



THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES À FINALITÉ APPROFONDIE

Étude d'une méthode numérique pour l'approximation du bassin d'attraction de systèmes dynamiques non linéaires au moyen de l'opérateur de Koopman

Bierwart, François-Grégoire

Award date:
2022

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



UNIVERSITÉ DE NAMUR
Faculté des Sciences

**Étude d'une méthode numérique pour l'approximation du
bassin d'attraction de systèmes dynamiques non linéaires
au moyen de l'opérateur de Koopman**

Promoteur : Alexandre Mauroy

Mémoire présenté pour l'obtention du grade académique de master en
sciences mathématiques à finalité approfondie

François-Grégoire BIERWART

Juin 2022

Remerciements

Cette page est ici pour moi l'occasion de remercier toutes les personnes qui ont contribué, d'une manière ou d'une autre, à l'élaboration de ce mémoire et sans qui ce dernier ne verrait probablement pas le jour.

Tout d'abord, je tiens à remercier mon promoteur Monsieur Alexandre Mauroy pour le temps qu'il a pu m'accorder lors de nos rendez-vous pour répondre à mes nombreuses questions mais également pour le temps passé à relire ce manuscrit. Ses conseils m'ont été d'une grande aide.

J'aimerais également prendre le temps de remercier ma famille et mes amis proches pour leur soutien constant et les bons moments passés. En particulier je remercie ma maman qui m'a toujours soutenu depuis le début des études et sur qui j'ai toujours pu compter.

Enfin, je tiens à remercier ma compagne pour son soutien dans les moments difficiles et pour ses relectures constantes de ce travail. Bien au-delà du mémoire, elle a toujours été là pour moi et a su me rassurer dans les moments de doute. Nous ne sommes pas sans savoir que le cursus universitaire est compliqué psychologiquement et je remercie à travers ces mots toutes les personnes qui ont été là pour moi pendant ces cinq années d'étude, qui arrivent à leur terme.

Et à tous ceux que j'ai oubliés, merci.

Résumé

Dans la théorie des systèmes dynamiques non linéaires, l'étude de la stabilité globale de points fixes est un sujet loin d'être évident. En effet, la stabilité de tels attracteurs nécessite en général de déterminer une fonction de Lyapunov particulière parmi une infinité de choix possibles, ce qui peut devenir assez fastidieux. Ce constat motive alors une approche fonctionnelle des systèmes dynamiques via l'opérateur de Koopman. Dans ce contexte, la stabilité de points fixes repose sur l'existence de fonctions propres particulières de l'opérateur et dont le support délimite les frontières du bassin d'attraction de l'équilibre. Cet opérateur étant de dimension infinie, le calcul de ses fonctions propres n'est pas évident et nous devons en général en déterminer une approximation sur un sous-espace fini de fonctions de base. Dans ce cas, l'approximation des fonctions propres est obtenue en calculant les vecteurs propres de la représentation matricielle de l'opérateur.

Dans ce mémoire, nous investiguons en particulier l'impact de la projection et du choix de la base sur l'approximation du bassin d'attraction. Pour ce faire, nous développons une méthode numérique générale que nous appliquons pour des fonctions de base radiales et monomiales et que nous testons également en trois dimensions. Nous montrons que l'utilisation de ces fonctions fournit une bonne estimation du bassin d'attraction en particulier pour les fonctions de base radiales qui n'avaient pas encore été traitées dans ce domaine. Parallèlement, nous traitons certaines questions théoriques découlant de cette étude.

Mots-clé : Opérateur de Koopman, systèmes dynamiques non linéaires, stabilité, points fixes, bassin d'attraction, fonctions propres, fonctions de base, méthode numérique.

Abstract

In the context of non linear dynamical systems theory, studying the stability of fixed points is a challenging topic. Indeed, determining stability properties of such attractors usually requires the identification of a particular Lyapunov function among an infinite number of possible choices, which may become quite a difficult task. This then motivates a functional approach of dynamical systems by means of the Koopman operator. More particularly, fixed points stability relies on the existence of particular eigenfunctions of the operator. As an infinite-dimensional operator, computing its eigenfunctions is not particularly obvious and we have to find out some approximations of them onto a finite-dimensional subspace. Therefore, the approximation of the eigenfunctions is obtained by computing right eigenvectors of the matrix representation of the operator.

In this master thesis, we investigate in particular the impact that the projection and the choice of the basis have on the approximation of the basin of attraction. For this purpose, we develop a general numerical method that we apply for radial and monomial basis functions, that we also test for three dimensional systems. We show that the use of these functions provides a good estimate of the basin of attraction, especially for radial basis functions which have not been treated so far. In parallel, we address some theoretical issues arising from this study.

Keywords : Koopman operator, non linear dynamical systems, stability, fixed points, basin of attraction, eigenfunctions, basis functions, numerical method.

Table des matières

Introduction	1
1 Systèmes dynamiques et opérateur de Koopman	3
1.1 Rappels et propriétés des systèmes dynamiques	3
1.2 Opérateur de Koopman	11
1.3 Caractérisation de la stabilité via l'opérateur de Koopman	16
2 Application numérique	21
2.1 Approximation de \mathcal{L} en dimension finie	21
2.1.1 Propriétés spectrales de \mathbf{L}_N	22
2.1.2 Cas particulier : Champ de vecteurs analytique	24
2.2 Approximation du bassin d'attraction avec une base de monômes	27
2.2.1 Application pour des systèmes de plus haute dimension	31
3 Étude du choix des fonctions de base et impact sur la stabilité	37
3.1 Choix des fonctions de base et de la projection	38
3.1.1 Implémentation d'un code générique	40
3.2 Application de la méthode et étude des paramètres	42
3.2.1 Étude des fonctions de base <i>thin plate spline</i>	42
3.2.2 Généralisation du Principe d'invariance de LaSalle	47
3.2.3 Étude des fonctions radiales de Gauss	51
3.2.4 Application en plus haute dimension	53
3.2.5 Qualité de l'approximation et temps de calcul	56
4 Fiabilité de l'approximation du bassin d'attraction	59
4.1 Explication du problème et solution proposée	60
4.1.1 Remarque générale sur l'implémentation	65
4.1.2 Application numérique	65
4.2 Vers l'implémentation d'une grille adaptative	68
Conclusion et Perspectives	73

Introduction

Dans quelques domaines comme la biologie ou encore la physique, les systèmes dynamiques sont utilisés pour décrire l'évolution d'un *phénomène* au cours du temps tel que le mouvement des planètes dans le Système solaire ou encore l'évolution de deux populations en interaction. Dans ce dernier cas, il est important d'étudier dans quelles mesures ces populations peuvent cohabiter, ce qui nécessite alors d'étudier la stabilité globale du système dynamique en question. Dans le cadre de la mécanique céleste, les questions de stabilité sont également importantes. L'étude de la stabilité des systèmes linéaires est un sujet qui est déjà bien connu et assez simple à traiter comme nous le verrons dans le Chapitre 1. Cependant, l'étude de la stabilité globale et l'évaluation d'un bassin d'attraction deviennent beaucoup plus complexes lorsque le champ de vecteurs du système n'est pas linéaire. Dans le cadre du mémoire, nous allons nous restreindre à l'étude de la stabilité d'attracteurs particuliers que sont les points d'équilibre. Dans ce cas, nous verrons qu'il existe très peu de méthodes systématiques pour déterminer la stabilité globale de tels points et que nous devons généralement déterminer une fonction de Lyapunov particulière parmi une infinité de choix possibles, ce qui rend alors le processus fastidieux et complexe. Ce constat motivera alors une approche fonctionnelle des systèmes dynamiques via la théorie des opérateurs, et plus particulièrement via l'opérateur de Koopman. Ce dernier est en réalité connu depuis bien longtemps mais cela fait seulement quelques années qu'il est remis au goût du jour afin d'étudier le comportement des systèmes dynamiques. Dans le cadre de ce mémoire, nous nous intéresserons en particulier à son utilisation pour approximer le bassin d'attraction de points fixes.

La suite du mémoire est alors organisée de la manière suivante. Dans un premier temps, nous définirons le concept général de systèmes dynamiques à temps discret et continu. Dans le cas particulier des systèmes continus, nous poserons les bases de la théorie de la stabilité en nous concentrant plus particulièrement sur le cas particulier de points fixes [8]. Dans ce contexte, nous verrons que la question de la stabilité est loin d'être évidente et qu'elle repose en général sur la recherche d'une fonction particulière appelée fonction de Lyapunov. Dès lors, à l'aide des références [5, 6, 10], nous verrons que l'opérateur de Koopman est un outil intéressant et que ses fonctions propres permettent de calculer de manière systématique une fonction de Lyapunov permettant ainsi d'induire la stabilité de ces points fixes.

Bien que cet opérateur soit intéressant dans l'étude de la stabilité, nous montrerons que ce dernier est de dimension infinie de sorte que le calcul de ses fonctions propres est compliqué à déterminer analytiquement. Dans le chapitre 2 et sur base des différentes références [7, 11], nous verrons qu'il est possible d'approximer les fonctions propres en projetant l'opérateur de Koopman sur un sous-espace fini de fonctions de base. Dans ce contexte, nous développerons une méthode numérique générale permettant d'approximer

ces fonctions propres peu importe le choix du sous-espace et nous étudierons alors son application sur différents systèmes avec différentes fonctions de base. Plus particulièrement, nous considérerons dans un premier temps le cas particulier des monômes en suivant un raisonnement similaire aux références [5, 6]. Bien que la méthode n'ait été appliquée que pour des systèmes de dimension deux dans le cadre de monômes, nous étendrons notre étude dans le chapitre 3 avec des fonctions de base radiales en testant également la méthode pour des systèmes de dimension trois. Dans ce cas, nous verrons que les fonctions radiales fournissent des approximations du bassin d'attraction tout à fait acceptables et ouvrent alors une application de ces dernières pour des systèmes de dimension plus élevée. Parallèlement à cette étude, nous étudierons certaines questions théoriques sur l'approximation du bassin d'attraction. En particulier nous obtiendrons des résultats théoriques lorsque la fonction de Lyapunov ne décroît pas partout le long des trajectoires.

Enfin, dans le chapitre 4, nous développerons un critère permettant de garantir rigoureusement la stabilité sur l'approximation du bassin d'attraction obtenue. Ce critère est basé sur une borne du gradient de la dérivée de la fonction de Lyapunov et sur la taille de la grille d'évaluation de cette dernière.

Chapitre 1

Systemes dynamiques et operateur de Koopman

Ce chapitre constitue une brève introduction à l'étude de la stabilité des systèmes dynamiques et à l'opérateur de Koopman. Dans un premier temps, nous allons définir la notion de systèmes dynamiques à temps discret et continu. Ensuite, nous nous pencherons plus en détails sur la stabilité de systèmes continus et nous en verrons différentes caractérisations, à l'aide de fonctions de Lyapunov par exemple. Dans un second temps, nous introduirons le semi-groupe d'opérateurs de Koopman et son générateur infinitésimal ainsi que ses principales propriétés. Ce chapitre est essentiellement basé sur les références [3, 5, 6, 7].

1.1 Rappels et propriétés des systèmes dynamiques

Un système dynamique est une équation permettant de décrire l'évolution d'un état au cours du temps. Pour ne citer qu'un exemple, nous pouvons penser à l'évolution du nombre d'individus au sein d'une population en fonction de différents paramètres tels que l'absence ou non de nourriture, de prédateurs, etc. L'évolution de l'état peut être soit décrite à l'aide d'un temps dit *discret* ou *continu*. Penchons nous dans un premier temps sur les systèmes à temps discret. Dans la suite du travail, nous noterons $\mathbb{X} \subset \mathbb{R}^m$ l'espace des phases où $m \in \mathbb{N}$ est fini et nous supposons que \mathbb{X} est un ensemble compact. Nous justifierons ce choix plus tard dans le travail.

Définition 1.1 (Système dynamique à temps discret)

Soit l'application $\mathbf{T} : \mathbb{X} \rightarrow \mathbb{X}$. Un système dynamique **à temps discret** est un système qui évolue dans un temps discret et pour lequel l'état courant dépend de l'état précédent via l'application \mathbf{T} , autrement dit

$$\mathbf{x}(n+1) = \mathbf{T}(\mathbf{x}(n)), \tag{1.1}$$

où $n \in \mathbb{N}$ représente le temps discret et $\mathbf{x}(0) \in \mathbb{X}$ est la condition initiale.

Ainsi, pour n'importe quelle condition initiale \mathbf{x}_0 fixée, l'évolution de l'état au cours du temps est donnée par l'ensemble des points

$$\mathcal{O}(\mathbf{x}(0)) := \{\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(n), \dots\},$$

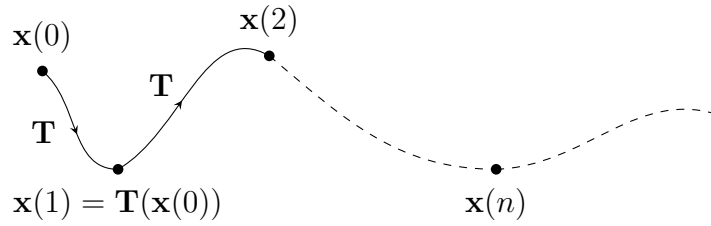


Figure 1.1 – Exemple d’une orbite issue d’une condition initiale \mathbf{x}_0 d’un système discret.

où $\mathcal{O}(\mathbf{x}(0))$ est ce qu’on appelle l’*orbite* du système (voir Figure 1.1). Par ailleurs, il n’est pas difficile de voir que cette orbite est uniquement déterminée par sa condition initiale étant donné que

$$\mathbf{x}(n) = \mathbf{T}(\mathbf{x}(n - 1)) = \mathbf{T}^n(\mathbf{x}(0)).$$

Dès lors, nous pouvons définir une fonction $\varphi : \mathbb{X} \times \mathbb{N} \rightarrow \mathbb{X}$ telle que pour tout $(\mathbf{x}_0, n) \in \mathbb{X} \times \mathbb{N}$, $\varphi^n(\mathbf{x}(0)) = \mathbf{T}^n(\mathbf{x}(0))$. Autrement dit, la fonction φ , généralement appelée le *flot* du système, permet de déterminer l’état au temps n à partir de la condition initiale $\mathbf{x}(0)$. Afin de rendre les choses plus claires, considérons l’exemple suivant.

Exemple 1.1 (Famille d’applications en tente). Considérons la famille des applications tente

$$\begin{aligned} \mathcal{G}_\mu &: [0, 1] \rightarrow [0, 1], \\ x &\mapsto \mathcal{G}_\mu(x) = \frac{\mu}{4} (1 - |2x - 1|), \end{aligned}$$

et considérons le système dynamique discret (1.1) avec $\mathbf{T} := \mathcal{G}_{\mu=4}$ et $\mathbb{X} = [0, 1]$. Dans ce cas, il vient que

$$x[n + 1] = 1 - |2x[n] - 1|. \tag{1.2}$$

Une illustration des trois premiers points de l’orbite issue de $x_0 = 0.9$ est donnée à la FIGURE 1.2. △

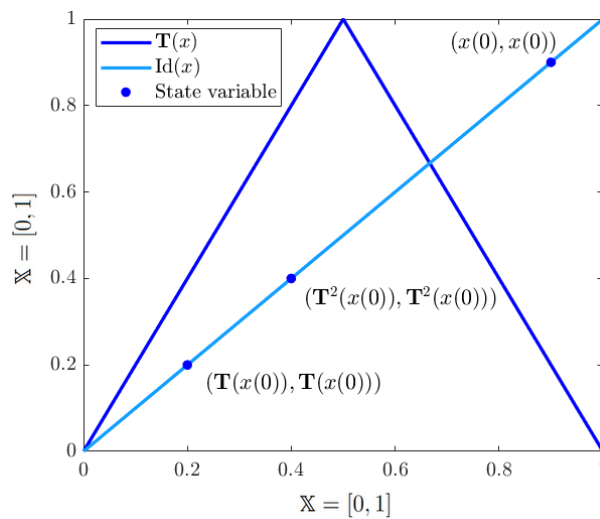


Figure 1.2 – Représentation du système (1.2). Les différents couples donnés par les points bleus sur la figure illustrent les coordonnées des trois premiers points de l’orbite. $\text{Id}(x) := x$ correspond à la fonction identité sur \mathbb{X} .

Précisons que les systèmes dynamiques à temps discret ont l'avantage, de par leur définition, de déterminer de manière exacte l'état du système en temps déterminé. Autrement dit, le flot est connu pour n'importe quelle condition initiale du système. En considérant les systèmes dynamiques en temps continu comme une généralisation des systèmes discrets, nous proposerions une définition analogue à celle donnée par la Définition 1.1. Cependant, lorsque le temps est défini sur un espace continu, le changement d'état entre deux temps successifs est plus flou et compliqué à formuler. En effet, dans un tel espace, il existe en réalité une infinité de valeurs entre deux temps quelconques t_1 et t_2 de sorte que la notion de temps successifs est infinitésimal. Dans ce cas, il devient alors compliqué de déterminer une expression de la forme (1.1). Ainsi, les systèmes dynamiques à temps continu sont définis par l'intermédiaire d'une équation différentielle ordinaire permettant de formaliser la variation de l'état de manière assez naturelle pour de tels systèmes.

Définition 1.2 (Système dynamique à temps continu)

Soit $\mathbf{F} : \mathbb{X} \times \mathbb{R}^+ \rightarrow \mathbb{X}$ et $(\mathbf{x}, t) \in \mathbb{X} \times \mathbb{R}^+$. Un système dynamique **à temps continu** est un système qui évolue dans un temps continu et qui est représenté par une équation différentielle ordinaire

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t).$$

Dans le cadre du mémoire, nous allons uniquement considérer des systèmes dynamiques de la forme

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t)), \quad (1.3)$$

c'est-à-dire, des systèmes *autonomes*. De manière assez semblable à ce qui est fait pour les systèmes discrets, nous pouvons définir le flot du système $\varphi : \mathbb{R}^+ \times \mathbb{X} \rightarrow \mathbb{X}$ tel que pour tout couple $(t, \mathbf{x}) \in \mathbb{R}^+ \times \mathbb{X}$, $\varphi(t, \mathbf{x}) = \mathbf{x}(t)$. Comme pour les systèmes discrets, le flot caractérise l'état au cours du temps. En effet, si nous fixons une condition initiale $\mathbf{x}_0 \in X$, nous voyons que $\varphi^t(\mathbf{x}_0)$ représente la trajectoire du système issue de \mathbf{x}_0 (voir Figure 1.3).

Exemple 1.2 (Système dynamique à temps continu en deux dimensions). Considérons le système dynamique

$$\begin{cases} \dot{x} &= -x, \\ \dot{y} &= -y. \end{cases} \quad (1.4)$$

Dans ce cas, pour une condition initiale $\mathbf{x}_0 = (x_0, y_0)$ fixée dans \mathbb{R}^2 , nous obtenons un problème de Cauchy assez simple à résoudre dont la solution livre

$$\begin{cases} x(t) &= e^{-t} x_0, \\ y(t) &= e^{-t} y_0. \end{cases}$$

Ainsi, le flot du système dynamique est donné par $\varphi^t(\mathbf{x}(0)) = e^{-t}\mathbf{x}(0)$. Notons qu'il est coutume de représenter un système dynamique par l'intermédiaire d'un *portrait de phase*. Ce dernier est un graphe particulier dans lequel nous représentons l'évolution de différentes conditions initiales au cours du temps, nous permettant ainsi d'avoir une première

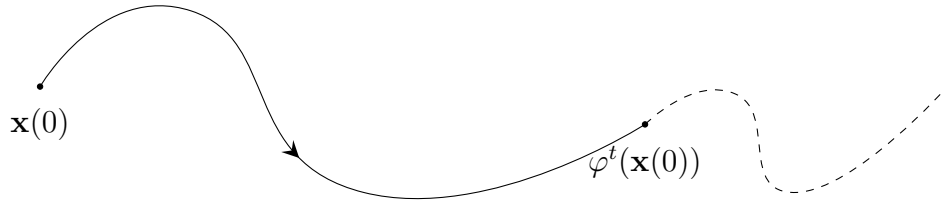


Figure 1.3 – Exemple d’une trajectoire issue d’une condition initiale x_0 d’un système continu. L’état $\mathbf{x}(0)$ est envoyé au temps t sur $\varphi(t, \mathbf{x}(0))$ que l’on notera $\varphi^t(\mathbf{x}(0))$ dans la suite.

compréhension générale du comportement du système. Ainsi, pour le système (1.4) nous remarquons à la Figure 1.4 que les trajectoires semblent toutes converger vers le point $(0, 0)$. Autrement dit, l’origine attire toutes les conditions initiales dans un voisinage plus ou moins grand. Nous verrons que cette notion est étroitement liée à la notion de stabilité établie plus tard dans le travail. Soulignons également que pour être tout à fait rigoureux, nous aurions dû considérer que le système évoluait dans le plus grand sous-ensemble compact contenu dans \mathbb{R}^2 .

△

Dans la suite, nous allons uniquement nous intéresser aux systèmes dynamiques à temps continu de la forme (1.3) où \mathbb{X} est de dimension $m \in \mathbb{N}$. Parmi ceux-ci, nous retrouvons les systèmes dynamiques linéaires, c’est-à-dire les systèmes de la forme

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad (1.5)$$

où $\mathbf{A} \in \mathbb{R}^{m \times m}$, comme c’est le cas pour le système (1.4) où $\mathbf{A} = -\mathbf{I}_{2 \times 2}$ avec $\mathbf{I}_{2 \times 2}$ la matrice identité 2×2 . Cependant, dans de nombreux cas, les applications \mathbf{T} et \mathbf{F} ne sont pas linéaires, ce qui rend les systèmes plus compliqués à étudier. Plus précisément, dans

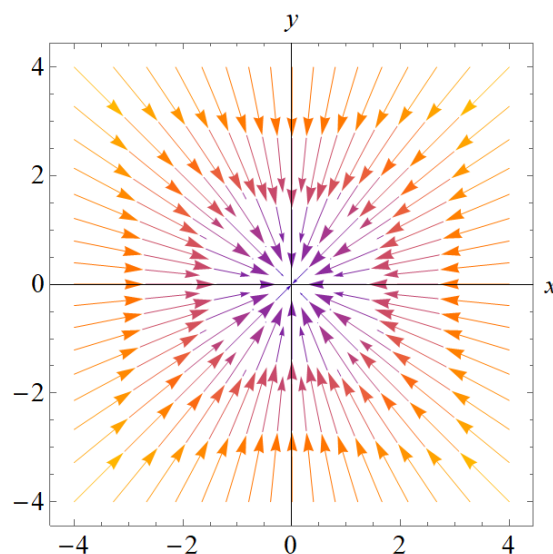


Figure 1.4 – Portraits de phase du système (1.4). Dans ce cas, nous constatons que l’origine du système attire un ensemble de conditions initiales dans son voisinage.

le cas particulier de la stabilité de systèmes dynamiques dont nous discutons ci-après, les propriétés spectrales de \mathbf{A} suffisent à induire la stabilité *globale* du système. Nous pouvons alors intuitivement comprendre qu'il est plus facile de traiter de tels systèmes. La complexité des systèmes non linéaires motive une approche fonctionnelle des systèmes dynamiques via la théorie des opérateurs, et plus particulièrement à l'aide de l'*opérateur de Koopman*. Dans le chapitre suivant, nous verrons en quoi cette approche est intéressante et quelles sont les conséquences que ce choix implique.

Définition 1.3 (Stabilité d'ensembles)

Considérons un ensemble A positivement invariant, c'est-à-dire tel que pour tout $t \geq 0$, $\varphi^t(A) \subseteq A$, et la distance

$$d(\mathbf{x}, A) = \min_{\mathbf{y} \in A} \|\mathbf{x} - \mathbf{y}\|,$$

où $\|\cdot\|$ dénote la norme euclidienne dans \mathbb{R}^m . Nous dirons que l'ensemble A est

(a) **stable** si pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$d(\mathbf{x}, A) < \delta \implies d(\varphi^t(\mathbf{x}), A) < \varepsilon, \quad \forall t \geq 0;$$

(b) **globalement uniformément attractif** dans X si pour tout $\varepsilon > 0$, il existe $T > 0$ tel que

$$t > T \implies d(\varphi^t(\mathbf{x}), A) < \varepsilon, \quad \forall \mathbf{x} \in X;$$

(c) **globalement (uniformément) asymptotiquement stable** dans X si A est stable et globalement uniformément attractif dans X .

Les notions de stabilité et d'attractivité uniforme sont reliées dans le sens où nous pouvons montrer que si un ensemble est uniformément attractif, alors il est stable [1, p. 15], mais la réciproque n'est généralement pas vérifiée comme le montre l'exemple suivant.

Exemple 1.3. Considérons le système dynamique suivant

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -x. \end{cases} \quad (1.6)$$

Au moyen du portrait de phase, nous remarquons (voir Figure 1.5) que $(0, 0)$ est un point fixe¹ stable du système au sens de la Définition 1.3. En effet, pour $\varepsilon > 0$ quelconque fixé, $\delta = \varepsilon$ convient. Cependant, il ne s'agit pas d'un ensemble globalement uniformément attractif, ce qui se vérifie assez directement.

△

Mis ensemble, les concepts de stabilité et d'attractivité uniforme vont nous permettre de définir le concept d'attracteur d'un système dynamique à temps continu.

1. Un point fixe du système (1.3) est un point $\mathbf{x}^* \in \mathbb{X}$ tel que $\mathbf{F}(\mathbf{x}^*) = 0$. Il s'agit d'un point particulier de l'espace des phases qui reste invariant sous l'action du flot.

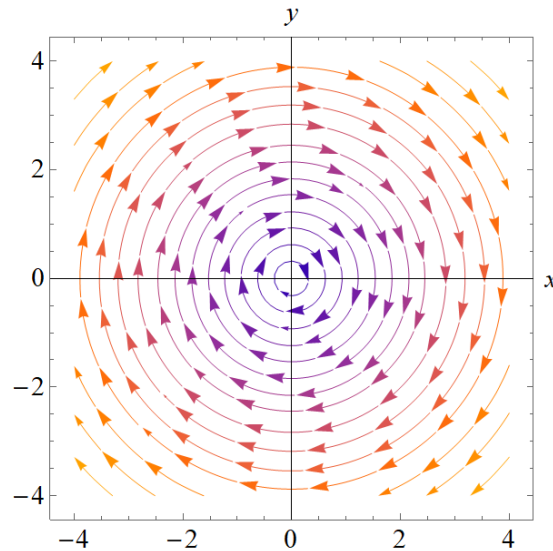


Figure 1.5 – Portrait de phase du système (1.6). Dans ce cas, nous constatons que l’origine du système est stable étant donné que $\delta = \varepsilon$ convient pour tout $\varepsilon > 0$.

Définition 1.4 (Attracteur)

Soit $A \subseteq \mathbb{X}$ positivement invariant. Nous dirons que A est un **attracteur** pour le système (1.3) si il existe un sous-ensemble $V \subseteq \mathbb{X}$ tel que $A \subset V$, et A est le plus petit ensemble globalement uniformément attractif dans V .

Par exemple, le système dynamique (1.4) admet un seul équilibre qui est globalement stable. Dans ce cas, il sera un attracteur pour le système puisqu’il s’agit du plus petit ensemble positivement invariant qui attire toutes les conditions initiales de l’espace des phases. Notons là qu’il s’agit d’un cas trivial d’attracteurs que sont les points fixes et pour de tels attracteurs, il est en général coutume de parler de leur *bassin d’attraction*.

Définition 1.5 (Bassin d’attraction)

Soit le système dynamique (1.2) dont $\mathbf{x}^* \in \mathbb{X}$ est un point fixe. Le **bassin d’attraction** de \mathbf{x}^* correspond à l’ensemble des points qui sont attirés par ce point fixe. Autrement dit, il s’agit du plus grand ensemble dans lequel le point fixe est globalement asymptotiquement stable.

Remarque 1.1. Les Définitions (1.3), (1.4) et (1.5) ne sont pas propres aux systèmes à temps continu et possèdent un équivalent pour les systèmes à temps discret mais nous n’en discuterons pas dans le cadre de ce mémoire.

Avant de conclure cette section introductive sur les systèmes dynamiques, nous allons rappeler quelques résultats classiques et bien connus de la théorie de la stabilité des systèmes dynamiques [8]. Penchons-nous dans un premier temps sur les systèmes dynamiques linéaires donnés par (1.5). Rappelons que dans ce cas, la matrice Jacobienne du système,

que l'on note \mathbf{J} , est donnée par

$$[\mathbf{J}]_{i,j=1}^m := \frac{\partial F_j(\mathbf{x})}{\partial \mathbf{x}_i} = \mathbf{A}_{ij}.$$

Cette matrice, ou plutôt son spectre, contient à lui seul toute l'information nécessaire pour induire la stabilité d'un tel système. A partir de maintenant et jusque le fin de ce mémoire, un ensemble *stable* désignera un ensemble globalement uniformément asymptotiquement stable.

Proposition 1.1 (Stabilité globale de systèmes linéaires)

Le système dynamique linéaire (1.5) est stable si $\operatorname{Re}(\lambda_j) < 0$ pour toutes les valeurs propres $\lambda_j \in \sigma(\mathbf{A})$ où $\sigma(\mathbf{A})$ désigne le spectre de la matrice Jacobienne.

Remarque 1.2. *Cette proposition est très intéressante car elle permet d'induire la stabilité de n'importe quel système sans passer par la définition, assez compliquée à établir en pratique.*

Notons que dans le cas particulier de systèmes dynamiques linéaires, nous parlons de la stabilité du système en général et non pas d'attracteurs en particulier. En réalité, ce n'est pas étonnant. En effet, il est assez immédiat que le vecteur nul de \mathbb{R}^N est un point fixe du système (1.5). Ainsi, deux cas de figure sont envisageables : Soit la matrice \mathbf{A} est inversible au quel cas 0 est le seul point fixe du système et nous ne perdons pas de généralité à parler de stabilité du système en général. Soit A est une matrice singulière et dans ce cas le système admet un *équilibre plan* formé par le sous-espace des vecteurs du noyau de A que l'on note $\mathcal{K}(\mathbf{A})$ dont l'origine fait partie. Dans ce cas, nous parlerons de la stabilité de cet équilibre plan. Ainsi, dans les deux cas de figure, il n'existera jamais deux points fixes isolés et c'est pourquoi nous ne perdons pas de généralité à discuter de la stabilité du système en lui-même. Nous référons à la source [8] pour plus de détails sur la stabilité des systèmes linéaires.

Lorsque le système n'est pas linéaire, nous devons spécifier de quel ensemble nous étudions la stabilité étant donné que dans ce cas, il peut exister plusieurs points fixes isolés par exemple (voir système (2.9)). En particulier, dans le cadre de ce mémoire, nous allons nous intéresser à la stabilité de tels points fixes isolés et la proposition suivante permet d'en donner une première caractérisation au moyen des valeurs propres de la matrice Jacobienne évaluée au point d'équilibre.

Proposition 1.2 (Stabilité locale de systèmes non linéaires)

Supposons que le système (1.3) admet un équilibre \mathbf{x}^* tel que $F(\mathbf{x}^*) = 0$ et notons \mathbf{J} , la matrice Jacobienne du système évaluée au point fixe. Dès lors, \mathbf{x}^* est localement stable si $\operatorname{Re}(\lambda_j) < 0$ pour toutes les valeurs propres λ_j de \mathbf{J} .

Comme nous l'avons mentionné, cette proposition nous permet d'avoir une idée du comportement local du point fixe mais nous ne savons, a priori, rien du comportement global de l'équilibre. À titre d'exemple, considérons le système dynamique de Van der Pol en temps inverse

$$\begin{cases} \dot{x}_1 &= -x_2, \\ \dot{x}_2 &= x_1 - x_2 + x_1^2 x_2, \end{cases} \quad (1.7)$$

dont l'origine est un point fixe. Comme nous pouvons le voir sur la Figure 1.6, ce système est caractérisé par un cycle limite instable (en rouge) délimitant les frontières du bassin d'attraction de l'équilibre à l'origine. En effet, nous remarquons que l'origine semble attirer toutes les conditions initiales à l'intérieur de cet ensemble. Par ailleurs, les valeurs propres de la matrice Jacobienne à l'équilibre sont données par $\lambda_1 = -0.5 + 0.87i$ et $\lambda_2 = -0.5 - 0.87i$, de sorte que $\mathbf{x}^* = 0$ est localement stable, ce qui confirme nos observations. Notons ici que le résultat permet d'assurer la stabilité locale du point fixe dans un voisinage de ce dernier et nous ne pouvons alors pas assurer que ce dernier est globalement stable dans l'ensemble défini par les frontières du cycle limite.

À l'aide de la Proposition 1.2, nous pouvons étudier le comportement local de point fixe mais nous ne savons, a priori, rien du comportement global de l'équilibre. En réalité, la stabilité globale de système dynamique non linéaire est plutôt difficile à établir. En effet, comme nous pouvons le voir avec l'exemple du système de Van der Pol, le bassin d'attraction est délimité par un cycle limite dont la géométrie est assez complexe. Cependant, à l'aide du Théorème de Lyapunov suivant, il est possible de caractériser la stabilité de tels systèmes en déterminant une fonction particulière appelée *fonction de Lyapunov*.

