in \mathbb{H}$ quelconque. On a

$$\begin{aligned} \dot{r} &= \dot{\hat{q}} * q^{-1} + q * (-q^{-1} * \dot{q} * q^{-1}) \\ &= \left(\frac{1}{2} q * \omega - E(r) * \hat{q} - \frac{1}{2} q * \omega \right) * q^{-1} \\ &= E(r) * r \end{aligned}$$

Donc

$$\dot{r} = E(r) * r \quad (\text{C.2})$$

avec $E_i(0,0) = 0$. La dynamique de l'erreur $r = \hat{q} * q^{-1}$ ne dépend pas de la trajectoire $t \mapsto q(t)$.

Approximation linéaire

Une erreur petite correspond à r proche de 1. Le choix des \vec{e}_3 est arbitraire et on choisit $\vec{e}_3 = \vec{G}$. On suppose de plus que le champ magnétique \vec{B} est horizontal en utilisant plutôt le vecteur $\vec{B} - (\vec{B} \cdot \mathbf{A}_{grav}) \mathbf{A}_{grav}$ que \vec{B} , où \cdot est le produit scalaire. L'utilisation de ce vecteur orthogonal à \mathbf{A}_{grav} rend la preuve de convergence globale plus simple (dernière section).

On peut donc choisir $\vec{e}_1 = \vec{B}$. On écrit l'erreur petite $r = 1 + \vec{\xi}$ avec $\vec{\xi}$ petit. On a jusqu'à l'ordre 2 en $\vec{\xi}$

$$\frac{d}{dt}\vec{\xi} = \sum_{1 \leq i \leq 3} \sum_{1 \leq j \leq 6} \frac{\partial E_i}{\partial x_j} (\vec{\xi} \times \vec{e}_3, \vec{\xi} \times \vec{e}_1)^j \vec{e}_i$$

où " \times " représente le produit vectoriel dans \mathbb{R}^3 . $\vec{\xi}$ est vu comme un vecteur $(\vec{\xi}, \vec{\eta})^j$ représente la j -ième coordonnée $(\vec{\xi}, \vec{\eta})$ sur la base canonique de $\mathbb{R}^3 \times \mathbb{R}^3$. On a donc

$$\frac{d}{dt} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \left(\frac{\partial E_i}{\partial x_j} \right)_{1 \leq i \leq 3} \begin{pmatrix} \xi^2 & 0 \\ -\xi^1 & \xi^3 \\ 0 & -\xi^2 \end{pmatrix}$$

On peut choisir les E_i de manière à ce que

$$\frac{d}{dt} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = -K_1 \begin{pmatrix} \xi^1 \\ \xi^2 \\ 0 \end{pmatrix} - K_2 \begin{pmatrix} 0 \\ \xi^2 \\ \xi^3 \end{pmatrix} \quad (\text{C.3})$$

On a convergence exponentielle autour de toute trajectoire avec les constantes de temps $1/K_1$ and $1/K_2$.

Une classe d'observateurs asymptotiques non-linéaires

On prend les observateurs

$$\begin{cases} \frac{d}{dt}\hat{q} = \frac{1}{2}\hat{q} * \omega \\ \quad - [K_1 \mathbf{A}_{grav} \times (\hat{q} * q^{-1} * \mathbf{A}_{grav} * q * \hat{q}^{-1} - \mathbf{A}_{grav}) \\ \quad + K_2 \vec{B} \times (\hat{q} * q^{-1} * \vec{B} * q * \hat{q}^{-1} - \vec{B})] * \hat{q} \\ \frac{d}{dt}\hat{q} = \frac{1}{2}\hat{q} * \omega - [K_1 \mathbf{A}_{grav} \times (\hat{q} * y_G * \hat{q}^{-1} - \mathbf{A}_{grav}) \\ \quad + K_2 \vec{B} \times (\hat{q} * y_B * \hat{q}^{-1} - \vec{B})] * \hat{q} \end{cases}$$

Cela correspond à (C.1). L'approximation linéaire pour r proche de 1 est exactement (C.3). De plus pour tout $K_1, K_2 > 0$:

$$\lim_{t \rightarrow +\infty} \hat{q}(t) - q(t) = 0 \text{ ou } \lim_{t \rightarrow +\infty} \hat{q}(t) + q(t) = 0$$

pour toute condition initiale. Ce résultat était partiellement montré dans [14] et il est montré dans [37]. On remarque que q et $-q$ représentent la même rotation de $SO(3)$. La preuve consiste à prendre comme fonction de Lyapunov fonction $V(r) = \|r * \mathbf{B} * r^{-1} - \mathbf{B}\|^2 + \|r * \mathbf{A}_{grav} * r^{-1} - \mathbf{A}_{grav}\|^2$ pour (C.2).

