

# THESE

présentée à

**L'UNIVERSITE ABOUBEKR BELKAID DE TLEMCEM**  
**Institut de Sciences Exactes**

pour obtenir le grade de

**DOCTEUR D'ETAT ES-SCIENCES PHYSIQUES**



## **CONTRIBUTION A L'APPROXIMATION ET LA SYNTHESE DE LA STABILITE D'UNE EQUATION DIFFERENTIELLE ORDINAIRE NON-LINEAIRE**

Le 12 octobre 1996

devant le jury

Président:

**A. BOUCHERIF**, Professeur à l'Université Aboubekr Belkaid, Tlemcen

Directeur de thèse en France

**O. ARINO**, Professeur à l'Université de Pau et des Pays de l'Adour

Directeur de thèse en Algérie

**B. BENYOUCEF**, Professeur à l'Université Aboubekr Belkaid, Tlemcen

Examineurs:

**H. MOKHTAR-KHARROUBI**, Professeur à l'Université d'Oran

**N. GHOUALI**, Maître de Conférences à l'Université Aboubekr Belkaid, Tlemcen

Rapporteurs extérieurs:

**E. SANCHEZ**, Professeur à l'Université Polytechnique de Madrid (Espagne)

**M. WITTEN**, Professeur à l'Université du Michigan (U.S.A)

à la mémoire de **HAMOU BENAMEUR**, mon père

à ma mère, pour ses sacrifices

à ma femme, pour son soutien sans failles

à mes enfants, Omar farouk et Asma soumia

à toute ma famille

## REMERCIEMENTS

**Ce travail a été réalisé au Laboratoire de Mathématiques Appliquées URA 1204, CNRS de l'Université de Pau et des Pays de l'Adour.**

**J'exprime ma profonde gratitude au Professeur et collègue Ovide Arino, qui m'a accueilli dans son laboratoire, et qui n'a lésiné sur aucun effort pour la réussite de cette thèse. Son soutien matériel et moral a été sans limite.**

**Le Professeur Boucherif Abdelkader m'a fait l'honneur de présider le jury de cette thèse, les Professeurs Mokhtar-Kharroubi Houcine, Benyoucef Boumedienne, Ghouali Nouredinne, ont accepté de faire partie de ce jury, je les en remercie respectueusement.**

**Enfin, je remercie l'ensemble des collègues du Laboratoire de Pau pour leur accueil et particulièrement mon ami Boussouar Ahmed.**

# SOMMAIRE

0.1	Introduction. . . . .	5
0.1.1	Note historique. . . . .	5
0.1.2	Position du problème. . . . .	5
0.1.3	Contribution. . . . .	9
<b>1</b>	<b>Présentation de la dérivée optimale.</b>	<b>14</b>
1.1	Introduction. . . . .	14
1.2	Problème posé. . . . .	14
1.3	Résolution du problème. . . . .	16
1.3.1	Rappels. . . . .	16
1.3.2	Résolution. . . . .	17
1.3.3	Condition nécessaire et suffisante pour que $\Gamma(x)$ soit in- versible. . . . .	18
1.4	Algorithme. . . . .	21
1.5	Propriétés de la procédure. . . . .	22
1.5.1	Cas où l'application est linéaire. . . . .	23
1.5.2	Cas général, où le système est la somme d'un terme linéaire et d'un terme non-linéaire. . . . .	23
1.6	Cas scalaire. . . . .	21
1.6.1	Introduction. . . . .	21
1.6.2	Calcul de la dérivée optimale. . . . .	25
1.6.3	Autre expression pour la dérivée optimale. . . . .	26

1.6.4	Propriétés. . . . .	27
1.7	Relation entre la dérivée optimale et celle obtenue par dérivation au sens de Fréchet en 0. . . . .	35
<b>2</b>	<b>Détermination de la dérivée optimale.</b>	<b>38</b>
2.1	Introduction. . . . .	38
2.2	Position du problème. . . . .	40
2.3	Relation de récurrence. . . . .	40
2.4	Existence de la dérivée optimale. . . . .	41
2.5	Unicité de la dérivée optimale. . . . .	46
2.6	Convergence de la dérivée optimale. . . . .	50
<b>3</b>	<b>Propriétés de la dérivée optimale.</b>	<b>53</b>
3.1	Introduction. . . . .	53
3.2	Ordre. . . . .	53
3.3	Stabilité. . . . .	57
3.3.1	Note historique . . . . .	57
3.3.2	Lien entre la dérivée optimale et la stabilité asymptotique. . . . .	58
3.4	Commentaires. . . . .	75
<b>4</b>	<b>Méthode d'approximation.</b>	<b>77</b>
4.1	Introduction. . . . .	77
4.2	Idée du problème. . . . .	78
4.3	Préliminaires. . . . .	78
4.3.1	Erreur sur les conditions initiales. . . . .	78
4.3.2	Changement de Variable et de fonction. . . . .	79
4.4	Procédure de la dérivée optimale dans l'intervalle $[\alpha, \beta]$ . . . . .	80
4.5	Procédure d'approximation optimale dans l'intervalle $[0, T]$ . . . . .	82
4.5.1	Algorithme. . . . .	83
4.6	Estimation de l'erreur. . . . .	85

4.6.1	Cas où la fonction $F$ n'est pas dissipative . . . . .	93
4.7	Commentaires. . . . .	95
<b>5</b>	<b>Mise en oeuvre numérique.</b>	<b>97</b>
5.1	Introduction. . . . .	97
5.2	Procédure de la dérivée optimale. . . . .	97
5.2.1	Schéma général. . . . .	97
5.2.2	Organigramme. . . . .	99
5.2.3	Applications. . . . .	101
5.3	Approximation et stabilité. . . . .	112
5.4	Méthode d'approximation optimale. . . . .	123
5.4.1	Commentaires. . . . .	125
	Conclusions générales . . . . .	129
	Références . . . . .	131

**INTRODUCTION.**

## 0.1 Introduction.

### 0.1.1 Note historique.

L'étude des équations différentielles est un domaine mathématique qui historiquement a fait l'objet de nombreuses recherches [4], [1], [5], [6], [10], ..., et continue cependant de rester d'actualité, par le fait qu'elle intéresse particulièrement des disciplines comme la mécanique, les sciences physiques et plus récemment la biologie, l'électronique, spécialités où de nombreux "modèles" conduisent à des équations du même type. Il faut savoir que la plupart de ces équations sont globalement de nature non-linéaire. La dénomination "non-linéaire" rassemble des systèmes extrêmement divers ayant peu de points communs dans leur comportement. Il en résulte qu'il n'existe pas, jusqu'à présent, de théorie d'ensemble des équations non-linéaires. Parmi ces problèmes non-linéaires, une classe importante est modélisée par les équations différentielles ordinaires non-linéaires de la forme

$$\begin{cases} \frac{dx}{dt} = F(x) \\ x(0) = x_0. \end{cases} \quad (0.1)$$

De nombreux travaux ont été consacrés à ce sujet, différant généralement par la motivation de l'auteur ( Mécanique, Géométrie, Physique, ...). Une liste exhaustive des travaux portant sur cette classe d'équations comprendrait des centaines d'articles et livres et déborderait largement sur le siècle passé. Nous nous bornerons à citer les ouvrages que nous avons utilisés dans notre travail: Coddington-Levinson (1955), Hale (1969, 1971), Rouche-Mawhin (1973), Poincaré (1975), Reinhart (1975), Sibony-Mardon (1988), Demailly (1991).

### 0.1.2 Position du problème.

Dans l'étude des équations différentielles ordinaires non-linéaires, les méthodes de linéarisation jouent un rôle important. Si beaucoup de systèmes peuvent admettre



un domaine de comportement linéaire, la linéarité est toujours une approximation de la réalité. L'approximation la plus classique est celle déterminée par la dérivée au sens de Fréchet de l'équation non-linéaire.

S'agissant de l'étude du comportement des solutions d'une équation non-linéaire autour d'un point singulier, la linéarisation classique ne permet pas de répondre, par exemple, dans les cas où la fonction n'est pas assez régulière et celui où elle est nulle. Pour mieux situer le problème, nous allons considérer brièvement les exemples suivants<sup>1</sup>[20], [22].

**Exemple 0.1.1**

$$\begin{cases} \frac{dx}{dt} = -x + \sin(|y|) \\ \frac{dy}{dt} = -y + \sin(|x|) \end{cases}, (x(0), y(0)) = (x_0, y_0)$$

*Ce système comporte une non-linéarité de type valeur absolue, donc non différentiable en zéro.*

**Exemple 0.1.2**

$$\begin{cases} \frac{dx}{dt} = -\frac{x^4}{x^2 + y^2} \\ \frac{dy}{dt} = -\frac{y^4}{x^2 + y^2} \end{cases}, (x(0), y(0)) = (x_0, y_0)$$

*La dérivée étant nulle au point (0,0), on ne peut déduire le comportement de la solution de l'équation non-linéaire de celle de la linéarisation en ce point.*

S'agissant de l'étude de la stabilité d'un point d'équilibre, la principale méthode est de considérer l'équation linéaire obtenue par dérivation au sens de Fréchet en ce point de la fonction non-linéaire définissant l'équation différentielle, c'est à dire

$$\frac{dx}{dt} = DF(\text{Pt d'équilibre}), x(0) = x_0.$$

---

<sup>1</sup>Ces exemples seront traités plus en détail dans le chapitre 5.

L'équation ainsi obtenue a le même comportement que l'équation non-linéaire dans le cas hyperbolique, quand les valeurs propres ne rencontrent pas l'axe des imaginaires. Il y a au moins quatre écueils importants à la bonne utilisation de cette méthode:

1) Impossibilité de situer les valeurs propres par rapport à l'axe des imaginaires, en particulier si une ou plusieurs valeurs propres sont proches de cet axe.

2) Si l'équilibre est un centre pour l'équation linéaire, c'est à dire, s'il y a des valeurs propres imaginaires ou si 0 est valeur propre. Le comportement de la solution de l'équation différentielle ordinaire non-linéaire au voisinage d'un tel point peut être très varié. C'est le cas de l'exemple (0.1.2), car le système admet une valeur propre double nulle.

3) Si la fonction n'est pas assez régulière au voisinage du point stationnaire. Dans ce cas, on ne peut en général calculer la dérivée au sens de Fréchet. Cela peut arriver par exemple si la fonction est seulement localement lipschitzienne, voir exemple (0.1.1).

4) Parfois, dans le cas où le système linéarisé présente un noeud et quand la non-linéarité n'est pas assez petite. Le problème est posé par le fait que dans ce cas en général, les courbes intégrales d'un champ de vecteurs quelconque, au voisinage d'un point singulier ne ressemblent pas toujours à celles du système qui lui est associé [17]. Pour voir plus clair, considérons l'exemple suivant [20].

### Exemple 0.1.3

$$\begin{cases} \frac{dx}{dt} = -x - \frac{2y}{\ln(x^2 + y^2)} \\ \frac{dy}{dt} = -y + \frac{2x}{\ln(x^2 + y^2)} \end{cases}, (x(0), y(0)) = (x_0, y_0)$$

sur le disque unité ouvert  $x^2 + y^2 < 1$ . On vérifie que l'origine est un point

singulier et le système linéaire associé s'écrit

$$\begin{cases} \frac{dx}{dt} = -x \\ \frac{dy}{dt} = -y \end{cases}, (x(0), y(0)) = (x_0, y_0).$$

Celui-ci présente un noeud stable ( $\lambda_1 = \lambda_2 = -1$ ), et l'espace propre est de dimension 2, alors que l'analyse des courbes intégrales du système non-linéaire montre qu'il présente un foyer stable.

Tous les problèmes que nous venons d'évoquer justifient amplement à nos yeux l'introduction et l'utilisation d'autres techniques, où méthodes de calculs pour l'étude du comportement des solutions ainsi que la synthèse de la stabilité des équations différentielles ordinaires non-linéaires.

Une méthode de linéarisation dite "optimale" basée sur le principe de moindres contraintes de Gauss a été introduite par Vujanovic, initialement pour l'étude des vibrations non-linéaires [11]. Elle a été appliquée à l'étude de la non-linéarité dans le cas de transfert de chaleur [7]. Récemment, elle a été utilisée par Jordan et al. [14] à l'analyse de systèmes non-linéaires régissant les régimes transitoires de circuits électroniques ainsi que la linéarisation d'équations d'état [15].

Le but recherché était d'améliorer la linéarisation, par rapport à celle donnée par la dérivée de Fréchet au point d'équilibre, en proposant l'application de la linéarisation optimale autour de la solution de l'équation non-linéaire, le principe étant de minimiser l'écart au sens des moindres carrés entre l'équation non-linéaire et l'équation linéaire, c'est à dire la quantité

$$\varepsilon(t) = F(x(t)) - Ax(t)$$

le long de la solution de l'équation  $A = DF(0)$ .

### 0.1.3 Contribution.

*Le point de départ réel de cette thèse, a coïncidé avec une visite de O. Arino à l'Institut de Sciences Exactes de l'Université de Tlemcen. Le travail a été effectué à partir de mon premier séjour au Laboratoire de Mathématiques Appliquées, C.N.R.S URA 1204 de l'Université de Pau et des Pays de l'Adour.*

Le travail présenté ici se veut une contribution à l'analyse des systèmes non-linéaires et, plus précisément les équations différentielles ordinaires non-linéaires, contribution sur au moins deux aspects [20], [22]:

1) Un aspect théorique. Nous associons une matrice optimale  $\hat{A}$  définissant une application linéaire à une équation différentielle ordinaire non-linéaire. Le système linéaire obtenu est une sorte de valeur moyenne des dérivées de la fonction non-linéaire en 0 le long des trajectoires partant de  $x_0$  et allant à l'origine, on l'appellera "dérivée optimale". Celle-ci sera vue comme une alternative à la dérivée au sens de Fréchet, indispensable dans le cas d'équations comportant des fonctions non régulières et en général non dérivables.