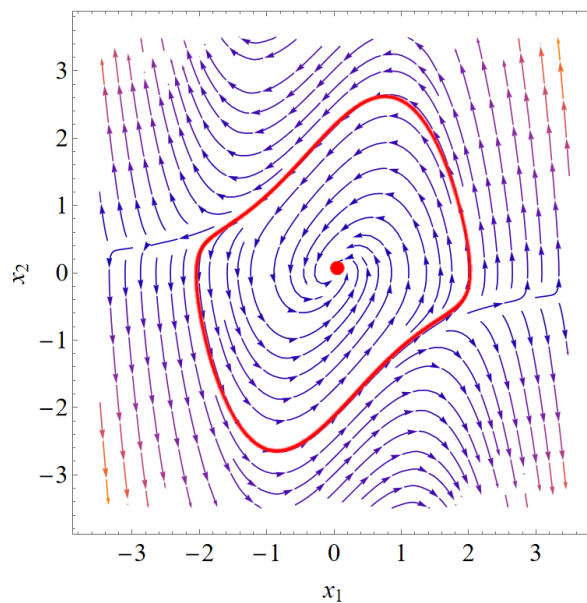


Figure 1.6 – Portrait de phase du système (1.7). Dans ce cas, nous constatons que l'origine est un point fixe globalement stable dans l'ensemble D délimité par le cycle limite tracé en rouge.

Théorème 1.1 (Stabilité globale de systèmes non linéaires)

Soit \mathbb{X} l'espace des phases du système dynamique (1.3) avec $\mathbf{F} \in C^1(\mathbb{X})$ possédant un équilibre $\mathbf{x}^* \in \mathbb{X}$. Supposons que $D \subseteq \mathbb{X}$ soit un sous-ensemble compact et positivement invariant tel que $\mathbf{x}^* \in D$. Dans ce cas, si il existe une fonction de Lyapunov, c'est-à-dire une fonction $V : D \rightarrow \mathbb{R}$ telle que

- (a) $V(\mathbf{x}^*) = 0$,
- (b) $V(\mathbf{x}) > 0, \quad \forall \mathbf{x} \in D \setminus \{\mathbf{x}^*\}$,
- (c) $V \in C^1(D)$ et $\dot{V}(\mathbf{x}) < 0, \quad \forall \mathbf{x} \in D \setminus \{\mathbf{x}^*\}$,

alors \mathbf{x}^* est globalement asymptotiquement stable dans D .

En réalité, déterminer une fonction de Lyapunov n'est pas une tâche aisée. Dans l'exemple du système de Van der Pol que nous avons considéré, nous pouvons intuitivement comprendre en quoi ce processus est fastidieux. En effet, pour ce système, nous devons déterminer une fonction positive mais décroissante le long des trajectoires pour toutes les conditions initiales à l'intérieur du cycle limite. Dès lors, au vu de la forme géométrique de ce dernier, nous comprenons qu'il n'est pas évident de trouver une telle fonction. Ce constat motive alors une approche fonctionnelle des systèmes par la théorie de l'opérateur de Koopman dont nous discuterons dans la section suivante. Dans la section 1.3 nous en verrons en quoi cet outil permet de surmonter ce problème.

1.2 Opérateur de Koopman

Comme nous l'avons mentionné, la section suivante sera dédiée à l'étude des systèmes dynamiques (et plus précisément leurs stabilités) via la théorie de l'opérateur de Koopman. Avant cela, nous devons introduire ce dernier. Pour ce faire, considérons un espace de Banach \mathcal{F} de fonctions $f : \mathbb{X} \rightarrow \mathbb{C}$ appelées des *observables*.

Définition 1.6 (Opérateur de Koopman)

Soit un système dynamique à temps continu (1.3) dont $\varphi^t(\mathbf{x})$ est le flot associé. Le **semi-groupe d'opérateurs de Koopman** $\{\mathcal{K}_t\}_{t \geq 0}$, associé au flot $\varphi^t(\cdot)$ et tel que $\mathcal{K}_t : \mathcal{F} \rightarrow \mathcal{F}$, est défini par

$$\mathcal{K}_t f = f \circ \varphi^t, \quad \forall f \in \mathcal{F}.$$

Dans la suite, nous supposerons que \mathbb{X} est positivement invariant pour que le semi-groupe d'opérateurs de Koopman soit bien défini. En effet, dans ce cas il vient que pour tout $\mathbf{x} \in \mathbb{X}$, $\varphi^t(\mathbf{x}) \in \mathbb{X}$ et nous avons alors que $(f \circ \varphi^t)(\mathbf{x})$ est bien défini.

Remarque 1.3. Dans le cas d'un système dynamique à temps discret défini par l'application \mathbf{T} , le semi-groupe d'opérateurs de Koopman est réduit à l'unique opérateur \mathcal{K} de sorte que $\mathcal{K}f = f \circ \mathbf{T}$ pour toute fonction $f \in \mathcal{F}$. Ceci n'est pas étonnant étant donné qu'un

système discret est uniquement déterminé par l'application \mathbf{T} . Sauf indication contraire, l'appellation « opérateur de Koopman » désignera le semi-groupe d'opérateurs ou l'unique opérateur de Koopman dans les systèmes continus ou discrets respectivement.

Dans le cadre du travail, nous allons considérer le semi-groupe d'opérateurs de Koopman défini sur l'espace des fonctions continues sur l'ensemble compact \mathbb{X} , que l'on notera $C^0(\mathbb{X})$. Cette condition assure que le semi-groupe d'opérateurs de Koopman est fortement continu [9, 10], nous permettant de définir son générateur infinitésimal.

Définition 1.7 (Générateur infinitésimal)

Soit Z un espace de Banach et $\{\mathcal{T}_t\}_{t \geq 0}$ un semi-groupe fortement continu où $\mathcal{T}_t : Z \rightarrow Z$. Le **générateur infinitésimal** de $\{\mathcal{T}_t\}_{t \geq 0}$, noté $\mathcal{L}_{\mathcal{T}}$, est défini par

$$\mathcal{L}_{\mathcal{T}}f = \lim_{t \rightarrow 0^+} \frac{\mathcal{T}_t f - f}{t},$$

pour tout $f \in D(\mathcal{L}_{\mathcal{T}})$, où $D(\mathcal{L}_{\mathcal{T}})$ correspond au domaine du générateur infinitésimal, c'est-à-dire l'ensemble des éléments $f \in Z$ pour lesquels la limite existe.

Dans le cas particulier de l'opérateur de Koopman défini sur $C^0(\mathbb{X})$, nous pouvons montrer [9] que le générateur est entièrement déterminé par le champ de vecteurs \mathbf{F} de sorte que

$$\mathcal{L}_{\mathcal{K}}f = \mathbf{F} \cdot \nabla f$$

pour toute fonction f continûment dérivable, c'est-à-dire $D(\mathcal{L}_{\mathcal{K}}) = C^1(\mathbb{X})$. Dans la suite et pour plus de facilité, nous noterons $\mathcal{L}_{\mathcal{K}} := \mathcal{L}$.

Remarque 1.4. *En général, déterminer le flot du système analytiquement est une tâche assez complexe et fastidieuse. Dans ce cas, nous travaillerons avec le générateur infinitésimal qui peut être obtenu plus facilement. Nous verrons également que travailler avec le générateur infinitésimal plutôt que le semi-groupe est en fait équivalent d'un point de vue de la stabilité de points fixes hyperboliques.*

L'opérateur de Koopman est particulièrement intéressant dans l'étude des systèmes dynamiques. Ce dernier permet d'étudier un système dynamique de dimension finie dans un espace *lifté* de fonctions de dimension infinie. Bien que ce changement nécessite d'étudier le système en dimension infinie, son avantage réside dans le fait que l'opérateur de Koopman est linéaire comme le montre la proposition suivante.

Proposition 1.3

Le semi-groupe d'opérateurs de Koopman donné par la Définition 1.6 et son générateur infinitésimal sont des opérateurs linéaires.

Démonstration. Soit $a, b \in \mathbb{R}$ et $f_1, f_2 \in \mathcal{F}$, montrons que pour tout $\mathbf{x} \in \mathbb{X}$,

$$\begin{cases} \mathcal{K}_t(af_1 + bf_2)(\mathbf{x}) &= a(\mathcal{K}_t f_1)(\mathbf{x}) + b(\mathcal{K}_t f_2)(\mathbf{x}), \\ \mathcal{L}(af_1 + bf_2)(\mathbf{x}) &= a(\mathcal{L} f_1)(\mathbf{x}) + b(\mathcal{L} f_2)(\mathbf{x}). \end{cases}$$

Soit $\mathbf{x} \in X$ quelconque arbitrairement fixé, nous avons que

• Opérateur de Koopman

$$\begin{aligned} [\mathcal{K}_t(af_1 + bf_2)](\mathbf{x}) &= (af_1 + bf_2)(\varphi^t(\mathbf{x})), \\ &= af_1(\varphi^t(\mathbf{x})) + bf_2(\varphi^t(\mathbf{x})), \\ &= a(\mathcal{K}_t f_1)(\mathbf{x}) + b(\mathcal{K}_t f_2)(\mathbf{x}). \end{aligned}$$

• Générateur infinitésimal

$$\begin{aligned} [\mathcal{L}(af_1 + bf_2)](\mathbf{x}) &= F \cdot [\nabla(af_1 + bf_2)(\mathbf{x})], \\ &= F \cdot [a(\nabla f_1)(\mathbf{x}) + b(\nabla f_2)(\mathbf{x})], \\ &= a(\mathcal{L} f_1)(\mathbf{x}) + b(\mathcal{L} f_2)(\mathbf{x}). \end{aligned}$$

□

Ainsi, l'opérateur de Koopman permet de concevoir une approche linéaire d'un système dynamique non linéaire, où nous savons que la stabilité est plus simple à établir. Étant donné que le semi-groupe d'opérateurs de Koopman ainsi que son générateur sont des opérateurs linéaires, il semble assez naturel d'étudier leurs spectres. Dans la suite du mémoire, nous considérerons uniquement le spectre ponctuel des deux opérateurs, c'est-à-dire

$$\begin{aligned} \sigma_p(\mathcal{K}_t) &= \{ \mu \in \mathbb{C} \mid \mathcal{K}_t - \mu I \text{ est non injectif} \}, \\ \sigma_p(\mathcal{L}) &= \{ \lambda \in \mathbb{C} \mid \mathcal{L} - \lambda I \text{ est non injectif} \}, \end{aligned}$$

où les éléments μ et λ sont respectivement appelés les *valeurs propres* de \mathcal{K}_t et \mathcal{L} . Ainsi, nous pouvons définir le concept de *fonctions propres de ces opérateurs*. Considérons dans un premier le spectre ponctuel du générateur infinitésimal \mathcal{L} .

Définition 1.8 (Fonctions propres du générateur infinitésimal)

Soit \mathcal{L} le générateur infinitésimal du semi-groupe d'opérateurs de Koopman $\{\mathcal{K}_t\}_{t \geq 0}$ et soit $\phi_\lambda \in \mathcal{F}$. Nous dirons que ϕ_λ est une **fonction propre** du générateur infinitésimal, associée à la **valeur propre** $\lambda \in \mathbb{C}$ si et seulement si

$$\mathcal{L}\phi_\lambda = \lambda\phi_\lambda. \tag{1.8}$$

En réalité, il s'avère que les spectres ponctuels du semi-groupe $\{\mathcal{K}_t\}_{t \geq 0}$ et de son générateur \mathcal{L} sont liés par le théorème de l'application spectrale [2, Théorème 2.6, p. 180] de la manière suivante :

$$\sigma_p(\mathcal{K}_t) \setminus \{0\} = \exp(\sigma_p(\mathcal{L})t), \quad \forall t \geq 0. \tag{1.9}$$

Dès lors, considérer le spectre du générateur ou celui du semi-groupe est équivalent. De plus, par la relation (1.9), si ϕ_λ est une fonction propre de \mathcal{L} associée à la valeur propre $\lambda \neq 0$, nous avons que

$$\mathcal{K}_t \phi_\lambda = e^{\lambda t} \phi_\lambda, \quad (1.10)$$

et λ désignera, par abus de langage, la valeur propre de l'opérateur de Koopman associée à la fonction propre ϕ_λ .

Remarque 1.5. *Par la relation (1.10), nous voyons que les différentes valeurs propres de l'opérateur de Koopman dépendent du temps t , ce qui n'est pas le cas de \mathcal{L} .*

Exemple 1.4 (Fonction propre de Koopman). Considérons le système dynamique unidimensionnel

$$\dot{x} = -x + x^2. \quad (1.11)$$

Dans ce cas, nous pouvons assez simplement montrer que la fonction

$$\phi_\lambda(x) = \frac{x}{1-x}, \quad (1.12)$$

est une fonction propre de l'opérateur associée à la valeur propre $\lambda = -1$. Étant donné que le spectre de \mathcal{K}_t et \mathcal{L} sont liés par le théorème de l'application spectrale, nous pouvons montrer de manière équivalente qu'il s'agit d'une fonction propre de \mathcal{L} . Il vient alors que

$$\begin{aligned} \mathcal{L}\phi_\lambda &= (-x + x^2) \left[\frac{d}{dx} \left(\frac{x}{1-x} \right) \right], \\ &= (-x + x^2) \frac{1}{(1-x)^2}, \\ &= x(-1 + x) \frac{1}{(1-x)^2}, \\ &= -\frac{x}{(1-x)}, \\ &= \lambda \phi_\lambda(x). \end{aligned}$$

Notons que nous pouvons déjà intuitivement comprendre en quoi les fonctions propres de l'opérateur de Koopman sont intéressantes pour la stabilité du système. En effet, dans cet exemple, il vient ici que $x = 0$ est un point fixe globalement stable sur l'intervalle $(-\infty, 1)$ comme observé à la Figure 1.7. Par ailleurs, nous remarquons que la fonction propre que nous avons calculée diverge à la limite du bassin d'attraction, ce qui n'est pas étonnant comme nous le verrons plus tard (voir Théorème 1.6). Ainsi, nous avons une première intuition de l'information de stabilité contenue dans les fonctions propres de l'opérateur de Koopman.

△

Afin de terminer l'introduction du spectre ponctuel de l'opérateur de Koopman, notons que par définition, si nous connaissons une fonction propre de l'opérateur, nous en connaissons une infinité comme le prouve la proposition suivante [6].

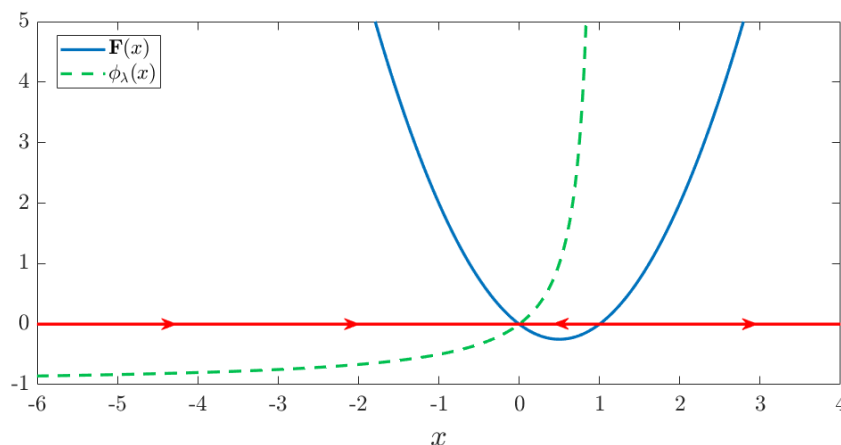


Figure 1.7 – Représentation du système (1.11) et de la fonction propre (1.12). Pour les systèmes unidimensionnels, il est courant de représenter le portrait de phase comme l'évolution de $\dot{\mathbf{x}}$ en fonction de \mathbf{x} . Dans ce cas, la ligne rouge fléchée représente le sens des différentes trajectoires. Nous voyons alors que l'origine est globalement stable sur l'intervalle $(-\infty, 1)$ et que la fonction propre explose à la frontière de ce dernier.

Proposition 1.4

Supposons que ϕ_{λ_1} et ϕ_{λ_2} sont deux fonctions propres de l'opérateur de Koopman associées aux valeurs propres λ_1 et λ_2 respectivement. Si $\phi_{\lambda_1}^{k_1} \phi_{\lambda_2}^{k_2} \in \mathcal{F}$, avec $k_1, k_2 \in \mathbb{R}$, alors $\phi_{\lambda_1}^{k_1} \phi_{\lambda_2}^{k_2}$ est une fonction propre associée à la valeur propre $k_1 \lambda_1 + k_2 \lambda_2$.

Démonstration. Par la relation (1.8) et la Définition 1.6, il vient que pour tout $x \in X$

$$\begin{aligned}
 [\mathcal{K}_t(\phi_{\lambda_1}^{k_1} \phi_{\lambda_2}^{k_2})](\mathbf{x}) &= (\phi_{\lambda_1}^{k_1} \phi_{\lambda_2}^{k_2})(\varphi^t(\mathbf{x})), \\
 &= (\phi_{\lambda_1}^{k_1}[\varphi^t(\mathbf{x})])(\phi_{\lambda_2}^{k_2}[\varphi^t(\mathbf{x})]), \\
 &= (\phi_{\lambda_1}[\varphi^t(\mathbf{x})])^{k_1} (\phi_{\lambda_2}[\varphi^t(\mathbf{x})])^{k_2}, \\
 &= e^{(k_1 \lambda_1 + k_2 \lambda_2)t} (\phi_{\lambda_1}^{k_1} \phi_{\lambda_2}^{k_2})(\mathbf{x}).
 \end{aligned}$$

□

Par la Proposition 1.4, il vient que si $\phi_{\lambda_i}^{k_i} \phi_{\lambda_j}^{k_j} \in \mathcal{F}$ pour $k_i = \lambda_j$ et $k_j = -\lambda_i$, alors il s'agit d'une fonction propre associée à la valeur propre $\lambda = 0$. Ceci est intéressant puisque ces fonctions propres sont constantes sous l'action du flot (par la relation (1.8)) et leurs courbes de niveau définissent alors des ensembles positivement invariants. Dans ce cas, les intersections non-dégénérées des courbes de niveau de $m - 1$ fonctions propres indépendantes associées à la valeur propre nulle donnent une famille d'ensembles de dimension 1 correspondant aux orbites du système dynamique. Cette propriété permet de montrer que les fonctions propres de l'opérateur de Koopman sont directement liées au système dynamique.

1.3 Caractérisation de la stabilité via l'opérateur de Koopman

Comme nous venons de le voir à travers l'Exemple 1.4, il semblerait que les fonctions propres de l'opérateur de Koopman contiennent de l'information sur la stabilité du système dynamique dont il est issu, mais nous ne savons pas encore comment. Cette section a pour objectif d'éclaircir ce point en considérant les différents arguments clés des références [4, 3, 5, 10]. Sauf indication contraire, nous supposons que $A \subseteq \mathbb{X}$ est un attracteur du système dynamique (1.3) et nous noterons

$$\mathcal{F}_{A_c} := \{f \in \mathcal{F} \mid f(\mathbf{x}) = 0, \forall \mathbf{x} \in A\}. \quad (1.13)$$

Remarquons que cet ensemble est invariant sous l'action de l'opérateur de Koopman, autrement dit $\mathcal{K}_t(\mathcal{F}_{A_c}) \subseteq \mathcal{F}_{A_c}$. En effet, pour tout $\mathbf{x} \in A$, nous avons que $\varphi^t(\mathbf{x}) \in A$ pour $t \geq 0$ puisque A est positivement invariant. Dès lors, il vient que

$$\mathcal{K}_t(f(\mathbf{x})) = f(\varphi^t(\mathbf{x})) = 0.$$

Dans la suite, nous noterons $\mathcal{K}_{A_c}^t : \mathcal{F}_{A_c} \rightarrow \mathcal{F}_{A_c}$ la restriction de l'opérateur de Koopman au sous-ensemble $\mathcal{F}_{A_c} \subseteq \mathcal{F}$.

Remarque 1.6. *Supposons que $A = \{\mathbf{x}^*\}$ ou \mathbf{x}^* est un point fixe du système (1.3). Dès lors, une fonction de Lyapunov n'est rien d'autre qu'une observable $\mathcal{V} \in \mathcal{F}_{A_c}$ telle que $\mathcal{L}\mathcal{V}(\mathbf{x}) < 0$ ou encore $\mathcal{K}_{A_c}^t \mathcal{V}(\mathbf{x}) < \mathcal{K}_{A_c}^s \mathcal{V}(\mathbf{x})$ pour tout $s < t < 0$ et $\mathbf{x} \in \mathbb{X}$.*

La restriction de l'opérateur de Koopman $\mathcal{K}_{A_c}^t$ au sous-ensemble des fonctions qui s'annulent en A permet de capturer des propriétés de stabilité du système dont il est associé.

Proposition 1.5 [6]

Soit $A \subset \mathbb{X}$ un attracteur pour le système (1.3). Nous avons alors que A est globalement attractif dans \mathbb{X} si et seulement si

$$\lim_{t \rightarrow \infty} \mathcal{K}_{A_c}^t f = 0, \quad \forall f \in \mathcal{F}_{A_c}$$

où $\mathcal{F} = C^0(\mathbb{X})$.

Ce résultat est très général et semble compliqué à mettre en pratique puisqu'il nécessite de se concentrer sur l'espace \mathcal{F}_{A_c} tout entier. Cependant, nous avons vu à travers la Proposition 1.3 que l'opérateur de Koopman est linéaire et que nous pouvons alors considérer son spectre. Ainsi, nous allons voir qu'en se concentrant uniquement sur les fonctions propres de l'opérateur, il est possible d'établir certains résultats sur la stabilité du système (1.3). Le théorème suivant en est un exemple.

Théorème 1.2 (Stabilité globale d'attracteurs, [6])

Considérons un sous-ensemble fini de valeurs propres $\Lambda \subseteq \sigma_p^-(L_U)$ tel que $\phi_\lambda \in C^0(X)$ pour tout $\lambda \in \Lambda$ avec $X \subset \mathbb{X}$. Alors, l'intersection des ensembles de niveau zéro

$$M := \bigcap_{\lambda \in \Lambda} \{\mathbf{x} \in X \mid \phi_\lambda(\mathbf{x}) = 0\} \quad (1.14)$$

est positivement invariant sous le flot φ^t et globalement asymptotiquement stable.

Essayons de comprendre en quoi le théorème précédent est intéressant. Supposons que l'ensemble $A \subset \mathbb{X}$ soit un ensemble positivement invariant, localement stable mais que nous ne savons rien de la stabilité globale de ce dernier. Supposons à présent que les dimensions respectives du système et de l'ensemble A sont données par m et q . Dès lors, si il existe $m - q$ fonctions propres différentes ϕ_{λ_i} telles que $\operatorname{Re}(\lambda) < 0$ pour tout $i = 1, \dots, m - q$ avec $M = A$, il vient que A est globalement stable dans Y , ce qui nous permet d'étendre notre intuition.

Exemple 1.5 (Application du Théorème 1.2). Considérons le système dynamique unidimensionnel $\dot{x} = -x^3$ où $x \in X := \mathbb{R}$. Dans ce cas, il est clair que $x^* = 0$ est un point fixe du système et que

$$\mathbf{J} = \left. \frac{dF(x)}{dx} \right|_{x=x^*} = 0.$$

Ainsi, nous n'avons aucune intuition sur la stabilité a priori de ce point fixe (voir Proposition 1.2). Cependant, comme nous pouvons le voir à la Figure 1.8, ce point fixe semble tout de même globalement stable dans \mathbb{R}^2 et nous aimerions confirmer cette intuition. Pour ce faire, remarquons que

$$\phi_\lambda(x) = \exp\left(\frac{-1}{x^2}\right)$$

est une fonction propre de l'opérateur de Koopman associée à la valeur propre $\lambda = -2$. En effet, nous pouvons montrer que

$$\begin{aligned} \mathcal{L}\phi_\lambda(x) &= -x^3 \left[\frac{2}{x^3} \exp\left(\frac{-1}{x^2}\right) \right], \\ &= \lambda \phi_\lambda(x). \end{aligned}$$

Dès lors, par le Théorème 1.2, il vient que l'ensemble

$$\{x \in X \mid \exp\left(\frac{-1}{x^2}\right) = 0\} = \{0\}$$

est asymptotiquement stable dans X et nous avons alors montré la stabilité globale de l'origine. △

2. Notons que pour être rigoureux, nous devrions considérer X comme étant le plus grand ensemble compact dans \mathbb{R} .

Dans le cadre de ce mémoire, nous allons en particulier nous intéresser aux attracteurs spécifiques que sont les points fixes et plus particulièrement aux points fixes hyperboliques³. Dans ce cas, la proposition suivante [6, 10] permet de montrer que les propriétés de stabilité locale d'attracteurs sont étendues globalement par le biais des fonctions propres de Koopman.

Proposition 1.6 (Stabilité globale de points fixes)

Soit $Y \subset \mathbb{R}^m$, un ensemble connexe, compact et positivement invariant. Supposons que le système dynamique (1.3) admet un point fixe $\mathbf{x}^* \in Y$ où $\mathbf{F} \in C^2(Y)$ et tel que la matrice Jacobienne \mathbf{J} au point fixe soit diagonalisable. Alors, l'équilibre est hyperbolique et globalement uniformément asymptotiquement stable dans Y si et seulement si l'opérateur de Koopman associé à ce système admet m fonctions propres $\phi_{\lambda_j} \in C^1(Y)$ avec $\text{Re}(\lambda_j) < 0$ et $\nabla \phi_{\lambda_j}(\mathbf{x}^*) \neq 0$. De plus, les valeurs propres λ_j correspondent aux valeurs propres de J .

Nous remarquons que la Proposition 1.6 reflète des résultats de stabilité locale (voir Proposition 1.2) qui sont valables globalement. De plus, tandis que la stabilité locale repose sur m valeurs propres de la matrice Jacobienne, la stabilité globale repose sur m fonctions propres de l'opérateur de Koopman. Notons par ailleurs que ces fonctions propres, ou plutôt leur support, permet de délimiter le *bassin d'attraction* du point fixe. En effet, par la relation (1.10), il vient que ces fonctions explosent de manière exponentielle à la limite du bassin d'attraction de l'équilibre. Notons que nous travaillons avec des points fixes hyperboliques étant donné que ces derniers nous permettent d'avoir une première intuition de la stabilité du système par la Proposition 1.2.

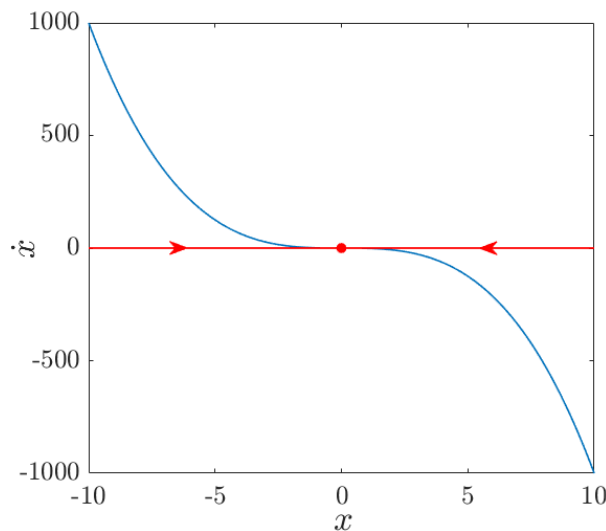


Figure 1.8 – Portrait de phase du système unidimensionnel $\dot{x} = -x^3$. Dans ce système, nous remarquons que l'origine semble globalement stable dans $X = \mathbb{R}$.

3. Nous dirons qu'un point fixe \mathbf{x}^* est hyperbolique si les valeurs propres $\{\lambda_i\}_{i=1}^n$ de la matrice Jacobienne au point fixe sont telles que $\text{Re}(\lambda_i) \neq 0$ pour tout $i = 1, \dots, n$.

Exemple 1.6. Si nous revenons sur le système dynamique (1.11), nous voyons que la fonction propre ϕ_λ admet une singularité à la limite du bassin d'attraction, c'est-à-dire en $x = 1$. Notons que cet exemple rentre dans les conditions de la proposition étant donné que $\mathbf{J}(0) = 0$ et que

$$\left. \frac{d}{dx} (\phi_\lambda(x)) \right|_{x=0} = 1 (\neq 0).$$

△

Remarque 1.7. Lorsque le point d'équilibre \mathbf{x}^* n'est pas hyperbolique, nous ne pouvons plus utiliser la Proposition 1.6. Cependant, dans certains cas, il se pourrait que l'usage du Théorème 1.2 nous fournisse tout de même des informations sur la stabilité du point fixe comme dans l'Exemple 1.5 où $\mathbf{J} = 0$. Dans ce cas, nous ne pouvons pas utiliser la Proposition 1.6 mais nous avons tout de même pu induire la stabilité globale du point d'équilibre grâce au Théorème 1.2.

Considérons le cas particulier des systèmes linéaires (1.5). Dans ce cas, nous pouvons montrer que $\phi_{\lambda_j}(\mathbf{x}) = \mathbf{x}^T \mathbf{w}_j$ est une fonction propre de l'opérateur de Koopman où \mathbf{w}_j correspond au vecteur propre à gauche de J associé à la valeur propre λ_j . En effet, nous avons que

$$\begin{aligned} (\mathcal{K}_t \phi_{\lambda_j})(\mathbf{x}) &= \phi_{\lambda_j}(\varphi^t(\mathbf{x})), \\ &= (\varphi^t(\mathbf{x}))^T \mathbf{w}_j. \end{aligned}$$

En utilisant le développement en série de l'exponentielle matricielle et le fait que \mathbf{w}_j est un vecteur propre à gauche de J , il vient que

$$\begin{aligned} (\mathcal{K}_t \phi_{\lambda_j})(\mathbf{x}) &= \mathbf{x}^T (e^{\lambda_j t}) \mathbf{w}_j, \\ &= e^{\lambda_j t} \phi_{\lambda_j}(\mathbf{x}). \end{aligned}$$

Dès lors, il est clair que $\phi_{\lambda_j} \in C^1(\mathbb{R}^N)$ et que $\nabla \phi_{\lambda_j}(\mathbf{x}^*) = \mathbf{w}_j \neq 0$. Dans ce cas, la Proposition 1.6 implique que le système est globalement stable si $\text{Re}(\lambda_j) < 0$ pour tout j , et nous retrouvons alors le résultat de la Proposition 1.1.

Comme nous avons pu le voir précédemment, la stabilité globale d'un point fixe \mathbf{x}^* peut également être établie en définissant une fonction de Lyapunov $\mathcal{V}(\mathbf{x})$ particulière. Dans ce cas, le support de cette dernière permet lui aussi de délimiter les frontières du bassin d'attraction, défini dans ce cas comme le plus grand ensemble de niveau pour lequel $\dot{\mathcal{V}}(\mathbf{x}) < 0$. En réalité, les fonctions propres de la Proposition 1.6 à parties réelles strictement négatives permettent de construire une telle fonction de Lyapunov de la manière suivante :

$$\mathcal{V}(x) = \left(\sum_{i=1}^m c_i |\phi_{\lambda_i}(x)|^p \right)^{1/p}, \quad (1.15)$$

avec $c_i > 0$, pour un certain entier $p \geq 1$. Nous vérifions assez facilement qu'il s'agit bien d'une fonction de Lyapunov via la proposition suivante [5, p. 5236, Propriété 2].

Proposition 1.7

Si $\phi_\lambda \notin \mathcal{F}_{A_c}$ est une fonction propre de l'opérateur de Koopman telle que $\phi_\lambda \in C^0(\mathbb{X})$ où $A \subset \mathbb{X}$ est un attracteur globalement asymptotiquement stable, alors $\operatorname{Re}(\lambda) = 0$.

Dès lors, en utilisant la contraposée de cette proposition, il vient que $\mathcal{V}(x)$ est une fonction positive qui s'annule au point fixe. De plus, nous vérifions également que

$$U^t V(x) = \left(\sum_{j=1}^m |U^t \phi_{\lambda_j}(x)|^p \right)^{1/p} = \left(\sum_{j=1}^m e^{p \operatorname{Re}\{\lambda_j\}t} |\phi_{\lambda_j}(x)|^p \right)^{1/p} < e^{\operatorname{Re}\{\lambda_1\}t} V(x),$$

où λ_1 est la valeur propre dont la partie réelle est la moins négative. Nous verrons plus tard en quoi il est intéressant de passer par la fonction de Lyapunov, notamment dans les approximations numériques.