Bibliographie

- [1] N. Aghannan. *Contrôle de Réacteurs de Polymérisation, observateur et invariance*. PhD thesis, Ecole des Mines de Paris, November 2003.
- [2] N. Aghannan and P. Rouchon. On invariant asymptotic observers. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 2, pages 1479–1484, 2002.
- [3] N. Aghannan and P. Rouchon. An intrinsic observer for a class of lagrangian systems. *IEEE AC*, 48(6) :936–945, 2003.
- [4] V. Andrieu and L. Praly. On the existence of a kazantzis-kravaris/luenberger observer. *SIAM Journal on Control and Optimization*, 45 :432–446, 2006.
- [5] R. Aris and N.R. Amundson. An analysis of chemical reactor stability and control-i,ii,iii. *Chem. Engng. Sci.*, 7 :121–155, 1958.
- [6] V. Arnold. *Méthodes Mathématiques de la Mécanique Classique*. Mir Moscou, 1976.
- [7] Didier Auroux. *Etude de différentes méthodes d’assimilation de données pour l’environnement*. PhD thesis, Université de Nice Sophia-Antipolis, 2003.
- [8] P. Becker and M. Busch. Modeling of ethylene copolymerizations with acrylate monomers. *Macromol. Theory Simul.*, 7 :435–446, 1998.
- [9] S. Bonnabel. Invariant extended kalman filter. In *Conference on Decision and Control 2007 (CDC07)*, 2007.
- [10] S. Bonnabel, Ph. Martin, and P. Rouchon. Groupe de lie et observateur non-linéaire. In *CIFA 2006 (Conference Internationale Francophone d’Automatique)*, Bordeaux, France., June 2006.
- [11] S. Bonnabel, Ph. Martin, and P. Rouchon. A non-linear symmetry-preserving observer for velocity-aided inertial navigation. In *American Control Conference (ACC06)*, pages 2910–2914, June 2006.
- [12] S. Bonnabel, Ph. Martin, and P. Rouchon. Symmetry-preserving observers. <http://arxiv.org/abs/math.OA/0612193>, Accepted for publication in *IEEE AC*, Dec 2006.
- [13] S. Bonnabel, M. Mirrahimi, and P. Rouchon. Observer-based hamiltonian identification for quantum systems. *Submitted*, 2007.

- [14] S. Bonnabel and P. Rouchon. *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems*, chapter On Invariant Observers, pages 53–66. Number 322 in Lecture Notes in Control and Information Sciences. Springer, 2005.
- [15] R.W. Brockett. Remarks on finite dimensional nonlinear estimation. *Asterisque*, 75-76 :47–55, 1980.
- [16] M. Buback and T. Dröge. High-pressure free-radical copolymerization of ethene and methyl acrylate. *Macromol. Chem. Phys.*, 198 :3627–3638, 1997.
- [17] M. Buback, T. Dröge, A. Van Herk, and F.-O Mähling. High-pressure free-radical copolymerization of ethene and n-butyl acrylate. *Macromol. Chem. Phys.*, 197 :4119, 1996.
- [18] M. Buback and J.Schweer. Chain-length dependence of free-radical rate coefficients, termination rate coefficient in pure ethylene polymerization. *Macromol. Chem., Rapid commun.*, 9 :699–704, 1988.
- [19] M. Buback and H. Lendle. The chemically initiated high pressure polymerization of ethylene. *Macromol. Chem.*, 184 :193–206, 1983.
- [20] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Mécanique Quantique*, volume I& II. Hermann, Paris, 1977.
- [21] G. Creamer. Spacecraft attitude determination using gyros and quaternion measurements. *Journal of Astronautical Sciences*, 44(3) :357–371, 1996.
- [22] F. Fagnani and J. Willems. Representations of symmetric linear dynamical systems. *SIAM J. Control and Optim.*, 31 :1267–1293, 1993.
- [23] J.P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.
- [24] J.W. Grizzle and S.I. Marcus. The structure of nonlinear systems possessing symmetries. *IEEE Trans. Automat. Control*, 30 :248–258, 1985.
- [25] G. E. Ham. *Copolymerization, High polymers Vol XVIII*. Interscience publishers, New York, 1964.
- [26] T Hamel and R. Mahony. Attitude estimation on $so(3)$ based on direct inertial measurements. In *International Conference on Robotics and Automation, ICRA2006*, 2006.
- [27] H. Hammouri and JP. Gauthier. Global time-varying linearization up to output injection. *SIAM J. Control Optim.*, 30 :1295–1310, 1992.
- [28] S. Haroche and J.M. Raimond. *Exploring the Quantum : Atoms, Cavities and Photons*. Oxford University Press, 2006.
- [29] S. Jiang and M. Ghil. Tracking nonlinear solutions with simulated altimetric data in a shallow water model. *Journal of physical oceanography*, 27 :72–95, 1997.
- [30] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Basic Eng., Trans. ASME, Ser. D.*, 83(3) :95–108, 1961.