2) Un deuxième aspect d'analyse numérique. Nous construisons une procédure numérique permettant la mise en oeuvre de l'aspect théorique, c'est à dire le développement d'une méthode.

En fait, le travail se compose de deux parties. La première partie comporte l'aspect théorique du problème posé, sa résolution ainsi que ses propriétés. Dans la seconde, on expose un autre type d'approximation basée sur la notion de "dérivée optimale" introduite dans la première, ainsi que l'aspect numérique.

**Première partie:** Chapitre 1, Chapitre 2, Chapitre 3.

Concernant l'aspect théorique: dans le chapitre 1, on définit un schéma d'approximation basé sur le principe des moindres carrés, qui va nous permettre d'associer une application linéaire à une équation différentielle ordinaire non-linéaire. Pour

cela, on considère la fonctionnelle suivante

$$G(A) = \int_0^{+\infty} \|\varepsilon(t)\|^2 dt.$$

La minimisation de celle-ci par rapport à une solution donnée nous permet de construire une procédure itérative permettant de déterminer, sous certaines conditions, une application linéaire définie comme la dérivée optimale de l'équation différentielle ordinaire non-linéaire (0.1) qui démarre avec la linéarisation par dérivation au sens de Fréchet  $DF(x)$  calculée en  $x_0$ , si celle-ci existe en  $x_0$ . Ainsi, on construit un schéma itératif qui comporte différents aspects et en particulier, le problème d'existence d'unicité et de convergence de l'application linéaire générée par la procédure à chaque pas. Dans les chapitres 2 et 3, des études ont été développées [19], [21] dans ce sens sur des classes de fonctions avec comme objectif la justification et l'approfondissement de la méthode. L'approximation que nous présentons possède certaines particularités, comme par exemple dans le cas scalaire, la convergence au premier pas vers la formule [20], [23]

$$\tilde{a}(x_0) = \frac{2}{x_0^2} \int_0^{x_0} f(z) dz \quad (0.2)$$

où  $f$  est la fonction définissant l'équation différentielle scalaire non-linéaire. On montre que cette formule converge vers la dérivée de  $f$ ,  $f'$  au point 0 quand la donnée initiale  $x_0 \rightarrow 0$ . Cette formule reste vraie même si la dérivée de  $f$  n'existe pas en 0.

Dans le chapitre 3, en supposant les conditions donnant a priori, l'existence, l'unicité et la convergence réunies, une étude des propriétés par rapport à l'équation non-linéaire a été faite [20], [22], concernant particulièrement l'ordre de l'approximation et son lien avec la stabilité asymptotique. Ainsi, concernant l'ordre, on a montré que l'approximation proposée est d'ordre 2 au moins en fonction de la donnée

initiale

$$\|x(t) - \tilde{y}(t)\| \leq o(\|x_0\|^2)$$

et en général dépend de la non-linéarité.  $x(t)$  représente la solution de l'équation non-linéaire et  $\tilde{y}(t)$  celle donnée par la dérivée optimale.

Concernant la stabilité, il a été prouvé que dans le cas scalaire, la dérivée optimale (0.2), permet la construction d'une fonction de Lyapunov pour l'équation non-linéaire. Dans le cas vectoriel, une étude détaillée sur le lien entre la dérivée optimale et la stabilité asymptotique d'une équation non-linéaire dont la linéarisée classique au point d'équilibre présente un centre a été proposée [21].

**Deuxième partie:** Chapitre 4, Chapitre 5.

Dans le chapitre 4, on propose un schéma d'approximation [25] utilisant la notion de "dérivée optimale" introduite dans la première partie. On considère un intervalle  $[0, T]$ . On le subdivise en une réunion d'intervalles  $[t_i, t_{i+1}]$  et on applique la dérivée optimale. On élabore ainsi une méthode d'approximation optimale permettant la résolution numérique du problème (0.1). Ceci permet de calculer une approximation  $x_{i+1}$  de  $x(t_{i+1})$  à partir de  $x_i$  une approximation de  $x(t_i)$ .

Enfin, la possibilité de savoir construire des méthodes de résolution numérique de ces problèmes qui soient à la fois simples et efficaces est un atout très important. Ce sera l'objet du dernier chapitre. Plusieurs exemples [20], [22], [24], ainsi qu'une comparaison avec les méthodes existantes seront présentés. Précisons que par la dérivée optimale, on obtient une approximation globale de la solution et par l'approximation optimale, on obtient une approximation locale de la solution en chaque point de la subdivision choisie. Dans les deux cas, la résolution numérique des équations se fait en utilisant la méthode de Runge-Kutta d'ordre 4.

De ce point de vue, la comparaison qu'on propose se fait entre l'équation non-linéaire et l'équation linéaire optimale, par la capacité de la linéarisation obtenue à remplacer l'équation initiale. On propose pour certains exemples un calcul d'erreur relative à chaque instant de l'intervalle choisie pour la résolution, entre

la solution du problème non-linéaire et son approximation.

Plus l'erreur relative est petite, meilleure est l'approximation. L'ordre de la dérivée optimale étant fonction de la donnée initiale, on remarque que tant que  $\|x_0\|$  est grand dans un certain sens, l'erreur relative est petite, et par conséquent, l'approximation est bonne. Les choses commencent à changer quand on approche de 0.

Enfin, les résultats numériques de la méthode d'approximation optimale introduite au chapitre 4 montrent que celle-ci est comparable aux méthodes déjà existantes (Euler, Runge-Kutta d'ordre 4).

Les travaux présentés ici ont été exposés dans plusieurs articles: un article "*Determination of the stability of a nonlinear ordinary differential equation by least square approximation. Computational procedure*" a été publié dans "Applied Mathematics and Computers Sciences", un deuxième "*Least square approximation of a nonlinear ordinary differential equation*" dans "Computers and Mathematics With Applications"; des résultats préliminaires ont fait l'objet de communications dans des conférences internationales: ainsi, un article est à paraître dans les "*Proceeding of the Fourth International Colloquium on Numerical Analysis*", Plovdiv, Bulgaria, un autre a été accepté pour publication dans les "*Proceeding of the 2nd International Conférence on Differential Equations in Marrakech*". La méthode d'approximation numérique a donné lieu à la rédaction d'un article "*Optimal approximation of the initial value problem*" actuellement soumis.

**PREMIER CHAPITRE  
PRESENTATION DE LA  
DERIVEE OPTIMALE.**



# Chapitre 1

## Présentation de la dérivée optimale.

### 1.1 Introduction.

Dans ce chapitre, nous allons introduire une certaine notion d'approximation d'une équation différentielle ordinaire non linéaire. Il s'agit en fait de déterminer une sorte d'approximation globale [20], [22], par opposition à la perturbation non-linéaire d'une équation linéaire, qui se distingue de l'approximation linéaire classique au voisinage d'un point stationnaire. L'approche suivie ici est de type optimisation.

### 1.2 Problème posé.

Considérons l'équation différentielle ordinaire non-linéaire de la forme

$$\begin{cases} \frac{dx}{dt} = F(x) \\ x(0) = x_0 \end{cases} \quad (1.1)$$

où:

$x \in \mathbb{R}^n$ ,  $F$  est définie dans un certain ouvert  $\Omega$ , à valeurs dans  $\mathbb{R}^n$ ,  
avec les hypothèses suivantes

H1)  $F(0) = 0$

H2) Le spectre  $\sigma(DF(x))$  est contenu dans l'ensemble  $\{z : \operatorname{Re} z < 0\}$  pour tout  $x \neq 0$ , dans un voisinage de 0, où  $DF(x)$  existe.

H3)  $F$  est continue, lipschitzienne, de constante de Lipschitz  $\gamma$ .

Le problème que nous nous posons peut-être traduit comme suit.

Trouver une équation différentielle ordinaire linéaire de la forme

$$\begin{cases} \frac{dx}{dt} = \tilde{A}x \\ x(0) = x_0 \end{cases} \quad (1.2)$$

approchant l'équation non-linéaire (1.1) aux mêmes conditions initiales telle que la fonctionnelle

$$G(A) = \int_0^{+\infty} \|F(x(t)) - Ax(t)\|^2 dt \quad (1.3)$$

soit minimale.  $F(x)$  étant comme au dessus, et  $\tilde{A} \in \mathcal{M}_n(\mathbb{R})$ , à déterminer.

Pour le moment,  $x$  est juste une fonction définie sur  $[0, +\infty[$ , bornée, continue et telle que  $x \in L^1(0, +\infty)$  et  $F(x(\cdot)) \in L^1(0, +\infty)$ . Par la suite, on introduira des Fonctions  $x(t)$  qui sont solutions de problèmes linéaires.

Le problème ainsi posé est de type optimisation au sens des moindres carrés, le but étant de chercher à remplacer l'équation initiale non-linéaire par une équation linéaire. Cela veut dire, chercher à approximer la solution du système (1.1) par celle de (1.2). Ainsi, la minimisation de la fonctionnelle  $G(A)$  par rapport à  $A$  se fait par rapport à des solutions issues de la valeur initiale  $x_0$  et tendant vers 0 quand  $t \rightarrow \infty$ .

**Remarque 1.2.1** Notons que l'avantage déterminant de l'approximation au sens des moindres carrés est son caractère linéaire.

## 1.3 Résolution du problème.

### 1.3.1 Rappels.

Nous allons brièvement rappeler des théorèmes<sup>1</sup> garantissant l'existence et l'unicité de la meilleure approximation au sens des moindres carrés. Ceci va permettre la résolution du problème posé dans la section (1.2).

On choisit dans  $\mathbb{R}^n$  la structure euclidienne canonique, on note  $\langle \cdot, \cdot \rangle$  le produit scalaire correspondant et  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  la norme associée.

Pour tout  $f$  donnée dans un espace vectoriel  $V$ , le théorème suivant nous assure l'existence dans un sous espace vectoriel  $E$  de  $V$ , de dimension finie, d'un élément  $g$  meilleure approximation de  $f$ .

**Théorème 1.3.1** [16] *Si  $E$  est un sous espace vectoriel de dimension finie, alors il existe au moins un élément  $g$  de  $E$  tel que*

$$\|f - g\| = \min_{h \in E} \|f - h\|$$

Ce théorème garantit en général l'existence d'une approximation.

**Théorème 1.3.2** [16] *Une condition nécessaire et suffisante pour que  $g$  soit une meilleure approximation de  $f$  est que*

$$\langle f - g, h \rangle = 0 \text{ pour tout } h \in E.$$

Cet élément  $g$  est appelé meilleure approximation de  $f$  au sens des moindres carrés.

**Théorème 1.3.3** [16] *L'approximation au sens des moindres carrés est unique.*

---

<sup>1</sup>La démonstration de ces théorèmes se trouve dans la référence [16].

### 1.3.2 Résolution.

La minimisation de  $G(A)$  est obtenue en calculant la différentielle par rapport à  $A$ . Donc, en différentiant (1.3) par rapport à  $A$  le long d'une fonction  $x$ , on obtient

$$DG(A) \alpha = 2 \int_0^{+\infty} \langle Ax(t) - F(x(t)), \alpha x(t) \rangle dt \quad (1.4)$$

pour toute matrice  $\alpha$ . En particulier, pour les matrices  $\alpha$  telles que  $\alpha_{l,m} = 1$ ;  $\alpha_{i,j} = 0$ , si  $(i,j) \neq (l,m)$ , nous avons

$$\int_0^{+\infty} \langle Ax(t) - F(x(t)), \alpha x(t) \rangle dt = \int_0^{+\infty} [Ax(t) - F(x(t))]_l x_m(t) dt. \quad (1.5)$$

Supposons que  $A$  minimise (1.3) le long d'une fonction donnée  $x$ . La relation (1.5) est égale à zéro, et

$$\int_0^{+\infty} [Ax(t) - F(x(t))]_l x_m(t) dt = 0, \forall 1 \leq l, m \leq n. \quad (1.6)$$

En notant  $(a_{i,j})$  le terme général de la matrice  $A$ , on obtient

$$\sum_{j=1}^n a_{l,j} \left( \int_0^{+\infty} x_j(t) x_m(t) dt \right)_{1 \leq j, m \leq n} = \left( \int_0^{+\infty} f_l(x(t)) x_m(t) dt \right)_{1 \leq l, m \leq n}. \quad (1.7)$$

En introduisant la fonction  $\Gamma$  définie par

$$\Gamma(x) = \int_0^{+\infty} [x(t)][x(t)]^T dt = \left( \int_0^{+\infty} x_j(t) x_m(t) dt \right)_{1 \leq j, m \leq n}, \quad (1.8)$$

on obtient

$$A = \left[ \int_0^{+\infty} [F'(x(t))][x(t)]^T dt \right] [\Gamma(x)]^{-1}. \quad (1.9)$$

$A$  est définie et unique à condition que

$$\Gamma(x) = \left[ \int_0^{+\infty} [x(t)][x(t)]^T dt \right] \quad (1.10)$$

soit inversible.

Cette condition sera vérifiée si la trajectoire de  $x(t)$  n'est contenue dans aucun sous espace strict de  $\mathbb{R}^n$ . Nous ne traiterons pas le cas où la fonction  $x$  serait à valeurs dans un sous espace strict de  $\mathbb{R}^n$ . Indiquons toutefois que dans ce cas, on peut se ramener à l'hypothèse précédente en ajoutant à  $x$  une perturbation de la forme

$$\varepsilon \begin{pmatrix} e^{-at} \\ \cdot \\ e^{-nat} \end{pmatrix}, \quad (1.11)$$

avec  $\varepsilon > 0$  assez petit, et  $a > 0$  choisi dans un ensemble générique de valeurs.

**Remarque 1.3.4** Notons que  $\Gamma(x)$  est une matrice symétrique. Elle est positive car

$$v^T \Gamma(x) v = \int_0^{+\infty} (\langle x(t), v \rangle)^2 dt \geq 0 \quad (1.12)$$

$\forall v \in \mathbb{R}^n$ .

Une condition nécessaire et suffisante pour que  $\Gamma$  soit inversible est donnée dans ce qui suit.