Chapitre 2

Approximation numérique du bassin d'attraction

Dans le chapitre précédent, nous avons investigué la question de la stabilité d'attracteur de systèmes dynamiques autonomes de la forme (1.2). Dans le cadre du mémoire, et sauf indication contraire, nous allons nous intéresser au cas particulier de points fixes hyperboliques. Dans ce cas, nous avons vu que les fonctions propres de l'opérateur de Koopman, associées à des valeurs propres à parties réelles strictement négatives, permettaient d'induire des résultats de stabilité de tels attracteurs. Les Exemples 1.4 et 1.5 nous ont d'ailleurs permis d'appliquer ces différents résultats en considérant certaines fonctions propres particulièrement bien choisies. Cependant, il n'est pas toujours évident de déterminer de telles fonctions propres. En effet, comme nous l'avons mentionné au préalable dans le mémoire, l'opérateur de Koopman est défini sur un espace de fonctions (continues) de dimension infinie. Ainsi, toute fonction propre de l'opérateur est identifiée par une infinité de coefficients et nous pouvons alors intuitivement comprendre pourquoi il est compliqué d'en déterminer une expression analytique. Pour pallier ce problème et malgré tout étudier la stabilité du système, nous allons approximer ces fonctions et analyser dans quelle mesure ces approximations nous permettent quand même d'induire des résultats de stabilité.

2.1 Approximation de \mathcal{L} en dimension finie

Dans cette section, nous allons voir comment l'opérateur le générateur infinitésimal \mathcal{L} peut être approximé en dimension finie afin de construire une approximation des fonctions propres pour étudier la stabilité du système. Pour ce faire, nous allons essentiellement nous baser sur les références [6, 7] que nous adapterons au cas du générateur.

Soit \mathcal{F} l'espace de fonctions sur lequel le générateur infinitésimal du semi-groupe d'opérateurs de Koopman est défini et le sous-espace $\mathcal{F}_N \subset \mathcal{F}$ engendré par des *fonctions de base* $\{\psi_j(\mathbf{x})\}_{j=1}^N$. Considérons également un opérateur de *projection* $\Pi : \mathcal{F} \rightarrow \mathcal{F}_N$, associant (ou projetant) n'importe quelle fonction $f \in \mathcal{F}$ dans le sous-espace \mathcal{F}_N de dimension finie. Dès lors, pour un sous-espace \mathcal{F} et un opérateur de projection Π fixés, une approximation en dimension finie de \mathcal{L} est donné par

$$\mathcal{L}_N := \Pi \mathcal{L}|_{\mathcal{F}_N} : \mathcal{F}_N \rightarrow \mathcal{F}_N,$$

où nous rappelons que $\mathcal{L}|_{\mathcal{F}_N}$ est la restriction du générateur infinitésimal au sous-espace \mathcal{F}_N . Soit $f \in \mathcal{F}_N$ une fonction arbitraire du sous-espace \mathcal{F}_N . Dès lors, il vient que

$$f(\mathbf{x}) = \sum_{i=1}^N a_i \psi_i(\mathbf{x}) := \mathbf{a}^T \Psi(\mathbf{x}) \quad \text{et} \quad (\mathcal{L}^N f)(\mathbf{x}) = \sum_{i=1}^N b_i \psi_i(\mathbf{x}) := \mathbf{b}^T \Psi(\mathbf{x}).$$

où $\mathbf{a} := (a_1, \dots, a_N)$ et $\mathbf{b} := (b_1, \dots, b_N)$ sont respectivement les coefficients de f et $\mathcal{L}^N f$ dans la base représentée par le vecteur $\Psi(\mathbf{x}) = (\psi_1(\mathbf{x}), \dots, \psi_N(\mathbf{x}))$. Ainsi, nous pouvons définir la matrice $\mathbf{L}_N \in \mathbb{R}^{N \times N}$ [11] telle que

$$\mathbf{L}_N \mathbf{a} = \mathbf{b}.$$

Dans ce cas, la matrice \mathbf{L}_N est une représentation du générateur infinitésimal \mathcal{L} projeté dans le sous-espace de fonctions \mathcal{F}_N . En effet, nous vérifions assez simplement que

$$[\mathcal{L}_N f](\mathbf{x}) = \mathcal{L}_N[\mathbf{a}^T \Psi(\mathbf{x})] = (\mathbf{L}_N \mathbf{a})^T \Psi(\mathbf{x}). \quad (2.1)$$

Ainsi, la matrice \mathbf{L}_N représente l'action du générateur sur la fonction $f \in \mathcal{F}_N$, exprimée dans la base $\{\psi_j\}_{j=1}^N$. Cependant, nous ne savons pas très bien à quoi correspondent les entrées de cette matrice. Afin d'éclaircir ce point, constatons que l'expression obtenue en (2.1) est générale et valide peu importe la fonction $f \in \mathcal{F}_N$. Ainsi, en particulier pour les fonctions de base $f = \psi_i$ il vient que

$$[\mathcal{L}_N \psi_i](\mathbf{x}) = (\mathbf{L}_N \mathbf{e}_i)^T \Psi(\mathbf{x}),$$

où nous avons utilisé le fait que $\mathbf{a} = \mathbf{e}_i$ est le i -ème vecteur de base de l'espace euclidien de \mathbb{R}^N . Nous remarquons alors que la i -ème colonne de la matrice \mathbf{L}_N contient les coefficients de $\mathcal{L}_N \psi_i$ exprimés dans la base de fonctions $\{\psi_j\}_{j=1}^N$.

Remarque 2.1. *Notons que le raisonnement que nous venons d'établir ne fait pas intervenir en particulier des propriétés du générateur infinitésimal et un raisonnement tout à fait similaire peut être établi pour l'opérateur de Koopman. En particulier, la référence [11] détail typiquement le cas particulier de l'opérateur de Koopman que nous avons adapté pour le générateur.*

Dans le cadre du mémoire, nous supposons que nous connaissons le champ de vecteurs $\mathbf{F}(\mathbf{x})$ du système dynamique étudié de sorte qu'il est beaucoup plus avantageux de travailler avec le générateur infinitésimal dont nous connaissons l'expression analytique. En effet, l'opérateur de Koopman nécessite le calcul du flot qui n'est pas toujours simple à extraire. Par souci de facilité, nous considérerons alors \mathbf{L}_N dans la suite du mémoire.

2.1.1 Propriétés spectrales de \mathbf{L}_N

Considérer une approximation de l'opérateur de Koopman en dimension finie implique également que l'on obtient une approximation de son spectre ponctuel et donc de ses valeurs propres. Remarquons dans un premier temps que les valeurs propres de \mathcal{L}_N et de \mathbf{L}_N coïncident¹. En effet, soit $\phi_{\lambda_j} \in \mathcal{F}_N$ de sorte que $\phi_{\lambda_j} = \mathbf{a}^T \Psi(x)$ pour un certain

1. Ici, il vient que $\mathcal{L}_N = \Pi \mathcal{L} |_{\mathcal{F}_N}$ avec $\mathcal{L} |_{\mathcal{F}_N}$, la restriction du générateur au sous-espace de fonctions \mathcal{F}_N .

$\mathbf{a} \in \mathbb{C}^N$. Par la relation (2.1) adaptée au cas du générateur infinitésimal, nous avons directement que

$$\lambda_j(\mathbf{a}^T \Psi(x)) = [\mathcal{L}_N \phi_{\lambda_j}](\mathbf{x}) = (\mathbf{L}_N \mathbf{a})^T \Psi(\mathbf{x}),$$

de sorte que le vecteur de coefficients \mathbf{a} de la fonction propre λ_j est un vecteur propre de \mathbf{L}_N associé à la valeur propre λ_j . Ainsi, nous venons de voir que la représentation matricielle \mathbf{L}_N de \mathcal{L}_N est caractérisée par les mêmes propriétés spectrales. Cependant, nous n'avons pas de garantie absolue que les fonctions propres de \mathbf{L}_N approximent au mieux celles de l'opérateur \mathcal{L} . En effet, nous n'avons pas forcément que les fonctions propres de l'opérateur projeté correspondent aux projections des fonctions propres exactes. Cependant, nous espérons intuitivement que ces dernières capturent en partie les propriétés spectrales nécessaires pour induire la stabilité du système dynamique.

Remarque 2.2. *Nous n'entrons pas dans les détails dans le cadre du mémoire mais l'approximation du générateur infinitésimal peut être vue comme la matrice d'un système linéaire lifté. En effet, nous avons [10] que*

$$\dot{\mathbf{z}} \approx \mathbf{L}_N^T \mathbf{z},$$

avec $\mathbf{z} = \Psi(\mathbf{x})$. Ainsi, nous comprenons intuitivement qu'il n'est pas étonnant que les propriétés spectrales du générateur soient utilisées pour induire des propriétés de stabilité.

Afin de mieux comprendre comment la matrice \mathbf{L}_N permet d'obtenir une approximation des fonctions propres de l'opérateur de Koopman, considérons l'exemple suivant [10].

Exemple 2.1. Soit le système dynamique unidimensionnel $\dot{x} = -x - x^3$. Pour cet exemple, considérons l'espace de fonctions de base $\{\psi_j\}_{j=1}^4 = \{x^{j-1}\}_{j=1}^4$ avec l'opérateur de projection défini par

$$\Pi x^j = \begin{cases} x^j & \text{if } j \leq 3, \\ 0 & \text{if } j > 3. \end{cases}$$

Dans ce cas, en considérant le fait que $\mathcal{L}\psi_j = (j-1)(-x^{j-1} - x^{j+1})$, nous avons que

$$\mathbf{L}_N = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & -1 & 0 & -3 \end{pmatrix}.$$

Ainsi, les valeurs propres de la matrice sont données par $0, -1, -2, -3$ et en calculant les vecteurs propres à droite de la matrice, nous retrouvons par exemple une approximation de la fonction propre $\phi_{\lambda=-1}(x) = 0.894x - 0.0447x^3$. Notons que dans cet exemple particulier, nous pouvons montrer que la fonction

$$\phi_{\lambda}(x) = \frac{x}{\sqrt{1+x^2}}$$

est une fonction propre² de l'opérateur de Koopman associé à la valeur propre $\lambda = -1$. Ainsi, en utilisant la Proposition 1.4, nous constatons que les valeurs propres de \mathbf{L}_N correspondent à un sous-ensemble exact du spectre ponctuel de \mathcal{L} .

△

Dans l'exemple que nous venons de considérer, nous avons vu que les valeurs propres de \mathbf{L}_N coïncident avec les valeurs propres exactes de \mathcal{L} . En réalité, lorsque nous travaillons avec une base de monômes et un champ de vecteurs analytique, les fonctions propres de l'opérateur sont analytiques dans un voisinage du point d'équilibre. Ainsi, il est possible d'en déterminer une approximation en tronquant ces dernières jusqu'à un certain ordre du développement de Taylor. Cette méthode est brièvement expliquée dans la section suivante mais nous référons à [5, 6] pour plus de détails techniques.

2.1.2 Cas particulier : Champ de vecteurs analytique

Dans cette section, nous allons brièvement développer la méthode numérique de la référence [5, p. 5238], permettant de déterminer une approximation des fonctions propres de l'opérateur de Koopman de manière systématique dans le cas où le champ de vecteurs F du système (1.3) est analytique. Considérons un système dynamique à temps continu quelconque (1.3) dont \mathbf{x}^* est un point fixe et supposons que les valeurs propres λ_i de la matrice Jacobienne J évaluée au point fixe \mathbf{x}^* sont non résonnantes³. Dès lors, comme conséquence du Théorème de linéarisation de Poincaré [4, p. 5238], les fonctions propres de l'opérateur de Koopman admettent une expansion de Taylor au moins dans un voisinage de \mathbf{x}^* . Autrement dit,

$$\phi_\lambda(\mathbf{x}) = \sum_{(k_1, \dots, k_m) \in \mathbb{N}^m} \phi_\lambda^{(k_1, \dots, k_m)} (x_1 - x_1^*)^{k_1} \cdots (x_m - x_m^*)^{k_m}, \quad (2.2)$$

avec les coefficients $\phi_\lambda^{(k_1, \dots, k_m)}$ définis par

$$\phi_\lambda^{(k_1, \dots, k_m)} = \frac{1}{k_1! \cdots k_m!} \left. \frac{\partial^{k_1 + \cdots + k_m} \phi_\lambda}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}} \right|_{\mathbf{x}^*}.$$

De manière similaire, étant donné que le champ de vecteurs $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_m(\mathbf{x}))$ est analytique, il vient que

$$F_l(\mathbf{x}) = \sum_{(k_1, \dots, k_m) \in \mathbb{N}^m} F_l^{(k_1, \dots, k_m)} (x_1 - x_1^*)^{k_1} \cdots (x_m - x_m^*)^{k_m},$$

où les coefficients $F_l^{(k_1, \dots, k_m)}$ sont définis par

$$F_l^{(k_1, \dots, k_m)} = \frac{1}{k_1! \cdots k_m!} \left. \frac{\partial^{k_1 + \cdots + k_m} F_l}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}} \right|_{\mathbf{x}^*}, \quad l = 1, \dots, m.$$

Ainsi, en injectant⁴ les expressions de Taylor de $\mathbf{F}(\mathbf{x})$ et de $\phi_\lambda(\mathbf{x})$ dans (1.8) et en identifiant les termes $(x_1 - x_1^*)^{k_1} \cdots (x_m - x_m^*)^{k_m}$, nous obtenons la relation

2. Pour ce faire, nous pouvons utiliser un raisonnement tout à fait similaire à celui développé dans l'Exemple 1.4 et le résultat en découle naturellement.

3. Les valeurs propres de la matrice Jacobienne, évaluée au point fixe, sont non résonnantes s'il n'existe pas d'entiers $c_k \geq 0$ tels que $\sum_k c_k \geq 2$ et $\lambda_i = \sum_k c_k \lambda_k$.

4. Nous n'entrons pas dans les détails calculatoires dans le cadre du mémoire et nous référons à [5, 6] pour plus de détails techniques.

$$\sum_{j_1=0}^{k_1} \cdots \sum_{j_m=0}^{k_m} \sum_{l=1}^m (k_l - j_l + 1) F_l^{(j_1, \dots, j_m)} \phi_\lambda^{(k_1-j_1, \dots, k_{l-1}-j_{l-1}, k_l-j_l+1, k_{l+1}-j_{l+1}, \dots, k_m-j_m)} = \lambda \phi_\lambda^{(k_1, \dots, k_m)} \quad (2.3)$$

pour tout m -uplet $(k_1, \dots, k_m) \in \mathbb{N}^m$. Dès lors, pour chaque entier $s \in \mathbb{N}$, nous pouvons considérer le vecteur $\Phi^{(s)}$ contenant les différents coefficients $\phi_\lambda^{(k_1, \dots, k_m)}$ pour lesquels $\sum_j k_j = s$. Pour $s = 0$, nous obtenons trivialement $\phi_\lambda^{(0, \dots, 0)} = 0$. Pour $s = 1$, la relation (2.3) revient à résoudre l'équation

$$\mathbf{J}^T \Phi^{(1)} = \lambda \Phi^{(1)}.$$

Autrement dit, $\Phi^{(1)} := \nabla \phi_\lambda(\mathbf{x}^*)$ est un vecteur propre à gauche de la matrice Jacobienne J associé à la valeur propre λ de sorte que les valeurs propres de l'opérateur de Koopman sont celles de la matrice Jacobienne au point d'équilibre. Pour des valeurs de $s \geq 1$, l'équation (2.3) définit le système d'équations

$$H^{(s)} \Phi_\lambda^{(s)} = \lambda \Phi_\lambda^{(s)} + G^{(s)} (\Phi^{(1)}, \Phi^{(2)}, \dots, \Phi^{(s-1)}), \quad (2.4)$$

où $H^{(s)}$ est une matrice et $G^{(s)}$ est un vecteur contenant des coefficients précédemment calculés. Dès lors, afin de déterminer les coefficients de $\phi_\lambda(x)$ dans la base des polynômes de Taylor, nous devons résoudre un système d'équations de manière récursive. Même si nous n'entrons pas dans les détails calculatoires, nous allons tout de même montrer que dans le cas particulier des systèmes de dimension 1, les relations (2.3) et (2.4) sont vérifiées.

Exemple 2.2 (Système de dimension 1). Considérons le cas particulier des systèmes de dimension 1, c'est-à-dire tels que $m = 1$. Dans ce cas, $\mathbf{F}(x) = F_1(x)$ et les développements de Taylor de $\mathbf{F}(x)$ et $\phi_\lambda(x)$ se simplifient et sont donnés par

$$\begin{cases} \phi_\lambda(x) = \sum_{k \in \mathbb{N}} \phi_\lambda^{(k)} (x - x^*)^k, \\ \mathbf{F}(x) = \sum_{k \in \mathbb{N}} F^{(k)} (x - x^*)^k. \end{cases} \quad (2.5)$$

Dès lors, le développement de Taylor de $\nabla \phi_\lambda(x)$ est donné par

$$\nabla \phi_\lambda(x) = \sum_{k \in \mathbb{N}_0} k \phi_\lambda^{(k)} (x - x^*)^{k-1}. \quad (2.6)$$

Montrons alors qu'en injectant les expressions obtenues en (2.5) et en (2.6) dans la relation (1.8), nous obtenons bien la relation (2.3). Pour ce faire, développons dans un premier temps l'expression $\mathcal{L} \phi_\lambda(x) = \mathbf{F}(x) \nabla \phi_\lambda(x)$. Il vient alors que

$$\begin{aligned}
(\mathcal{L}\phi_\lambda)(x) &= F(x)\nabla\phi_\lambda(x), \\
&= \phi_\lambda^{(1)} \left(\sum_{k \in \mathbb{N}} F^{(k)}(x-x^*)^k \right) + 2\phi_\lambda^{(2)}(x-x^*) \left(\sum_{k \in \mathbb{N}} F^{(k)}(x-x^*)^k \right) + \dots, \\
&= F^{(0)}\phi_\lambda^{(1)} + F^{(1)}\phi_\lambda^{(1)}(x-x^*) + \dots + 2F^{(0)}\phi_\lambda^{(2)}(x-x^*) + 2F^{(1)}\phi_\lambda^{(2)}(x-x^*)^2 \\
&\quad + \dots, \\
&= F^{(0)}\phi_\lambda^{(1)} + \left[F^{(1)}\phi_\lambda^{(1)} + 2F^{(0)}\phi_\lambda^{(2)} \right] (x-x^*)^1 + \dots, \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^k (k+1-j) F^{(j)}\phi_\lambda^{(k+1-j)}(x-x^*)^k.
\end{aligned}$$

Dès lors, par définition d'une fonction propre du générateur, il vient que

$$\sum_{k=0}^{\infty} \sum_{j=0}^k (k+1-j) F^{(j)}\phi_\lambda^{(k+1-j)}(x-x^*)^k = \lambda \sum_{k=0}^{\infty} \phi_\lambda^{(k)}(x-x^*)^k.$$

Par identification, nous obtenons finalement que pour tout $k \in \mathbb{N}$,

$$\sum_{j=0}^k (k+1-j) F^{(j)}\phi_\lambda^{(k+1-j)}(x-x^*)^k = \lambda\phi_\lambda^{(k)}, \quad (2.7)$$

ce qui correspond bien à la relation (2.3) appliquée au système de dimension 1. Montrons à présent que le système obtenu en (2.7) nous permet de déterminer les coefficients $\phi_\lambda^{(k)}$ de manière itérative. Soit $\lambda = F^{(1)}$, la valeur propre de la matrice Jacobienne au point fixe⁵.

$k = 0$

Considérons que $k = 0$. Dans ce cas, nous avons

$$F^{(0)}\phi_\lambda^{(0)} = \lambda\phi_\lambda^{(0)}.$$

Autrement dit, nous avons que $\phi_\lambda^{(0)} = 0$ étant donné que $F^{(0)} = F(x^*) = 0$.

$k = 1$

Considérons à présent le cas où $k = 1$. Dans ce cas, nous devons résoudre

$$F^{(1)}\phi_\lambda^{(1)} + \underbrace{2F^{(0)}\phi_\lambda^{(2)}}_{=0} = \lambda\phi_\lambda^{(1)}.$$

Dès lors, il vient que

$$\left. \frac{dF(x)}{dx} \right|_{x^*} \phi_\lambda^{(1)} = \lambda\phi_\lambda^{(1)} \iff J\phi_\lambda^{(1)} = \lambda\phi_\lambda^{(1)},$$

5. En réalité la méthode peut être utilisée pour déterminer les autres fonctions propres de la forme $(\phi_\lambda(x))^n$ associées aux valeurs propres λn pour n'importe quel entier $n \geq 1$.

et $\phi_\lambda^{(1)} = \nabla \phi_\lambda(x^*)$ est un vecteur propre à gauche et à droite de la matrice Jacobienne J (puisque nous sommes dans le cas de systèmes de dimension 1). De plus, nous remarquons que pour $H^{(1)} = J^T (= J)$, $\Phi^{(1)} = \phi_\lambda^{(1)}$ et pour $G^{(1)} = 0$, nous obtenons la relation (2.4). Notons que dans ce cas, nous avons que $\phi_\lambda^{(1)}$ est quelconque. Ceci semble cohérent puisque les fonctions propres sont définies à une constante multiplicative près. Dès lors, les différentes valeurs non nulles de $\phi_\lambda^{(1)}$ vont déterminer toutes les fonctions propres possibles associées à la valeur propre λ .

$k = 2$

Par la relation (2.7), il vient que

$$F^{(2)}\phi_\lambda^{(1)} + 2F^{(1)}\phi_\lambda^{(2)} + \underbrace{3F^{(0)}\phi_\lambda^{(3)}}_{=0} = \lambda\phi_\lambda^{(2)}.$$

De manière équivalente, nous trouvons que

$$\begin{aligned} 2F^{(1)}\phi_\lambda^{(2)} &= \lambda\phi_\lambda^{(2)} - F^{(2)}\phi_\lambda^{(1)}, \\ \iff H^{(2)}\Phi^{(2)} &= \lambda\Phi^{(2)} + G^{(2)}, \end{aligned}$$

où $H^{(2)} = 2F^{(1)}$, $G^{(2)} = -F^{(2)}\phi_\lambda^{(1)}$ et $\Phi^{(2)} = \phi_\lambda^{(2)}$. Notons que les coefficients qui apparaissent dans $G^{(2)}$ sont bien d'un ordre inférieur à 2 et ont déjà été calculés.

$k \geq 3$

Dès lors, en continuant de cette manière pour des valeurs de k plus grandes ou égales à 3, nous pouvons calculer les différents coefficients $\phi_\lambda^{(k)}$ nous permettant ainsi de déterminer $\phi_\lambda(x)$. △

Ainsi, en utilisant la méthode que nous venons de détailler, nous pouvons également construire une approximation des fonctions propres de l'opérateur de Koopman dans une base de monômes. Dans ce cas, les différents coefficients de l'approximation sont alors obtenus en tronquant la fonction propre exacte jusqu'à un certain ordre de la série de Taylor. Notons que si nous tronquons la fonction propre exacte, nous avons bon espoir que l'approximation ainsi obtenue soit assez proche de la vraie fonction pour un ordre suffisamment important, ce qui n'était pas spécialement garanti par la méthode du générateur. Cependant, dans le cas particulier d'une base de monômes avec un champ de vecteurs analytique, cette méthode est équivalente à celle développée dans la section précédente. En réalité, nous sommes dans un cas particulier où la projection des fonctions propres correspondent aux fonctions propres de l'opérateur projeté. Par ailleurs, les valeurs propres de la matrice \mathbf{L}_N contiennent les valeurs propres de la matrice Jacobienne au point d'équilibre ainsi que leurs multiples à coefficients entiers. Autrement dit, le spectre de \mathbf{L}_N contient les N premières valeurs propres de l'opérateur de Koopman.

2.2 Approximation du bassin d'attraction avec une base de monômes

Dans cette section, nous allons appliquer la méthode développée dans la section 2.1 dans le cas particulier où le sous-espace de fonctions de base est constitué de monômes. Nous y étudierons en particulier la qualité de l'approximation du bassin d'attraction de

différents systèmes à travers les deux exemples principaux de la référence [6]. Ces exemples étant tous les deux à deux dimensions, nous verrons enfin dans la dernière partie de cette section comment la méthode s'applique pour deux exemples en trois dimensions, où nous relèverons les inconvénients d'une telle base, nous amenant ainsi progressivement au Chapitre 3.

Considérons dans un premier temps le système dynamique de Van der Pol défini en (1.7) et fréquemment utilisé dans la littérature pour tester et valider des méthodes numériques. Rappelons que ce système est caractérisé par un point fixe stable dont le bassin d'attraction est délimité par un cycle limite instable (voir Figure 1.6). Afin d'appliquer la méthode, considérons un ensemble de fonctions contenant des monômes jusqu'à l'ordre $s \in \mathbb{N}$, c'est-à-dire

$$\{\psi_j\}_j = \{x_1^i x_2^j \mid i + j \leq s\},$$

où nous supposons que l'opérateur de projection de Π est donné par

$$\Pi(x_1^i x_2^j) = \begin{cases} x_1^i x_2^j & \text{si } i + j \leq s, \\ 0 & \text{si } i + j > s. \end{cases}$$

Dès lors, pour $s = 40$, nous remarquons à la Figure 2.1 (a) que la méthode numérique semble fournir une bonne estimation du bassin d'attraction de \mathbf{x}^* , délimité par le cycle limite instable en rouge. En effet, le module⁶ de la fonction propre, ou du moins de son approximation $\tilde{\phi}_{\lambda_1}(\mathbf{x})$, diverge à la frontière de celui-ci, ce qui est attendu. Cependant, nous ne sommes pas tout à fait rigoureux. En effet, les fonctions propres que nous considérons sont en réalité des approximations. Ainsi, leur support est lui aussi approximé et nous ne sommes dès lors plus sûrs que le support de l'approximation délimite exactement le bassin d'attraction. Dans notre exemple, nous remarquons que les fonctions on

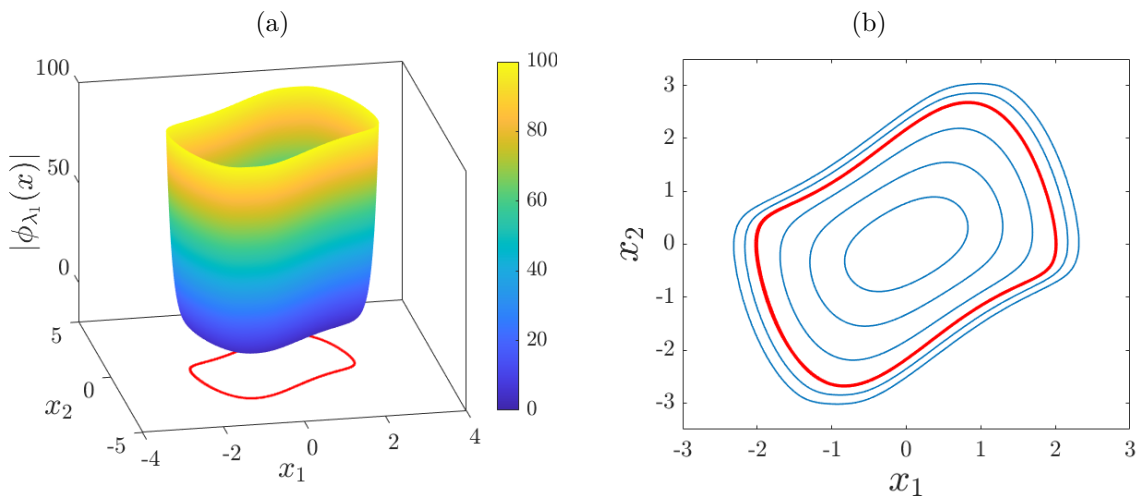


Figure 2.1 – (a) Représentation du module l'approximation de $\phi_{\lambda_1}(\mathbf{x})$, calculée sur l'espace de phase $\mathbb{X} = [-3, 3]^2$. (b) Représentation de différentes courbes de niveau de l'approximation $|\tilde{\phi}_{\lambda_1}(\mathbf{x})|$ de $|\phi_{\lambda_1}(\mathbf{x})|$. Dans ce cas, nous remarquons que la fonction propre explose à la limite du bassin d'attraction mais également en dehors.

6. Dans le cas où la fonction propre est complexe, il suffit d'utiliser son module pour déterminer le bassin d'attraction puisque $|\phi_\lambda(x)| = \exp(\operatorname{Re}\{\lambda\}t) |\phi_\lambda(x)|$.

des valeurs finies même en dehors du cycle limite (voir Figure 2.1 (b)). Pour pallier ce problème, rappelons que les fonctions propres de la Proposition 1.6 permettent de définir une fonction de Lyapunov

$$\mathcal{V}(x) = \left(\sum_{i=1}^m c_i |\phi_{\lambda_i}(x)|^p \right)^{1/p}, \quad (2.8)$$

pour un certain $p \in \mathbb{N}$ avec les différents coefficients $c_i > 0$. Dans l'exemple que nous étudions, si nous construisons cette fonction à l'aide de l'approximation $\tilde{\phi}_\lambda(\mathbf{x})$, nous obtenons également une approximation de la fonction de Lyapunov de sorte que son support ne permet plus non plus d'assurer une délimitation exacte du bassin d'attraction. En réalité, nous avons toujours que $\mathcal{V}(\mathbf{x}^*) = 0$ et $\mathcal{V}(\mathbf{x}) \geq 0$ pour tout $\mathbf{x} \in D$ avec D l'ensemble des points sur lequel la fonction est définie. Cependant, nous n'avons plus nécessairement que la fonction décroît le long des trajectoires. Afin de contourner ce problème, nous pouvons définir l'ensemble

$$\bar{X} = \{\mathbf{x} \in \mathbb{X} \mid \dot{V}(\mathbf{x}) < 0\} \cup \{\mathbf{x}^*\},$$

c'est-à-dire le plus grand ensemble sur lequel l'approximation de $\mathcal{V}(\mathbf{x})$, que l'on note $V(\mathbf{x})$, est strictement décroissante le long des trajectoires. Dans ce cas, si nous considérons le plus grand sous-ensemble de niveau

$$\Psi_\delta = \{x \in \mathbb{R}^m \mid V(x) \leq \delta\},$$

tel que $\Psi_\delta \subseteq \bar{X}$, il vient que Ψ_δ fournit une approximation exacte du bassin d'attraction par l'intérieur. Autrement dit, à partir de l'approximation des fonctions propres de l'opérateur de Koopman, l'ensemble Ψ_δ détermine, de manière exacte, un sous-ensemble du bassin d'attraction.

Notons qu'en pratique, la région \bar{X} est calculée sur un certain domaine discrétisé \mathbb{X} . Dans ce cas, nous n'obtenons pas exactement \bar{X} mais plutôt un ensemble discret de points

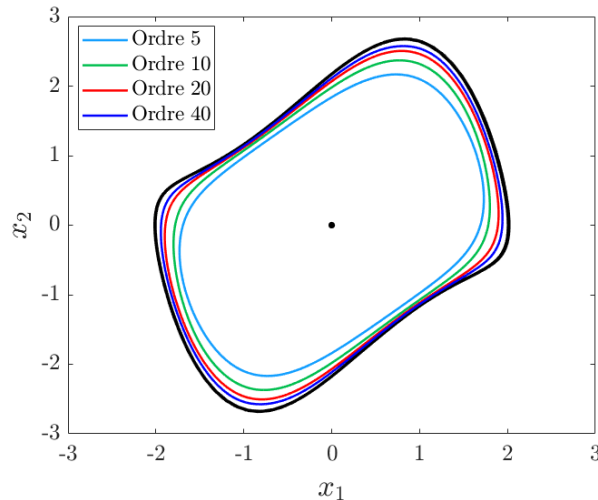


Figure 2.2 – Approximation du bassin d'attraction en calculant le plus grand ensemble de niveau pour lequel $\Psi_\delta \subset \bar{X}$ avec $\mathcal{V}(x) = |\tilde{\phi}_{\lambda_1}(\mathbf{x})|^2$ étant donné que les valeurs propres de la matrice Jacobienne sont complexes. Dans ce cas, nous remarquons que plus le nombre de fonctions de base augmente, meilleure est l'approximation. Ceci n'est pas étonnant au vu de l'analyticité des fonctions propres dans ce cas-ci.

dans cet ensemble. Cependant, rien ne nous garantit qu'entre ces différents points, le signe de \dot{V} demeure inchangé. Ainsi, il pourrait très bien exister des régions pour lesquelles la fonction change de signe entre les mailles de la grille de sorte que les résultats de stabilité ainsi obtenus sont alors biaisés. Ce manque de rigueur sera en particulier traité dans le chapitre 4 dans le cas particulier où la base de fonctions est monomiale. Cependant, nous ne tiendrons pas compte de ce manque de rigueur dans ce chapitre et nous interpréterons les résultats dans le cas idéal où le signe ne change pas.

Calculons à présent cet ensemble \bar{X} et Ψ_δ dans le cadre du système de Van der Pol. Dans ce cas, nous remarquons à la Figure 2.2 qu'au plus nous augmentons l'ordre de la décomposition de Taylor, et donc le nombre de fonctions de base, au plus le bassin d'attraction se rapproche du bassin exact délimité par le cycle limite en noir. Notons par ailleurs que ces résultats sont en accord avec ceux obtenus dans [6] avec la méthode basée sur la décomposition de Taylor.