- [31] A. Krener. *Algebraic and Geometric Methods in Nonlinear Control Theory*, chapter The intrinsic geometry of dynamic observations, pages 77–87. D.Reidel Publishing Company, 1986.
- [32] Schwartz L. Opérateur invariants par rotation. fonctions métaharmoniques. In *Séminaire Schwarz, tome 2, exp 9*, pages 1–5, 1954.
- [33] L. Landau and E. Lifshitz. *Mechanics*. Mir, Moscow, 4th edition, 1982.
- [34] W. Lohmiller and J.J.E. Slotine. On metric analysis and observers for nonlinear systems. *Automatica*, 34(6) :683–696, 1998.
- [35] D. Luenberger. An introduction to observers. *IEEE Transaction on Automatic Control*, AC-16(6) :596-602, 16(6) :596–602, 1971.
- [36] F. Stein Luft G. and Maximilian Dorn. The free-radical terpolymerisation of ethylene, methyl acrylate and vinyl acetate at high pressure. *Die Angewandte Makromolekulare Chemie*, 211 :131–140, 1993.
- [37] R. Mahony, T. Hamel, and J-M Pflimlin. Complimentary filter design on the special orthogonal group $so(3)$. In *Proceedings of the IEEE Conference on Decision and Control, CDC05, Seville*, 2005.
- [38] D. H. S. Maithripala, W. P. Dayawansa, and J. M. BERG. Intrinsic observer-based stabilization for simple mechanical systems on lie groups. *SIAM J. Control and Optim.*, 44 :1691–1711, 2005.
- [39] S.I. Marcus. Algebraic and geometric methods in nonlinear filtering. *SIAM J. Control Optimization*, 22 :817–844, 1984.
- [40] Ph. Martin, P. Rouchon, and J. Rudolph. Invariant tracking. *ESAIM : Control, Optimisation and Calculus of Variations*, 10 :1–13, 2004.
- [41] T. McKenna. *Polymer Reaction Engineering*. cours de l’ENSPM, 2003.
- [42] M. Mirrahimi and P. Rouchon. Mesure continue d’un ensemble statistique de systèmes quantiques. In *CIFA 2006 (Conférence Internationale Francophone d’Automatique)*, Bordeaux, France, June 2006.
- [43] M. Mirrahimi and P. Rouchon. Observer-based hamiltonian identification for quantum systems. *arXiv :math-ph/0703024v1*, 2007.
- [44] P. J. Olver. *Equivalence, Invariants and Symmetry*. Cambridge University Press, 1995.
- [45] P. J. Olver. *Classical Invariant Theory*. Cambridge University Press, 1999.
- [46] W. Respondek and I.A. Tall. Nonlinearizable single-input control systems do not admit stationary symmetries. *Systems and Control Letters*, 46 :1–16, 2002.
- [47] A.J. van der Schaft. Symmetries in optimal control. *SIAM J. Control Optim.*, 25 :245–259, 1987.

- [48] Y.K. Song and J.W. Grizzle. The extended kalman filter as a local asymptotic observer. *Estimation and Control*, 5 :59–78, 1995.
- [49] M.W. Spong and F. Bullo. Controlled symmetries and passive walking,. *IEEE Trans. Automat. Control*, 50 :1025–1031, 2005.
- [50] Malcom P. Stevens. *Polymer Chemistry, An introduction 3rd Ed.* Oxford University Press, 1999.
- [51] J.J. Stoker. *Differential Geometry.* Wiley-Interscience, 1969.
- [52] N.G Van Kampen. *Stochastic process in physics and chemistry.* Elsevier, North Holland Personal Library, 1992.
- [53] N.G. Van Kampen. *Stochastic processes in physics and chemistry.* Elsevier, 1992.
- [54] J. Villermaux. *Génie de la réaction chimique : conception et fonctionnement des réacteurs.* 2e éd.revue et augmentée edition, 1993.
- [55] N. K. Read Zhang, S.X. and W.H. Ray. Runaway phenomena in low-density polyethylene autoclave reactors. *AIChE*, 42 :10, 1996.