### 1.3.3 Condition nécessaire et suffisante pour que $\Gamma(x)$ soit inversible.

**Lemme 1.3.5** [22] La matrice  $\Gamma(x)$  est inversible si et seulement si l'espace engendré par  $x(\mathbb{R}^+)$  est dense dans  $\mathbb{R}^n$ .

**Preuve:** D'après la relation (1.12),  $\Gamma(x)$  étant une matrice symétrique positive, une condition nécessaire et suffisante pour que  $\Gamma(x)$  soit inversible est qu'elle soit définie positive, c'est à dire

$$v^T \Gamma(x) v > 0, \text{ pour tout } v \in \mathbb{R}^n, v \neq 0, \quad (1.13)$$

c'est à dire

$$v^T \Gamma(x) v = \int_0^{+\infty} (\langle x(t), v \rangle)^2 dt > 0. \quad (1.14)$$

Cela est vrai si et seulement si  $\forall v \neq 0, \langle x(\cdot), v \rangle \neq 0$ , c'est à dire, si et seulement si l'espace engendré par  $x(\mathbb{R}^+)$  est dense dans  $\mathbb{R}^n$ .

En effet, si  $\{x(\mathbb{R}^+)\}$  est dense dans  $\mathbb{R}^n$ , alors

pour tout  $v \in \mathbb{R}^n, v \neq 0$ , il existe  $t \in \mathbb{R}^+$  tel que  $\langle x(\cdot), v \rangle \neq 0$ .

Sinon,  $\{x(\mathbb{R}^+)\}$  serait contenu dans  $\{v\}^\perp$ , et comme cet ensemble est fermé, il contient aussi la fermeture de  $\{x(\mathbb{R}^+)\}$ , ce qui contredit la densité.

Maintenant, si  $\{x(\mathbb{R}^+)\}$  n'est pas dense, alors

il existe un vecteur  $v \neq 0, v \in \{x(\mathbb{R}^+)\}^\perp$  tel que  $\langle x(\cdot), v \rangle = 0, \forall t \geq 0$ , en contradiction avec l'hypothèse.  $\square$

Supposons maintenant que  $x$  soit la solution de l'équation linéaire

$$\begin{cases} \frac{dx}{dt} = A_0 x \\ x(0) = x_0 \end{cases} \quad (1.15)$$

où  $A_0 = DF(x_0)$ .

De l'hypothèse H2) dans la section 1.2, nous trouvons que  $x(t)$  tend vers zéro exponentiellement, quand  $t$  tend vers  $+\infty$ , et par conséquent, la fonction  $\Gamma(x)$  est bien définie. Dans ce cas, le lemme (1.3.5) donne la condition suivante pour que  $\Gamma(x)$  soit inversible.

**Lemme 1.3.6** [22] *Supposons que  $x(t) = \exp(tA_0)x_0$ , où  $A_0$  satisfait H2) dans la section 1.2. Alors, une condition nécessaire et suffisante pour que la matrice  $\Gamma(x)$  soit inversible est que le rang de la famille  $x_0, A_0x_0, \dots, A_0^{n-1}x_0$  soit égal à  $n$ ,*

$$\text{Rg} [x_0, A_0x_0, \dots, A_0^{n-1}x_0] = n. \quad (1.16)$$

**Preuve:** D'après le théorème de Cayley-Hamilton, les itérés de  $A_0$  sont dans le sous-espace vectoriel de l'ensemble des matrices  $M_n(\mathbb{R})$  engendré par la famille

$\{Id, A_0, \dots, A_0^{n-1}\}$ . On a donc  $e^{tA_0}x_0 \in \{x_0, A_0x_0, \dots, A_0^{n-1}x_0\}$ , et donc

$$\text{adh} \{e^{tA_0}x_0 : t \geq 0\} = \mathbb{R}^n \subset \{x_0, A_0x_0, \dots, A_0^{n-1}x_0\},$$

d'où  $\text{Rg} [x_0, A_0x_0, \dots, A_0^{n-1}x_0] = n$ .

Maintenant, si  $\text{Rg} [x_0, A_0x_0, \dots, A_0^{n-1}x_0] = n$ , on construit alors, l'ensemble

$$\left\{ x_0, \frac{d}{dt} (e^{tA_0}x_0) \Big|_{t=0}, \dots, \frac{d^{n-1}}{dt} (e^{tA_0}x_0) \Big|_{t=0} \right\} = \mathbb{R}^n.$$

On pose  $y_j = \frac{d^j}{dt} (e^{tA_0}x_0) \Big|_{t=0}$ . En utilisant la formule de Taylor pour  $t > 0$  petit,

on a

$$e^{tA_0}x_0 = \sum_{j=0}^{n-1} \frac{t^j}{j!} y_j + o(t^{n-1})$$

avec  $\{y_j\}_{0 \leq j \leq n-1} = \mathbb{R}^n$ .

On peut alors choisir  $n$  valeurs de  $t$ ,  $t_0, \dots, t_{n-1}$  assez proches de zéro pour que

$$[e^{t_0A_0}x_0, \dots, e^{t_{n-1}A_0}x_0] = [y_0, \dots, y_{n-1}] \mathcal{M}(t_0, \dots, t_{n-1}) + o(|t|^{n-1})$$

$|t| = \max_{j=0}^{n-1} |t_j|$ , et où  $\mathcal{M}(t_0, \dots, t_{n-1})$  est une matrice de Vandermonde inversible.

Donc

$$\{e^{t_0A_0}x_0, \dots, e^{t_{n-1}A_0}x_0\} = \mathbb{R}^n,$$

et donc la famille  $\{e^{tA_0}x_0, t \geq 0\}$  est dense.  $\square$

**Remarque 1.3.7** 1)-Bien évidemment, dans le cas de dimension finie, dire que l'espace vectoriel engendré par la famille  $\{e^{tA_0}x_0, t \geq 0\}$ , est dense revient à dire que cet espace est l'espace tout entier, autrement dit, qu'il existe  $n$  valeurs de  $t$  telles que les  $n$  vecteurs correspondants forment une famille génératrice.

2)-La condition du Lemme 1.3.6 exclut notamment la possibilité pour  $x_0$  d'être un vecteur propre de  $A_0$ , ou d'appartenir à un espace invariant de dimension inférieur à  $n$ . Dans le cas où  $A_0$  a seulement des valeurs propres simples, le

résultat est vrai si et seulement si  $x_0$  a des composantes non nulles sur tous les espaces propres de  $F$ .

## 1.4 Algorithme.

On va utiliser le calcul précédent de manière itérative. On supposera que les matrices successives  $A_j$  sont de type stable, leur spectre est contenu dans  $\{z : \operatorname{Re} z < 0\}$ . La matrice initiale est la matrice jacobienne de  $F$  en  $x_0$  qui en principe est un point arbitraire pris dans un voisinage de 0 et tel que  $F$  soit différentiable en  $x_0$ . Puisque  $F$  est lipschitzienne, on sait qu'elle est presque partout différentiable [12].

Considérons le système (1.1)

$$\begin{cases} \frac{dx}{dt} = F(x) \\ x(0) = x_0 \end{cases}$$

**1ère étape:**

Calcul de  $A_0 = DF(x_0)$ .

**2ème étape:**

Calcul de  $A_1$  à partir de la solution de l'équation

$$\begin{cases} \frac{dy}{dt} = A_0 y \\ y(0) = x_0 \end{cases} \quad (1.17)$$

en minimisant la fonctionnelle

$$G(A) = \int_0^{+\infty} \|F(y(t)) - Ay(t)\|^2 dt. \quad (1.18)$$

$A_1$  est déterminée de manière unique par la formule (1.9), où  $x$  est solution de l'eq.(1.17). On voit ainsi que mis à part la matrice initiale, les matrices



déterminées par la procédure ne sont pas les matrices jacobienes de  $F$  en un point donné. Par la suite, il est nécessaire que les conditions du début de cette section soient satisfaites à chaque pas.

Si nous supposons que cela est vrai, alors, la procédure marche comme suit:

**3ème étape:**

Pour calculer  $A_j$  à partir de  $A_{j-1}$ , on doit d'abord résoudre

$$\begin{cases} \frac{dy}{dt} = [A_{j-1}]y \\ y(0) = x_0. \end{cases} \quad (1.19)$$

On notera  $y_j(t)$  la solution de l'équation (1.19).

La minimisation de la fonctionnelle

$$G_j(A) = \int_0^{+\infty} \|F(y_j(t)) - Ay_j(t)\|^2 dt \quad (1.20)$$

conduit à  $A_j$ .

La relation entre  $A_j$  et  $A_{j-1}$  s'écrit, alors

$$A_j \Gamma(y_j) = \int_0^{+\infty} [F(y_j)] [y_j]^T dt. \quad (1.21)$$

En supposant encore que  $\Gamma(y_j)$  est inversible,  $A_j$  s'écrit

$$A_j = \left[ \int_0^{+\infty} [F(y_j)] [y_j]^T dt \right] [\Gamma(y_j)]^{-1}. \quad (1.22)$$

**Définition 1.4.1** *Si la suite  $A_j$  converge, alors la limite  $\hat{A}$  est appelée la dérivée optimale de  $F$  en  $x_0$ .*

## 1.5 Propriétés de la procédure.

Nous considérons maintenant la situation où la procédure converge.

### 1.5.1 Cas où l'application est linéaire.

Si  $F$  est linéaire et le spectre  $\sigma(F)$  est contenu dans le plan complexe à partie réelle strictement négative, alors la procédure donne  $F$  à la première itération.

Dans ce cas l'éq.(1.7) s'écrit

$$A\Gamma(x) = F\Gamma(x). \quad (1.23)$$

Il est clair que  $A = F$  est une solution. Elle est unique si  $\Gamma(x)$  est inversible. Ceci est vrai si et seulement si la condition (1.16) est satisfaite. Cette solution est obtenue à la première itération, avec  $A_0 = F$ .

On peut donc affirmer que la dérivée optimale d'un système linéaire est le système lui même.

### 1.5.2 Cas général, où le système est la somme d'un terme linéaire et d'un terme non-linéaire.

Considérons un système général d'équations non-linéaires avec une non-linéarité de la forme

$$F(x) = Mx + F^*(x), x(0) = x_0 \quad (1.24)$$

où  $M$  est linéaire.

Le calcul de la matrice  $A_1$  donne

$$A_1 = \left[ \int_0^{+\infty} [F^*(x(t))] [x(t)]^T dt \right] [\Gamma(x)]^{-1}, \quad (1.25)$$

qui peut s'écrire comme suit

$$A_1 = \left[ M\Gamma(x) + \left( \int_0^{+\infty} [F^*(x(t))] [x(t)]^T dt \right) \right] [\Gamma(x)]^{-1} \quad (1.26)$$

et finalement

$$A_1 = M + \left[ \int_0^{+\infty} [F^*(x(t))] [x(t)]^T dt \right] [\Gamma(x)]^{-1}. \quad (1.27)$$

Donc,  $A_1 = M + A_1^*$  avec

$$A_1^* = \left[ \int_0^{+\infty} [F^*(x(t))] [x(t)]^T dt \right] [\Gamma(x)]^{-1}. \quad (1.28)$$

Alors, pour tout  $j$  nous avons

$$A_j = M + A_j^* \quad (1.29)$$

avec

$$A_j^* = \left[ \int_0^{+\infty} [F^*(x_j(t))] [x_j(t)]^T dt \right] [\Gamma(x_j(t))]^{-1}. \quad (1.30)$$

Si en particulier, certaines composantes de  $F$  sont linéaires, alors les composantes correspondantes de  $F^*$  sont égales à zéro, et les composantes correspondantes de  $A_j$  sont celles de  $F$ .

Si  $f_k$  est linéaire, alors la  $k$ ième ligne de la matrice  $A_j$  est égale à  $f_k$ .

## 1.6 Cas scalaire.

### 1.6.1 Introduction.

Considérons l'équation non-linéaire suivante

$$\frac{dx}{dt} = f(x), x(0) = x_0 \quad (1.31)$$

$f : \mathbb{R} \rightarrow \mathbb{R}$ , on suppose que:

H1)  $f(0) = 0$ ,

H2)  $f'(x) < 0$  en tout point où  $f'$  existe, dans un intervalle  $]-\alpha, +\alpha[$ ,  $\alpha > 0$ .

H3)  $f$  est absolument continue par rapport à la mesure de Lebesgue,  
 On veut approcher l'équation différentielle ordinaire non-linéaire (1.31), par une  
 équation linéaire de la forme

$$\begin{cases} \frac{dx}{dt} = \tilde{a}x \\ x(0) = x_0 \end{cases} \quad (1.32)$$

en minimisant la fonctionnelle

$$G(a) = \int_0^{+\infty} |f(x(t)) - ax(t)|^2 dt. \quad (1.33)$$

Ce problème est résolu en utilisant la procédure exposée dans la section 1.4.