Remarque 2.3. *Précisons que dans cet exemple, nous ne nous sommes pas particulièrement penchés sur le poids des différents coefficients c_i de la fonction de Lyapunov. Dans l'exemple que nous avons considéré à la Figure 2.2, la fonction $\mathcal{V}(\mathbf{x})$ est calculée avec $c_i = 1/2 = c_2$ avec $p = 2$. Cependant, nous sommes dans un cas particulier où les valeurs propres de l'opérateur sont complexes de sorte que $|\phi_{\lambda_1}| = |\phi_{\lambda_2}|$. Ainsi, pour n'importe quelle valeur de c_i pour $p = 2$, nous obtenons la fonction de Lyapunov⁷ $\mathcal{V}(x)^2 = (c_1 + c_2)|\phi_{\lambda_1}|$ et les c_i n'ont alors aucun impact. En revanche, lorsque que les fonctions propres de l'opérateur sont à valeurs réelles, le poids des coefficients peut être utilisé comme tuning afin d'obtenir une meilleure approximation du bassin d'attraction (voir section 2.2.1).*

Penchons-nous à présent sur l'exemple de dimension deux suivant

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -2x_1 + \frac{1}{3}x_1^3 - x_2, \end{cases} \quad (2.9)$$

caractérisé par un point fixe localement stable à l'origine ainsi que par deux points de selle $\mathbf{x}^* = (\pm\sqrt{6}, 0)$ (voir Figure 2.3 (a)). Le point fixe à l'origine est en réalité globalement stable dans une certaine région délimitée par les variétés stables des deux points de selle que nous avons tracées en rouge sur la Figure. Dans ce cas, l'approximation des fonctions propres de \mathcal{L} permet également de définir une bonne estimation du bassin d'attraction comme nous pouvons le voir à la Figure 2.3 (b)⁸. Cependant, il semblerait qu'à partir d'un certain ordre, augmenter le nombre de fonctions de base n'est plus intéressant. En réalité, nous pouvons justifier cela à l'aide du Théorème de linéarisation de Poincaré. En effet, ce dernier affirme [5] que les fonctions propres seront analytiques sur le plus grand disque qui ne contient pas d'autres points fixes (éventuellement complexes). Or, dans cet exemple, nous avons deux autres points fixes de sorte que le support des fonctions propres ne permet pas de déterminer toute la géométrie du bassin d'attraction. Une manière de pallier ce problème serait d'étudier le système par l'intermédiaire d'autres fonctions de base, ce qui sera abordé et discuté dans le chapitre suivant.

7. Notons que si $\mathcal{V}(\mathbf{x})$ est une fonction de Lyapunov, le carré de cette fonction est aussi une fonction de Lyapunov. Ainsi, par abus de langage, $\mathcal{V}(\mathbf{x})$ désignera parfois $\mathcal{V}(\mathbf{x})^2$.

8. Notons que dans cet exemple, nous obtenons également les mêmes résultats en appliquant la deuxième méthode basée sur le développement en série.

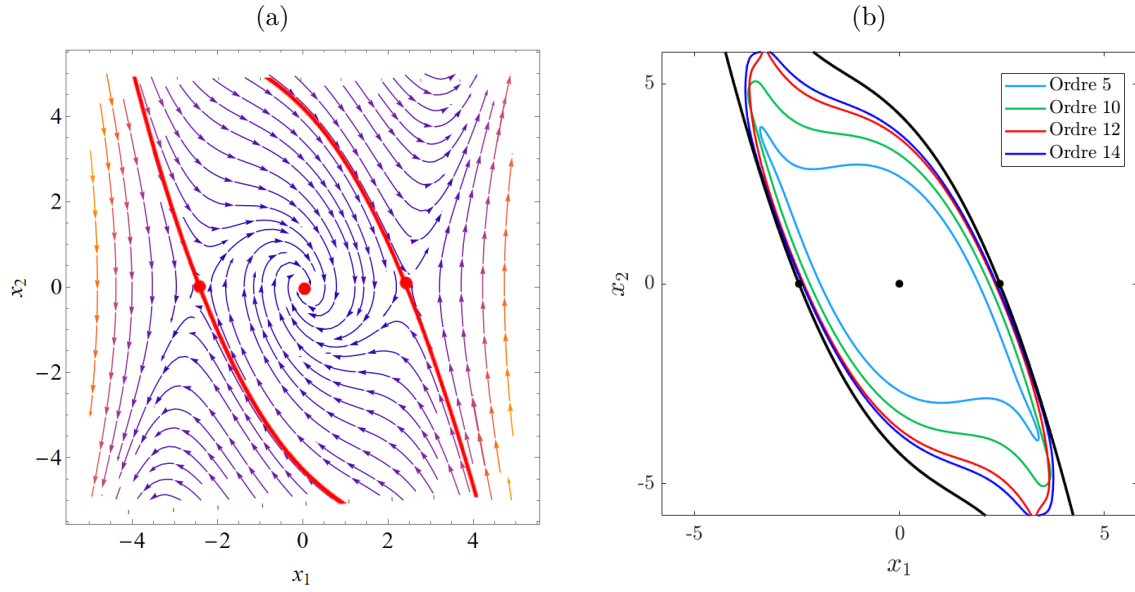


Figure 2.3 – (a) Représentation du portrait de phase du système (2.9). Nous observons que ce système est caractérisé par un point fixe stable à l’origine dont le bassin d’attraction est délimité par les deux variétés instables des deux autres points fixes $(\pm\sqrt{6}, 0)$. (b) Approximation du bassin d’attraction avec $\mathcal{V}(x) = |\tilde{\phi}_{\lambda_1}(\mathbf{x})|^2$ étant donné que les valeurs propres de la matrice Jacobienne sont également complexes. Dans ce cas, nous remarquons que plus le nombre de fonctions de base augmente, meilleure est l’approximation.

Notons que les différents exemples que nous avons considérés jusque maintenant sont bi-dimensionnels. Dans la section suivante, nous allons voir comment appliquer la méthode sur deux exemples de dimension trois, ce qui n’avait pas encore été fait dans la littérature. Nous y verrons notamment en quoi l’utilisation de monômes peut devenir problématique lorsque la dimension du système est importante.

2.2.1 Application pour des systèmes de plus haute dimension

Lorsque nous travaillons avec des systèmes de plus haute dimension, la méthode développée dans le cas d’un champ de vecteurs analytique peut vite devenir fastidieuse à implémenter. Ainsi, nous préférons utiliser une approximation du générateur infinitésimal par souci de facilité. Par ailleurs, cette approche permet de généraliser notre code pour n’importe quelle dimension, et n’importe quelles fonctions de base comme nous le verrons dans le chapitre suivant.

Considérons dans un premier temps le système dynamique à trois dimensions

$$\begin{cases} \dot{x}_1 = -x_2, \\ \dot{x}_2 = -x_1 + x_2 - x_1^2 x_2, \\ \dot{x}_3 = x_3^3 - x_3, \end{cases} \quad (2.10)$$

dont l’origine est un point fixe stable. Ce système est en réalité une combinaison entre le système dynamique de Van der Pol et le système unidimensionnel $\dot{x} = x^3 - x$, dont l’origine et ± 1 sont des points fixes. Dès lors, comme nous le voyons sur le portrait de

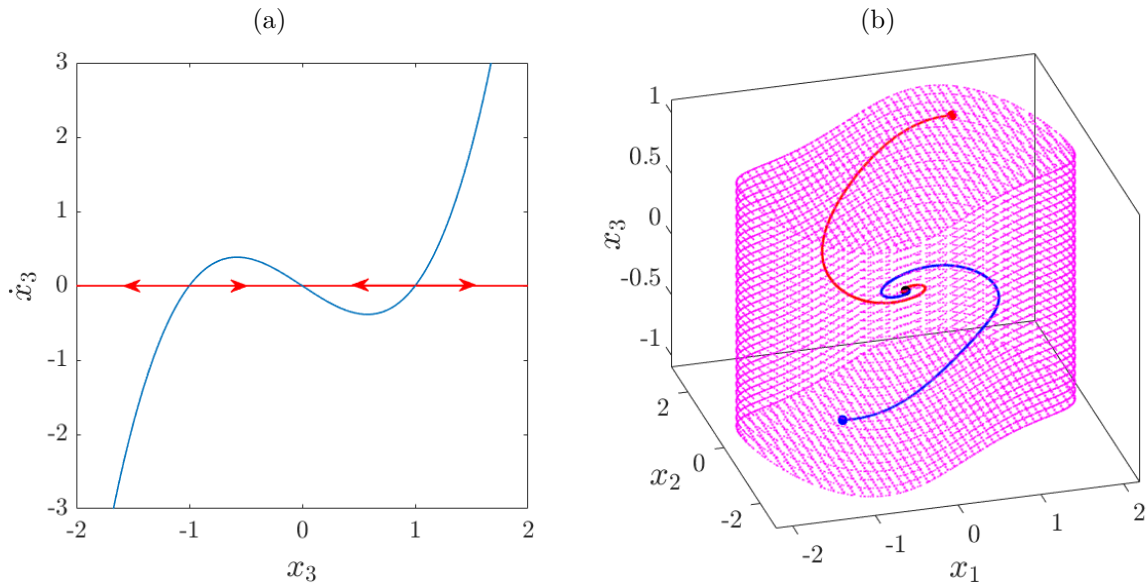


Figure 2.4 – (a) Représentation du portrait de phase du système $\dot{x}_3 = x_3^3 - x_3$. Dans ce cas, nous voyons que $x_3 = 0$ est un point fixe stable dont le bassin d’attraction est délimité par le segment de droite $(-1, 1)$. (b) Bassin d’attraction du système (2.10) représenté par la variété en mauve. En rouge et bleu, nous avons représenté deux trajectoires, toutes convergentes vers l’équilibre $\mathbf{x}^* = 0$.

phase du système à la Figure 2.4 (a), le point d’équilibre 0 est un point stable et son bassin d’attraction est délimité par le segment de droite ouvert $(-1, 1)$. Dès lors, combiné avec le système de Van der Pol, nous avons que 0 est un point fixe stable, pour le système (2.10), dont le bassin d’attraction est délimité par la variété représentée à la Figure 2.4 (b). Ainsi, en appliquant la méthode avec une base de monômes de degré au plus 5 et 10, nous obtenons une bonne estimation du bassin d’attraction donnée par les Figures 2.5 (a) - (b). À nouveau, nous obtenons un meilleur résultat pour un nombre de fonctions de base plus élevé. Notons que dans cet exemple, les valeurs propres de la matrice Jacobienne au point d’équilibre sont données par

$$\lambda_1 = -0.5 + 0.866i, \quad \lambda_2 = -0.5 - 0.866i \quad \text{et} \quad \lambda_3 = -1.$$

de sorte que l’approximation de la fonction propre ϕ_{λ_3} est à valeurs réelles. Dans ce cas, le poids des coefficients peut être utilisé comme *tuning* afin d’obtenir une meilleure approximation du bassin d’attraction comme nous pouvons l’observer aux Figures 2.5 (c) - (d).

Remarque 2.4. Dans cet exemple, nous avons considéré des monômes de degré au plus 5 et 10. En réalité, au plus la dimension du système augmente, au plus le nombre de monômes nécessaire pour atteindre un certain ordre fixe devient important, ce qui motive ce choix. Par ailleurs, nous remarquons à la Figure 2.5 (d) qu’augmenter le nombre de fonctions de base n’est plus tellement intéressant étant donné que nous capturons déjà la majeure partie du bassin d’attraction à l’ordre 10. Ainsi, il y a un trade-off entre le temps d’exécution de l’algorithme et la précision des résultats obtenus, bien que le temps de calcul reste tout à fait raisonnable dans le cadre des systèmes en trois dimensions.

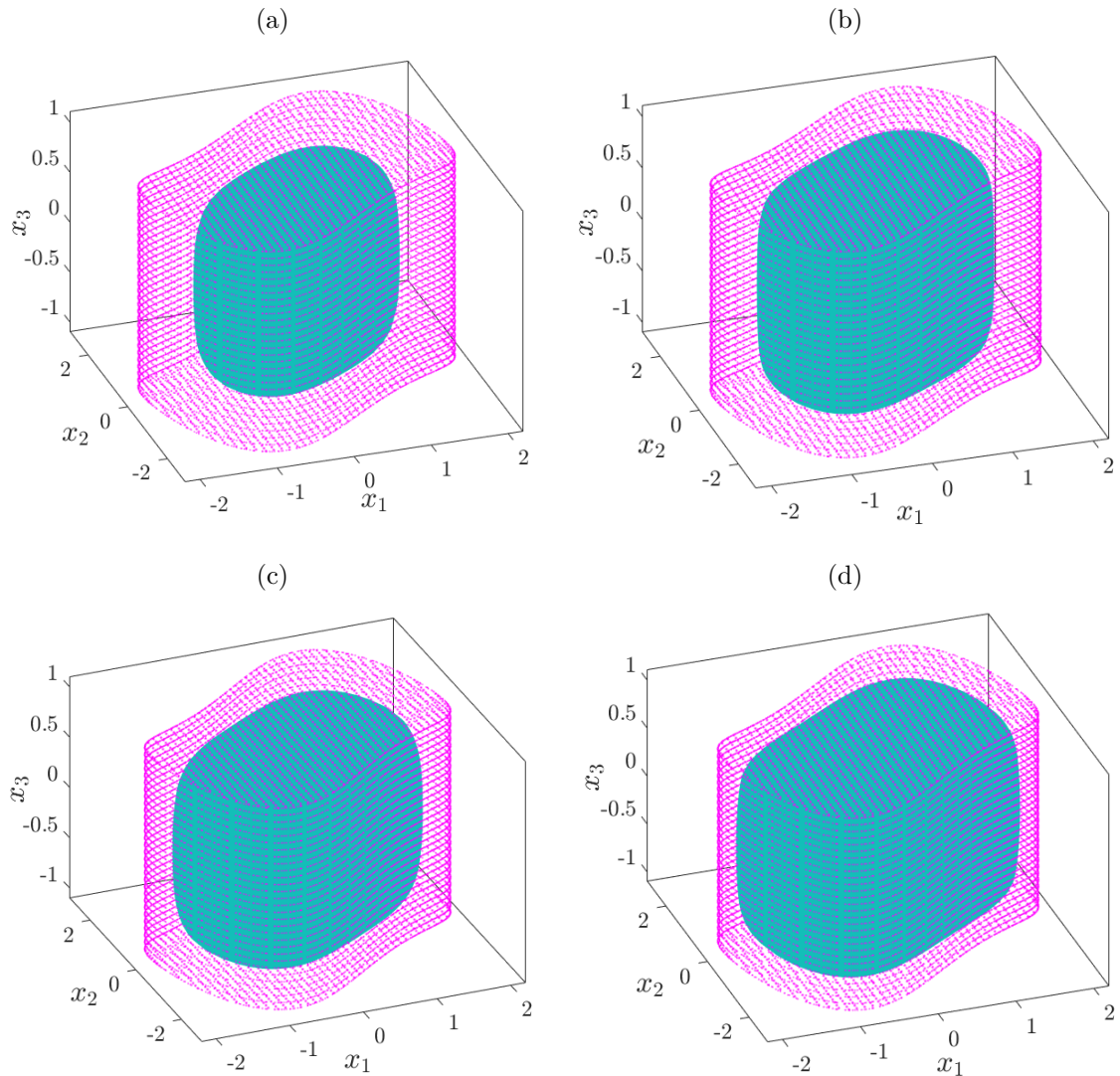


Figure 2.5 – Approximation du bassin d’attraction pour le système (2.10) en utilisant des monômes de degré au plus 5 et 10 avec (a)-(b) $c_i = 1$ pour tout $i = 1, \dots, 3$, (c)-(d), $c_1 = c_2 = 0.4$ et $c_3 = 1$.

L’exemple que nous venons de traiter est un peu simple étant donné que nous avons ici couplé deux systèmes dynamiques qui évoluent indépendamment l’un de l’autre. Ainsi, afin clôturer ce chapitre, considérons un exemple un peu plus complexe et adapté de la référence [12]. Soit le système dynamique

$$\begin{cases} \dot{x}_1 = -x_2, \\ \dot{x}_2 = x_2 x_1^2 + x_1 - x_2 + x_3, \\ \dot{x}_3 = x_3^3 - x_3, \end{cases} \quad (2.11)$$

caractérisé par un point fixe localement stable à l’origine et deux points fixes instables en $(\pm 1, 0, \mp 1)$. De nouveau, le point fixe à l’origine est en réalité globalement stable à l’intérieur de la variété représentée à la Figure 2.6. Notons que cet exemple est plus complexe étant donné que la dynamique de la troisième variable d’état n’est pas découplée comme dans l’exemple précédent. Appliquons alors la méthode avec la même base de monômes que précédemment, c’est-à-dire des monômes de degré au plus 5 et 10. Dans cas

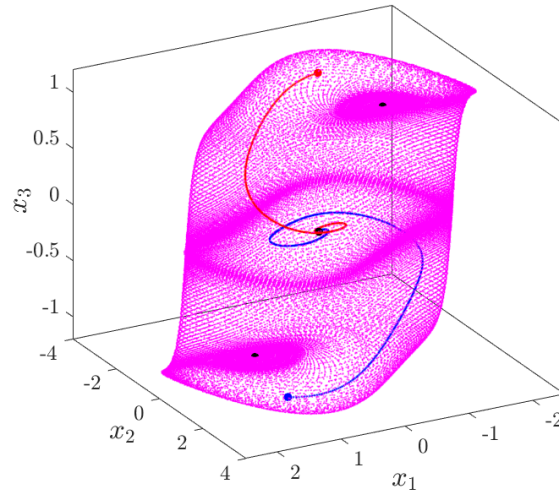


Figure 2.6 – Bassin d’attraction du point d’équilibre à l’origine pour le système dynamique (2.11). En rouge et bleu, nous avons représenté deux trajectoires du système convergeantes vers le point fixe afin de visualiser la géométrie du bassin d’attraction délimité par la variété en mauve.

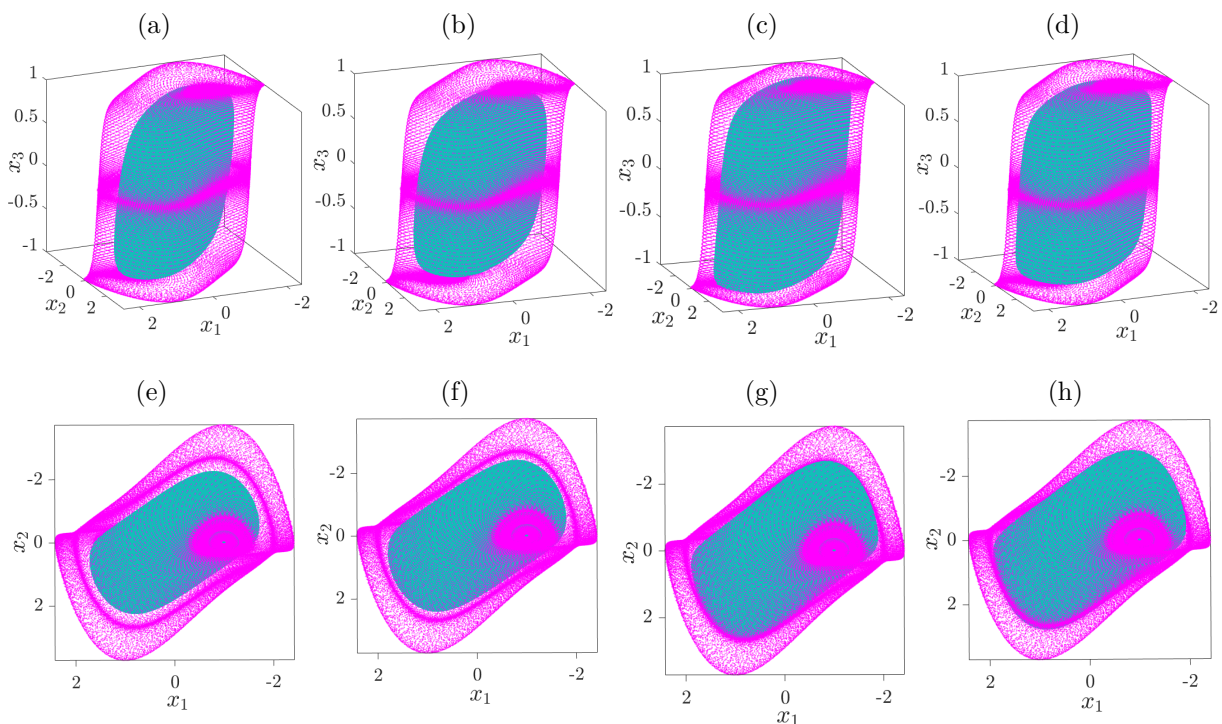


Figure 2.7 – Approximation du bassin d’attraction du point d’équilibre à l’origine du système (2.11) pour des monômes de degré au plus 5 et 10 avec (a)-(b) $c_i = 1$ pour tout $i = 1, \dots, 3$, (c)-(d), $c_1 = c_2 = 0.4$ et $c_3 = 1$. (e) - (h) Vue aérienne des différentes figures (a)-(d). Dans ce cas, nous remarquons que pour un ordre plus élevé, nous obtenons une meilleure approximation du bassin d’attraction mais que jouer sur le poids des coefficients c_i n’est pas particulièrement avantageux.

cas, nous obtenons une bonne estimation du bassin en utilisant la fonction de Lyapunov avec les différents coefficients $c_i = 1$ pour tout $i = 2, \dots, 3$ (voir les Figures 2.7 (a) - (b)). Cependant, il semblerait que dans cet exemple, le poids des coefficients c_i n'améliore pas les résultats de manière significative comme nous pouvons le voir aux Figures 2.7 (c) - (d). Notons que ce n'est pas particulièrement étonnant au vu de la géométrie complexe du bassin d'attraction. Par ailleurs, si nous regardons attentivement le bassin d'attraction à la Figure 2.6 nous pouvons remarquer que le bassin d'attraction semble être creusée par endroit de sorte qu'il semble finalement difficile d'obtenir une meilleure approximation que celle obtenue à la Figure 2.7 (d).

Remarque 2.5. *Précisons qu'ici encore, le théorème de linéarisation de Poincaré nous assure que les fonctions propres seront analytiques dans le plus grand disque qui ne contient pas d'autres points d'équilibre. Cependant, cet exemple est caractérisé par deux autres points fixes que l'origine, de sorte qu'il n'est pas étonnant que les fonctions propres ne capturent pas toute la géométrie du bassin d'attraction.*

Lorsque nous travaillons avec des systèmes de dimensions beaucoup plus importantes, nous avons besoin de beaucoup de fonctions de base pour un degré fixé. En effet, rappelons que le nombre de monômes de base à p variables de degré au plus n est donné par $(n+m)!/(n!p!)$. Ainsi, si nous travaillons avec un système de dimension $p = 17$, le nombre de monômes de degré au plus 10 s'élève à 8436285, ce qui devient beaucoup trop coûteux numériquement. Cette constatation motive alors le choix de nouvelles fonctions de base, qui restent encore une question ouverte à l'heure actuelle. Dans le chapitre suivant, nous étudierons l'impact du choix des fonctions de base et nous verrons comment ce choix impacte les résultats pour les différents exemples que nous avons traités dans ce chapitre. En particulier, nous nous pencherons sur certains aspects théoriques concernant les régions de stabilité ainsi obtenues.

Chapitre 3

Étude du choix des fonctions de base et impact sur la stabilité

Dans le chapitre précédent, nous avons vu à travers différents exemples numériques comment l'approximation des fonctions propres de l'opérateur de Koopman permettaient d'induire la stabilité globale de points fixes hyperboliques. Plus particulièrement, nous nous sommes penchés sur une approximation des fonctions propres de l'opérateur dans une base de monômes avec l'opérateur de projection Π donné par un opérateur de troncature. Dans ce cas, nous avons vu qu'il était possible d'approximer les fonctions propres exactes de l'opérateur de Koopman à l'aide d'une simple troncature de ces dernières. Dans ce contexte, nous obtenions des résultats tout à fait satisfaisants avec, dans certains cas, une approximation du bassin d'attraction presque parfaite. Cependant, les exemples que nous avons considérés sont assez simples de sorte que l'utilisation des monômes soit avantageuse et fournisse de bons résultats de stabilité en temps de calcul tout à fait raisonnable. Toutefois, lorsque nous manipulons des systèmes de grandes dimensions, comme c'est le cas pour beaucoup d'applications réelles des systèmes dynamiques, l'utilisation d'une telle base peut devenir coûteuse d'un point de vue computationnel. En effet, rappelons que le nombre de monômes de base nécessaire pour un ordre fixé devient astronomiquement grand lorsque la dimension du système est suffisamment élevée. Ce constat motive alors de considérer une approche de la stabilité par l'intermédiaire d'autres fonctions de base telles que les fonctions radiales de Gauss, utilisées notamment dans des méthodes numériques basées sur des données [13, 14]. Dans ce chapitre, nous allons nous intéresser à une telle question du choix de la base et de son impact sur l'inférence de régions stables pour les différents systèmes que nous avons considérés depuis le début du mémoire. Plus particulièrement, nous considérerons des fonctions de base radiales et nous comparerons les résultats avec ceux obtenus dans le cadre des monômes. Nous préciserons par ailleurs que ces résultats auront été générés à l'aide d'une méthode générale implémentée en Matlab et permettant à n'importe quel utilisateur d'étudier la stabilité du système de son choix pour n'importe quel choix de fonctions de base. Ensuite, nous verrons que le choix d'une telle base ne permet pas d'assurer directement la stabilité de l'origine et nous verrons comment adapter la preuve du théorème de LaSalle pour pallier ce problème. Enfin, nous terminerons ce chapitre en appliquant la méthode sur les deux exemples en trois dimensions que nous avons considérés à la fin du chapitre précédent.

3.1 Choix des fonctions de base et de la projection

Comme nous l'avons mentionné dans l'introduction de ce chapitre, nous allons en particulier travailler avec des fonctions de base *radiales*. La particularité de ces dernières, contrairement aux monômes, est qu'elles dépendent de certains points appelés *centres*.

Définition 3.1 (Fonctions de base radiales)

Une **fonction de base radiale** est une fonction ψ_c à valeurs réelles qui vérifie l'égalité $\phi_c(\mathbf{x}) = \varphi(\|\mathbf{x} - c\|)$ avec c le **centre** de la fonction. Il existe différents types de fonctions radiales dont les plus courantes :

- (a) Les **fonctions radiales de Gauss** définies par

$$\psi_c(\mathbf{r}) := e^{-(\gamma \mathbf{r})^2}$$

avec $\mathbf{r} = \|\mathbf{x} - c\|$ et $\gamma > 0$ un paramètre permettant d'influencer la forme de la fonction ;

- (b) Les fonctions radiales dites **thin plate spline** (TPS) définies par

$$\psi_c(\mathbf{r}) := r^2 \ln(\mathbf{r}).$$

À la Figure 3.1, nous avons représenté ces deux fonctions de base centrées en zéro. Dans ce cas, nous voyons que le paramètre γ permet d'influencer la forme des fonctions de base en construisant des fonctions plus ou moins raides. Précisons que le choix d'une base radiale en particulier est motivé par sa simplicité à mettre en pratique mais également par les différents paramètres qu'elle implique, comme le choix des centres ou du paramètre

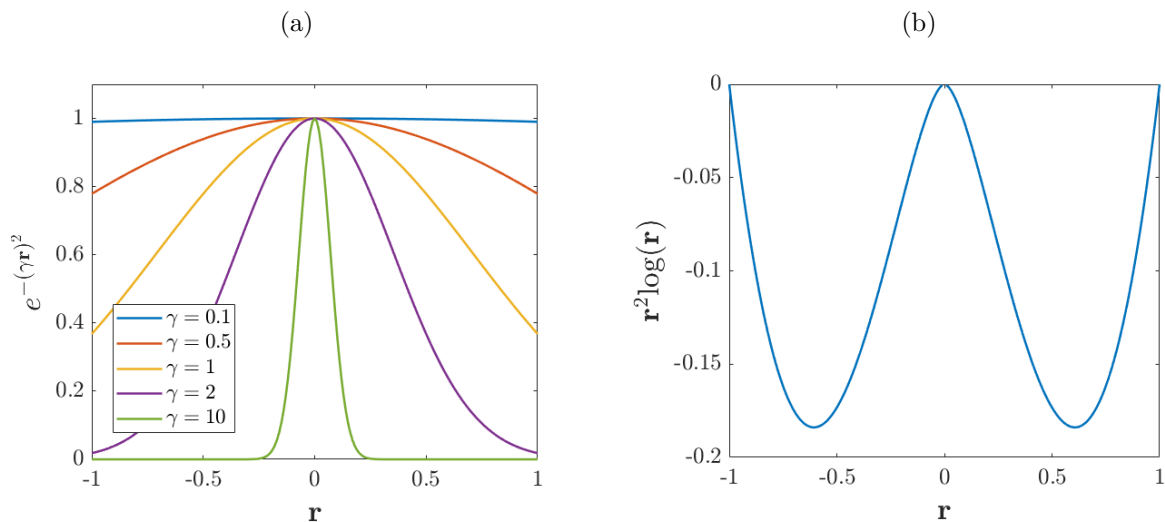


Figure 3.1 – (a) Fonctions radiales de Gauss pour différentes valeurs de γ . Dans ce cas, nous voyons que le paramètre γ influence la courbure de la fonction. (b) Représentation d'une fonction radiale TPS. Notons que dans les deux exemples, nous avons considéré que les fonctions étaient centrées en zéro, c'est-à-dire $c = 0$.

de *tuning* γ pour les fonctions de Gauss. En effet, ces paramètres offrent la possibilité d'adapter la base contrairement à une base de monômes que l'on pourrait qualifier comment étant *fixe*. Précisons cependant que ces paramètres peuvent, a contrario, être très sensibles et compliqués à utiliser.

Maintenant que nous avons fixé les différentes fonctions de base que nous allons utiliser dans ce chapitre, nous devons définir un nouvel opérateur de projection. En effet, il est assez immédiat de constater que $\mathcal{L}\psi_1$ n'est plus un polynôme et que la troncature n'est pas bien définie. Ainsi, nous devons trouver une autre manière de projeter les différentes fonctions $\mathcal{L}\psi_i$ sur le sous-espace de fonctions $\{\psi_i\}_{i=1}^N$ et nous opterons pour une projection (orthogonale) au sens des moindres carrés.

Pour ce faire, considérons dans un premier temps le cas général d'une fonction $f \in \mathcal{Y}$ avec $\mathcal{Y}(\Omega)$ un espace de fonctions continues définies sur l'ensemble compact Ω . Ainsi, la projection orthogonale de cette fonction sur le sous-espace sous-tendu par les fonctions de bases $\{\psi_i\}_{i=1}^N \subset \mathcal{Y}$, avec $N \in \mathbb{N}$, est donnée par l'unique fonction $\tilde{g}(\cdot)$ telle que

$$\begin{aligned}\tilde{g}(\mathbf{x}) &= \arg \min_{g \in \mathcal{Y}} \|f - g\|_2, \\ &= \arg \min_{g \in \mathcal{Y}} \left[\int_{\Omega} (f - g)^2(\mathbf{x}) d\mathbf{x} \right]^{1/2},\end{aligned}$$

Or, rappelons que la norme $\|\cdot\|_2$ permet de définir le produit scalaire entre deux fonctions f et g de la manière suivante :

$$\langle f, g \rangle = \int_{\Omega} f(x)g(x)dx.$$

Dès lors, à l'aide de différents calculs intermédiaires que nous laissons au soin du lecteur, nous pouvons induire une *condition nécessaire d'optimalité* permettant de déterminer l'approximation $\tilde{g}(\cdot)$ en résolvant le système matriciel suivant :

$$\begin{pmatrix} \langle \psi_1, \psi_1 \rangle & \cdots & \langle \psi_1, \psi_N \rangle \\ \vdots & \ddots & \vdots \\ \langle \psi_N, \psi_1 \rangle & \cdots & \langle \psi_N, \psi_N \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} \langle f, \psi_1 \rangle \\ \vdots \\ \langle f, \psi_N \rangle \end{pmatrix}, \quad (3.1)$$

où les coefficients a_i correspondent aux coefficients de $\tilde{g}(\cdot)$ dans la base de fonctions $\Psi(\mathbf{x})$ ¹. Ainsi, nous pouvons tout aussi bien appliquer ce raisonnement en remplaçant la fonction f dans l'expression (3.1) par les différentes fonctions $\mathcal{L}\psi_i$ et nous n'avons qu'à calculer les intégrales présentes dans la relation (3.1) afin de déterminer la matrice \mathbf{L}_N . Cependant, ces intégrales sont en général compliquées à obtenir numériquement et c'est pourquoi il est coutume de les approximer. Dans le cadre du mémoire, nous calculerons ces dernières

1. Précisons tout de même que la condition d'optimalité est obtenue en dérivant la norme de l'erreur et en annulant l'expression ainsi obtenue, permettant de définir le système matriciel (3.1).