### 1.6.2 Calcul de la dérivée optimale.

On choisit  $x_0 \in ]-\alpha, +\alpha[$ ,  $x_0 \neq 0$ , tel que  $f'(x_0)$  existe, on pose  $a_0 = f'(x_0)$ , et  
 on résoud l'équation linéaire

$$\begin{cases} \frac{dx}{dt} = a_0 x \\ x(0) = x_0 \end{cases} \quad (1.34)$$

qui donne

$$x(t) = \exp(a_0 t) x_0. \quad (1.35)$$

En substituant, dans l'expression (1.9)  $f$  à  $F$ , on obtient

$$a_1 = \frac{\left( \int_0^{+\infty} f(x(t)) e^{a_0 t} dt \right)}{\left( \int_0^{+\infty} e^{2a_0 t} dt \right)} \frac{1}{x_0}. \quad (1.36)$$

On vérifie que, avec  $x_0 \neq 0$ ,  $f(x(t))$  est presque partout dérivable [12] et

$$\frac{df(x(t))}{dt} = f'(x(t)) e^{a_0 t} x_0 a_0, \quad (1.37)$$

d'où

$$\frac{1}{a_0} \int_0^{+\infty} f(x(t)) e^{a_0 t} dt = \frac{1}{a_0} [f(x(t)) e^{a_0 t}]_0^{+\infty} - \frac{1}{a_0} \int_0^{+\infty} (f'(x(t)) e^{2a_0 t} dt) x_0 a_0. \quad (1.38)$$

Enfin  $a_1$  s'écrit

$$a_1 = 2 \left( \frac{f(x_0)}{x_0} + a_0 \int_0^{+\infty} f'(x(t)) e^{2a_0 t} dt \right). \quad (1.39)$$

En changeant la variable  $t$  en  $x(t)$  dans l'intégrale et après calculs, on obtient

$$a_1 = \frac{2}{x_0^2} \int_0^{x_0} f(z) dz. \quad (1.40)$$

La formule donnant  $a_1$  montre que celle-ci ne dépend pas de  $a_0$ . En répétant la procédure comme indiquée au dessus, on obtient le même résultat. Dans ce cas, la procédure donne la dérivée optimale au premier pas *i.e.*

$$\tilde{a}(x_0) = \frac{2}{x_0^2} \int_0^{x_0} f(z) dz. \quad (1.41)$$

### 1.6.3 Autre expression pour la dérivée optimale.

Si nous supposons  $f$  analytique en 0

$$f(x) = \sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{n!} x^n, \quad (1.42)$$

nous pouvons donner une autre expression pour  $\tilde{a}(x_0)$ .

En utilisant la relation (1.41), et en remplaçant,  $f(z)$  par l'expression donnée par la relation (1.42), il vient que

$$\tilde{a}(x_0) = \frac{2}{x_0^2} \int_0^{x_0} \sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{n!} z^n dz = \frac{2}{x_0^2} \sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{n!} \int_0^{x_0} z^n dz, \quad (1.43)$$

et

$$\tilde{a}(x_0) = 2 \sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{(n+1)!} x_0^{n-1}. \quad (1.44)$$

Enfin

$$\tilde{a}(x_0) = f'(0) + \frac{1}{3} x_0 f''(0) + \dots + \frac{2}{(n+1)!} x_0^{n-1} f^{(n)}(0) + \dots, \quad (1.45)$$

formule valable dans l'intervalle de convergence de la série de Taylor en 0. Plus généralement, si  $f$  est de classe  $C^k$  dans un voisinage de 0,  $f(0) = 0$ , avec  $k \geq 1$ , alors  $\tilde{a}$  est de classe  $C^{k-1}$  dans ce voisinage et on a

$$\tilde{a}^{(j)}(0) = \frac{2}{(j+1)!} x_0^{j-1} f^{(j)}(0); 0 \leq j \leq k-1.$$

#### 1.6.4 Propriétés.

Nous allons donner deux propriétés intéressantes de la dérivée optimale dans le cas scalaire. La première est que la dérivée optimale converge vers la dérivée de la fonction  $f$  en 0 quand  $x_0 \rightarrow 0$ . La seconde est que dans certains cas la dérivée optimale converge quand  $x_0 \rightarrow 0$ , même si la dérivée de  $f$  en 0 n'existe pas. D'où, l'importance du résultat, c'est à dire, la possibilité d'utiliser  $\tilde{a}(x_0)$  pour la description du comportement de la solution au voisinage de 0 quand la dérivée en ce point n'existe pas.

**Lemme 1.6.1** *Si  $f$  est continue, et si la dérivée de  $f$  en 0 existe, alors*

$$\lim_{x_0 \rightarrow 0} \tilde{a}(x_0) = f'(0).$$

**Preuve:** Avec  $f(z) = z f'(0) + z \varepsilon(z)$ , l'éq.(1.41) peut s'écrire

$$\tilde{a}(x_0) = f'(0) + \frac{2}{x_0^2} \int_0^{x_0} z \varepsilon(z) dz. \quad (1.46)$$

le second terme de l'éq.(1.46), satisfait l'inégalité

$$\left| \frac{2}{x_0^2} \int_0^{x_0} z \varepsilon(z) dz \right| \leq |\varepsilon(x_0)| \frac{2}{x_0^2} \left| \int_0^{x_0} z dz \right| = |\varepsilon(x_0)| \left( \frac{2}{x_0^2} \right) \left( \frac{x_0^2}{2} \right) \quad (1.47)$$

et converge vers 0 quand  $x_0 \rightarrow 0$ . Donc  $\lim_{x_0 \rightarrow 0} \tilde{a}(x_0) = f'(0)$ .  $\square$

On peut voir que la dérivée optimale définie par l'éq.(1.41) dépend de la valeur initiale  $x_0$  et converge vers  $f'(0)$  quand  $x_0 \rightarrow 0$  si  $f'(0)$  existe.

### Cas où la dérivée de la fonction non-linéaire n'existe pas en 0.

Dans ce qui suit, on va voir qu'il est possible de trouver une limite même si la dérivée de  $f$  en 0 n'existe pas, comme nous allons le montrer pour une certaine classe de fonctions.

En écrivant  $f(z)$  sous la forme

$$f(z) = -zg(z) \quad (1.48)$$

dans la relation (1.41), il vient que

$$\tilde{a}(x_0) = -\frac{2}{x_0^2} \int_0^{x_0} zg(z) dz. \quad (1.49)$$

On choisit la fonction

$$g_r(z) = p(|\ln z|^r) \quad (1.50)$$

où  $p$  est une fonction périodique de période 1, bornée,  $p \geq 0$ ; et  $\bar{p} = \int_0^1 p(z) dz > 0$ .

La relation (1.49) s'écrit pour  $r = 1$  et  $0 < x_0 < 1$

$$\tilde{a}(x_0) = -\frac{2}{x_0^2} \int_0^{x_0} zp(|\ln z|) dz. \quad (1.51)$$

Pour tout  $\alpha$ ,  $0 < \alpha < 1$ , on a

$$\begin{aligned}\tilde{a}(\alpha x_0) &= -\frac{2}{\alpha^2 x_0^2} \int_0^{\alpha x_0} z p(-\ln z) dz \\ &= -\frac{2}{\alpha^2 x_0^2} \int_0^{x_0} \alpha^2 z p(-\ln \alpha - \ln z) dz \\ &= -\frac{2}{x_0^2} \int_0^{x_0} z p(-\ln \alpha - \ln z) dz.\end{aligned}\tag{1.52}$$

Si en particulier, on a  $\ln \alpha = -1 \implies \alpha = e^{-1}$ , alors  $\tilde{a}\left(\frac{1}{e}x_0\right) = \tilde{a}(x_0)$ .

Dans ce cas,  $\tilde{a}(x_0)$  n'a pas de limite quand  $x_0 \rightarrow 0^+$ .

Dans le cas où  $r > 1$ ,

$$\tilde{a}_r(x_0) = -2 \int_0^1 z p\{(-\ln x_0 - \ln z)^r\} dz.\tag{1.53}$$

On note que si  $\theta$  est une fonction  $C^\infty$ , à support compact, alors  $\theta * p$  est  $C^\infty$ , 1-périodique. Si  $\theta_\varepsilon(x) = \frac{1}{\varepsilon} \theta\left(\frac{x}{\varepsilon}\right)$ , alors  $\theta_\varepsilon * p \rightarrow p$ , uniformément sur  $\mathbb{R}$ . Si  $p_1$  et  $p_2$  sont deux fonctions telles que

$$\|p_1 - p_2\|_\infty \leq \delta\tag{1.54}$$

et  $\tilde{a}_r(p_1)(x_0)$ ,  $\tilde{a}_r(p_2)(x_0)$  sont les fonctions associées respectivement à  $p_1, p_2$ , alors

$$\|\tilde{a}_r(p_2)(x_0) - \tilde{a}_r(p_1)(x_0)\| \leq \delta.\tag{1.55}$$

En effet, d'après la relation (1.53), on a

$$\begin{aligned}\|\tilde{a}_r(p_2)(x_0) - \tilde{a}_r(p_1)(x_0)\| &\leq \left\| 2 \int_0^1 z (p_2(-\ln x_0 - \ln z) - p_1(-\ln x_0 - \ln z)) dz \right\| \\ &\leq 2 \int_0^1 z \|p_2(-\ln x_0 - \ln z) - p_1(-\ln x_0 - \ln z)\|_\infty dz\end{aligned}\tag{1.56}$$

et puisque  $\|p_1 - p_2\|_\infty \leq \delta$ , donc

$$\|\tilde{a}_r(p_2)(x_0) - \tilde{a}_r(p_1)(x_0)\| \leq \delta.$$



Soit  $(p_j)$  une famille de fonction 1-périodique telle que  $p_j$  converge uniformément sur  $\mathbb{R}$  vers une fonction  $p$ . Si pour chaque  $j$ ,  $\lim_{x_0 \rightarrow 0} \tilde{a}_r(p_j)(x_0)$  existe, alors, il en est de même pour  $\tilde{a}_r(p)$  et de plus on a

$$\lim_{x_0 \rightarrow 0} \tilde{a}_r(p)(x_0) = \lim_{j \rightarrow +\infty} \left( \lim_{x_0 \rightarrow 0} \tilde{a}_r(p_j)(x_0) \right). \quad (1.57)$$

Considérons maintenant la relation

$$\tilde{a}_r(x_0) = -\frac{2}{x_0^2} \int_0^{x_0} u g_r(u) du \quad (1.58)$$

où  $g_r(u) = p(|\ln u|^r)$ , on note que  $g_r(u)$  n'est pas dérivable en 0.

En faisant le changement de variable  $v = |\ln u|^r$ , on obtient

$$\tilde{a}_r(x_0) = -\frac{2}{rx_0^2} \int_{+\infty}^{|\ln x_0|^r} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv. \quad (1.59)$$

On commence par remplacer  $x_0$  par une suite particulière de réels tendant vers 0, la suite:  $e^{-k^{\frac{1}{r}}}$ , avec  $k \in \mathbb{N}, k \rightarrow \infty$ . Pour cela, on note que si  $x_0$  est assez petit, il existe un entier  $k$  unique tel que

$$k < |\ln x_0|^r < k+1. \quad (1.60)$$

Après calculs, on trouve qu'il existe une constante  $C \geq 0$  indépendante de  $x_0$  telle que

$$e^{-k^{\frac{1}{r}} - Ck^{\frac{1}{r}-1}} < x_0 < e^{-k^{\frac{1}{r}}}. \quad (1.61)$$

De cet encadrement, on déduit en particulier que

$$\frac{x_0^2}{e^{-2k^{\frac{1}{r}}}} \rightarrow 1 \text{ quand } x_0 \rightarrow 0. \quad (1.62)$$

Nous allons maintenant calculer la limite quand  $x_0 \rightarrow 0$  du rapport

$$\frac{\tilde{a}_r \left( e^{-k\frac{1}{r}} \right)}{\tilde{a}_r (x_0)} = \frac{-\frac{2}{r e^{-2k\frac{1}{r}}} \int_k^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}{-\frac{2}{r x_0^2} \int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}. \quad (1.63)$$

Du fait de la relation (1.62), le rapport des termes hors intégrales, dans le membre de droite de l'égalité tend vers 1. On a donc

$$\frac{\tilde{a}_r \left( e^{-2k\frac{1}{r}} \right)}{\tilde{a}_r (x_0)} \sim \frac{\int_k^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}. \quad (1.64)$$

Le terme de droite peut encore s'écrire

$$\frac{\int_k^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv} = 1 + \frac{\int_k^{k+1} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}. \quad (1.65)$$

En supposant que  $0 < m \leq p(v) \leq M < +\infty$ , (quitte pour cela à modifier  $p$  en lui ajoutant un  $\varepsilon > 0$ ),

$$\frac{\int_k^{k+1} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} p(v) v^{\frac{1}{r}-1} dv} \leq \frac{M \int_k^{k+1} e^{-2v\frac{1}{r}} v^{\frac{1}{r}-1} dv}{m \int_{k+1}^{+\infty} e^{-2v\frac{1}{r}} v^{\frac{1}{r}-1} dv}. \quad (1.66)$$

Les intégrales dans le terme de droite sont calculables. On a la formule

$$\begin{aligned} \int_a^b e^{-2v\frac{1}{r}} v^{\frac{1}{r}-1} dv &= r \int_{a\frac{1}{r}}^{b\frac{1}{r}} e^{-2w} dw \\ &= \frac{r}{2} [-e^{-2w}]_{a\frac{1}{r}}^{b\frac{1}{r}} \\ &= \frac{r}{2} \left( e^{-2a\frac{1}{r}} - e^{-2b\frac{1}{r}} \right). \end{aligned} \quad (1.67)$$

En utilisant la relation (1.67), il vient que

$$\frac{\int_k^{k+1} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv} \leq \frac{M e^{-2k^{\frac{1}{r}}} - e^{-2(k+1)^{\frac{1}{r}}}}{e^{-2(k+1)^{\frac{1}{r}}}} \quad (1.68)$$

et donc

$$\begin{aligned} \frac{\int_k^{k+1} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv}{\int_{k+1}^{+\infty} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv} &\leq \frac{M}{m} \left( e^{-2\left((k+1)^{\frac{1}{r}} - k^{\frac{1}{r}}\right)} - 1 \right) \\ &\leq \frac{M}{m} \left( e^{-2k^{\frac{1}{r}} \left(1 + \frac{1}{k}\right)^{\frac{1}{r}} - 2k^{\frac{1}{r}}} - 1 \right) \\ &\leq \frac{M}{m} \left( e^{-\frac{2}{r} k^{\frac{1}{r}-1}} - 1 \right) \end{aligned}$$

qui tend vers 0 quand  $k \rightarrow \infty$ . On en déduit que

$$\frac{\tilde{a}_r \left( e^{-k^{\frac{1}{r}}} \right)}{\tilde{a}_r(x_0)} \rightarrow 1, \text{ quand } x_0 \rightarrow 0. \quad (1.69)$$

Conclusion:  $\tilde{a}_r \left( e^{-k^{\frac{1}{r}}} \right)$  et  $\tilde{a}_r(x_0)$  ont même limite quand  $x_0 \rightarrow 0$ .

Ceci nous amène à l'étude du comportement de

$$\tilde{a}_r \left( e^{-k^{\frac{1}{r}}} \right) = -\frac{2}{r e^{-2k^{\frac{1}{r}}}} \int_k^{+\infty} e^{-2v^{\frac{1}{r}}} p(v) v^{\frac{1}{r}-1} dv. \quad (1.70)$$

En utilisant la série de Fourier de la fonction  $p(v) = \bar{p} + \tilde{p}(v)$ , où  $\bar{p} = \int_0^1 p(v) dv$  désigne la valeur moyenne non-nulle de  $p(v)$ , alors

$$\begin{aligned} \tilde{a}_r \left( e^{-k^{\frac{1}{r}}} \right) &= -\bar{p} \frac{2}{r e^{-2k^{\frac{1}{r}}}} \int_k^{+\infty} e^{-2v^{\frac{1}{r}}} v^{\frac{1}{r}-1} dv \\ &\quad - \frac{2}{r e^{-2k^{\frac{1}{r}}}} \int_k^{+\infty} e^{-2v^{\frac{1}{r}}} \tilde{p}(v) v^{\frac{1}{r}-1} dv. \end{aligned} \quad (1.71)$$

D'après la relation (1.67),

$$-\bar{p} \frac{2}{r e^{-2k^{\frac{1}{r}}}} \int_k^{+\infty} e^{-2v^{\frac{1}{r}}} v^{\frac{1}{r}-1} dv = -\bar{p} \frac{2}{r e^{-2k^{\frac{1}{r}}}} \frac{r}{2} e^{-2k^{\frac{1}{r}}} = -\bar{p} \quad (1.72)$$

et donc

$$\tilde{a}_r \left( e^{-2k^{\frac{1}{r}}} \right) = -\bar{p} - 2 \int_k^{+\infty} e^{2\left(k^{\frac{1}{r}} - v^{\frac{1}{r}}\right)} \tilde{p}(v) v^{\frac{1}{r}-1} dv. \quad (1.73)$$

Le deuxième terme de la relation (1.73), s'écrit

$$B_k = -2 \int_k^{+\infty} e^{2\left(k^{\frac{1}{r}} - v^{\frac{1}{r}}\right)} \tilde{p}(v) v^{\frac{1}{r}-1} dv. \quad (1.74)$$

En faisant le changement de variable  $v = k + w$ , on obtient

$$B_k = -2 \int_0^{+\infty} e^{2\left(k^{\frac{1}{r}} - (k+w)^{\frac{1}{r}}\right)} \tilde{p}(w) (k+w)^{\frac{1}{r}-1} dw \quad (1.75)$$

et avec le changement  $w = kz$ , on obtient

$$B_k = -2k^{\frac{1}{r}} \int_0^{+\infty} e^{2k^{\frac{1}{r}}\left(1 - (1+z)^{\frac{1}{r}}\right)} \tilde{p}(kz) (1+z)^{\frac{1}{r}-1} dz. \quad (1.76)$$

Pour l'étude de  $B_k$ , nous décomposons l'intégrale en une somme de deux:

$$\begin{aligned} B_k = & -2k^{\frac{1}{r}} \int_0^\varepsilon e^{2k^{\frac{1}{r}}\left(1 - (1+z)^{\frac{1}{r}}\right)} \tilde{p}(kz) (1+z)^{\frac{1}{r}-1} dz \\ & -2k^{\frac{1}{r}} \int_\varepsilon^{+\infty} e^{2k^{\frac{1}{r}}\left(1 - (1+z)^{\frac{1}{r}}\right)} \tilde{p}(kz) (1+z)^{\frac{1}{r}-1} dz. \end{aligned} \quad (1.77)$$

Nous étudions d'abord la deuxième intégrale et nous montrons qu'elle tend vers 0. En utilisant l'inégalité suivante

$$1 - (1+z)^{\frac{1}{r}} \leq -C(1+z)^{\frac{1}{r}}, \text{ pour } z > \varepsilon > 0, C > 0, C = C(\varepsilon), \quad (1.78)$$

il vient que

$$|B_k| \leq 2Mk^{\frac{1}{r}} \int_\varepsilon^{+\infty} e^{-2k^{\frac{1}{r}}C(1+z)^{\frac{1}{r}}} (1+z)^{\frac{1}{r}-1} dz. \quad (1.79)$$

Si on fait le changement de variable  $v = k^{\frac{1}{r}} C (1+z)^{\frac{1}{r}}$ , on obtient

$$2Mk^{\frac{1}{r}} \int_{\varepsilon}^{+\infty} e^{-2k^{\frac{1}{r}} c(1+z)^{\frac{1}{r}}} (1+z)^{\frac{1}{r}-1} dz \leq \frac{2CMk^{\frac{1}{r}}}{r} \int_{Ck^{\frac{1}{r}}(1+\varepsilon)^{\frac{1}{r}}}^{+\infty} (e^{-2v}) dv, \quad (1.80)$$

et  $\int_{Ck^{\frac{1}{r}}(1+\varepsilon)^{\frac{1}{r}}}^{+\infty} (e^{-2v}) dv \rightarrow 0$  quand  $k \rightarrow +\infty$ .

Il reste à évaluer

$$2k^{\frac{1}{r}} \int_0^{\varepsilon} e^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} \tilde{p}(kz) (1+z)^{\frac{1}{r}-1} dz. \quad (1.81)$$

On choisit une primitive de  $\tilde{p}$  notée  $\tilde{P}$  obtenue en intégrant formellement la série de Fourier de  $\tilde{p}$  (dont la moyenne est nulle). L'intégration de  $\tilde{p}(kz)$  donne  $\frac{1}{k} \tilde{P}(kz)$ .

L'intégration par partie de la relation (1.81) donne

$$\begin{aligned} & 2 \left[ \frac{1}{k} \tilde{P}(kz) k^{\frac{1}{r}} e^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} (1+z)^{\frac{1}{r}-1} \right]_0^{\varepsilon} \\ & - 4k^{\frac{2}{r}-1} \int_0^{\varepsilon} \tilde{P}(kz) \left\{ -\frac{1}{r} (1+z)^{\frac{1}{r}-1} e^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} (1+z)^{\frac{1}{r}-1} \right\} dz \\ & - 4k^{\frac{2}{r}-1} \int_0^{\varepsilon} \tilde{P}(kz) \left\{ c^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} \left( \frac{1}{r} - 1 \right) (1+z)^{\frac{1}{r}-2} \right\} dz. \end{aligned} \quad (1.82)$$

Le premier terme tend vers 0 quand  $k \rightarrow +\infty$ , car

$$e^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} \leq e^{-2Ck^{\frac{1}{r}}(1+z)^{\frac{1}{r}}} \rightarrow 0.$$

De la même manière puisque  $e^{2k^{\frac{1}{r}} (1-(1+z)^{\frac{1}{r}})} \leq e^{-2Ck^{\frac{1}{r}}(1+z)^{\frac{1}{r}}} \rightarrow 0$ , quand  $k \rightarrow +\infty$ , le deuxième et troisième terme convergent vers 0, si  $k^{\frac{2}{r}-1} \rightarrow 0$ . Ceci est possible à condition que

$$\frac{2}{r} - 1 < 0, \quad (1.83)$$

c'est à dire pour  $r > 2$ .

Pour le cas  $r < 2$ , on itère le calcul précédent. On peut définir des primitives

successives de  $\tilde{p}$ , que nous notons

$$\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_n = \tilde{P} \quad (1.84)$$

obtenues en intégrant formellement la série de Fourier de  $\tilde{p}$  (dont la moyenne est nulle).

L'intégration successive de  $\tilde{p}(kz)$  donne

$$\frac{1}{k} \tilde{P}_1(kz), \frac{1}{k^2} \tilde{P}_2(kz), \dots, \frac{1}{k^n} \tilde{P}_n(kz), \dots \quad (1.85)$$

On s'arrête dès que l'on a  $r > \frac{n}{n-1}, \dots$

Finalement, avec  $r > 1$ , on a la convergence vers 0 du terme

$$2k^{\frac{1}{r}} \int_0^\varepsilon e^{2k^{\frac{1}{r}}(1-(1+z)^{\frac{1}{r}})} \tilde{p}(kz) (1+z)^{\frac{1}{r}-1} dz. \quad (1.86)$$

D'où,

$$\tilde{a}_r(x_0) \longrightarrow -\bar{p}, \text{ quand } x_0 \longrightarrow 0, \text{ pour tout } r > 1. \quad (1.87)$$

Ceci permet d'affirmer que la dérivée optimale peut exister même si la dérivée de la fonction  $f$  en 0 n'existe pas.

## 1.7 Relation entre la dérivée optimale et celle obtenue par dérivation au sens de Fréchet en 0.

Dans le cas vectoriel et en supposant que la suite définie par la relation (1.22) converge vers

$$\hat{A} = \left[ \int_0^{+\infty} [F(x)] [x]^T dt \right] [\Gamma(x)]^{-1}, \quad (1.88)$$

il est possible d'écrire cette limite en fonction de la dérivée de  $F$  en 0.

Si la dérivée de  $F$  en 0,  $DF(0)$  existe, on peut écrire

$$F(x) = DF(0)x + o(|x|) \quad (1.89)$$

et en remplaçant dans l'expression (1.88), on obtient

$$\tilde{A} = \left[ \int_0^{+\infty} [DF(0)x(t) + o(|x(t)|)] [x(t)]^T dt \right] \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right]^{-1}. \quad (1.90)$$

En utilisant les propriétés de la section (1.5), il vient que

$$\begin{aligned} \tilde{A} = & DF(0) \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right] \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right]^{-1} \\ & + \left[ \int_0^{+\infty} [o(|x(t)|)] [x(t)]^T dt \right] \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right]^{-1}, \end{aligned}$$

enfin

$$\tilde{A} = DF(0) + \left[ \int_0^{+\infty} [o(|x(t)|)] [x(t)]^T dt \right] \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right]^{-1}$$

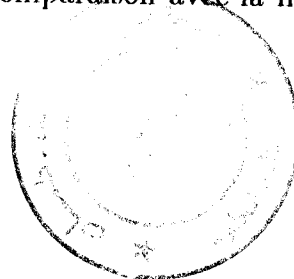
où

$$\left[ \int_0^{+\infty} [o(|x(t)|)] [x(t)]^T dt \right] \left[ \int_0^{+\infty} [x(t)] [x(t)]^T dt \right]^{-1} = o(1)$$

c'est à dire, une quantité qui tend vers 0 quand  $x_0 \rightarrow 0$ , en supposant que  $|x(t)|$  reste de l'ordre de  $|x_0|$ .

Là aussi, il est possible comme dans le cas scalaire, d'associer une équation linéaire au problème non-linéaire même si la dérivée en 0 n'existe pas.

Dans le chapitre V consacré à la mise en oeuvre numérique, plusieurs exemples (5.2.1), (5.2.2), (5.2.3), (5.2.5), ainsi qu'une comparaison avec la linéarisation classique (5.2.6), seront présentés.



**DEUXIEME CHAPITRE  
DETERMINATION DE LA  
DERIVEE OPTIMALE.**



## Chapitre 2

# Détermination de la dérivée optimale.

### 2.1 Introduction.

Dans ce chapitre, on aborde l'étude de la dérivée optimale et particulièrement la procédure de calcul donnant celle-ci.

On fixe  $x_0 \neq 0$ , et on définit dans  $M_n(\mathbb{R})$ , l'application  $A \longrightarrow \Phi(A)$  telle que

$$\Phi(A) = \left[ \int_0^{+\infty} [F(e^{tA}x_0)] [e^{tA}x_0]^T dt \right] [\Gamma(A)]^{-1} \quad (2.1)$$

avec

$$\Gamma(A) = \int_0^{+\infty} [e^{tA}x_0] [e^{tA}x_0]^T dt.$$

Maintenant, on construit la suite d'approximation

$$\begin{cases} A_j = \Phi(A_{j-1}) \\ A_0 = DF(x_0), \end{cases} \quad (2.2)$$

$x_0$  est en principe un point arbitraire pris dans un voisinage de 0 et tel que  $F$  soit différentiable en  $x_0$ . La limite de la suite  $A_j$ , si elle existe est ce que nous

appelons la dérivée optimale. Celle-ci nous permet d'avoir une équation linéaire optimale.

En fait, c'est un problème de point fixe que nous sommes amené à résoudre, puisque, dans les conditions de convergence de la suite  $A_j$  on aura à la fois l'existence et l'unicité si l'on peut montrer que l'application  $\Phi$  est lipschitzienne, de rapport plus petit que 1. Le problème ainsi posé est très délicat compte tenu de la nature de l'opérateur  $\Phi$ . Nous avons cherché à le résoudre dans le cas de la perturbation non-linéaire d'une équation linéaire. On a développé deux approches:

Dans la première, on considère la perturbation d'une matrice diagonale [19], qui s'écrit

$$\Delta = \begin{bmatrix} \delta_1 & 0 & . & . & . & 0 & 0 \\ 0 & \delta_2 & & & 0 & 0 & 0 \\ . & & . & & 0 & . & \\ . & & & . & & . & \\ . & 0 & & . & 0 & . & \\ 0 & 0 & 0 & 0 & \delta_{n-1} & 0 & \\ 0 & 0 & . & . & . & 0 & \delta_n \end{bmatrix}; 0 < \delta_1 < \delta_2 < \dots < \delta_n.$$

Cette étude est proposée dans le présent chapitre. La deuxième approche correspond au cas où on perturbe une matrice dont les valeurs propres sont imaginaires [21]. Celle-ci sera présentée au troisième chapitre. La situation envisagée dans ce chapitre est générique. On peut s'y ramener dans le cas où la matrice est diagonalisable et a toutes ses valeurs propres réelles et distinctes. Nous n'avons pas abordé le cas générique à valeurs propres réelles ou complexes. La deuxième situation, examinée au chapitre 3 est non générique.