à l'aide d'approximation de Monte-Carlo, c'est-à-dire

$$\begin{aligned}\langle \psi_i, \psi_j \rangle &\approx \frac{1}{m} \sum_{k=1}^m \psi_i(x_k) \psi_j(x_k), \\ \langle \mathcal{L}\psi_i, \psi_j \rangle &\approx \frac{1}{m} \sum_{k=1}^m \mathcal{L}\psi_i(x_k) \psi_j(x_k),\end{aligned}$$

où les différents points $\{x_i\}_{i=1}^m$ sont des points distribués de manière uniforme sur Ω . En définissant les deux matrices Ψ_x et $\dot{\Psi}_x$ par

$$\Psi_x = \begin{pmatrix} \psi_1(x_1) & \cdots & \psi_1(x_m) \\ \vdots & \ddots & \vdots \\ \psi_N(x_1) & \cdots & \psi_N(x_m) \end{pmatrix} \quad \text{et} \quad \dot{\Psi}_x = \begin{pmatrix} \mathcal{L}\psi_1(x_1) & \cdots & \mathcal{L}\psi_1(x_m) \\ \vdots & \ddots & \vdots \\ \mathcal{L}\psi_N(x_1) & \cdots & \mathcal{L}\psi_N(x_m) \end{pmatrix},$$

il vient assez directement que

$$\langle \psi_i, \psi_j \rangle = [\Psi_x \Psi_x^T]_{ij} \quad \text{et} \quad \langle \mathcal{L}\psi_i, \psi_j \rangle = [\dot{\Psi}_x \Psi_x^T]_{ij}. \quad (3.2)$$

Dès lors, en rassemblant les expressions obtenues en (3.1) et (3.2), il vient que l'approximation du générateur \mathbf{L}_N est donnée par

$$\mathbf{L}_N^T = (\dot{\Psi}_x \Psi_x^T) (\Psi_x \Psi_x^T)^+,$$

où $(\Psi_x \Psi_x^T)^+$ dénote la matrice pseudo-inverse de Penrose² de $\Psi_x \Psi_x^T$. Dans ce cas, une approximation des fonctions propres de l'opérateur de Koopman est obtenue en prenant les vecteurs propres à droite de la matrice \mathbf{L}_N .

Notons que la projection que nous venons de présenter définit une projection *orthogonale*, de sorte que la solution obtenue $\tilde{g}(\cdot)$ est la meilleure approximation de $f(\cdot)$ sur Ω . Cependant, rien ne nous empêche de définir la projection sur un sous-ensemble de Ω , disons $\tilde{\Omega}$. Dans ce cas, l'approximation $\tilde{g}(\cdot)$ sera la meilleure projection de $f(\cdot)$ dans $\tilde{\Omega}$. Nous pouvons voir cette projection particulière comme un opérateur qui enverrait toutes les fonctions à support dans l'ensemble $\tilde{\Omega}^c$, sur la fonction nulle³. Nous verrons en réalité que cette considération particulière est intéressante dans le cas particulier des fonctions radiales et engendre de meilleurs résultats pour des régions assez petites. Sauf indication contraire, l'ensemble $\Omega \subset \mathbb{X}$ désignera le domaine du produit scalaire définissant la projection.

3.1.1 Implémentation d'un code générique

Dans cette section, nous allons brièvement discuter de l'implémentation d'une méthode générale pour le calcul des fonctions propres de l'opérateur de Koopman. En effet, une fois que les fonctions de base et la projection sont fixées, nous devons calculer une fonction de Lyapunov pour induire la stabilité du système étudié. Ainsi, une grosse partie du travail

2. Notons que dans la littérature nous retrouvons parfois la régularisation de Tikhonov au lieu de la matrice pseudo-inverse. L'inverse de Tikhonov d'une matrice A est la matrice inversible donnée par $(A + \varepsilon I)^{-1}$ avec I la matrice identité de même taille que A .

3. $\tilde{\Omega}^c$ désigne ici le complémentaire de l'ensemble $\tilde{\Omega}$ au sens topologique usuel.

est en réalité cachée derrière l'implémentation générale d'une telle méthode ainsi que des réflexions sous-jacentes. La méthode que nous avons développée permet de calculer une approximation d'une fonction de Lyapunov et de sa dérivée temporelle sur un domaine discrétisé avec n'importe quelles fonctions de base encodées par un utilisateur. Pour ce faire, nous avons notamment utilisé le calcul symbolique à disposition dans Matlab et qui permet justement d'être générale dans l'implémentation⁴. Il serait assez fastidieux d'expliquer la manière avec laquelle cette méthode a été implémentée et c'est pourquoi nous n'en discuterons pas en détail dans le cadre de ce rapport. Cependant, nous tenons tout de même à préciser que l'implémentation du code a été repensée plusieurs fois afin d'obtenir des résultats en un temps de calcul suffisamment raisonnable. Par exemple, le calcul de la fonction de Lyapunov et de sa dérivée temporelle ont été considérés *vectoriellement* permettant ainsi de jouir des facultés de Matlab à exécuter de telles opérations. Tentons d'éclaircir ce point à travers le calcul de la fonction de Lyapunov. Pour ce faire, rappelons que dans le cadre de la stabilité à travers l'opérateur de Koopman, une fonction de Lyapunov est donnée par

$$\mathcal{V}(x) = \left(\sum_{i=1}^m c_i |\phi_{\lambda_i}(x)|^p \right)^{1/p}.$$

Par souci de facilité dans le cadre du mémoire, nous avons considéré le cas particulier où $p = 2$. Supposons que nous cherchons à évaluer la fonction sur le domaine \mathbb{X} et considérons que ce dernier ait été discrétisé en s points x_i pour $i = 1, \dots, s$. Soit $\tilde{\phi}_{\lambda_1}(\cdot)$ l'approximation de $\phi_{\lambda_1}(\cdot)$ dans la base de fonctions $\{\psi_i\}_{i=1}^N$. Dans ce cas, il vient que

$$\begin{pmatrix} \tilde{\phi}_{\lambda_1}(x_1) \\ \vdots \\ \tilde{\phi}_{\lambda_1}(x_s) \end{pmatrix} = \alpha_1 \begin{pmatrix} \psi_1(x_1) \\ \vdots \\ \psi_N(x_s) \end{pmatrix} + \dots + \alpha_N \begin{pmatrix} \psi_1(x_1) \\ \vdots \\ \psi_N(x_s) \end{pmatrix},$$

où les différents coefficients α_i correspondent aux coefficients de $\tilde{\phi}_{\lambda_1}(\cdot)$ dans la base $\{\psi_i\}_{i=1}^N$. Ainsi, en se rappelant que

$$|\tilde{\phi}_{\lambda_1}(\mathbf{x})|^2 = [\mathcal{R}(\tilde{\phi}_{\lambda_1}(\mathbf{x}))]^2 + [\mathcal{I}(\tilde{\phi}_{\lambda_1}(\mathbf{x}))]^2$$

où $\mathcal{R}(\tilde{\phi}_{\lambda_1}(\mathbf{x}))$ et $\mathcal{I}(\tilde{\phi}_{\lambda_1}(\mathbf{x}))$ désigne respectivement les parties réelles et imaginaires de $\tilde{\phi}_{\lambda_1}(\mathbf{x})$, nous pouvons calculer la fonction de Lyapunov de manière vectorielle selon l'Algorithme 2.

Grâce à cette approche, nous obtenons ainsi une approximation de la fonction de Lyapunov plus rapidement comparé à un calcul ponctuel, plus long en Matlab pour une grille suffisamment grande. Précisons que nous avons uniquement détaillé le cas de la fonction de Lyapunov car la démarche est assez simple. Cependant, le calcul de sa dérivée temporelle par exemple a demandé une attention plus particulière pour en concevoir une évaluation vectorielle qui soit optimale au niveau du temps de calcul. Néanmoins, nous n'entrerons pas dans les détails dans le cadre du mémoire par souci de facilité et afin d'éviter l'introduction de différentes variables temporaires établies dans le code qui entacheraient la compréhension du calcul si nous devions l'exposer dans ce rapport.

4. Précisons que la méthode générale a en réalité été implémentée pour la projection *orthogonale* mais également pour la troncature dans le cadre des monômes.

Algorithme 1 CALCUL DE $\mathcal{V}(\mathbf{x})$ **Require:** Domaine \mathbb{X} discrétisé en un ensemble de points $\{x_i\}_{i=1}^s$.**Ensure:** Approximation de la fonction de Lyapunov $\mathcal{V}(\mathbf{x})$ sur le domaine discrétisé.

- 1: Soit v la matrice construite par colonne et contenant les coefficients des approximations de $\tilde{\phi}_{\lambda_j}(\mathbf{x})$ pour $j = 1, \dots, m$ avec m la dimension du système.
- 2: **for** $i = 1 \rightarrow N$ **do**
- 3: $f \leftarrow$ évaluation de ψ_i sur la grille. On a alors $f \in \mathbb{R}^s$.
- 4: **for** $k = 1 \rightarrow m$ **do**
- 5: $\tilde{\phi}_{\lambda_k} \leftarrow \tilde{\phi}_{\lambda_k} + v(i, k)f$
- 6: **end for**
- 7: **end for**
- 8: **for** $k = 1 \rightarrow m$ **do**
- 9: $\mathcal{V} \leftarrow \mathcal{V} + c_k \left\{ \left[\mathcal{R}(\tilde{\phi}_{\lambda_k}) \right]^2 + \left[\mathcal{I}(\tilde{\phi}_{\lambda_k}) \right]^2 \right\}$
- 10: **end for**

3.2 Application de la méthode et étude des paramètres

Maintenant que nous avons défini tout le bagage nécessaire afin de déterminer une approximation des fonctions propres dans une autre base, nous allons voir comment le mettre en pratique dans le cas particulier des fonctions de base radiales.

Comme nous avons pu le voir dans la section précédente, la projection orthogonale que nous avons définie nécessite de calculer des approximations de Monte-Carlo par l'intermédiaire de points de données x_i aléatoirement générés sur un certain domaine Ω . Cependant, il n'est pas spécialement évident ou trivial de définir un tel ensemble, au même titre que les centres des fonctions de base, dont le domaine d'échantillonnage sera noté Θ . Dès lors et afin de pouvoir mesurer les performances de la méthode au cas particulier des fonctions de base radiales, nous nous pencherons dans un premier temps sur les fonctions TPS où nous étudierons l'impact du choix respectif de Θ et Ω sur l'approximation du bassin d'attraction. Lors de cette étude, nous verrons que le Théorème de Lyapunov 1.1 ne tient plus forcément et nous généraliserons alors le principe d'invariance de LaSalle pour pouvoir induire la stabilité des systèmes que nous étudions. Ensuite, une fois que nous aurons établi comment les fonctions radiales permettent d'induire la stabilité des différents points fixes, nous nous pencherons sur les fonctions radiales de Gauss où nous jugerons de la qualité de l'approximation du bassin d'attraction en tenant compte du paramètre de *tuning* γ . Enfin, nous appliquerons la méthode à deux systèmes en trois dimensions et nous discuterons des approximations obtenues.

3.2.1 Étude des fonctions de base *thin plate spline*

Commençons dans un premier temps avec les fonctions radiales TPS étant donné que ces dernières ne dépendent d'aucun paramètre supplémentaire si ce n'est Θ et Ω . Comme nous l'avons mentionné, la méthode numérique que nous étudions nécessite de définir des centres et des points de données x_i . Concentrons-nous dans un premier temps sur l'impact

du choix des centres pour un domaine de projection Ω fixé. Définir les centres des fonctions radiales n'est pas particulièrement évident. En effet, il n'existe pas de méthode systématique nous permettant de les déterminer et leur choix dépend généralement du problème étudié ou de sa géométrie⁵. Dans cette section, nous allons voir comment le choix de ces derniers impacte l'approximation du bassin d'attraction pour les deux systèmes à deux dimensions (1.7) et (2.9). Plus particulièrement, nous nous pencherons dans un premier temps sur le système de Van der Pol où nous effectuerons une comparaison détaillée. Sur base des différents résultats obtenus, nous passerons brièvement sur le système (2.9) en mettant en exergue les résultats les plus concluants.

Variation du domaine d'échantillonnage des centres Θ pour Ω fixé (Van der Pol)

Avant toute chose, rappelons que la qualité de l'approximation du bassin d'attraction est complètement caractérisée par le signe de $\dot{\mathcal{V}}$. Ainsi, afin de pouvoir illustrer et comparer les différents résultats entre eux, il semble tout à fait naturel de définir les deux ensembles

$$\dot{\mathcal{V}}_{\geq 0} = \{\mathbf{x} \in \mathbb{X} \mid \dot{\mathcal{V}}(\mathbf{x}) \geq 0\} \quad \text{et} \quad \dot{\mathcal{V}}_{< 0} = \{\mathbf{x} \in \mathbb{X} \mid \dot{\mathcal{V}}(\mathbf{x}) < 0\},$$

où par abus de langage, $\mathcal{V}(\mathbf{x})$ désigne l'approximation de la fonction de Lyapunov obtenue en (1.15) avec les fonction propres estimées. À la Figure 3.2⁶, nous avons représenté ces régions en fonction du domaine d'échantillonnage Θ pour $\Omega = [-3, 3]$ fixé. Comme nous pouvons l'observer, les résultats que nous obtenons ne sont pas particulièrement bons. En effet, la géométrie de ces différentes régions ne permet pas de définir un ensemble de niveau Ψ_δ autour de l'origine qui soit suffisamment grand et ce, peu importe le domaine

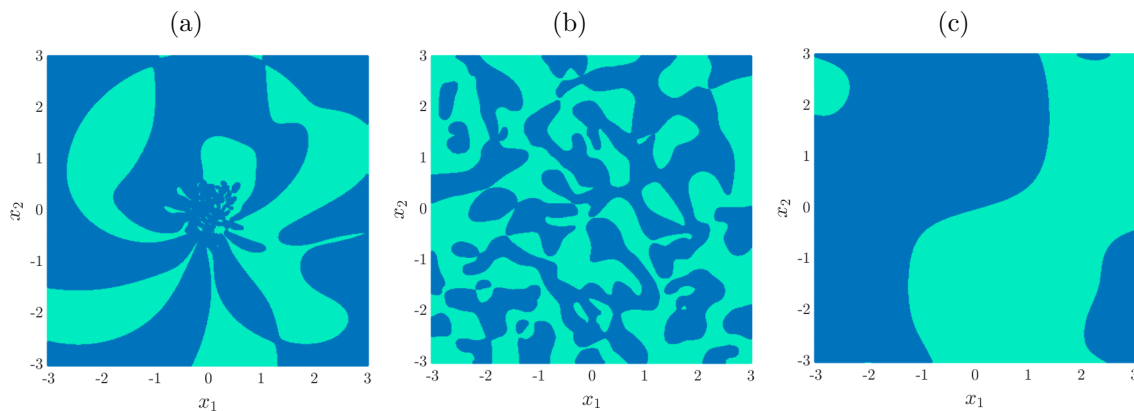


Figure 3.2 – Approximation des régions $\dot{\mathcal{V}}_{< 0}$ et $\dot{\mathcal{V}}_{> 0}$ obtenues à l'aide de 100 fonctions de base radiales TPS pour $\Omega = [-3, 3] \times [-3, 3]$ avec (a) $\Theta = [-0.5, 0.5] \times [-0.5, 0.5]$, (b) $\Theta = [-3, 3] \times [-3, 3]$ et (c) $\Theta = [-50, 50] \times [-50, 50]$. Pour chacune des figures, nous avons respectivement représenté en bleu clair et en bleu foncé les ensembles $\dot{\mathcal{V}}_{< 0}$ et $\dot{\mathcal{V}}_{\geq 0}$. Dans ce cas, les résultats ne permettent pas d'obtenir une bonne approximation. Cependant, les ensembles $\dot{\mathcal{V}}_{\geq 0}$ et $\dot{\mathcal{V}}_{< 0}$ sont plus *lisses* en (c). Les résultats ont été ici obtenus à l'aide de 100 fonctions de base radiales TPS.

5. Notons également que le choix d'une base radiale particulière peut également affecter celui des centres, au vu des formes différentes de ces fonctions (voir Figure 3.1).

6. Précisons que les domaines de la Figure 3.2 ont été choisis sur base de différents tests numériques afin d'extraire différents comportements sans devoir étudier et analyser une multitude de domaines différents.

des centres. Néanmoins, nous remarquons que pour Θ suffisamment large, ces régions sont plus distinguables et plus *lisses*, ce qui laisse penser qu'en changeant le domaine de la projection, nous pourrions obtenir de meilleurs résultats. Notons par ailleurs que le choix de la projection est en réalité tout à fait arbitraire et est intuitivement et uniquement motivé par la Figure 1.6 où nous remarquons que le point fixe ainsi que son bassin d'attraction sont entièrement capturés sur ce domaine.

Avant de nous lancer dans une étude comparative du domaine d'intégration Ω , nous tenons tout de même à effectuer la remarque suivante concernant les résultats obtenus. Lorsque nous utilisons des fonctions de base radiales, nous ne sommes pas sûrs que les fonctions propres que nous approximons soient associées aux valeurs propres de la matrice Jacobienne au point d'équilibre et c'est ce que nous observons lorsque le domaine d'intégration est assez grand. En effet, nous ne sommes plus forcément dans le cas particulier où la projection des fonctions propres coïncident avec les fonctions propres du générateur infinitésimal projeté de sorte que les valeurs propre de \mathbf{L}_N contiennent celles de la matrice Jacobienne au point d'équilibre. Dans ce cas, nous devons alors *choisir* les fonctions propres que nous allons utiliser pour construire la fonction de Lyapunov $\mathcal{V}(\mathbf{x})$ et par convention, nous choisirons les fonctions propres qui s'en rapprochent le plus, pour autant que leurs parties réelles soient négatives. Nous reviendrons sur ce point notamment dans l'application de la méthode pour le deuxième exemple.

Remarque 3.1. *Notons que nous imposons que la partie réelle des approximations numériques soient positives pour rester en accord avec la relation (2.8), affirmant que $\mathcal{V}(\mathbf{x})$ est une fonction de Lyapunov pour autant que les fonctions propres soient à parties réelles négatives.*

Variation de l'ensemble Ω pour un domaine d'échantillonnage des centres Θ fixé (Van der Pol)

À travers la partie précédente, nous avons vu que le choix des centres ne semblait pas particulièrement impacter l'approximation du bassin d'attraction. En effet, nous obtenons une approximation du bassin très mauvaise et quasi inexistante bien que les ensembles de niveau soient plus *lisses* pour un large domaine Θ . Nous pouvons alors essayer de pallier ce problème en tournant notre attention cette fois-ci sur le domaine de la projection Ω . Pour ce faire, nous avons reproduit la Figure 3.2 en faisant varier cette fois-ci les bornes du domaine Ω avec des centres fixés dans l'intervalle $\Theta = [-3, 3] \times [-3, 3]$. De nouveau, ce choix est tout à fait arbitraire. Nous remarquons alors à la Figure 3.3 que l'approximation du bassin est meilleure en considérant les bornes de l'ensemble Ω assez petites. Dans ce cas, il est possible d'extraire une courbe de niveau que nous avons représentée en mauve sur la Figure 3.3 (a). Cependant, pour induire la stabilité du point fixe nous considérons uniquement les points de l'ensemble strictement inclus dans $\dot{\mathcal{V}}_{<0}$. Néanmoins, même si nous obtenons une approximation du bassin, cette dernière ne nous permet pas de capturer toute la géométrie du bassin d'attraction délimité par le cycle limite en rouge. Ceci est en particulier dû à la géométrie complexe des régions $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$. Une manière de contourner ce problème serait alors de considérer $\Theta = [-50, 50]$ pour lequel nous avons vu à travers la Figure 3.2 (c) que ces régions admettaient justement une géométrie plus *lisse*. Dans ce cas, nous obtenons une approximation du bassin vraiment satisfaisante représentée à la Figure 3.4 où nous avons ici choisi $\Omega = [-0.1, 0.1]$ au lieu de $\Omega = [-0.5, 0.5]$

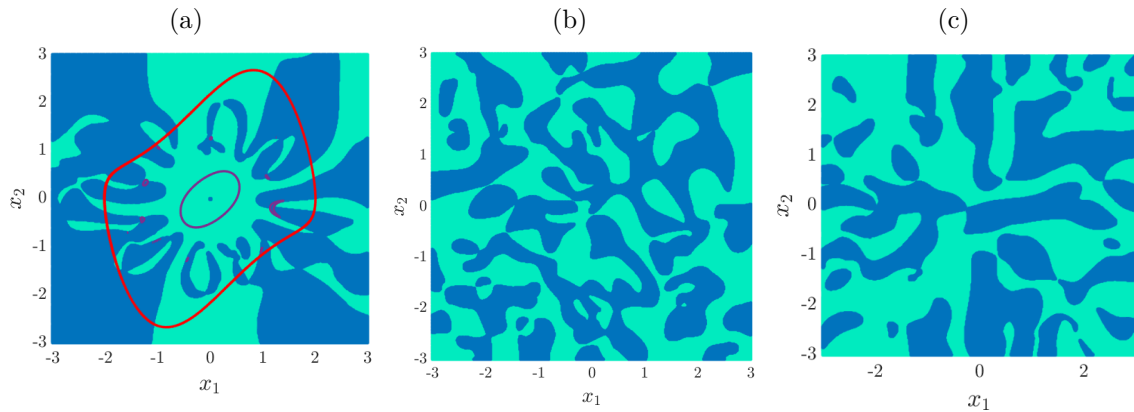


Figure 3.3 – Approximation des régions $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$ obtenues à l'aide de 100 fonctions de base radiales TPS pour $\Theta = [-3, 3] \times [-3, 3]$ avec (a) $\Omega = [-0.5, 0.5] \times [-0.5, 0.5]$, (b) $\Omega = [-3, 3] \times [-3, 3]$ et (c) $\Omega = [-50, 50] \times [-50, 50]$. Pour chacune des figures, nous avons respectivement représenté en bleu clair et en bleu foncé les ensembles $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$. Dans ce cas, nous nous obtenons une meilleure approximation pour Ω assez petit où nous pouvons extraire un sous-ensemble de niveau dans la région $\dot{\mathcal{V}}_{<0}$. Les résultats ont été ici obtenus à l'aide de 100 fonctions de base radiales TPS.

car les résultats étaient un peu meilleurs. Notons que même si l'approximation n'est pas aussi bien que celle obtenue dans le cadre des monômes, elle reste tout à fait acceptable et prometteuse.

Application des fonctions radiales pour le deuxième exemple

Penchons-nous à présent sur le système (2.9). Sur base des constatations précédentes, considérons directement le cas où $\Theta = [-50, 50]$ et $\Omega = [-0.1, 0.1]$ à la Figure 3.5 (a).

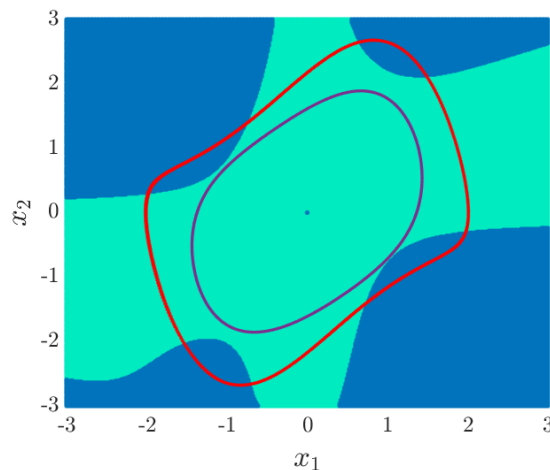


Figure 3.4 – Approximation du bassin d'attraction avec $\Theta = [-50, 50]$ et $\Omega = [-0.1, 0.1]$ où nous avons représenté en bleu clair et en bleu foncé les ensembles $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{>0}$. Dans ce cas, nous remarquons que nous obtenons une bonne estimation du bassin d'attraction donnée par l'ensemble de niveau en mauve. Ce résultat été ici obtenu à l'aide de 100 fonctions de base radiales TPS.

Dans ce cas, nous remarquons que l'approximation du bassin d'attraction n'est pas très satisfaisante pour ce choix de Θ . Cependant, lorsque nous élargissons ce domaine, nous remarquons à la Figure 3.5 (b) que l'approximation obtenue est bien meilleure et nous arrivons à capturer une bonne partie du bassin d'attraction. En réalité, il est possible d'affiner davantage l'approximation. En effet, depuis le début de cette section, nous supposons que les différents centres c_i sont générés aléatoirement dans l'ensemble Θ à l'aide de la fonction `rand` en Matlab. Dès lors, nous pourrions tout aussi bien considérer une répartition uniforme des centres selon une discrétisation du domaine Θ en une grille dont les centres en seraient les points. Dans ce cas, nous remarquons à la Figure 3.5 (c) que l'approximation obtenue est encore meilleure et nous couvrons une grosse partie du bassin. Notons que ce résultat en particulier semble légèrement meilleur que celui obtenu pour les monômes à la Figure 2.3.

Remarque 3.2. *Précisons qu'en général, les résultats sont sensiblement identiques selon que les données soient obtenues aléatoirement ou selon une grille. Cependant, dans certains cas, une des deux considérations peut entraîner de meilleurs résultats comme nous l'avons observé pour le système (2.9). De nouveau, il n'y pas de choix meilleur qu'un autre et ce dernier dépend bien souvent du système étudié. Par exemple, pour le système de Van der Pol, les deux approches ont été testées et les résultats sont similaires.*

En ce qui concerne le deuxième exemple (2.9), nous n'avons pas considéré une étude comparative détaillée comme pour le système de Van der pol et nous avons directement considéré les résultats les plus concluants après différents tests numériques. En réalité, une deuxième étude détaillée serait redondante du point de vue des conclusions que nous pouvons tirer et n'est pas spécialement pertinente dans le cadre du mémoire. En effet, ce chapitre a pour but de montrer que l'utilisation de fonctions de base radiales permet également d'obtenir de bonnes estimations du bassin d'attraction, ouvrant la voie à leur utilisation pour des systèmes de plus grande dimension. Par ailleurs, au vu de l'approximation du bassin d'attraction obtenue pour cet exemple, il semble compliqué d'obtenir un résultat plus concluant compte tenu de la géométrie des régions $\dot{V}_{<0}$ et $\dot{V}_{\geq 0}$.

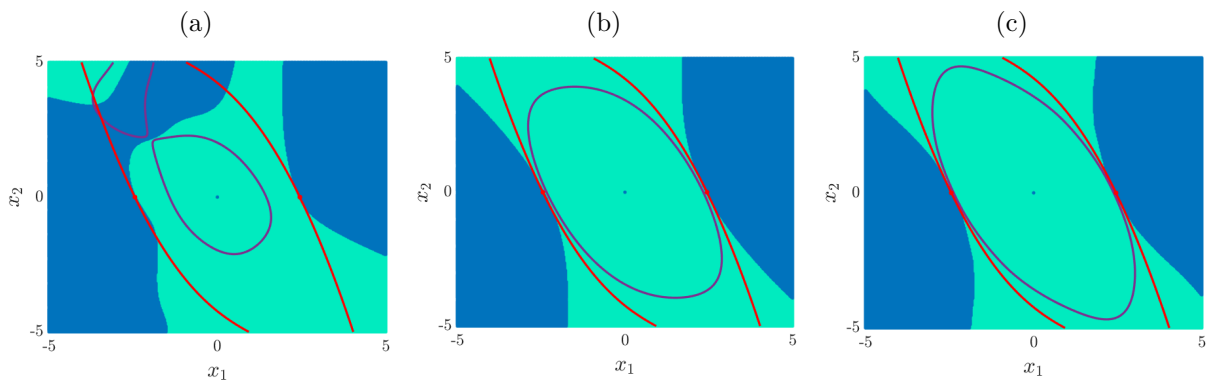


Figure 3.5 – Approximation du bassin d'attraction avec $\Omega = [-0.1, 0.1]$ pour (a) $\Theta = [-50, 50]$ et les centres c_i générés aléatoirement dans ce domaine, (b) $\Theta = [-150, 150]$ et les centres c_i générés aléatoirement dans ce domaine, (c) $\Theta = [-5, 5]$ et les centres c_i distribués uniformément. Dans chacun de ces exemples, nous avons considéré 100 fonctions de base radiales de sorte que le domaine Θ en (c) est subdivisé en une grille de 10×10 points.

Notons que même si nous n’entrons pas dans les détails de cet exemple, nous avons également observé à travers différents tests numériques que nous obtenions de meilleurs résultats dès que le domaine de la projection Ω est suffisamment petit (comme c’est le cas à la Figure 3.5). Cependant, cette condition reste encore un peu heuristique. Une explication intuitive que l’on pourrait en donner serait qu’un domaine trop grand capture des zones d’instabilité comme le cycle limite ou les deux points fixes instables pour le système (2.9) par exemple. Dans ce cas, le spectre de la matrice \mathbf{L}_N pourrait contenir des valeurs propres associées à des propriétés instables du système dynamique sous-jacent. Dès lors, les fonctions propres estimées ne contiendraient plus toute l’information nécessaire pour caractériser la partie stable du système, et pourraient alors traduire la géométrie complexe des ensembles $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$. Cette explication est notamment soutenue par le fait que lorsque le domaine de projection Ω est trop grand, nous ne retrouvons pas du tout les valeurs propres de la matrice Jacobienne au point d’équilibre. Par ailleurs, cette explication est corroborée par le fait que les meilleurs résultats que nous avons obtenus (voir Figure 3.4, 3.5 (b) et (c)) ont été calculés avec des fonctions propres dont les valeurs propres estimées étaient très proches de celles de la matrice Jacobienne au point d’équilibre. Dans ce cas, il n’est pas si étonnant que ces fonctions propres engendrent de bons résultats étant donné que nous entrons dans certaines conditions de la Proposition 1.6 où nous savons que le support des fonctions exactes, associées aux valeurs propres de la matrice Jacobienne, capturent la géométrie du bassin d’attraction. Rappelons toutefois que ce n’est pas parce que les valeurs propres ne sont pas proches de celles de la matrice Jacobienne au point fixe que nous devons nous alarmer étant donné qu’il n’y a pas de garantie que ce soit le cas. Cependant, comme les valeurs propres de la matrice Jacobienne encodent des propriétés de stabilité de l’équilibre, certes locales, nous pouvons supposer que les fonctions propres associées à ces valeurs propres tiennent compte de ces informations de stabilité locale pour induire la stabilité globale du système. Notons néanmoins que nous devons rester prudents face à cet argument étant donné qu’il est possible d’obtenir des fonctions propres associées à des valeurs propres proches de celles de la matrice Jacobienne au point d’équilibre et pour lesquelles les résultats ne sont pas concluants (voir section 3.2.3).