## 2.2 Position du problème.

On considère l'équation différentielle ordinaire non-linéaire

$$\begin{cases} \frac{dx}{dt} = F(x) \\ x(0) = x_0 \end{cases} \quad (2.3)$$

avec

$$F(x) = -\Delta x + G(x),$$

$x \in \mathbb{R}^n$ ,

$F$  est définie dans un certain ouvert  $\Omega$  à valeurs dans  $\mathbb{R}^n$ .

On supposera que:

$$F(0) = 0,$$

$F$  est continue, Lipschitzienne,

$x_0$  est choisi à composantes toutes non-nulles.

Nous nous intéressons particulièrement à l'existence, l'unicité et la convergence de la dérivée optimale obtenue par minimisation au sens des moindres carrés et associée à l'équation (2.3) de la forme

$$\begin{cases} \frac{dx}{dt} = \tilde{A}x \\ x(0) = x_0, \end{cases} \quad (2.4)$$

$\tilde{A}$  étant la limite, si elle existe, de la suite définie par (2.2).

## 2.3 Relation de récurrence.

En utilisant l'expression de  $F$ , la relation de récurrence entre  $A_{j-1}$  et  $A_j$  s'écrit:

$$\begin{aligned} A_j &= -\Delta \left[ \int_0^{+\infty} [e^{tA_{j-1}} x_0] [e^{tA_{j-1}} x_0]^T dt \right] [\Gamma(A_{j-1})]^{-1} \\ &\quad + \left[ \int_0^{+\infty} G(e^{tA_{j-1}} x_0) [e^{tA_{j-1}} x_0]^T dt \right] [\Gamma(A_{j-1})]^{-1} \end{aligned} \quad (2.5)$$

ou encore

$$A_j = -\Delta + \left[ \int_0^{+\infty} G(e^{tA_{j-1}}x_0) [e^{tA_{j-1}}x_0]^T dt \right] [\Gamma(A_{j-1})]^{-1}, \quad (2.6)$$

à condition que  $\Gamma$  soit inversible.

Si on pose

$$\varphi(A) = \left[ \int_0^{+\infty} G(e^{tA}x_0) [e^{tA}x_0]^T dt \right] [\Gamma(A)]^{-1}, \quad (2.7)$$

alors on obtient l'équation de récurrence

$$\begin{cases} A_j = -\Delta + \varphi(A_{j-1}) \\ A_0 \in B(-\Delta, \rho). \end{cases} \quad (2.8)$$

$B(-\Delta, \rho)$  est une boule de centre  $-\Delta$ , et de rayon  $\rho > 0$  dans l'espace des matrices.

La matrice initiale étant

$$A_0 = -\Delta + DG(x_0). \quad (2.9)$$

## 2.4 Existence de la dérivée optimale.

On cherche  $A$  dans un voisinage de la matrice  $-\Delta$ , c'est à dire dans la boule de centre  $-\Delta$  et de rayon  $\rho > 0$  dans l'espace des matrices. On choisit  $\rho$  assez petit pour que  $A$  ait ses valeurs propres distinctes et à partie réelle négative.

**Lemme 2.4.1** [19] *Sous les hypothèses sur  $-\Delta$ , pour tout  $\rho > 0$ , il existe  $\gamma > 0$  tel que, pour toute fonction  $G$  continue, vérifiant de plus*

$$\|G(x)\| \leq \gamma \|x\|, \quad (2.10)$$

*et pour tout  $A \in \mathcal{M}_n(\mathbb{R})$ , vérifiant  $\|A + \Delta\| \leq \rho$ , on a*

$$\|\varphi(A)\| \leq \gamma C'''', \quad (2.11)$$

$$\text{avec } C''' = \frac{k^2 \|x_0\|^2}{(1 - C''\rho)^2} \frac{1}{2\eta} \frac{1}{m(x_0, \delta)}.$$

**Preuve:** Avec

$$\varphi(A) = \left[ \int_0^{+\infty} G(e^{At}x_0) [e^{At}x_0]^T dt \right] [\Gamma(A)]^{-1}, \quad (2.12)$$

et d'après la relation (2.10), on a

$$\|\varphi(A)\| \leq \gamma \left[ \int_0^{+\infty} \|e^{At}x_0\|^2 dt \right] \|[\Gamma(A)]^{-1}\|. \quad (2.13)$$

**Estimation de  $\|[\Gamma(A)]^{-1}\|$ .**

Pour cela, calculons

$$v^T \Gamma(A) v = \int_0^{+\infty} (v^T e^{At} x_0)^2 dt. \quad (2.14)$$

$v^T \Gamma(A) v$  est l'intégrale du carré d'une fonction. Donc,  $v^T \Gamma(A) v$  sera non-nulle si la fonction est non-nulle. On prend  $A = -\Delta$ , la fonction dont on calcule le carré est donc la fonction

$$\psi(v(t)) = \sum_{j=1}^n v_j x_{0,j} e^{-\delta_j t}.$$

Les  $\delta_j$  étant tous distincts,  $\psi(v(t))$  est équivalent quand  $t \rightarrow +\infty$ , au terme correspondant au plus petit entier  $j$  tel que

$$v_j x_{0,j} \neq 0,$$

c'est à dire

$$v_j \neq 0.$$

Par conséquent, on a  $v^T \Gamma(A) v > 0$  pour tout  $v \neq 0$ . Donc, la forme quadratique associée à  $\Gamma(A)$  est définie positive, et donc:

pour tout  $(x_0, \delta)$ , il existe une constante  $m(x_0, \delta)$  telle que

$$v^T \Gamma(A) v \geq m(x_0, \delta) \|v\|^2 \quad (2.15)$$

avec

$$m(x_0, \delta) > 0, \quad (2.16)$$

et  $0 < \delta_1 < \delta_2 < \dots < \delta_n$ .

On a donc

$$\|\Gamma(A)\| \geq m(x_0, \delta). \quad (2.17)$$

Avec  $\lambda_{\min} = \min \{\lambda : \lambda \in \sigma(\Gamma(A))\}$ , il vient que

$$\lambda_{\min} \geq m(x_0, \delta),$$

ce qui implique  $\forall \lambda \in \sigma([\Gamma(A)]^{-1})$ ,  $(\lambda \leq \lambda_{\min}^{-1})$ , et

$$\lambda \leq \frac{1}{m(x_0, \delta)}. \quad (2.18)$$

Ce qui donne

$$v^T [\Gamma(A)]^{-1} v \leq \frac{1}{m(x_0, \delta)} \|v\|^2. \quad (2.19)$$

D'où, l'on déduit qu'avec la matrice  $[\Gamma(A)]^{-1}$  symétrique positive,

$$\|[\Gamma(A)]^{-1}\| \leq \frac{1}{m(x_0, \delta)}. \quad (2.20)$$

**Estimation du terme  $\|e^{At}\|$ .**

On choisit un contour  $\Gamma$  dans le plan complexe, qui entoure les nombres  $-\delta_i$  et aussi les valeurs propres de  $A$  pour  $\|A + \Delta\| < \rho$ . On peut prendre un contour

dont la distance aux  $[-\delta_i]_{1 \leq i \leq n}$  soit égale à  $\rho$ . Si on pose  $\eta = \delta_1 - \rho$ , dans ce cas

$$e^{At} = \frac{1}{2i\pi} \int_{\Gamma} e^{zt} (zI - A)^{-1} dz \quad (2.21)$$

et

$$\begin{aligned} \|e^{At}\| &\leq \frac{1}{2\pi} \int_{\Gamma} |e^{zt}| \|(zI - A)^{-1}\| |dz| \\ &\leq \frac{1}{2\pi} \int_{\Gamma} e^{-\eta t} \|(zI - A)^{-1}\| |dz|. \end{aligned} \quad (2.22)$$

Evaluons maintenant le terme  $\|(zI - A)^{-1}\|$ . On a

$$(zI - A)^{-1} = (zI + \Delta)^{-1} (I - (zI + \Delta)^{-1} (A + \Delta))^{-1} \quad (2.23)$$

$$\|(zI - A)^{-1}\| \leq \frac{C}{\text{dist}(z, (-\delta_j))} \frac{1}{1 - \|(zI + \Delta)^{-1} (A + \Delta)\|} \quad (2.24)$$

$$\|(zI - A)^{-1}\| \leq C' \frac{1}{1 - \|(zI + \Delta)^{-1}\| \|(A + \Delta)\|} \quad (2.25)$$

$$\|(zI - A)^{-1}\| \leq C' \frac{1}{1 - \frac{\|(A + \Delta)\|}{\text{dist}(z, (-\delta_j))}} \quad (2.26)$$

$$\|(zI - A)^{-1}\| \leq C' \frac{1}{1 - C'' \|(A + \Delta)\|} \quad (2.27)$$

enfin

$$\|e^{At}\| \leq \frac{ke^{-\eta t}}{1 - C'' \|(A + \Delta)\|}. \quad (2.28)$$

En reportant les inégalités (2.20) et (2.28), dans la relation (2.13), on obtient:

$$\|\varphi(A)\| \leq \frac{\gamma \|x_0\|^2}{(1 - C'' \|(A + \Delta)\|)^2} \frac{k^2}{m(x_0, \delta)} \int_0^{+\infty} e^{-2\eta t} dt \quad (2.29)$$

et donc

$$\|\varphi(A)\| \leq \gamma C''' \quad (2.30)$$

$$\text{avec } C''' = \frac{k^2 \|x_0\|^2}{(1 - C'' \rho)^2} \frac{1}{2\eta} \frac{1}{m(x_0, \delta)}. \quad \square$$

**Proposition 2.4.2** [19] *Sous les hypothèses du lemme (2.4.1), il existe  $\rho > 0$  telle que l'équation*

$$A = -\Delta + \varphi(A) \quad (2.31)$$

*admette au moins une solution dans  $B(-\Delta, \rho)$ .*

**Preuve:** Si on impose  $\gamma C''' < \rho$ , on a

$$\|\varphi(A)\| < \rho.$$

Alors, l'application  $\Phi : A \rightarrow -\Delta + \varphi(A)$ , définie de  $B(-\Delta, \rho)$  dans  $\mathcal{M}_n(\mathbb{R})$  est en fait à valeurs dans  $B(-\Delta, \rho)$ .

Cette application est continue. La continuité de  $\Phi$  est conséquence de la continuité de chacune des applications  $A \rightarrow (\Gamma(A))^{-1}$  et  $A \rightarrow \int_0^{+\infty} G(x) [e^{tA} x_0]^T dt$ . Pour la première application, il suffit de montrer la continuité de l'application qui à  $A \rightarrow \Gamma(A)$ , puisque nous savons que  $\Gamma(A)$  est inversible uniformément (dans un voisinage de  $A = 0$ ). Or

$$\Gamma(A) = \int_0^{+\infty} [x][x]^T dt = \int_0^\tau [x][x]^T dt + \int_\tau^{+\infty} [x][x]^T dt$$

l'intégrale de 0 à  $\tau$  dépend continument de  $A$ , et l'intégrale de  $\tau$  à  $+\infty$  est uniformément petite quand  $\tau$  est assez grand. De la même manière pour la deuxième application

$$\varphi(A) = \int_0^{+\infty} G(x) [e^{tA} x_0]^T dt = \int_0^\tau G(x) [e^{tA} x_0]^T dt + \int_\tau^{+\infty} G(x) [e^{tA} x_0]^T dt$$

l'intégrale de 0 à  $\tau$  dépend continument de  $A$ , et l'intégrale de  $\tau$  à  $+\infty$  est uniformément petite quand  $\tau$  est assez grand. D'où l'on peut écrire

$$\begin{aligned} \Gamma(A) &= \lim_{\tau \rightarrow +\infty} \text{unif} \int_0^\tau (e^{tA} x_0) (e^{tA} x_0)^T dt \\ \varphi(A) &= \lim_{\tau \rightarrow +\infty} \text{unif} \int_0^\tau G(e^{tA} x_0) [e^{tA} x_0]^T dt. \end{aligned}$$



$\Gamma(A)$  et  $\varphi(A)$  sont limite uniforme d'une suite de fonctions continues et sont donc continues.

$\Phi$  applique la boule fermée dans elle même et  $\Phi$  est continue. D'après le théorème de Brouwer [13], on conclut à l'existence d'un point fixe.  $\square$

## 2.5 Unicité de la dérivée optimale.

**Proposition 2.5.1** [19] *Sous les mêmes hypothèses sur  $-\Delta$ , pour tout  $\rho > 0$ , il existe  $\gamma > 0$  tel que pour toute fonction  $G$  vérifiant  $G(0) = 0$  et*

$$\|G(x) - G(x')\| \leq \gamma \|x - x'\|, \forall x, x' \in \mathbb{R}^n, \quad (2.32)$$

*l'application  $\Phi$  est lipschitzienne, de rapport strictement inférieur à 1 de la boule  $B(-\Delta, \rho)$  dans elle même. Dans ce cas, l'équation (2.31) a une et une seule solution dans  $B(-\Delta, \rho)$ .*

**Preuve:** Pour montrer que  $A$  est unique, il suffit de vérifier que  $\varphi$  est lipschitzienne de rapport strictement plus petit que 1. Par rapport au résultat précédent, il suffit de renforcer les conditions sur  $G$  pour assurer que  $\Phi$  soit une contraction stricte. Pour cela, considérons l'application qui à la matrice  $B$  fait correspondre  $\varphi(B)$  tel que

$$\varphi(B) = \left[ \int_0^{+\infty} G(e^{Bt}x_0) [e^{Bt}x_0]^T dt \right] [\Gamma(B)]^{-1}. \quad (2.33)$$

Estimons la différence

$$\begin{aligned} \varphi(B_2) - \varphi(B_1) = & \left[ \int_0^{+\infty} G(e^{B_2 t}x_0) [e^{B_2 t}x_0]^T dt - \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] [\Gamma(B_2)]^{-1} \\ & + \left[ \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] \left[ [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right]. \end{aligned} \quad (2.34)$$

En posant

$$R = \int_0^{+\infty} G(e^{B_1 t} x_0) [e^{B_1 t} x_0]^T dt [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1}, \quad (2.35)$$

l'estimation de  $R$  donne,

$$\|R\| \leq \left\| \int_0^{+\infty} G(e^{B_1 t} x_0) [e^{B_1 t} x_0]^T dt \right\| \left\| [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right\|.$$

On a

$$\left\| \int_0^{+\infty} G(e^{B_1 t} x_0) [e^{B_1 t} x_0]^T dt \right\| \leq \gamma \frac{k^2 \|x_0\|^2}{(1 - C'' \|B + \Delta\|)^2 2\eta} \leq \gamma \frac{k^2 \|x_0\|^2}{(1 - C'' \rho)^2 2\eta}. \quad (2.36)$$

Avec

$$[\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} = [\Gamma(B_2)]^{-1} [\Gamma(B_1) - \Gamma(B_2)] [\Gamma(B_1)]^{-1} \quad (2.37)$$

et

$$S = [\Gamma(B_1)] - [\Gamma(B_2)],$$

on obtient l'estimation

$$\left\| [\Gamma(B_2)]^{-1} [[\Gamma(B_1)] - [\Gamma(B_2)]] [\Gamma(B_1)]^{-1} \right\| \leq \frac{1}{m(x_0, \delta)} \|S\| \frac{1}{m(x_0, \delta)}. \quad (2.38)$$

Evaluons maintenant le terme

$$S = \left[ \int_0^{+\infty} e^{B_2 t} x_0 x_0^T e^{B_2^T t} dt - \int_0^{+\infty} e^{B_1 t} x_0 x_0^T e^{B_1^T t} dt \right] \quad (2.39)$$

$$= \int_0^{+\infty} [e^{B_2 t} - e^{B_1 t}] x_0 x_0^T e^{B_2^T t} dt + \int_0^{+\infty} e^{B_1 t} x_0 x_0^T [e^{B_2^T t} - e^{B_1^T t}] dt. \quad (2.40)$$

**Estimation de  $e^{B_2 t} - e^{B_1 t}$ .**

D'après la relation (2.21)

$$\begin{aligned} e^{B_2 t} - e^{B_1 t} &= \frac{1}{2i\pi} \int_{\Gamma} e^{zt} \left\{ (zI - B_2)^{-1} - (zI - B_1)^{-1} \right\} dz \\ &= \frac{1}{2i\pi} \int_{\Gamma} e^{zt} (zI - B_2)^{-1} \left\{ (zI - B_1) - (zI - B_2) \right\} (zI - B_1)^{-1} dz. \end{aligned} \quad (2.41)$$

En passant à la norme on obtient

$$\begin{aligned} \|e^{B_2 t} - e^{B_1 t}\| &\leq \frac{1}{2\pi} \|B_2 - B_1\| \int_{\Gamma} e^{-\eta t} \|(zI - B_2)^{-1}\| \|(zI - B_1)^{-1}\| |dz| \\ &\leq \|B_2 - B_1\| e^{-\eta t} \frac{C'}{1 - C'' \|B_2 + \Delta\|} \frac{C'}{1 - C'' \|B_1 + \Delta\|}. \end{aligned} \quad (2.42)$$

Si  $\|B_1 + \Delta\| \leq \rho$ ,  $\|B_2 + \Delta\| \leq \rho$ , alors

$$\|e^{B_2 t} - e^{B_1 t}\| \leq \|B_2 - B_1\| e^{-\eta t} \frac{C'^2}{(1 - C''\rho)^2}. \quad (2.43)$$

L'estimation de  $S$ , donne

$$\begin{aligned} \|\Gamma(B_1) - \Gamma(B_2)\| &\leq \|B_2 - B_1\| e^{-\eta t} \frac{C'^2}{(1 - C''\rho)^2} \left[ \int_0^{+\infty} e^{-\eta t} \|x_0 x_0^T\| \|e^{B_2 t}\| dt \right] \\ &\quad + \frac{C'^2}{(1 - C''\rho)^2} \left[ \int_0^{+\infty} e^{-\eta t} \|x_0 x_0^T\| \|e^{B_1 t}\| dt \right] \end{aligned} \quad (2.44)$$

et avec  $\|e^{B_2 t}\|$ ,  $\|e^{B_1 t}\|$  évalués en (2.28), on obtient

$$\|\Gamma(B_1) - \Gamma(B_2)\| \leq \|B_2 - B_1\| \frac{C'^3}{(1 - C''\rho)^3} \|x_0 x_0^T\| \frac{1}{2\eta}. \quad (2.45)$$

La relation (2.45), peut s'écrire

$$\|\Gamma(B_1) - \Gamma(B_2)\| \leq C_4 \|B_2 - B_1\|, \quad (2.46)$$

avec

$$C_4 = \frac{C'^3}{(1 - C''\rho)^3} \|x_0\|^2 \frac{1}{2\eta}. \quad (2.47)$$

Enfin, l'estimation de la relation (2.37), s'écrit

$$\|\Gamma(B_2)^{-1} - \Gamma(B_1)^{-1}\| \leq \frac{1}{(m(x_0, \delta))^2} C_4 \|B_2 - B_1\|. \quad (2.48)$$

Finalement

$$\|R\| \leq \gamma C_5 \|B_2 - B_1\| \quad \text{avec } C_5 = C_4 \frac{1}{2\eta} \frac{1}{(m(x_0, \delta))^2} \frac{k^2 \|x_0\|^2}{(1 - C''\rho)^2}. \quad (2.49)$$

Maintenant, posons

$$T = \left[ \int_0^{+\infty} G(e^{B_2 t} x_0) [e^{B_2 t} x_0]^T dt - \int_0^{+\infty} G(e^{B_1 t} x_0) [e^{B_1 t} x_0]^T dt \right] [\Gamma(B_2)]^{-1}.$$

L'estimation de  $T$  donne

$$\begin{aligned} \|T\| \leq & \left\| \int_0^{+\infty} [G(e^{B_2 t} x_0) - G(e^{B_1 t} x_0)] [e^{B_2 t} x_0]^T dt [\Gamma(B_2)]^{-1} \right\| \\ & + \left\| \int_0^{+\infty} G(e^{B_1 t} x_0) \left[ [e^{B_2 t} x_0]^T - [e^{B_1 t} x_0]^T \right] dt [\Gamma(B_2)]^{-1} \right\|. \end{aligned} \quad (2.50)$$

Avec l'hypothèse sur  $G$

$$\|G(x) - G(x')\| \leq \gamma \|x - x'\|, \quad \forall x, x' \in \mathbb{R}^n,$$

et les relations (2.28) et (2.42), on obtient

$$\begin{aligned} \|T\| \leq & \int_0^{+\infty} \gamma \|B_2 - B_1\| e^{-\eta t} \frac{(C')^2}{(1 - C''\rho)^2} \|x_0\| \frac{C' e^{-\eta t}}{(1 - C''\rho)} \frac{1}{m(x_0, \delta)} dt \\ & + \int_0^{+\infty} \gamma \frac{C' e^{-\eta t}}{(1 - C''\rho)} \|x_0\| \cdot \|B_2 - B_1\| \cdot e^{-\eta t} \frac{(C')^2}{(1 - C''\rho)^2} \|x_0\| \frac{1}{m(x_0, \delta)} dt. \end{aligned} \quad (2.51)$$

Enfin

$$\|T\| \leq 2\gamma \frac{(C')^3}{(1 - C''\rho)^3} \|x_0\|^2 \frac{1}{m(x_0, \delta)} \frac{1}{2\eta} \|B_2 - B_1\| \leq \gamma C_6 \|B_2 - B_1\|. \quad (2.52)$$

Finalement

$$\|\varphi(B_2) - \varphi(B_1)\| \leq \gamma(C_5 + C_6) \|B_2 - B_1\|. \quad (2.53)$$

En prenant  $\gamma(C_5 + C_6) < 1$ , et  $\gamma C''' < \rho$ , on obtient que  $\Phi$  est lipschitzienne, de rapport  $< 1$  de la boule  $B(-\Delta, \rho)$  dans elle même. On conclut que  $A$  est unique.  $\square$

## 2.6 Convergence de la dérivée optimale.

**Théorème 2.6.1** *Sous les hypothèses du Lemme (2.4.1) et des propositions (2.4.2) (2.5.1), on a:*

$$\Phi : B(-\Delta, \rho) \longrightarrow B(-\Delta, \rho),$$

*$\Phi$  est une contraction stricte.*

*Alors, une condition pour que la suite  $A_j$  converge vers l'unique point fixe dans la boule est que  $A_0 \in B(-\Delta, \rho)$ .*

**Preuve:** On a

$$A_0 \in B(-\Delta, \rho) \iff A_0 + \Delta = DG(x_0) \in B(0, \rho) \quad (2.54)$$

D'après la relation (2.10), on a

$$\|DG(x_0)\| \leq \gamma. \quad (2.55)$$

Si l'on impose la condition

$$\gamma < \rho, \quad (2.56)$$

on aura

$$A_0 \in B(-\Delta, \rho). \quad (2.57)$$

Dans ce cas, la suite

$$\begin{cases} A_j = -\Delta + \Phi(A_{j-1}) \\ A_0 \in B(-\Delta, \rho) \end{cases} \quad (2.58)$$

converge vers  $\tilde{A}$ .

Par conséquent, pour tout  $A_0 \in B(-\Delta, \rho)$ , la suite  $A_j$  est bien définie, à valeurs dans  $B(-\Delta, \rho)$  et de plus elle converge géométriquement vers l'unique point fixe dans la boule.  $\square$

**TROISIEME CHAPITRE  
PROPRIETES DE LA  
DERIVEE OPTIMALE.**

# Chapitre 3

## Propriétés de la dérivée optimale.

### 3.1 Introduction.

Nous allons nous intéresser à l'étude des propriétés de l'équation linéaire obtenue par approximation, par rapport à l'équation non-linéaire. Plus généralement, nous nous poserons la question suivante:

Quelles relations existent-ils entre les propriétés qualitatives (stabilité, ordre ...) de l'équation non-linéaire et celles de l'équation optimale?

Ainsi dans ce chapitre, on précise l'ordre de l'approximation [22]. Ensuite, on aborde la question fondamentale du lien entre la dérivée optimale et la stabilité asymptotique.

### 3.2 Ordre.

Nous allons estimer l'écart définissant l'erreur. Ce qui va permettre de préciser l'ordre de l'approximation. Ceci revient en fait à évaluer  $\|x(t) - \tilde{y}(t)\|$ , où  $x$  est la solution de l'éq.(1.1) et  $\tilde{y}$  est la solution de la dérivée optimale, les deux ayant



la même valeur initiale. Il s'agit là d'une estimation globale.

**Lemme 3.2.1** [22] *Sous les hypothèses H1), H2), H3) du chapitre 1, section 1.2, sur la fonction  $F$ , et si en plus*

$$\|F(x) - DF(0)x\| = O(\|x\|^2), \text{ quand } x_0 \longrightarrow 0,$$

$$\|\tilde{y}(t)\| \leq C \|x_0\|,$$

alors

$$\int_0^{+\infty} \|F(\tilde{y}(t)) - \hat{A} \tilde{y}(t)\|^2 dt \leq O(\|x_0\|^2)^2. \quad (3.1)$$

**Preuve:** Nous allons évaluer la fonctionnelle définie par la relation (1.3)

$$\int_0^{+\infty} \|F(y(t)) - Ay(t)\|^2 dt.$$

Pour calculer la première matrice  $A_1$  de la suite définie par la relation (1.22), on part d'une matrice  $A_0$  quelconque, et on minimise la fonctionnelle

$$\int_0^{+\infty} \|F(y_0(t)) - A_1 y_0(t)\|^2 dt \quad (3.2)$$

où  $y_0(t)$  est la solution de l'équation

$$\begin{cases} \frac{dy}{dt} = A_0 y(t) \\ y(0) = x_0. \end{cases} \quad (3.3)$$

$A_1$  minimise la relation (1.3) au sens des moindres carrés le long de la solution  $y_0(t)$  du système (3.3).  $A_1$  sera tel que

$$\int_0^{+\infty} \|F(y_0(t)) - A_1 y_0(t)\|^2 dt \leq \int_0^{+\infty} \|F(y_0(t)) - Ay_0(t)\|^2 dt, \quad (3.4)$$

$A$  étant une matrice quelconque appartenant à l'espace des matrices admissibles, c'est à dire  $Re(\sigma(A)) < 0$ .

De la même manière  $A_{j+1}$  minimise la relation (1.3) au sens des moindres carrés le long de la solution  $y_j(t)$  du système

$$\begin{cases} \frac{dy}{dt} = A_j y \\ y(0) = x_0 \end{cases} \quad (3.5)$$

et  $A_{j+1}$  sera tel que

$$\int_0^{+\infty} \|F(y_j(t)) - (A_{j+1})y_j(t)\|^2 dt \leq \int_0^{+\infty} \|F(y_j(t)) - A_j y_j(t)\|^2 dt, \quad (3.6)$$

pour tout  $A$ ,  $\operatorname{Re}(\sigma(A)) < 0$ .

A la limite ( $j \rightarrow +\infty$ ), on obtient

$$\int_0^{+\infty} \|F(\tilde{y}(t)) - \hat{A}\tilde{y}(t)\|^2 dt \leq \int_0^{+\infty} \|F(\tilde{y}(t)) - A\tilde{y}(t)\|^2 dt. \quad (3.7)$$

Ainsi

$$\begin{aligned} & \int_0^{+\infty} \|F(\tilde{y}(t)) - \hat{A}\tilde{y}(t)\|^2 dt \\ &= \inf_{\forall A \in \mathcal{M}_n(\mathbb{R}), \operatorname{Re} \sigma(A) \subset ]-\infty, 0[} \int_0^{+\infty} \|F(\tilde{y}(t)) - A\tilde{y}(t)\|^2 dt. \end{aligned} \quad (3.8)$$

En particulier pour  $A = DF(0)$ , on a

$$\int_0^{+\infty} \|F(\tilde{y}(t)) - \hat{A}\tilde{y}(t)\|^2 dt \leq \int_0^{+\infty} \|F(\tilde{y}(t)) - DF(0)\tilde{y}(t)\|^2 dt. \quad (3.9)$$

Avec  $\|F(x) - DF(0)x\| = O(\|x\|^2)$ , quand  $x_0 \rightarrow 0$ , et  $\|\tilde{y}(t)\| \leq C\|x_0\|$ , on conclut que

$$\int_0^{+\infty} \|F(\tilde{y}(t)) - \hat{A}\tilde{y}(t)\|^2 dt \leq O(\|x_0\|^2)^2. \quad \square \quad (3.10)$$

**Proposition 3.2.2** [22] *Sous les hypothèses du lemme (3.2.1), et pour tout  $T > 0$ , il existe  $M \geq 0$  tel que*

$$\|x(t) - \tilde{y}(t)\| \leq M(\|x_0\|^2), \quad (3.11)$$

pour  $0 \leq t \leq T$ , et tout  $x_0$  dans un voisinage de 0, indépendant de  $T$ .