3.2.2 Généralisation du Principe d’invariance de LaSalle

À travers la section 3.2.1, nous avons vu que pour un choix de paramètres bien précis, les fonctions radiales *thin plate spline* permettent d’obtenir une bonne estimation du bassin d’attraction de l’origine donné par les différents ensembles de niveau illustrés en mauve aux Figures 3.4, 3.5 (b) et (c). Cependant, nous ne sommes pas tout à fait rigoureux lorsque nous affirmons que ces régions correspondent au bassin d’attraction du point d’équilibre. En effet, lorsque nous utilisons des fonctions de base radiales, le Théorème 1.1 n’est pas valide étant donné que nous n’avons pas forcément que la fonction de Lyapunov vérifie

$$\mathcal{V}(\mathbf{x}^*) = 0.$$

Ainsi, nous ne pouvons pas vraiment affirmer que les courbes de niveau (en mauve sur les figures) définissent rigoureusement le bassin d’attraction de l’équilibre à l’origine. Néanmoins, nous pourrions penser qu’en définissant la fonction de Lyapunov

$$\mathcal{H}(\mathbf{x}) := \mathcal{V}(\mathbf{x}) - \mathcal{V}(\mathbf{x}^*),$$

le problème serait résolu étant donné que cette nouvelle fonction s'annule désormais au point d'équilibre et nous avons par ailleurs $\dot{\mathcal{H}}(x) = \dot{\mathcal{V}}(x)$ de sorte que les régions $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$ demeurent inchangées. Autrement dit, en appliquant le Théorème 1.1 avec cette fonction de Lyapunov, nous pourrions induire la stabilité de l'origine. Cependant, un problème persiste toujours. En effet, rappelons que le Théorème de Lyapunov nécessite que $\dot{\mathcal{V}}(\mathbf{x}) < 0$ pour tout \mathbf{x} dans un certain domaine D . Toutefois, la preuve de ce théorème requiert initialement que la fonction de Lyapunov \mathcal{V} soit décroissante le long des trajectoires, c'est à dire que

$$\mathcal{V}(\mathbf{x}(t)) < \mathcal{V}(\mathbf{x}(0)), \quad \forall t \geq 0.$$

Dès lors, en imposant $\dot{\mathcal{V}}(\mathbf{x}(t)) < 0$ pour tout $\mathbf{x} \in D$, nous avons nécessairement que cette condition est satisfaite pour peu que D soit un ensemble positivement invariant. Plus précisément, la preuve du théorème fait intervenir la relation suivante [8, Théorème 3.1, pp. 100-102]

$$\dot{\mathcal{V}}(\mathbf{x}(t)) < 0 \implies \mathcal{V}(\mathbf{x}(t)) < \mathcal{V}(\mathbf{x}(0)), \quad \forall t \geq 0 \quad (3.3)$$

qui n'est satisfaite que si D est positivement invariant. Cependant, nous ne satisfaisons pas ce critère pour les deux exemples étudiés. En utilisant une discrétisation du domaine assez fine pour les systèmes (1.7) et (2.9), nous remarquons à la Figure 3.6 qu'il existe une petite région autour de l'équilibre $\Gamma \subseteq \dot{\mathcal{V}}_{\geq 0}$ entouré par $\dot{\mathcal{V}}_{<0}$. Dans ce cas, le théorème de Lyapunov devrait alors se restreindre à l'ensemble

$$D := \Xi \setminus \Gamma \cup \{\mathbf{x}^*\}, \quad (3.4)$$

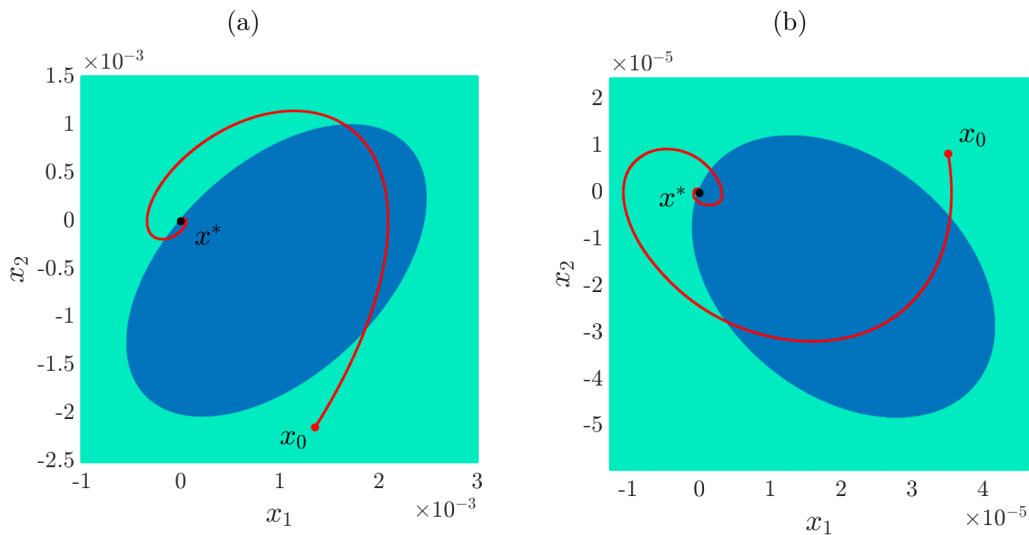


Figure 3.6 – Zoom dans un voisinage de l'origine pour (a) la Figure 3.4, (b) la Figure 3.5 (c). Dans ce cas, nous remarquons qu'il existe une petite région $\Gamma \subseteq \dot{\mathcal{V}}_{\geq 0}$ avec Γ entouré par $\dot{\mathcal{V}}_{<0}$. Pour chacun de ces deux exemples nous avons représenté une trajectoire (en rouge) pour une condition initiale dans l'ensemble $\Xi \setminus \Gamma \cup \{\mathbf{x}^*\}$ avec Ξ l'ensemble dont les frontières sont délimité par la plus grande courbe de niveau incluse dans $\dot{\mathcal{V}}_{<0}$ pour ces deux systèmes respectifs (voir courbes en mauves aux figures associées). Pour chacun des deux exemples, nous remarquons que cet ensemble n'est pas positivement invariant de sorte que les Théorèmes 1.1 et 3.1 ne s'applique pas.

avec $\Gamma \subset \Xi$ et Ξ correspond à l'ensemble dont les frontières sont délimité par la plus grande courbe de niveau incluse dans $\dot{\mathcal{V}}_{<0}$. Cependant, pour les deux systèmes étudiés, la relation (3.3) n'est plus forcément satisfaite étant donné que l'ensemble D n'est plus invariant comme nous l'observons à la Figure 3.6. En effet, nous remarquons que la trajectoire issue de la condition initiale $x_0 \in D$ ne reste pas identiquement dans cet ensemble pour tout $t \geq 0$. Dès lors et afin de pouvoir induire la stabilité de tels points fixes, nous pouvons dans un second temps penser au principe d'invariance de LaSalle suivant [8, Théorème 3.4, pp. 115].

Théorème 3.1 (Principe d'invariance de LaSalle)

Soit $Y \subset \mathcal{H}$ un ensemble compact positivement invariant et considérons une fonction $\mathcal{V} : \mathcal{H} \rightarrow \mathbb{R}$ continuellement dérivable telle que $\dot{\mathcal{V}}(\mathbf{x}) \leq 0$ dans Y . Soit E l'ensemble des points dans Y telle que $\dot{\mathcal{V}}(\mathbf{x}) = 0$ et M le plus grand ensemble invariant dans E . Dans ce cas, n'importe quelle solution dont la condition initiale appartient Y tend vers M lorsque $t \rightarrow \infty$.

Malheureusement, l'ensemble $Y := D$ donné en (3.4) dans les deux exemples n'est de nouveau pas positivement invariant de sorte que le théorème ne s'applique plus et nous ne pouvons rien dire de la stabilité de l'origine pour nos deux exemples. Cependant, dans le cadre du mémoire, nous en avons fourni une généralisation, permettant ainsi d'induire la stabilité des points d'équilibre à l'origine.

Théorème 3.2 (Généralisation du principe d'invariance de LaSalle)

Soit $\mathcal{V} : D \rightarrow \mathbb{R}$ une fonction continue et positive avec D un sous-ensemble compact positivement invariant de \mathbb{R}^n et définissons les sous-ensembles non vides de D suivants :

$$\dot{\mathcal{V}}_{<0} = \{\mathbf{x} \in D \mid \dot{\mathcal{V}}(\mathbf{x}) < 0\}, \quad \dot{\mathcal{V}}_{>0} = \{\mathbf{x} \in D \mid \dot{\mathcal{V}}(\mathbf{x}) > 0\}, \quad \dot{\mathcal{V}}_0 = \{\mathbf{x} \in D \mid \dot{\mathcal{V}}(\mathbf{x}) = 0\}.$$

Dès lors, pour tout $\mathbf{x}_0 \in D$, $L^+(\mathbf{x}_0) \subseteq M \subseteq \dot{\mathcal{V}}_0$ avec M le plus grand ensemble invariant contenu dans $\dot{\mathcal{V}}_0$. L'ensemble $L^+(\mathbf{x}_0)$ désigne ici l'ensemble limite de \mathbf{x}_0 et est défini comme l'ensemble des points $p \in D$ pour lesquelles il existe une suite $\{t_n\}_n$ telle que [8]

$$\varphi^{t_n}(\mathbf{x}_0) \xrightarrow{n \rightarrow \infty} p.$$

Démonstration. En tant que résultats préliminaires à cette démonstration, rappelons le Lemme 3.1 de la référence [8, Lemme 3.1, pp. 114].

Lemme 3.1. *Si une solution $\varphi^t(\mathbf{x}_0)$ de (1.2) est bornée et appartient à l'ensemble \mathbb{X} pour tout $t \geq 0$, alors dans ce cas son ensemble limite $L^+(\mathbf{x}_0)$ est un ensemble non-vide, compact et positivement invariant. De plus,*

$$\varphi^t(\mathbf{x}_0) \xrightarrow{m \rightarrow \infty} L^+(\mathbf{x}_0).$$

Soit $\mathbf{x}_0 \in D \subset \mathbb{X}$ arbitrairement fixé. Étant donné que D est un ensemble compact et que la fonction $\varphi^t(\cdot)$ est une fonction continue, nous avons qu'elle est bornée pour tout $t \geq 0$. Par ailleurs, nous avons $\varphi^t(\mathbf{x}_0) \in D$ étant donné que ce dernier est positivement invariant. Dans ce cas, par le Lemme 3.1 nous avons que $L^+(\mathbf{x}_0)$ est un ensemble non-vide, compact et positivement invariant. Nous allons dans un premier temps montrer que pour tout $p \in L^+(\mathbf{x}_0)$, $p \notin \dot{\mathcal{V}}_{<0}$. Pour ce faire, supposons par l'absurde que $p \in L^+(\mathbf{x}_0)$ et $p \in \dot{\mathcal{V}}_{<0}$. Par définition de $L^+(\mathbf{x}_0)$, nous qu'il existe une suite $(t_n)_{n \in \mathbb{N}}$ telle que

$$\varphi^{t_n}(\mathbf{x}_0) \xrightarrow{n \rightarrow \infty} p.$$

Cependant, par continuité de $\mathcal{V}(\mathbf{x})$ nous avons que

$$\lim_{n \rightarrow \infty} \mathcal{V}[\varphi^{t_n}(\mathbf{x}_0)] = \mathcal{V}(p) := c > 0. \quad (3.5)$$

Par hypothèse absurde, il vient que $p \in \dot{\mathcal{V}}_{<0}$ de sorte qu'il existe $m \in \mathbb{N}$ et $r > 0$ tels que pour tout $n > m$ et pour toutes conditions initiales dans $\overline{B}_r(p)$ ⁷, les trajectoires issues de ces conditions restent dans la boule et convergent vers p . Autrement dit,

$$\varphi^{t_n}(\mathbf{x}_0) \xrightarrow[n > m]{n \rightarrow \infty} p,$$

où $\varphi^{t_n}(\mathbf{x}_0) \in \overline{B}_r(p)$ pour tout $n > m$. Cependant, nous savons que pour tout $s \in \mathbb{N}$,

$$\mathcal{V}[\varphi^{t_m+s}(\mathbf{x}_0)] = \mathcal{V}[\varphi^{t_m}(\mathbf{x}_0)] + \int_{t_m}^{t_m+s} \dot{\mathcal{V}}[\varphi^\tau(\mathbf{x}_0)] d\tau. \quad (3.6)$$

Posons alors $-\gamma = \max_{\mathbf{x} \in \overline{B}_r(p)} \dot{\mathcal{V}}(\mathbf{x}) < 0$. Nous avons alors que pour tout $s \in \mathbb{N}$, la relation (3.6) livre

$$\mathcal{V}[\varphi^{t_m+s}(\mathbf{x}_0)] < \mathcal{V}[\varphi^{t_m}(\mathbf{x}_0)] - \gamma t_m - \gamma t_s,$$

et en particulier pour $s \rightarrow \infty$, nous avons

$$\lim_{s \rightarrow \infty} \mathcal{V}[\varphi^{t_m+s}(\mathbf{x}_0)] < -\infty,$$

ce qui contredit la relation (3.5).

Supposons à présent que nous voulons montrer que pour tout $p \in L^+(\mathbf{x}_0)$, $p \notin \dot{\mathcal{V}}_{>0}$. De nouveau, nous allons utiliser un raisonnement par l'absurde et considérer que $p \in L^+(\mathbf{x}_0)$ et $p \in \dot{\mathcal{V}}_{>0}$. En réalité, tout ce qui a été fait précédemment reste valide si ce n'est que nous considérons ici $\overline{B}_r(p) \subset \dot{\mathcal{V}}_{>0}$ avec $\gamma = \min_{\mathbf{x} \in \overline{B}_r(p)} \dot{\mathcal{V}}(\mathbf{x}) > 0$. Dans ce cas, la relation (3.6) livre

$$\mathcal{V}[\varphi^{t_m+s}(\mathbf{x}_0)] > \mathcal{V}[\varphi^{t_m}(\mathbf{x}_0)] + \gamma t_m + \gamma t_s$$

de sorte que lorsque s tend vers l'infini, il tient que

7. $\overline{B}_r(p)$ désigne ici la boule fermée centrée en p de rayon r au sens de la topologie induite par la norme euclidienne dans \mathbb{R}^m .

$$\lim_{s \rightarrow \infty} \mathcal{V} [\varphi^{t_m+s}(x_0)] > \infty,$$

et nous arrivons également à une contradiction puisque par continuité, $\mathcal{V}[\varphi^{t_n}(\mathbf{x}_0)]$ converge vers p lorsque n tend vers l'infini. Dans ce cas, nous venons de montrer que $L^+(\mathbf{x}_0)$ ne peut être contenu que dans l'ensemble $\dot{\mathcal{V}}_0$. Ainsi, étant donné que $L^+(\mathbf{x}_0)$ est un ensemble positivement invariant, nous avons enfin que

$$L^+(\mathbf{x}_0) \subseteq M \subseteq \dot{\mathcal{V}}_0,$$

où M est le plus grand ensemble positivement invariant contenu dans $\dot{\mathcal{V}}_0$, ce qui clôture la preuve. □

Ainsi, pour les différents systèmes (1.7) et (2.9), nous pouvons définir l'ensemble D du Théorème 3.2 comme le plus grand ensemble de niveau contenu dans la région $\dot{\mathcal{V}}_{<0}$, que nous avons représenté en mauve lorsque c'était possible. Dans ce cas, il est direct que cet ensemble est invariant par définition étant donné que la fonction de Lyapunov décroît pour toute trajectoire sur l'ensemble et il n'est donc pas possible d'en sortir une fois dedans. En d'autres mots, toute trajectoire issue de cet ensemble appartient à un sous-ensemble de niveau inférieur, ce qui en traduit l'invariance. Ainsi, par le Théorème 3.2, nous avons que pour tout $\mathbf{x}_0 \in \Xi$, $L^+(\mathbf{x}_0) \subset M$ avec M le plus grand ensemble invariant contenu dans $\dot{\mathcal{V}}_0$. Cependant, au vu du portrait de phase de ces systèmes, il est assez direct de remarquer que le plus grand ensemble positivement invariant est alors réduit à l'origine $\{\mathbf{x}^*\}$ puisque nous n'avons pas d'autres points fixes ou de cycle limite dans ces régions par exemple. Ainsi, nous avons montré que toute solution associée à la condition initiale $\mathbf{x}_0 \in \Xi$, $\varphi^t(\mathbf{x}_0)$ converge vers \mathbf{x}^* de sorte que nous ne perdons plus de généralité à considérer les ensembles de niveau Ξ comme les bassins d'attraction pour ces systèmes.

3.2.3 Étude des fonctions radiales de Gauss

À travers la section précédente et à l'aide du Théorème 3.2, nous avons vu que les fonctions de base radiales *thin plate spline* permettaient d'obtenir une bonne estimation du bassin d'attraction du point d'équilibre à l'origine pour les deux systèmes (1.7) et (2.9). Dans cette section, nous allons à présent tourner notre attention vers les fonctions radiales de Gauss et nous verrons comment ces dernières améliorent ou corroborent les approximations obtenues préalablement.

Dans cette section, nous n'allons pas refaire une étude détaillée du choix des centres et du domaine de la projection, ce qui prendrait un temps considérable si nous voulions étudier l'impact du paramètre γ en parallèle. Par ailleurs, après différents tests numériques, nous avons assez rapidement pu expérimenter que la qualité de l'approximation était de nouveau nettement meilleure en considérant le domaine de la projection assez petit. Dès lors, par souci de facilité, nous nous plaçons directement dans le cas où $\Omega = [-0.1, 0.1] \times [-0.1, 0.1]$. En ce qui concerne les centres, il s'agit là encore d'une question assez ouverte et leur choix n'est pas évident. Dans un premier temps, concentrons-nous sur le système de Van der Pol défini en (1.7) où nous avons fait varier le domaine de ces derniers ainsi que le paramètre γ comme illustré à la Figure 3.7. Dans ce cas, mis à part pour $\gamma = 0.001$ où les meilleurs résultats sont obtenus pour un domaine assez large, il

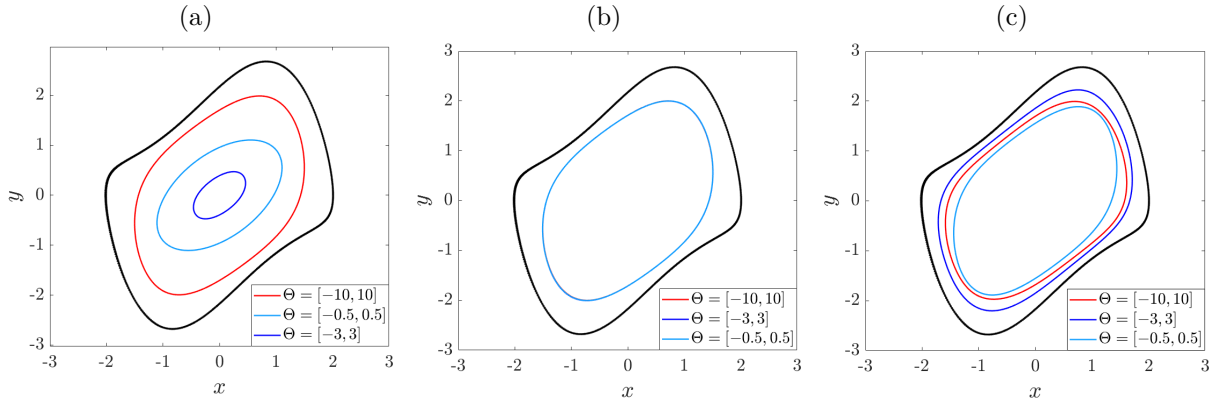


Figure 3.7 – Approximation du bassin d’attraction obtenu avec les fonctions radiales de Gauss dont le domaine de projection $\Omega = [-0.1, 0.1]$. (a) $\gamma = 0.001$, (b) $\gamma = 0.01$ et (c) $\gamma = 0.1$. Pour chacune de ces figures, nous avons représenté la plus grande courbe de niveau contenue dans la région $\dot{\mathcal{V}}_{<0}$. Pour chacune de ces figures, nous obtenons également une petite région $\dot{\mathcal{V}}_{>0}$ mais qui n’est pas possible de représenter selon notre représentation. Notons que pour la figure (b) il existe bien trois courbes de niveau différentes mais elles sont confondues. Précisons également que ces résultats ont notamment été obtenus à l’aide de 100 fonctions de bases dans chacun des cas.

semblerait que peu importe le domaine d’échantillonnage des centres, les fonctions radiales de Gauss permettent également d’obtenir une approximation du bassin d’attraction assez satisfaisante. Nous devons tout de même préciser que nous n’avons ici représenté que les cas idéaux du paramètre γ de sorte que les approximations soient plus au moins bonnes et assez simples à visualiser. Cependant, le choix de ce dernier n’est pas particulièrement trivial et d’autres choix de γ pourraient entraîner de mauvais résultats, comme nous pouvons notamment l’observer à la Figure 3.8 avec $\gamma = 1$. Par ailleurs, notons que cet exemple est typiquement un cas particulier où les valeurs propres sont données par

$$\tilde{\lambda}_1 = -0.500308 + 0.866126i, \quad \tilde{\lambda}_2 = -0.500308 - 0.866126i$$

mais pour lesquelles les résultats ne sont pas bons.

De par la multitude de combinaisons possibles des centres et du paramètre γ , il semble compliqué de trouver le choix optimal permettant d’obtenir la meilleure approximation. Par ailleurs, une telle étude semble compliquée à mettre en pratique et ne garantit pas nécessairement de meilleurs résultats. Cette situation nous motive donc à uniquement montrer les résultats les plus concluants sur base de différents tests numériques réalisés au préalable afin de garder et de dégager les principales tendances pour ces fonctions de base. Cette approche permet ainsi de motiver l’utilisation d’une telle base pour des systèmes de plus grande dimension, que nous considérons à travers deux exemples particuliers dans la section suivante.

Remarque 3.3. *Notons que pour chacune des approximations numériques obtenues à la Figure 3.7, il peut toujours exister une région $\dot{\mathcal{V}}_{\geq 0}$ autour de l’origine comme nous l’avons observé dans le cas des fonctions TPS. Dans ce cas, nous pouvons appliquer le Théorème 3.2 de sorte que nous ne perdons pas de généralité à directement parler d’approximation du bassin d’attraction donné par le plus grand ensemble de niveau contenu dans $\dot{\mathcal{V}}_{<0}$.*

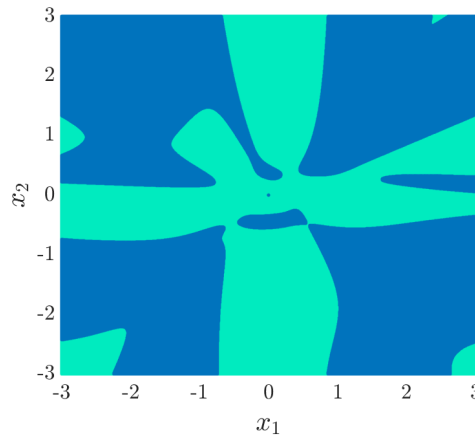


Figure 3.8 – Approximation des régions $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$ obtenues à l’aide de 100 fonctions de base radiales TPS pour $\Theta = [-10, 10] \times [-10, 10]$, $\Omega = [-0.1, 0.1] \times [-0.1, 0.1]$ et $\gamma = 1$. Pour chacune des figures, nous avons respectivement représenté en bleu clair et en bleu foncé les ensembles $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$. Dans ce cas, nous remarquons que l’approximation n’est pas particulièrement satisfaisante.

Penchons-nous à présent sur le système dynamique (2.9). De manière similaire à ce que nous avons fait pour le système de Van der Pol, nous allons regarder comment le paramètre γ ainsi que le domaine d’échantillonnage Θ impactent la qualité de l’approximation. Pour ce faire, considérons dans un premier temps $\gamma = 0.1$ pour lequel nous avons généré des centres aléatoirement selon les différents ensembles Θ des Figures 3.9 (a)-(b). Dans ce cas, nous remarquons que les approximations du bassin d’attraction données par les deux courbes de niveau mauves complètement à l’intérieur de l’ensemble $\dot{\mathcal{V}}_{<0}$ ne sont pas particulièrement bonnes. Cependant, en considérant les centres sur une grille discrétisée comme aux Figures 3.9 (c)-(d), nous pouvons améliorer ces approximations et nous capturons alors une plus grande partie du bassin d’attraction. Même si l’approximation est meilleure, elle n’est pas parfaite. Pour pallier ce problème, nous pouvons diminuer le paramètre γ et nous remarquons alors à la Figure 3.10 que l’approximation est meilleure⁸. Plus particulièrement, après différents tests empiriques, nous avons constaté qu’il n’est pas possible d’obtenir une meilleure approximation pour cet exemple.

3.2.4 Application en plus haute dimension

Dans la section précédente, nous avons vu que les fonctions de base radiales permettaient d’approximer le bassin d’attraction des points d’équilibre des systèmes (1.7) et (2.9). Plus particulièrement, au vu des résultats concluants que nous avons obtenus, les fonctions de base radiales semblent d’être une bonne alternative aux monômes, coûteux pour de grandes dimensions. Dans cette section, nous allons tenter d’appuyer cet argument en montrant que les fonctions de base radiales permettent également d’approximer le bassin d’attraction des systèmes (2.10) et (2.11) en trois dimensions. Cependant, nous n’effectuerons pas d’étude comparative et nous nous contenterons de directement donner des résultats concluants pour un choix de paramètres que nous aurons établi numériquement en amont. Précisons que pour ces exemples, nous verrons de nouveau comment le choix des coefficients c_i de la fonction de Lyapunov affecte la qualité des résultats.

⁸. Précisons que nous obtenons un résultat tout à fait similaire à ceux obtenus pour la Figure 3.10 en considérant des centres sur une grille discrétisée.

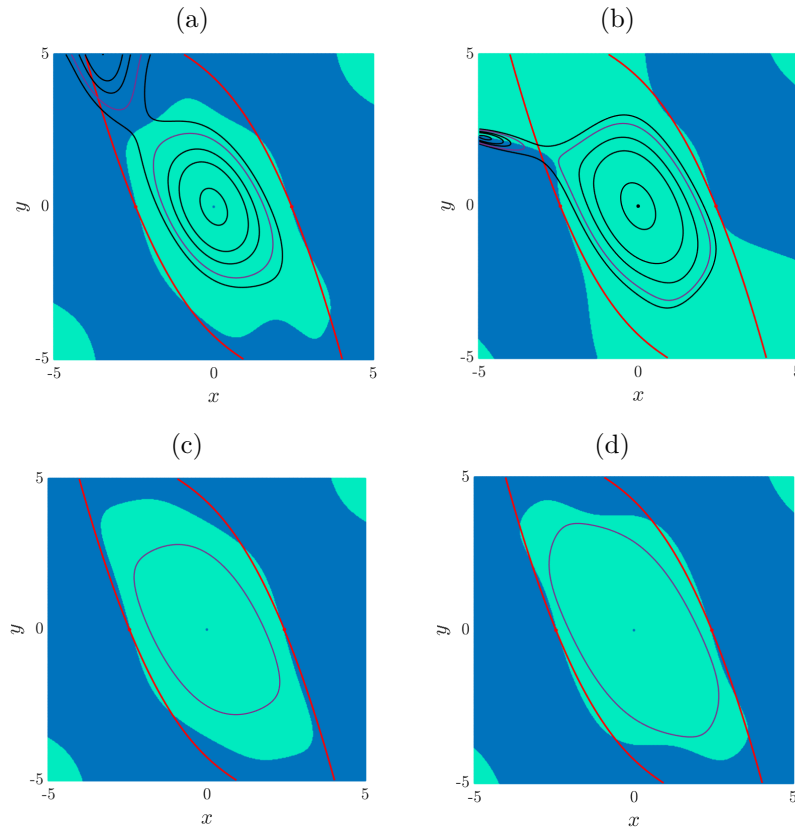


Figure 3.9 – Approximation du bassin d’attraction obtenu avec les fonctions radiales de Gauss pour $\Omega = [-0.1, 0.1]$ et $\gamma = 0.1$ pour des centres générés (a) aléatoirement sur le domaine $\Theta = [-5, 5]$, (b) aléatoirement sur le domaine $\Theta = [-10, 10]$, (c) selon une discrétisation de l’ensemble $\Theta = [-5, 5]$, (d) selon une discrétisation de l’ensemble $\Theta = [-10, 10]$. Dans ce cas, nous remarquons que pour ce choix de γ les meilleurs résultats sont obtenus pour des centres générés selon une discrétisation du domaine Θ . Précisons que ces résultats ont notamment été obtenus à l’aide de 100 fonctions de base.

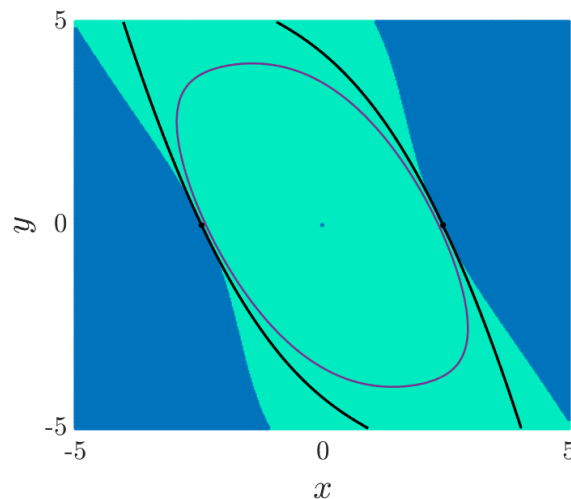


Figure 3.10 – Approximation du bassin d’attraction obtenu avec les fonctions radiales de Gauss pour $\Omega = [-0.1, 0.1]$, $\gamma = 0.01$ et $\Theta = [-5, 5]$. Dans ce cas, nous remarquons que l’approximation du bassin est satisfaisante et nous en capturons une bonne partie. À nouveau, ce résultat a été obtenu à l’aide de 100 fonctions de bases.

Concentrons-nous dans un premier temps sur les fonctions radiales TPS et sur base des résultats obtenus dans la section précédente, supposons que $\Omega = [-0.1, 0.1]$. En effet, rappelons que pour les deux systèmes que nous avons étudiés, ce choix de domaine pour la projection induit de meilleurs résultats peu importe le système. Ainsi, nous avons alors de bonnes raisons de penser que ce choix est judicieux. En ce qui concerne la génération des centres, il n'y a pas vraiment de manière optimale de les calculer et après quelques tests, nous avons fixé le domaine d'échantillonnage des centres à $\Theta = [-10, 10]$. Ainsi, pour ces deux systèmes, nous remarquons aux Figures 3.11 (a) - (b) que l'approximation du bassin d'attraction n'est pas particulièrement satisfaisante puisque nous ne capturons qu'une brève partie de ce dernier. Cependant, notons que l'approximation des valeurs propres obtenues pour ce choix de paramètres livre

$$\begin{cases} \tilde{\lambda}_1 = -0.5000087 + 0.8660i, & \tilde{\lambda}_2 = -0.5000087 - 0.8660i, & \tilde{\lambda}_3 = -1.00004, \\ \tilde{\lambda}_1 = -0.5000138 + 0.8660i, & \tilde{\lambda}_2 = -0.5000138 - 0.8660i, & \tilde{\lambda}_3 = -1.00004, \end{cases}$$

respectivement pour les systèmes (2.10) et (2.11). Ainsi, un des vecteurs propres de \mathbf{L}_N est à coefficients réels et nous pouvons alors changer le poids des fonctions propres de la

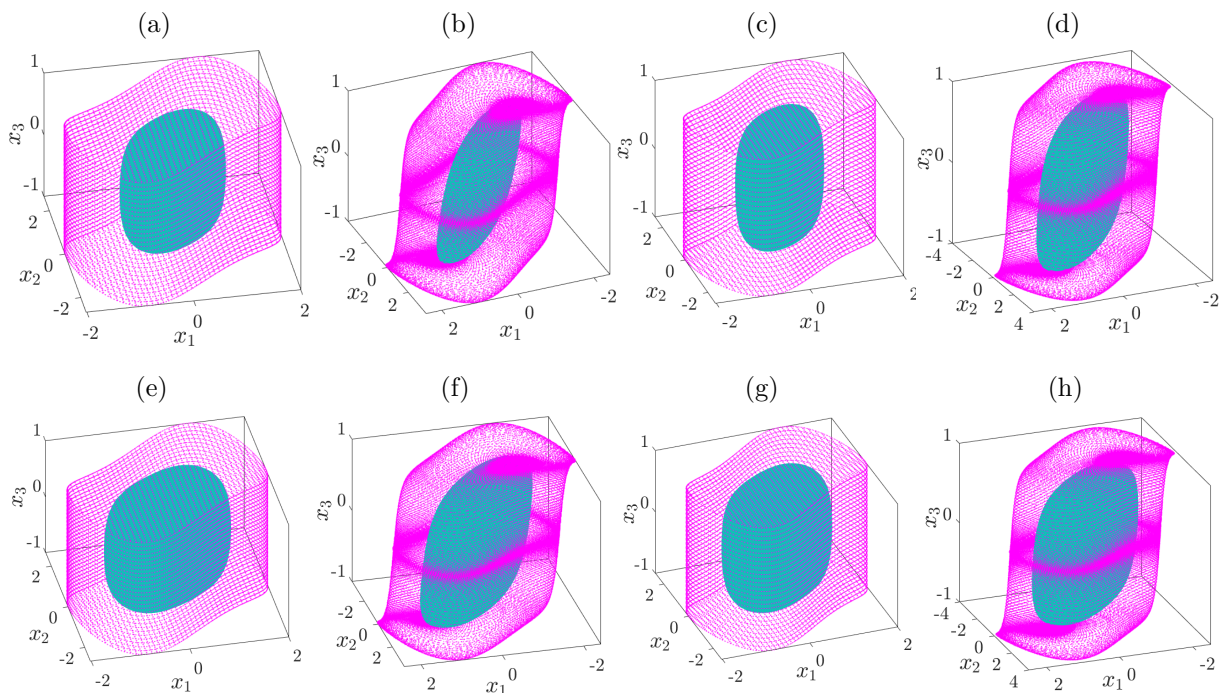


Figure 3.11 – Approximation du bassin d'attraction pour les systèmes à trois dimensions (2.10) et (2.11) avec $\Omega = [-0.1, 0.1]^2$ pour des fonctions de base radiales. Dans ces exemples, les centres ont été générés aléatoirement sur le domaine $\Theta = [-10, 10]$. **Figures du dessus :** (a) - (b) Résultats obtenus avec les fonctions radiales TPS pour (a) le système (2.10) et (b) le système (2.11). (c) - (d) Même résultats mais obtenus avec des fonctions radiales de Gauss avec $\gamma = 0.01$. **Figures du dessous :** (e) - (f) Représentation des mêmes figures (a) - (d) en changeant les coefficients c_i par $c_1 = c_2 = 0.4$ et $c_3 = 1$ avec c_1 et c_2 les deux coefficients associés aux valeurs propres complexes. Ces résultats ont notamment été obtenus à l'aide de 100 fonctions de bases dans chacun des cas.

fonction de Lyapunov afin d'améliorer la qualité du bassin d'attraction comme observé aux Figures 3.11 (e)-(f). Cependant, même si les résultats obtenus restent acceptables, nous ne capturons toujours qu'une petite partie du bassin d'attraction. Malheureusement, nous tenons à préciser qu'après certains tests numériques, nous n'avons en réalité pas trouvé de meilleures approximations en faisant varier Θ et les coefficients c_i pour ce choix de fonctions de base. Afin de déterminer une meilleure approximation du bassin d'attraction, nous pourrions alors nous pencher sur le choix des fonctions radiales de Gauss avec le même choix de $\Theta = [-10, 10]$. Cependant, nous obtenons le même genre de résultats que pour le cas des fonctions radiales TPS, ce qui n'améliore pas la qualité de l'approximation. Ici encore, nous avons réalisé différents tests numériques de sorte que les résultats obtenus à la Figure 3.11 sont en réalité les meilleurs que nous puissions obtenir pour ces deux systèmes. En particulier, nous avons fait varier le paramètre γ mais nous n'obtenions pas spécialement de meilleurs résultats et nous avons par exemple observé que pour $\gamma = 0.1$ les approximations étaient très mauvaises. Ce constat nous permet notamment de confirmer que l'utilisation des fonctions radiales de Gauss est plutôt sensible. Ainsi, il semblerait que les fonctions radiales fournissent des approximations tout à fait acceptables du bassin d'attraction de l'équilibre à l'origine⁹ mais restent plus conservatives comparativement aux monômes.

Remarque 3.4. *Précisons que les centres sont générés aléatoirement sur le domaine Θ . Ainsi, il se peut que pour un certain échantillonnage des centres, les approximations soient déjà bonnes de sorte que les coefficients de tuning c_i ont peu d'impact sur la qualité de l'approximation.*

3.2.5 Qualité de l'approximation et temps de calcul

Dans cette section, nous allons donner certaines précisions concernant le calcul des approximations que nous avons obtenues dans la section précédente. Plus particulièrement, nous avons mentionné à deux reprises que sur base de différents tests numériques, les approximations du bassin d'attraction que nous avons obtenues étaient les *meilleures*, ou en faisaient du moins partie, pour d'autres choix de paramètres. En réalité, nous devons nuancer ces propos. En effet, lorsque nous étudions des systèmes de plus haute dimension, une étude comparative de l'impact des paramètres sur l'approximation devient assez coûteuse numériquement et nous devrions alors y consacrer beaucoup de temps pour espérer obtenir des résultats de meilleures factures. En effet, il n'est pas garanti qu'une étude détaillée engendre de meilleurs résultats si nous pouvons constater des tendances similaires et invariantes pour seulement quelques tests numériques. Ainsi, nous *pensons* que les résultats obtenus à la Figure 3.11 sont les meilleurs peu importe le choix des paramètres impliqués, mais cet argument reste heuristique et uniquement fondé sur quelques tests que l'on espère représentatifs du comportement général de la méthode pour ces systèmes. Par ailleurs, ces résultats pourraient très probablement être encore améliorés avec d'autres fonctions de base. Cependant, précisons que la section précédente a pour unique but de montrer que les fonctions de base radiales produisent des approximations acceptables en trois dimensions et sont alors prometteuses pour une utilisation à de plus grandes dimensions.

9. Notons qu'ici encore, il pourrait de nouveau exister une région $\dot{\mathcal{V}}_{\geq 0} \subset \dot{\mathcal{V}}_{< 0}$. Dans ce cas, nous pouvons toujours utiliser le Théorème 3.2 et tirer des conclusions similaires qu'au préalable.

Afin de clôturer cette section, nous tenions à effectuer une dernière remarque d'un point de vue computationnel tout à fait général. Dans le cadre du mémoire, les différents codes implémentés ont été réalisés dans le langage de programmation Matlab afin de pouvoir construire un code générique (voir section 3.1.1). Dans ce cas, les différentes approximations que nous avons obtenues jusque maintenant ont été réalisées avec ce logiciel. Cependant, lorsque nous travaillons en plus haute dimension, les calculs deviennent assez coûteux. Que ce soit le calcul de la projection par l'intermédiaire de la méthode de Monte-Carlo ou celui de la fonction de Lyapunov et de sa dérivée temporelle, ces opérations peuvent devenir assez conséquentes pour obtenir des résultats satisfaisants. Ainsi, pour de plus grandes dimensions, nous pensons que ce code peut devenir très compliqué à utiliser en pratique bien que ce dernier ait été optimisé selon les préférences de Matlab à travailler vectoriellement. Une des perspectives du mémoire serait alors d'adapter ce code en C, ou en python par exemple, pour lesquels les opérations numériques sont en général plus rapides. Cette considération permettrait sans doute de réduire le temps de calcul mais serait peut-être compliquée à mettre en pratique pour garder un code général et nécessiterait sûrement une réflexion théorique conséquente.

Chapitre 4

Fiabilité de l'approximation du bassin d'attraction

Dans le Chapitre 2, nous avons vu que l'approximation des fonctions propres de l'opérateur de Koopman permettait de définir une estimation d'une fonction de Lyapunov $\mathcal{V}(\cdot)$. Dans ce cas, nous avons expliqué que son support ne permettait pas de définir exactement le bassin d'attraction de l'équilibre et qu'il fallait en réalité calculer le plus grand ensemble de niveau $\Xi \subset \dot{\mathcal{V}}_{<0}$ afin d'en approximer une région exacte. Cependant, nous manquons de rigueur dans notre analyse. En effet, nous ne calculons $\dot{\mathcal{V}}_{<0}$ seulement qu'en différents points répartis sur une grille uniforme. Dès lors, rien ne garantit que la fonction $\dot{\mathcal{V}}(\cdot)$ garde le même signe dans chacune des mailles de cette grille. Ce point est notamment illustré à travers l'exemple unidimensionnel représenté à la Figure 4.1 où nous observons que la fonction change de signe entre les deux points de la grille x et $x + \Delta x$ (où $\Delta x > 0$ représente le pas de discrétisation spatiale). Dans ce cas, il vient que la région $\dot{\mathcal{V}}_{<0}$ n'est plus particulièrement bien définie et les résultats de stabilité sont alors biaisés. Dans ce chapitre, nous allons voir comment il est possible de traiter ce manque de rigueur dans le cas particulier des monômes. Plus particulièrement, nous verrons dans la section suivante qu'une des pistes de solution consiste à déterminer une borne sur la norme du gradient de $\dot{\mathcal{V}}(\cdot)$ permettant ainsi de construire un algorithme dont l'objectif est de déterminer les différentes cellules de l'ensemble $\dot{\mathcal{V}}_{<0}$ au sein desquelles nous pouvons garantir que le signe

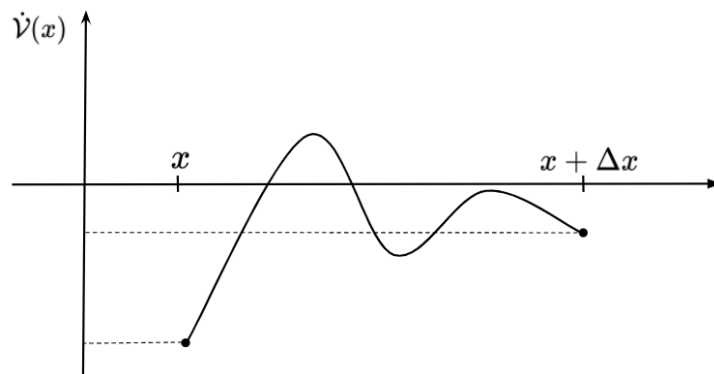


Figure 4.1 – Illustration de $\dot{\mathcal{V}}(x)$ entre deux points x et $x + \Delta x$ d'une cellule de la maille pour lesquels $\dot{\mathcal{V}}(x) < 0$ et $\dot{\mathcal{V}}(x + \Delta x) < 0$. Dans ce cas, nous remarquons que la fonction change de signe entre ces deux points.

de la dérivée demeure inchangé. Cependant, nous verrons que cet algorithme peut rapidement devenir coûteux numériquement lorsque le pas de discrétisation devient trop fin et nous verrons alors comment l'utilisation d'une grille adaptative se révèle efficace dans ce contexte. Sauf indication contraire, la notion de cellule désignera dans la suite tout cube ou hypercube de longueur $\Delta \mathbf{x}$ formé par les points de la grille. Notons que ce choix est en réalité arbitraire et que nous aurions tout aussi bien pu considérer des cellules dont les côtés sont de longueurs différentes. Cependant, nous laissons cette extension possible de la méthode proposée dans les perspectives de ce travail.

4.1 Explication du problème et solution proposée

Comme nous l'avons mentionné, l'idée de ce chapitre est de déterminer dans quelle mesure le signe de $\dot{\mathcal{V}}(\cdot)$ demeure inchangé au sein de chaque cellule dont les extrémités x_i appartiennent à l'ensemble $\dot{\mathcal{V}}_{<0}$. Pour ce faire, nous nous placerons dans le cas des monômes, qui sont plus faciles à traiter étant donné que la fonction de Lyapunov ainsi que sa dérivée temporelle sont des polynômes. En effet, rappelons que

$$\dot{\mathcal{V}}(\mathbf{x}) = \mathbf{F}(x)\nabla\mathcal{V}(\mathbf{x}),$$

de sorte que $\dot{\mathcal{V}}(\mathbf{x})$ est un polynôme pour peu que le champ de vecteurs soit lui-même un polynôme, ce que nous considérons dans cette partie. Essayons à présent de formaliser le problème afin d'en trouver une solution. À cette fin, rappelons dans un premier temps que pour toute fonction f à p variables, le théorème des accroissements finis donne

$$\dot{\mathcal{V}}(\mathbf{x} + \delta) = \dot{\mathcal{V}}(\mathbf{x}) + \nabla\dot{\mathcal{V}}(w)^T \delta,$$

avec $\mathbf{x}, \delta \in \mathbb{R}^p$ et $w \in [\mathbf{x}, \mathbf{x} + \delta]^p$. Notons que cette expression nous permet de dire qu'à partir de $\dot{\mathcal{V}}(\mathbf{x})$ et en suivant la direction $\nabla\dot{\mathcal{V}}(w)$ d'un pas de δ , nous arrivons au point $\dot{\mathcal{V}}(\mathbf{x} + \delta)$. Dès lors, en considérant la pente positive la plus raide, et à partir de $\dot{\mathcal{V}}(\mathbf{x})$, la valeur maximale que nous pouvons atteindre entre \mathbf{x} et $\mathbf{x} + \delta$ est donnée par

$$\dot{\mathcal{V}}(\mathbf{x}) + \max_{w \in [\mathbf{x}, \mathbf{x} + \delta]^p} \nabla\dot{\mathcal{V}}(w)^T \delta, \quad (4.1)$$

et correspond alors au pire scénario possible. Notons que ce cas de figure est intéressant lorsque que nous ne connaissons pas particulièrement la fonction ailleurs qu'aux points \mathbf{x} et $\mathbf{x} + \delta$ et que nous cherchons à caractériser sa variation le long de ce domaine. Dans ce cas, la relation (4.1) permet de déterminer le point le plus haut auquel il est possible d'aller depuis $\dot{\mathcal{V}}(\mathbf{x})$.

Essayons de voir à présent comment utiliser ce que nous venons de voir pour notre problème. À cette fin, rappelons que nous cherchons à déterminer dans quelle mesure la fonction $\dot{\mathcal{V}}(\cdot)$ demeure de signe constant à l'intérieur de chaque maille de la grille sur l'ensemble $\dot{\mathcal{V}}_{<0}$ ¹. Par souci de facilité dans les explications, considérons le cas unidimensionnel où chaque cellule est donnée par un intervalle $C = [\mathbf{x}, \mathbf{x} + \Delta \mathbf{x}]$. L'idée est alors la suivante. Afin d'assurer que la fonction reste négative dans cette cellule, nous pouvons imposer que

1. Notons que nous sommes précisément dans le cas de figure où nous ne connaissons la fonction $\dot{\mathcal{V}}$ qu'en certains points d'un domaine discrétisé et que nous cherchons à mesurer sa variation.

$$\dot{\mathcal{V}}(\mathbf{x}) + \max_{w \in \mathcal{C}} \nabla \dot{\mathcal{V}}(w) \Delta \mathbf{x} < 0. \quad (4.2)$$

Cette condition permet alors d'assurer que dans le pire scénario possible, $\dot{\mathcal{V}}(\cdot)$ reste tout de même négative entre \mathbf{x} et $\mathbf{x} + \Delta \mathbf{x}$. À la Figure 4.2, nous avons représenté ce cas de figure où la droite en rouge correspond à la direction de plus forte montée, encodée par la tangente au point bleu de la courbe en noir. Notons que cette condition est peut-être trop conservative. En effet, à l'aide de la droite en rouge à la Figure 4.2, nous savons que dans le pire scénario possible, la fonction restera quoiqu'il arrive en dessous de cette droite. Cependant, il en va de même pour la droite en bleu où il est assez direct de remarquer qu'en prenant la même pente que la droite en rouge (à un signe près), la fonction restera également *en-dessous* de cette droite dans le pire scénario possible. Dans ce cas, nous remarquons qu'il n'est pas spécialement judicieux d'imposer la condition (4.2) étant donné que certains points sur la droite bleue, comme le point mauve, seront en réalité bornés par une valeur plus petite donnée sur la droite bleue. Ainsi, afin d'alléger la condition (4.2), nous pouvons en réalité imposer les deux conditions

$$\dot{\mathcal{V}}(\mathbf{x}_i) + \max_{w \in \mathcal{C}} \nabla \dot{\mathcal{V}}(w) \frac{\Delta \mathbf{x}}{2} < 0. \quad (4.3)$$

pour $i = 1, 2$ et où les différents points x_1 et x_2 sont respectivement donnés par \mathbf{x} et $\mathbf{x} + \Delta \mathbf{x}$. Notons que l'on pourrait dans un premier temps se demander pourquoi en $x + \Delta \mathbf{x}$, nous obtenons un signe *plus* au lieu d'un signe *moins* alors que nous prenons la pente opposée (voir Figure 4.2). Ceci provient du fait que les différentes valeurs obtenues seront les mêmes en allant dans la direction du gradient ou dans celle opposée. Nous pouvons aisément le

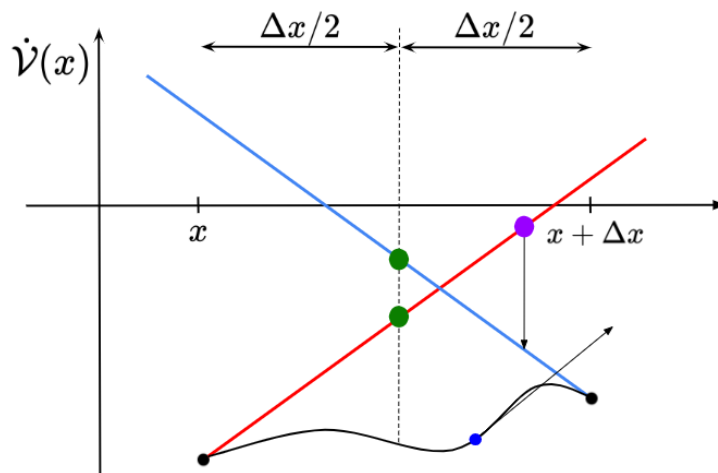


Figure 4.2 – Illustration de la condition (4.2). La courbe en noir représente $\dot{\mathcal{V}}$. La courbe rouge représente la direction de plus forte montée translatée au point $\dot{\mathcal{V}}(x)$. La courbe en bleu représente la courbe en rouge mais avec la pente opposée. Notons que les points sur cette dernière sont les mêmes que ceux obtenus à partir de $\dot{\mathcal{V}}(x + \Delta x)$ en longeant la direction de plus forte montée. Ce point est assez évident à remarquer sur la figure de sorte que si nous translations la tangente au point bleu au point $\dot{\mathcal{V}}(x + \Delta x)$, nous obtenons les mêmes points. Les deux points en vert représentent les points obtenus en avançant d'un pas $\Delta x/2$ le long de la direction de plus forte montée.

remarquer sur la Figure 4.2 en prolongeant la droite selon la direction de la plus forte montée. Ensuite, il peut paraître étrange de choisir $\Delta\mathbf{x}/2$ au lieu de $\Delta\mathbf{x}$ dans l'expression (4.3). Éclaircissons cela. Si nous considérons par exemple le point mauve sur la droite en rouge de la Figure 4.2, nous voyons que ce dernier a été *envoyé* plus loin qu'en $\Delta\mathbf{x}/2$. Cependant, pour ce même point, il existe également une autre borne donnée par la droite en bleu de sorte que ce point sera toujours borné par le maximum des valeurs en vert sur la figure où les points en vert correspondent aux différentes valeurs

$$\dot{\mathcal{V}}(\mathbf{x}_i) + \max_{w \in C} \nabla \dot{\mathcal{V}}(w) \frac{\Delta\mathbf{x}}{2}.$$

Ainsi, si les deux conditions en (4.3) sont satisfaites, nous sommes sûrs que la fonction restera négative dans l'intervalle $[\mathbf{x}, \mathbf{x} + \Delta\mathbf{x}]$ mais nous avons affiné la condition en (4.2) qui était peut-être trop conservative.

Remarque 4.1. *Précisons que le choix de $\Delta\mathbf{x}/2$ est en réalité tout à fait arbitraire et une autre combinaison telle que $(\Delta\mathbf{x}/3, 2\Delta\mathbf{x}/2)$ aurait tout aussi bien fonctionné. Par exemple, il semblerait qu'une combinaison encore moins conservative serait donnée à l'intersection des deux droites rouge et bleue de la Figure 4.2. Cependant, ce cas de figure n'a pas été considéré dans le cadre du mémoire.*

Avant de poursuivre, nous tenons à préciser que le raisonnement que nous venons d'établir en une dimension tient également en plusieurs dimensions. La seule chose qui défère c'est le pas $\Delta\mathbf{x}/2$. En effet, ce choix est un cas particulier en une dimension où le centre de chaque maille est donné par la médiatrice de cette dernière. Lorsque nous considérons des systèmes de plusieurs dimensions, nous ne pouvons pas en dire autant. Ainsi, pour un domaine de dimension $m \in \mathbb{N}$, le centre de chaque maille se situe à une distance donnée² par

$$\tilde{\Delta}\mathbf{x} = \sqrt{\sum_{k=1}^m \left(\frac{\Delta\mathbf{x}}{2}\right)^2},$$

à partir de chaque extrémité de la maille. Dans ce cas, nous obtenons l'ensemble de conditions générales donné par

$$\dot{\mathcal{V}}(\mathbf{x}_i) + \max_{w \in C} \nabla \dot{\mathcal{V}}(w) \tilde{\Delta}\mathbf{x} < 0. \quad (4.4)$$

Ainsi, en utilisant l'inégalité de Cauchy-Schwartz nous obtenons alors

$$\dot{\mathcal{V}}(\mathbf{x}_i) + \max_{w \in C} \|\nabla \dot{\mathcal{V}}(w)\| \|\tilde{\Delta}\mathbf{x}\| < 0,$$

Notons qu'il est en réalité nécessaire de faire apparaître la norme du gradient étant donné que nous ne connaissons pas la direction du gradient sur toute la maille. Finalement, en remarquant que l'ensemble de conditions obtenu en (4.4) peut se réécrire en une seule condition, nous obtenons la condition suivante

$$\mathcal{C} \equiv \max_{w \in C} \|\nabla \dot{\mathcal{V}}(w)\| \|\tilde{\Delta}\mathbf{x}\| < \min_{i \in I} [-\dot{\mathcal{V}}(\mathbf{x}_i)] \quad (4.5)$$

2. Notons que cette formule découle de l'application successive du Théorème de Pythagore dans un hypercube dont les côtés sont égaux et donnés par $\Delta\mathbf{x}$.

où les points $\{\mathbf{x}_i\}_{i \in I}$ correspondent aux différentes extrémités de la maille C . Dès lors, l'idée va être de trouver une certaine borne sur le gradient de $\dot{\mathcal{V}}(\mathbf{x})$ et remplacer la relation $\max_{w \in C} \|\nabla \dot{\mathcal{V}}(w)\|$ par cette borne obtenue dans la relation (4.5). Dans ce cas, si la condition est vérifiée avec la borne, elle le sera en particulier pour $\max_{w \in C} \|\nabla \dot{\mathcal{V}}(w)\|$ et nous aurons alors un critère, certes conservatif, assurant que la fonction ne change pas de signe pour chaque maille considérée.

À présent, toute la question réside dans le calcul d'une borne sur $\|\nabla \dot{\mathcal{V}}(w)\|$. Comme nous l'avons mentionné au début du chapitre, nous supposons ici que le champ de vecteurs est analytique et que les fonctions de base sont des monômes. Dans ce cas, il vient que $\dot{\mathcal{V}}(\mathbf{x})$ est également un polynôme. Cependant, ce qui nous intéresse en particulier ce n'est pas ce polynôme mais plutôt la norme de son gradient. Néanmoins, comme nous travaillons avec des polynômes, il vient de nouveau que $\nabla \dot{\mathcal{V}}(\mathbf{x})$ est un vecteur dont les entrées sont des polynômes. Par souci de facilité, supposons le cas à deux dimensions³ avec

$$\nabla \dot{\mathcal{V}}(\mathbf{x}) = \begin{pmatrix} \sum_{i=1}^{N_1} \alpha_i v_i(\mathbf{x}) \\ \sum_{i=1}^{N_2} \beta_i \omega_i(\mathbf{x}) \end{pmatrix}, \quad (4.6)$$

où $v_i(\mathbf{x})$ et $\omega_i(\mathbf{x})$ sont des monômes. Notons que les différents monômes impliqués pour chacune des composantes du gradient en (4.6) ne correspondent pas forcément aux monômes de base. En effet, lorsque nous calculons la fonction de Lyapunov $\mathcal{V}(\mathbf{x})$, les fonctions propres sont prises au carré de sorte que de nouveaux monômes sont créés. Ainsi, il n'est pas du tout étonnant que les composantes du gradient fassent intervenir d'autres monômes de base en quantités différentes N_1 et N_2 . Dans ce cas particulier, il vient que

$$\|\nabla \dot{\mathcal{V}}(\mathbf{x})\|_2^2 = \left(\sum_{i=1}^{N_1} \alpha_i v_i(\mathbf{x}) \right)^2 + \left(\sum_{i=1}^{N_2} \beta_i \omega_i(\mathbf{x}) \right)^2 := H_1(\mathbf{x}) + H_2(\mathbf{x}).$$

Concentrons-nous dans un premier temps sur H_1 . Dans ce cas, il est assez direct de constater que

$$\begin{aligned} H_1(\mathbf{x}) &= \sum_{i=1}^{N_1} \alpha_i^2 v_i^2(\mathbf{x}) + 2 \sum_{i,j|i < j} (\alpha_i v_i(\mathbf{x}))(\alpha_j v_j(\mathbf{x})), \\ &= \sum_{i=1}^N \alpha_i^2 v_i^2(\mathbf{x}) + 2 \sum_{i,j|i < j} (\alpha_i \alpha_j) (v_i(\mathbf{x}) v_j(\mathbf{x})), \end{aligned}$$

où $v_i(\mathbf{x}) v_j(\mathbf{x})$ est également un polynôme. Dès lors, par un raisonnement tout à fait similaire, nous obtenons également

$$H_2(\mathbf{x}) = \sum_{i=1}^{N_2} \alpha_i^2 \omega_i^2(\mathbf{x}) + 2 \sum_{i,j|i < j} (\alpha_i \alpha_j) (\omega_i(\mathbf{x}) \omega_j(\mathbf{x})).$$

3. Notons que le raisonnement qui suit peut tout à fait être appliqué en d'autres dimensions. Nous considérons ici un cas simple pour expliquer la démarche effectuée.

Notons alors qu'en identifiant les termes communs des deux expressions H_1 et H_2 , il vient que $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|_2^2$ est un polynôme et par souci de facilité, nous noterons

$$\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|_2^2 = \sum_i \gamma_i \phi_i(\mathbf{x}), \quad (4.7)$$

où $\{\phi_i(\mathbf{x})\}_{i=1}^n$ est un ensemble de monômes avec n un entier fixé et connu. Rappelons que nous cherchons ici à déterminer une borne sur l'ensemble $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|_2^2$ pour tout $\mathbf{x} \in C$ où C est une maille particulière du domaine discrétisé dans l'ensemble $\dot{\mathcal{V}}_{<0}$. L'idée la plus naïve et que nous avons considérée dans le cadre du mémoire consiste à borner chaque terme un par un. Pour un monôme quelconque p sur une maille C , il est assez évident de constater que ce dernier atteint respectivement sa valeur minimale et maximale au point \mathbf{x} et $\mathbf{x} + \Delta\mathbf{x}$ où \mathbf{x} désigne ici le point qui représente la maille C ⁴. Dans ce cas, selon que les coefficients des différents monômes de (4.7) soient positifs ou négatifs, nous pouvons borner la norme du gradient en considérant pour chaque terme la valeur maximale ou minimale du monôme sur la grille que nous multiplions par le coefficient associé. Dès lors, en bornant $\max_{w \in C} \|\nabla\dot{\mathcal{V}}(\mathbf{x})\|$ dans l'expression (4.5), nous obtenons le critère

$$\mathcal{C} \equiv K_C \|\tilde{\Delta}\mathbf{x}\| < \min_{i \in I} [-\dot{\mathcal{V}}(\mathbf{x}_i)] \quad (4.8)$$

où K_C correspond à la borne du gradient calculée sur la cellule C . Ainsi, pour une grille fixée avec un pas de discrétisation $\Delta\mathbf{x}$ nous pouvons établir un algorithme (voir Algorithme 2) qui permet de tester la condition (4.8) pour chaque cellule de l'ensemble $\dot{\mathcal{V}}_{<0}$.

Algorithme 2 VÉRIFICATION DU SIGNE DES MAILLES DANS $\dot{\mathcal{V}}_{<0}$

Require: Domaine \mathbb{X} discrétisé en un ensemble de s points $\{\mathbf{x}_i\}_{i=1}^s$ et construction de différentes cellules C_i . Soit n le nombre de cellules.

Ensure: Vecteur S dont les composantes sont égales à 1 si le signe de $\dot{\mathcal{V}}$ reste négatif sur chaque cellule associée aux points dans l'ensemble $\dot{\mathcal{V}}_{<0}$.

- 1: Calcul de $\dot{\mathcal{V}}$ sur la grille.
 - 2: Calcul des monômes et des coefficients associés de $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|_2^2$.
 - 3: **for** $i = 1 \rightarrow n$ **do**
 - 4: Calcul de K_C sur la cellule C_i .
 - 5: Définir I comme l'ensemble d'indice contenant les extrémités de la cellule C_i .
 - 6: **if** $(\sqrt{K_C} \|\tilde{\Delta}\mathbf{x}\| < \min_{i \in I} [-\dot{\mathcal{V}}(\mathbf{x}_i)])$ et $(\dot{\mathcal{V}}(\mathbf{x}_i) < 0$ pour tout $i \in I)$ **then**
 - 7: $S(i) = 1$
 - 8: **else**
 - 9: $S(i) = 0$
 - 10: **end if**
 - 11: **end for**
-

4. Le fait qu'il représente la cellule veut simplement dire que c'est le point à partir duquel les extrémités de la cellule sont calculées.

4.1.1 Remarque générale sur l'implémentation

De nouveau, une majeure partie de ce chapitre est en réalité cachée derrière l'implémentation d'un code Matlab. Même si nous n'entrerons pas dans les détails de ce dernier par souci de facilité afin de garder un rapport clair et concis, nous tenons tout de même à spécifier l'un ou l'autre élément. En réalité, la partie la plus complexe dans cette partie correspond au calcul de la borne, c'est-à-dire K_C . En effet, le calcul de cette dernière nécessite de connaître les différents monômes que constituent $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|_2^2$. Plus précisément, ce calcul nécessite successivement celui de $\mathcal{V}(x)$, $\nabla\mathcal{V}(x)$, $\dot{\mathcal{V}}$, et $\nabla\dot{\mathcal{V}}$ de sorte que le calcul de K_C n'est pas direct et nécessite le stockage en mémoire de toutes ces expressions calculées. Le stockage est en réalité obligatoire étant donné que le calcul de la borne de $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|$ nécessite les différents monômes qui le composent. Dès lors, tous ces calculs ont dû être implémentés et ont nécessité un certain temps de réflexion. Néanmoins, une fois que le polynôme $\|\nabla\dot{\mathcal{V}}(\mathbf{x})\|$ est identifié, l'implémentation de l'Algorithme 2 est plus simple à mettre en pratique.

4.1.2 Application numérique

Dans cette section, nous allons appliquer l'Algorithme 2 aux cas particuliers des systèmes en deux dimensions (1.7) et (2.9). Pour ce faire, nous nous pencherons dans un premier temps sur le système de Van der Pol où nous verrons comment l'ordre de l'approximation des fonctions propres ainsi que le pas de discrétisation impactent la qualité des approximations obtenues dans la section 2.2 du chapitre 2. Forts de ces analyses, nous nous pencherons ensuite sur le système (2.9) où nous verrons que les résultats ne sont pas aussi concluants que pour le système de Van der Pol et que l'algorithme peut devenir assez coûteux. Dès lors et afin de pallier ce problème, nous proposerons enfin l'implémentation d'une grille adaptative dont le code est essentiellement basé sur l'Algorithme 2.

Penchons-nous dans un premier temps sur le cas particulier du système de Van der Pol. Afin de pouvoir appliquer l'Algorithme 2 et étudier l'impact du choix de l'ordre et du pas de la grille, nous avons calculé les fonctions propres de l'opérateur de Koopman tronquées aux ordres 5 et 10 respectivement. Dès lors, pour chacune de ces deux approximations, nous avons appliqué l'Algorithme 2 avec le domaine $\mathbb{X} = [-3, 3]$ discrétisé avec deux pas différents $\Delta\mathbf{x} = 0.02$ et $\Delta\mathbf{x} = 0.006$. À la Figure 4.3 nous avons représenté les différentes régions $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$ respectivement en bleu clair et bleu foncé où l'ensemble rose correspond aux points de $\dot{\mathcal{V}}_{<0}$ pour lesquels le signe de $\dot{\mathcal{V}}$ reste constant à l'intérieur des différentes cellules associées. Dans ce cas, nous remarquons que les résultats sont plutôt satisfaisants lorsque le pas de la grille est assez petit, ce qui n'est pas étonnant. En effet, au plus le pas est petit, au plus l'écart entre les points du domaine est petit de sorte que si une fonction est négative aux extrémités d'une cellule, elle le reste entre ces points. Autrement dit, par continuité de la fonction, il y a peu de chance que cette dernière change de signe si le pas de la grille est petit. D'un autre côté, nous remarquons qu'il n'est pas spécialement avantageux d'augmenter le nombre de fonctions de base. En effet, nous observons de meilleurs résultats à l'ordre 5 où la courbe de niveau obtenue en mauve est quasi identique à celle obtenue à la Figure 2.3 que nous avons tracée en pointillés noirs. Une explication intuitive serait qu'en ajoutant des fonctions de base, nous augmentons le nombre d'évaluations nécessaires pour le calcul de K_C , ce qui pourrait entraîner des erreurs numériques conséquentes à grande échelle de sorte que la condition (4.8) ne soit plus satisfaite.

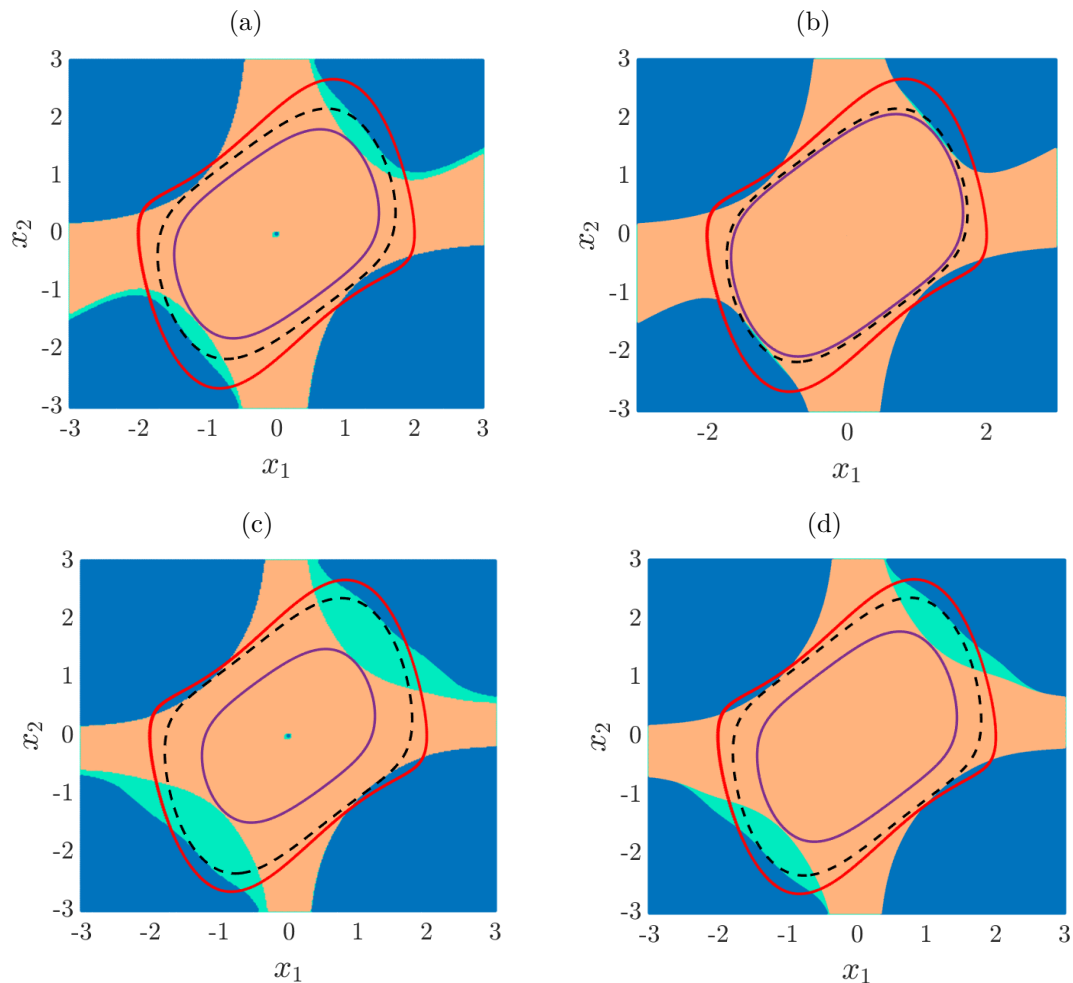


Figure 4.3 – Application de l’Algorithme 2 au système de Van der Pol. (a) - (b) Pour ces figures, les fonctions propres de l’opérateur de Koopman ont été calculées à l’ordre 5 où le domaine $\mathbb{X} = [-3, 3] \times [-3, 3]$ a été discrétisé selon (a) $\Delta \mathbf{x} = 0.02$ et (b) $\Delta \mathbf{x} = 0.006$. (c) - (d) Mêmes figures mais avec les fonctions propres calculées jusqu’à l’ordre 10. Pour chacune de ces figures, nous avons représenté en mauve la plus grande courbe de niveau que l’on peut extraire de l’ensemble rose et en pointillés noirs celle obtenue à la Figure 2.3. Peu importe le nombre de fonctions de base, nous remarquons que les approximations sont meilleures en discrétisant le pas. Cependant, augmenter le nombre de fonctions de base n’améliore pas la qualité des résultats et tend même à la baisser.

Revenons quelques instants sur les résultats obtenus à la Figure 4.3. Même si cela n’est pas très visible, il existe en réalité une petite région $\Gamma \subset \dot{\mathcal{V}}_{<0}$ autour de l’origine dont les points (ou cellules) la constituant ne satisfont pas la condition (4.8). Dès lors, plusieurs cas de figure sont envisageables ; Soit l’origine est le seul point pour lequel $\dot{\mathcal{V}}(\mathbf{x}) = 0$ pour tout $\mathbf{x} \in \dot{\mathcal{V}}_{<0}$ et nous obtenons la stabilité à l’aide du Théorème 1.1 de Lyapunov. Soit, dans le pire scénario possible, il existe une certaine région non vide autour de l’origine $\Upsilon \subset \dot{\mathcal{V}}_{>0} \subset \dot{\mathcal{V}}_{<0}$. Dans ce cas, nous pouvons toujours utiliser le Théorème 3.2 permettant d’assurer que pour toute condition initiale \mathbf{x}_0 dans le plus grand ensemble de niveau Ξ , tracé en mauve dans l’ensemble rose, la solution y étant issue converge dans le plus grand ensemble positivement invariant dans la région $\dot{\mathcal{V}}_0$.

Penchons nous à présent sur le deuxième exemple (2.9) pour lequel nous avons reproduit exactement les mêmes résultats que ceux de la Figure 4.3 pour $\Delta \mathbf{x} = 0.02$ et $\Delta \mathbf{x} = 0.00625$. Dans ce cas, nous remarquons à la Figure 4.4 que nous pouvons tirer des conclusions assez similaires à celles obtenues pour le système de Van der Pol où nous remarquons à la Figure 4.3 (b) que la courbe de niveau que nous pouvons extraire est assez proche de celle obtenue à la la Figure 2.3. Notons toutefois que le résultat n'est pas

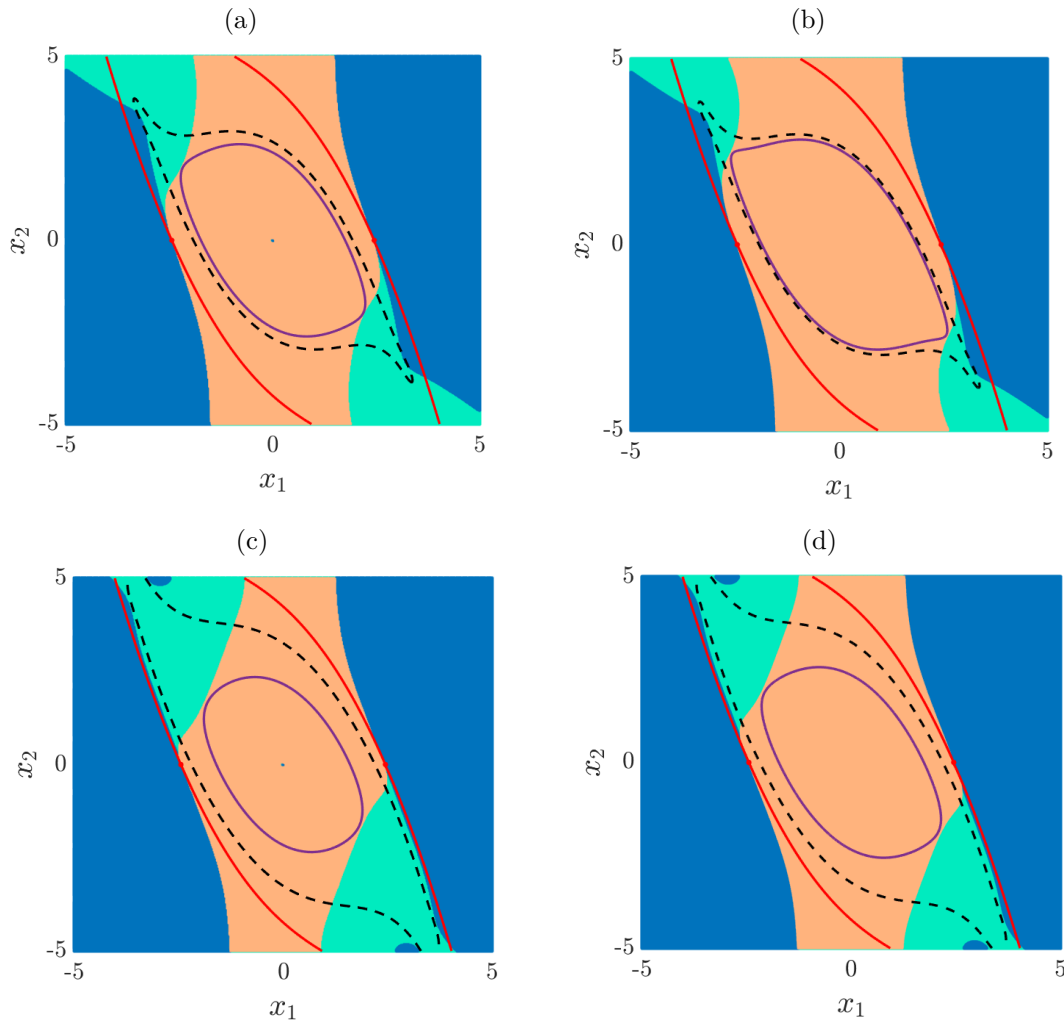


Figure 4.4 – Application de l'Algorithme 2 au système (2.9). (a) - (b) Pour ces figures, les fonctions propres de l'opérateur de Koopman ont été calculées à l'ordre 5 où le domaine $\mathbb{X} = [-3, 3] \times [-3, 3]$ a été discrétisé selon (a) $\Delta \mathbf{x} = 0.02$ et (b) $\Delta \mathbf{x} = 0.006$. (c) - (d) Même figures mais avec les fonctions propres calculées jusqu'à l'ordre 10. Pour chacune de ces figures, nous avons représenté en mauve la plus grande courbe de niveau que l'on peut extraire de l'ensemble rose et en pointillés noirs celle obtenue à la Figure 2.3. Peu importe le nombre de fonctions de base, nous remarquons de nouveau que les approximations sont meilleures en discrétisant le pas mais qu'augmenter le nombre de fonctions de base n'améliore pas la qualité des résultats. De nouveau ici, nous obtenons un résultat assez satisfaisant à l'ordre 5 ou nous remarquons que la courbe de niveau que nous pouvons extraire de l'ensemble rose (en mauve) colle plutôt bien à la courbe de niveau que nous avons extraite à la Figure 2.3.

parfait étant donné que nous n'arrivons pas à capturer les deux extrémités de la courbe de niveau en pointillés. Précisons également que pour cet exemple, il existe aussi une petite région $\Gamma \subset \dot{\mathcal{V}}_{<0}$ autour de l'origine dont les points (ou cellule) la constituant ne satisfont pas la condition (4.8). Cependant, nous pouvons appliquer un raisonnement similaire à celui qui a été décrit pour le système de Van der Pol.

Remarque 4.2. *À travers cette remarque, nous tenons à préciser que la taille du domaine considéré a un impact assez conséquent sur le temps de calcul de l'Algorithme 2. En effet, remarquons par exemple que le système (2.9) est en général discrétisé sur le domaine $\mathbb{X} = [-5, 5]$ comparé à $\mathbb{X} = [-3, 3]$ pour le système de Van der Pol. Dans ce cas, pour un pas de discrétisation égale, nous avons beaucoup plus de points pour le domaine $\mathbb{X} = [-5, 5]$, ce qui ralentit de manière considérable le temps d'exécution du code et peut, à une certaine échelle, entraîner la gestion de grille beaucoup trop grande pour un ordinateur classique.*

4.2 Vers l'implémentation d'une grille adaptative

Bien que les résultats obtenus pour le système de Van der Pol à la Figure 4.3 (b) soient pratiquement parfaits à l'ordre 5, nous pourrions vouloir chercher à améliorer ceux du système (2.9) afin de capturer les extrémités de l'ensemble de niveau en pointillés à la Figure 4.4 (b). Pour ce faire, nous pourrions simplement diminuer le pas de discrétisation davantage jusqu'à satisfaction. Cependant, réduire le pas nécessite de travailler avec un domaine qui devient impossible à gérer d'un point de vue numérique. En effet, si nous considérons le deuxième exemple (2.9), nous remarquons qu'un pas de discrétisation $\Delta \mathbf{x} = 0.00625$ engendre une grille contenant 2 563 201 points. Afin de pallier ce problème, nous pouvons alors imaginer construire une grille itérative permettant d'adapter le pas et réduire ainsi le temps de calcul. Plus précisément, nous pourrions partir avec une grille initiale fixée et pour chacune des cellules la constituant, nous testons la condition (4.8). Si cette condition est vérifiée, nous gardons en mémoire ces cellules et leur pas de discrétisation. Sinon, nous divisons les cellules restantes et nous réitérons ce dernier point jusqu'à un certain seuil que nous aurons fixée en amont. Dernièrement, nous avons typiquement implémenté une telle méthode faisant l'objet de cette dernière section. Dans cette section, nous allons l'appliquer aux deux systèmes (1.7) et (2.9) où nous verrons qu'il est possible d'affiner légèrement la précision des résultats obtenus à la Figure 2.3 pour le système (2.9), ouvrant ainsi la voie à une utilisation pour d'autres systèmes.

Penchons nous dans un premier temps sur le système de Van der Pol. Pour ce faire, nous avons considéré une grille initiale $\mathbb{X} = [-3, 3]$ avec un certain pas de discrétisation $\Delta \mathbf{x} = 0.6$. Afin d'appliquer la méthode itérative, nous avons ici divisé chaque cellule de manière dichotomique pour deux valeurs seuils $\Delta \mathbf{x} = 0.075$ et $\Delta \mathbf{x} = 0.0023$. Dans ce cas, nous remarquons à la Figure 4.5 que la grille adaptative livre de meilleurs résultats lorsque le nombre d'itérations augmente. En effet nous remarquons que la grille est plus fine de sorte que l'approximation du bassin d'attraction est meilleure et nous obtenons en réalité un ensemble très proche de celui en rose à la Figure 4.3 (b). Notons également que la grille adaptative est intéressante car elle permet d'éliminer une bonne partie du domaine assez rapidement, ce que nous observons à travers les cellules de tailles plus importantes. Cependant, nous tenons à préciser que les temps d'exécution sont en réalité semblables à une grille de taille fixée avec $\Delta \mathbf{x} = 0.0023$ si bien qu'il n'est peut-être pas si avantageux d'utiliser une grille adaptative dans ce cas, compte tenu en réalité des résultats très bons

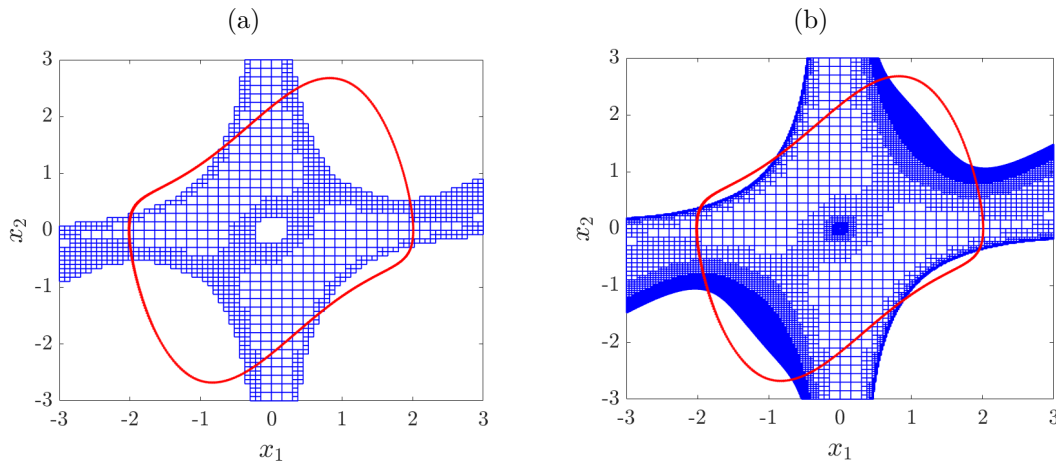


Figure 4.5 – Application de la grille adaptative au système de Van der Pol où l’algorithme a été stoppé lorsque le pas de discrétisation était plus petit que (a) 0.075, (b) 0.0023. Dans ce cas, nous remarquons que si nous laissons évoluer l’algorithme suffisamment longtemps, nous obtenons de meilleurs résultats avec une région assez satisfaisante. Dans ce cas, le plus grand ensemble de niveau dans cette région serait très proche de celui obtenu à la Figure 2.2 à l’ordre 5.

obtenus pour une grille fixe. Afin de terminer cette analyse sur le système de Van der Pol, nous tenons également à faire remarquer qu’au plus la grille est fine, au plus nous réduisons la taille de la région près de l’origine pour laquelle la condition n’est pas satisfaite. Ce constat est alors un bon indicateur selon lequel la fonction a effectivement peu de chance de changer de signe au voisinage de l’origine. Ainsi, même si nous ne sommes pas totalement sûrs que la fonction ne change pas de signe, nous avons à présent un argument intuitif nous permettant de penser que le Théorème de Lyapunov est bien applicable dans ce cas, permettant ainsi de justifier plus rigoureusement les résultats du Chapitre 2.

Penchons nous à présent sur le deuxième exemple où nous avons reproduit la Figure 4.5 cette fois-ci avec un domaine de départ $\mathbb{X} = [-5, 5] \times [-5, 5]$ discrétisé selon $\Delta \mathbf{x} = 1$ et pour lequel nous avons arrêté l’algorithme pour des valeurs de $\Delta \mathbf{x}$ inférieures à 0.125 et 0.0039⁵ avec $\mathbb{X} = [-5, 5]$ (voir Figure 4.6). Dans ce cas, nous pouvons tirer plus au moins les mêmes conclusions que pour le système de Van der Pol.

Cependant, contrairement au système de Van der Pol, l’utilisation de la grille adaptative peut devenir avantageuse pour cet exemple. En effet, compte tenu des cellules obtenues à la Figure 4.6 (b), nous remarquons qu’il faudrait une discrétisation encore plus fine du domaine pour capturer les zones du bassin d’attraction en haut à gauche et en bas à droite respectivement. Cependant, si nous considérons une discrétisation du domaine avec un pas de discrétisation inférieur à $\Delta \mathbf{x} = 0.0039$ pour ce système, l’algorithme 2 pour le cas d’une grille fixe devient beaucoup trop coûteux numériquement. Pour être plus précis, nous manipulons des données tellement volumineuses qu’un ordinateur classique ne peut pas traiter. Dans ce cas, la grille adaptative peut s’avérer utile. Si nous considérons le domaine de départ $\mathbb{X} = [-5, 5]$ avec $\Delta \mathbf{x} = 1$, nous pourrions rencontrer

5. Ce choix est en réalité dû au fait que nous avons considéré le même nombre de points dans ce domaine que pour le système de Van der Pol.

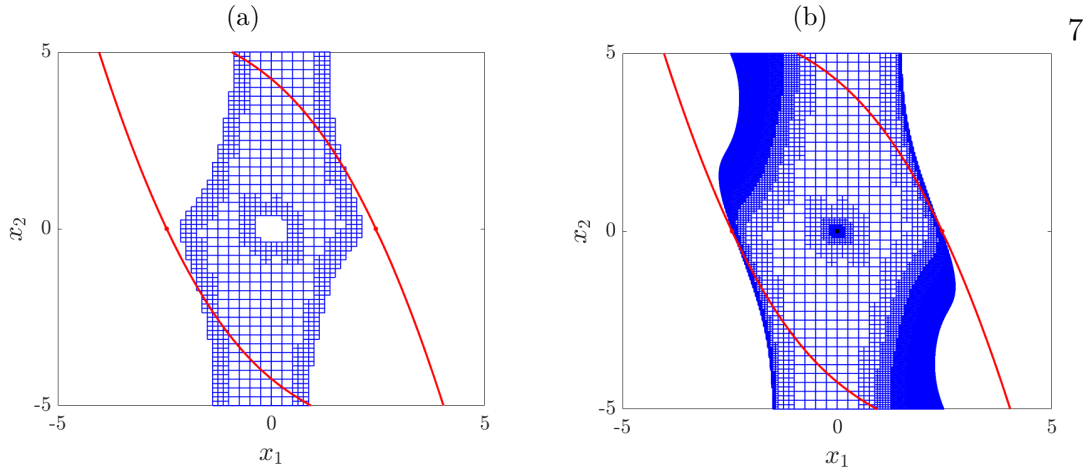


Figure 4.6 – Application de la grille adaptative au système (2.9) où l’algorithme a été stoppé lorsque le pas de discrétisation était plus petit que (a) 0.125, (b) 0.0039. Dans ce cas, même si nous obtenons de meilleurs résultats lorsque nous laissons évoluer l’algorithme suffisamment longtemps, ces derniers restent assez mauvais et nous ne capturons qu’une petite partie du bassin d’attraction.

un problème. En effet, à travers la Figure 4.6 (b), nous nous sommes rendus compte qu’il était nécessaire d’atteindre un pas de discrétisation suffisamment petit. Cependant, si le domaine de départ est discrétisé de la sorte, nous devons effectuer énormément de subdivision dichotomique avant de pouvoir retirer des cellules pour lesquelles nous sommes sûrs que le signe de $\dot{\mathcal{V}}$ demeure inchangé. Dans ce cas, nous nous retrouverions dans le cas similaire d’une grille fixe où nous manipulerions un nombre de données beaucoup trop élevé. A contrario, si nous considérons une discrétisation du domaine initiale assez fine, dès la première itération il sera possible d’enlever une grosse partie des cellules, laissant ainsi l’opportunité à l’algorithme de détecter d’autres régions en diminuant le pas à des valeurs que l’Algorithme 2 ne peut pas traiter numériquement. Autrement dit, la grille adaptative permet dans ce cas de considérer des pas $\Delta \mathbf{x}$ encore plus petits que pour une grille fixe. Dans le cas de l’exemple 2, nous avons typiquement appliqué ce raisonnement en considérant initialement le domaine $\mathbb{X} = [-5, 5]$ avec $\Delta \mathbf{x} = 0.01$. Dans ce cas, lorsque nous laissons évoluer l’algorithme suffisamment longtemps, nous obtenons le résultat illustré à la Figure 4.7. Dès lors, nous remarquons que nous avons réussi à couvrir une région plus importante que précédemment, permettant ainsi d’améliorer la qualité de l’approximation donnée par la courbe pleine en mauve sur le Figure. Notons cependant que le temps d’exécution de l’algorithme pour cet exemple est particulièrement long de sorte qu’il y a de nouveau un compromis entre le temps d’exécution de l’algorithme par rapport à la précision des résultats. Pour ce système par exemple, nous remarquons que les courbes de niveau sont telles qu’il n’est peut-être pas si avantageux d’envisager une grille adaptative puisque les résultats ne sont pas significativement différents. Ainsi, la question du compromis dépend également du problème étudié et de sa géométrie. Pour le système (2.9), nous avons typiquement une géométrie plus complexe. Cependant, au vu du résultat obtenu, nous avons bon espoir que son utilisation soit intéressante lorsque le domaine doit être discrétisé assez finement pour capturer la majeure partie des zones au sein desquelles le signe de $\dot{\mathcal{V}}$ est négatif et le demeure.

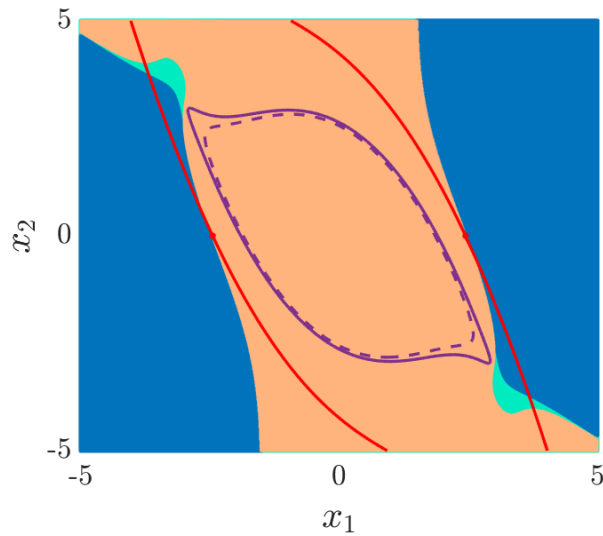


Figure 4.7 – Application de la grille adaptative au système (2.9) où l’algorithme a été stopper lorsque le pas de discrétisation était plus petit que (a) 0.00125 avec une domaine initiale $\mathbb{X} = [-5, 5]$ discrétisé selon $\Delta \mathbf{x} = 0.01$. Dans ce cas, nous avons représenté en bleu clair et bleu foncé respectivement les zone $\dot{\mathcal{V}}_{<0}$ et $\dot{\mathcal{V}}_{\geq 0}$ obtenu à l’ordre 5 selon un domaine arbitraire afin de voir à quoi ressemble ces régions. Notons que pour cet exemple, l’ensemble des points de la grille adaptative a été représenté en rose. Dans ce cas, il semblerait que l’algorithme livre des résultats très satisfaisant où nous retrouvons la quasi totalité de la région $\dot{\mathcal{V}}_{<0}$. En mauve pointillé nous avons représenté la courbe de niveau obtenue pour le cas d’une grille fixe à la Figure 4.4 (b), et en mauve (trait plein) la courbe de niveau pour ce domaine adapté. Dans ce cas, nous remarquons que nous obtenons un meilleur résultats mais que l’amélioration est très légère. Ceci est notamment dû à la géométrie de la fonction de Lyapunov (voir courbes de niveau de la Figure 2.3).

Conclusion et Perspectives

L'objet de ce mémoire a été d'étudier la théorie de la stabilité des systèmes dynamiques. Plus précisément, dans le cadre du chapitre 1, nous avons rappelé certaines propriétés fondamentales à ce propos dans le cas particulier des systèmes à temps continu, linéaires et non linéaires. En ce qui concerne les systèmes linéaires, nous avons vu que la question de la stabilité était déjà bien établie et reposait sur les propriétés spectrales de la matrice associée au système. En revanche, pour les systèmes non linéaires, il n'existe pas réellement de méthodes systématiques pour induire la stabilité globale d'attracteurs du système et dans le cas spécifique de points fixes, nous devons généralement déterminer une fonction de Lyapunov parmi une infinité de choix possibles. Ceci a donc motivé une approche fonctionnelle de la théorie des systèmes dynamiques au moyen de l'opérateur de Koopman. Nous avons vu que cet opérateur jouissait de propriétés très intéressantes dont la linéarité, ce qui nous a notamment permis d'étudier son spectre. Nous avons en particulier constaté que les fonctions propres de l'opérateur de Koopman permettaient de caractériser la stabilité globale des systèmes.

Cependant, bien que le spectre de l'opérateur de Koopman soit particulièrement intéressant dans le contexte de la stabilité de points fixes, il s'agit d'un opérateur de dimension infinie de sorte que le calcul de ses fonctions propres n'est pas évident et compliqué à déterminer analytiquement. Dans ce cas, nous avons vu dans le chapitre 2 qu'il était possible de les approximer en considérant une projection de l'opérateur (ou de son générateur) sur un sous-espace fini de fonctions. Dans le cadre du mémoire, nous avons alors développé une méthode numérique permettant d'approximer les fonctions propres de l'opérateur peu importe le choix de la base et pour un opérateur de projection fixé. Précisons que dans ce travail, n'ont été considérés que des opérateurs de troncature et de projection orthogonale au sens des moindres carrés. Une fois la méthode développée, nous l'avons dans un premier temps testée aux cas particuliers des monômes afin de reproduire les résultats de la référence [5, 6]. Dans ce cas, nous avons vu que l'utilisation des monômes permettaient d'obtenir de très bonnes approximations du bassin d'attraction des deux systèmes de référence étudiés tout au long du mémoire. Forts de ces analyses, nous avons également testé la méthode sur deux systèmes en trois dimensions où nous avons observé des résultats tout aussi satisfaisants et qui n'avaient pas encore été abordés dans la littérature.

À la fin du Chapitre 2, nous avons vu que l'utilisation des monômes pouvait devenir assez coûteuse pour des systèmes de plusieurs dimensions, ce qui nous a alors motivé à considérer la stabilité des systèmes dynamiques par l'intermédiaire d'autres fonctions de base. Dans le chapitre 3, nous nous sommes intéressés à cette question à travers l'utilisation de fonctions radiales souvent utilisées dans les méthodes numériques basées sur les données [13, 14]. À travers les deux systèmes de référence, nous montrons que les fonctions radiales étaient en réalité très sensibles à certains paramètres tels que le domaine

des centres et de la projection. Cependant, pour un choix particulier de paramètres, nous avons observé des résultats assez satisfaisants et prometteurs pour une application de ces fonctions en plusieurs dimensions. Nous avons notamment appuyé cet argument en testant la méthode sur deux exemples de systèmes de dimension trois qui nous permettaient également d'approximer la bassin d'attraction. Malgré que les approximations obtenues ne soient pas aussi satisfaisantes comparativement aux monômes, elles restent toutefois motivantes pour une application en plus hautes dimensions. Notons que parallèlement à cette étude, nous avons vu qu'il fallait rester assez prudent lorsque nous parlions du bassin d'attraction dans le cas particulier des fonctions de base radiales étant donné que nous avons observé l'existence de régions où la fonction de Lyapunov croit le long des trajectoires de sorte que ni le théorème de LaSalle ni celui de Lyapunov n'étaient adaptés à cette situation et ne permettaient pas d'induire la stabilité du point fixe. Afin de pallier ce problème, nous avons adapté la preuve du théorème de LaSalle et de Lyapunov où nous avons vu que nous pouvions alors obtenir un résultat de stabilité globale.

Dans le chapitre 2, nous avons vu que nous manquions de rigueur sur l'approximation du bassin d'attraction. En effet, cette dernière est obtenue en calculant le plus grand ensemble de niveau de la fonction de Lyapunov dans l'ensemble des points pour lesquels la fonction décroît le long des trajectoires. Cependant, cet ensemble de points est déterminé selon une grille discrétisée et rien ne garantissait que la fonction restait décroissante entre les mailles de la grille. Dans le chapitre 4, nous avons vu que dans le cadre des monômes, il était possible de développer un critère basé sur une borne du gradient de la dérivée de la fonction de Lyapunov. Ce critère nous a permis de construire un algorithme afin d'extraire toutes les mailles de la grille pour lesquelles le signe de la dérivée restait négatif. Dans ce cas, nous avons pu garantir rigoureusement la stabilité des approximations du bassin d'attraction. À travers les différents exemples en deux dimensions que nous avons étudiés tout au long du mémoire, nous avons observé que l'algorithme permettait d'obtenir d'assez bons résultats lorsque le pas de la grille ainsi que l'ordre des monômes de base étaient suffisamment petits. Malgré que les approximations soient tout à fait acceptables pour une grille fixe, nous avons vu qu'il était possible d'améliorer les résultats du système (2.9) en considérant une grille adaptative avec un pas de discrétisation variable. Cependant, nous avons expliqué que cet algorithme pouvait devenir assez coûteux d'un point de vue numérique et qu'il existe en réalité un compromis entre le temps d'exécution de l'algorithme et la précision des résultats obtenus.

Du point de vue des perspectives de ce mémoire, nous pensons qu'il pourrait être tout à fait intéressant d'appliquer la méthode développée à des systèmes en plus haute dimension. Dans ce contexte, nous pensons qu'il pourrait être assez judicieux de questionner l'adaptation de la méthode, ou d'une moins une partie, dans un autre langage de programmation. En effet, certains calculs peuvent rapidement devenir coûteux en Matlab et le choix d'un autre langage de programmation permettrait de pallier ce problème pour une application en plus haute dimension. Par ailleurs, nous n'avons ici testé que des fonctions de base radiales et il serait possible d'envisager d'autres bases telles que des bases hybrides où nous combinerions certains monômes avec des fonctions radiales. Nous pourrions également utiliser des méthodes de *machine learning* pour choisir ces fonctions de base de manière automatique. En réalité, compte tenu de la diversité du choix de fonctions de base possibles, cette question est encore ouverte et il n'était bien entendu pas possible de la couvrir entièrement dans ce mémoire. Dans le chapitre 4, nous avons étudié une borne

K_C sur la norme du gradient de $\mathcal{V}(\cdot)$. Cependant, nous avons conclu que cette borne était conservative. Ainsi, une perspective du mémoire serait d'obtenir une borne moins conservative en considérant éventuellement une autre approche que celle utilisant la norme du gradient. Enfin, vers la fin du chapitre 4, nous avons développé une grille adaptative mais il s'agit d'un travail préliminaire. Ainsi, une autre perspective du mémoire serait de tester l'algorithme obtenu dans ce contexte sur d'autres exemples afin d'analyser la qualité des approximations et éventuellement de l'adapter pour d'autres fonctions de base.

Bibliographie

- [1] ENGEL K.J. & NAGEL R. *One-Parameter Semigroups for Linear Evolution Equations*. Science & Business Media, vol. 194, Springer, Berlin, 1999.
- [2] KÜHNER V. *What can Koopmanism do for attractors in dynamical systems?*. The Journal of Analysis, 2019.
- [3] MAUROY A. *Méthodes avancées pour les systèmes non linéaires*. Université de Namur, Année académique 2020-2021.
- [4] MAUROY A. *Modélisation et analyse des systèmes dynamiques*. Université de Namur, Année académique 2018-2019.
- [5] MAUROY A. & MEZIĆ I. *A spectral operator-theoretic framework for global stability*. 52nd IEEE Conference on Decision and Control, pp. 5234-5239, 2013.
- [6] MAUROY A. & MEZIĆ I. *Global stability analysis using the eigenfunctions of the Koopman operator*. IEEE Transactions on Automatic Control, vol. 61(11), pp. 3356-3369, 2018.
- [7] MAUROY A., MEZIĆ I. & SOOTLA A. *Koopman framework for global stability analysis*. The Koopman operator in systems and control, Springer, pp. 35-58, 2020.
- [8] KHALIL H. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, 2002.
- [9] LASOTA A. & MACKEY M.C. *Chaos, fractals, and noise : stochastic aspects of dynamics*. Science & Business Media, vol. 97, Springer, 1998.
- [10] MAUROY A., SUSUKI Y. & MEZIĆ I. *Koopman operator in systems and control*. Berlin : Springer International Publishing, 2020.
- [11] MAUROY A. & GONCALVES J. *Koopman-based lifting techniques for nonlinear systems identification*. IEEE Transactions on Automatic Control, vol. 65(6), pp. 2550-2565, 2019.
- [12] KORDA M., HENRION D. & JONES C.N. *Controller design and region of attraction estimation for nonlinear dynamical systems*. IFAC Proceedings Volumes, vol.47(3), pp. 2310-2316, 2014.
- [13] KLUS S., NÜSKE F., KOLTAI P., WU H., KEVREKIDIS I., SCHÜTTE C. & NOÉ F. *Data-driven model reduction and transfer operator approximation*. Journal of Nonlinear Science, vol. 28(3), pp. 985-1010, 2018.
- [14] KLUS S., HUSIC B.E., MOLLENHAUER M. & NOÉ F. *Kernel methods for detecting coherent structures in dynamical data*. Chaos : An Interdisciplinary Journal of Nonlinear Science, vol. 29(12) :123112, 2019.