**Preuve:** Evaluons maintenant la différence

$$\begin{aligned} \frac{dx}{dt} - \frac{d\tilde{y}}{dt} &= F(x(t)) - \tilde{A}\tilde{y}(t) \\ &= F(x(t)) - F(\tilde{y}(t)) + F(\tilde{y}(t)) - \tilde{A}\tilde{y}(t). \end{aligned} \quad (3.12)$$

De l'hypothèse H3), chap I, section 1.2, il vient

$$\frac{d}{dt} \|x(t) - \tilde{y}(t)\| \leq \gamma \|x(t) - \tilde{y}(t)\| + \|F(\tilde{y}(t)) - \tilde{A}\tilde{y}(t)\|, \quad (3.13)$$

et en utilisant le lemme de Gronwall, on obtient

$$\begin{aligned} \|x(t) - \tilde{y}(t)\| &\leq \int_0^t e^{\gamma t} \|F(\tilde{y}(s)) - \tilde{A}\tilde{y}(s)\| ds \\ &\leq \left( \int_0^t e^{2\gamma(t-s)} ds \right)^{\frac{1}{2}} \left( \int_0^t \|F(\tilde{y}(s)) - \tilde{A}\tilde{y}(s)\|^2 ds \right)^{\frac{1}{2}}. \end{aligned} \quad (3.14)$$

Alors, pour tout  $T > 0$ , il existe  $M \geq 0$  tel que

$$\|x(t) - \tilde{y}(t)\| \leq M (\|x_0\|^2), \text{ pour } 0 \leq t \leq T. \quad \square \quad (3.15)$$

Donc, l'approximation proposée est d'ordre deux au moins en fonction de la donnée initiale, en général, elle est de l'ordre de la non-linéarité.

**Remarque 3.2.3** [22] L'estimation donnée dans la proposition (3.2.2) a au moins un point faible, elle dépend de  $T$  et est donc limitée à l'intégration sur un intervalle de temps fini. Toutefois, cette approximation peut être étendue à  $\mathbb{R}^+$  si nous supposons que  $F$  est dissipative, c'est à dire si pour un  $\alpha > 0$ , nous avons

$$\langle F(x) - F(y), x - y \rangle \leq -\alpha \|x - y\|^2 \quad (3.16)$$

pour tout  $x, y$ . Alors, avec  $F(0) = 0$ , on obtient

$$\|x(t) - \tilde{y}(t)\| \leq \left( \int_0^{+\infty} e^{-2\alpha(t-s)} ds \right)^{\frac{1}{2}} \left( \int_0^{+\infty} \|F(\tilde{y}(s)) - \tilde{A}\tilde{y}(s)\|^2 ds \right)^{\frac{1}{2}}. \quad (3.17)$$

Finalement, il existe  $M \geq 0$ , tel que

$$\|x(t) - \tilde{y}(t)\| \leq \left[ \frac{1}{(2\alpha)^{\frac{1}{2}}} \right] [M(\|x_0\|^2)]. \quad (3.18)$$

### 3.3 Stabilité.

#### 3.3.1 Note historique

Routh (1884), Thomson et Tait (1879), Joukovsky (1882) ne retenaient des équations différentielles que leur approximation linéaire au premier ordre. Tout en reconnaissant que le procédé n'était pas rigoureux, et sans en proposer aucune justification, l'étude de la stabilité du système non-linéaire était alors menée sur le système linéarisé.

Un apport majeur de Lyapunov et simultanément, de Poincaré, a été de donner des conditions de validité de cette approximation, basées sur les propriétés des solutions du modèle linéarisé. Cette méthode permet de conclure localement sans avoir à donner de renseignements quantitatifs. En fait, depuis l'apparition du célèbre mémoire de Lyapunov "Problème général de la stabilité du mouvement" en (1892), beaucoup d'études ont été menées Hahn (1963), LaSalle et Lefschetz (1961), ..., etc. Cependant, la détermination d'une fonction de Lyapunov constitue une difficulté majeure. Plusieurs méthodes de recherche de cette fonction ont été proposées, méthode du gradient variable Schultz et Gibson (1962), Zoubov (1957); à partir des formes quadratiques Aizerman et Gantmacher (1964), ...etc.

### 3.3.2 Lien entre la dérivée optimale et la stabilité asymptotique.

Nous examinons, dans ce qui suit, le lien entre la notion de fonction de Lyapunov ou plus généralement la stabilité d'un point d'équilibre et les propriétés de la dérivée optimale en ce point. Pour mieux illustrer le lien qui peut exister entre la dérivée optimale et l'équation non-linéaire, on va présenter un exemple dans lequel la jacobienne du champ linéaire associée à l'équation non-linéaire a un spectre à partie réelle strictement négative au voisinage de 0, sauf en  $x = 0$ , et les solutions ne tendent pas vers 0.

Il n'existe pas de résultat théorique permettant de conclure à partir de cette hypothèse sur la nature de la stabilité de l'origine, et l'exemple tends à illustrer le fait que le résultat dépend de l'équation. La procédure de dérivée optimale permet de détecter celles des équations pour lesquelles il y a stabilité.

#### Exemple 3.3.1

$$\begin{aligned}\frac{dx}{dt} &= -xg(x) \\ \frac{dy}{dt} &= -|x|^n y + |x|^{n-1}\end{aligned}\tag{3.19}$$

$n \geq 2$ , et  $g(x) > 0$ .

Si  $y \rightarrow 0$ ,  $\frac{dy}{dt} > 0$ , et donc si  $y(0) > 0 \Rightarrow y(t) > 0$ , et cela veut dire que  $y(t)$  est non bornée.

La jacobienne du champ linéaire associé à l'équation (3.19) s'écrit, pour  $n = 2$ , et  $g(x) = \frac{x^2}{2}$

$$DF(x) = \begin{bmatrix} -\frac{3x^2}{2} & 0 \\ a(x, y) & -|x|^n \end{bmatrix}$$

matrice dont le déterminant est positif  $\det(DF(x)) > 0$ , et la trace est négative  $\text{tr}(DF(x)) < 0$ . On en conclue que le spectre de la jacobienne associée à l'équation non-linéaire est à partie réelle strictement négative, donc asymptotiquement stable sauf en  $x = 0$ .

Le calcul de la dérivée optimale donne au premier pas avec  $(x_0, y_0) = (1, 0)$

$$A_1 = \begin{bmatrix} -0.340938 & 0.227282 \\ 0.909119 & -0.17040710^{-4} \end{bmatrix}$$

dont les valeurs propres sont

$$\lambda_1 = -0.65595$$

$$\lambda_2 = 0.314995.$$

La dérivée optimale appliquée à cet exemple, donne au premier pas une matrice avec une valeur propre  $> 0$  et une valeur propre  $< 0$ . La conclusion obtenue par utilisation de la dérivée optimale est donc conforme au résultat observé.

### Cas scalaire.

**Proposition 3.3.2** [20], [22] Sous les hypothèses  $H1$ ),  $H2$ ) chapitre I, section 1.6 sur la fonction  $f$ , alors

$$v(x) = x^2 \tilde{a}(x) \tag{3.20}$$

est une fonction de Lyapunov pour l'équation non-linéaire.

**Preuve:** En effet, si  $x(t)$  est une solution de l'équation (1.31), en dérivant

$$v(x(t)) = [(x(t))^2 \tilde{a}(x(t))] \tag{3.21}$$

par rapport à  $t$ , on obtient

$$\frac{d}{dt} [(x(t))^2 \tilde{a}(x(t))] = (f(x(t)))^2 > 0. \tag{3.22}$$

Puisque, au vu des hypothèses  $H1$  et  $H2$ , chap I, section 1.6

$$v(x(t)) < 0, \tag{3.23}$$

on obtient que  $v(x(t)) \rightarrow 0$  quand  $t \rightarrow +\infty$ , donc  $x(t) \rightarrow 0$  quand  $t \rightarrow +\infty$ .  $\square$

Donc dans le cas scalaire, la dérivée optimale permet la construction d'une fonction de Lyapunov pour l'équation non-linéaire.

### Cas vectoriel.

La réponse à la question y a-t-il un lien entre la propriété de stabilité de l'équation linéaire obtenue par la dérivée optimale et celle de l'équation non-linéaire dans le cas vectoriel est très délicate. La procédure de calcul nous a permis de résoudre des problèmes où la linéarisée classique tombe en défaut. Dans le chapitre 5, consacré à la mise en oeuvre numérique, les exemples (5.3.1), (5.3.2), (5.3.3), (5.3.4), illustrent parfaitement le problème concernant le cas où la linéarisée présente un centre. L'exemple (5.3.5) illustre le cas où l'équation linéaire classique présente un noeud, sachant que c'est le type de stabilité qui est en défaut. Enfin, l'exemple (5.2.1) montre la possibilité offerte par la procédure pour l'étude de la stabilité dans le cas où la linéarisée n'existe pas. D'une manière générale, lorsque la procédure converge, la matrice obtenue est stable. Toutes ces considérations nous amènent à la conjecture suivante.

**Conjecture 3.3.3** *Si la procédure converge, et la limite de la suite  $A_{j+1}$  est exponentiellement stable (ou, si la suite  $A_{j+1} = M + r(A_j)$  a un point fixe stable) alors, le système non-linéaire est stable.*

**Commentaires.** Dans ce qui va suivre, on va présenter une étude concernant la perturbation d'une équation linéaire d'ordre 2 qui admet des valeurs propres imaginaires. Ceci va nous permettre de montrer que les conditions sous lesquelles la conjecture a été formulée peuvent être satisfaites, c'est à dire l'existence d'un point fixe stable. Le but des résultats qui seront présentés au chapitre 5 est de montrer que la conjecture (3.3.3) est vérifiée dans des exemples. Ces résultats ont été obtenus à partir de la procédure de calcul de la dérivée optimale exposé

dans le chapitre 1.

Avant de commencer l'étude, il nous a semblé utile de rappeler brièvement la perturbation non-linéaire d'une équation linéaire d'ordre deux.

### **Perturbation non-linéaire d'une équation linéaire.**

Considérons le système non-linéaire d'ordre 2 perturbé [1]

$$\begin{cases} \frac{dx}{dt} = ax + by + g_1(x, y) \\ \frac{dy}{dt} = cx + dy + g_2(x, y) \end{cases} \quad (3.24)$$

où,  $a, b, c, d$  sont des constantes réelles tel que  $ad - bc \neq 0$ ,  $g_1, g_2$  des fonctions continues définies dans un disque autour de l'origine  $(x, y) = (0, 0)$ .

$g_1, g_2$  sont appelés perturbations, et le système non-linéaire (3.24) système perturbé correspondant au système linéaire

$$\begin{cases} \frac{dx}{dt} = ax + by \\ \frac{dy}{dt} = cx + dy. \end{cases} \quad (3.25)$$

Si les perturbations  $g_1, g_2$  sont petites dans un certain sens [1], et satisfont un minimum d'hypothèses, on peut dire que les courbes intégrales du système (3.24) au voisinage de l'origine dans l'espace des phases sont similaires à celles du système linéaire (3.25). Mais en général, ce n'est pas vrai lorsque le système linéaire présente un centre et parfois un noeud [17], [1].

Nous allons considérer le cas où  $a = d = 0, b = -1, c = 1$ , c'est à dire

$$\begin{cases} F(x) = Mx + G(x) \\ x(0) = x_0 \end{cases} \quad (3.26)$$



où

$$M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Les valeurs propres de  $M$  sont égales à  $\pm i$

i)  $x = (x_1, x_2) \in \mathbb{R}^2$ ,

ii)  $F = (f_1, f_2)$  est une fonction définie sur un certain ouvert  $\Omega$  à valeurs dans  $\mathbb{R}^2$  contenant  $(0, 0)$ , avec les hypothèses suivantes

H1)  $F(0) = 0$ ,

H2)  $G$  est de classe  $C^1$ .

On se place dans le cas où  $G$  est une fonction qui vérifie

$$G(x) = -\Phi(\|x\|)x + G_3(x). \quad (3.27)$$

On supposera que dans tout voisinage de zéro,

$$\Phi(\|x\|) = O(\|x\|)$$

$$G_3(x) = o(\|x\|).$$

On en déduit d'après l'hypothèse H1), que  $G(0) = 0$ , et  $DG(0) = 0$ .

Dans la suite, on utilise une technique dite de localisation, qui nous permet d'obtenir un résultat d'approximation global par l'introduction d'une fonction qu'on notera  $G_\varepsilon$ . On en déduit un résultat local pour l'équation initiale. Celle-ci est définie par

$$G_\varepsilon(x) = \chi_\varepsilon(x)G(x) \quad (3.28)$$

où

$$\chi_\varepsilon(x) = \chi\left(\frac{x}{\varepsilon}\right), 0 \leq \chi \leq 1 \text{ et } \chi \in C^1(\mathbb{R}^2)$$

$\chi = 1$  dans un voisinage de 0

$\chi = 0$  en dehors de la boule unité

$G_\varepsilon$  est à support dans la boule de centre 0 et de rayon  $\varepsilon$ ; son support se concentre

sur  $\{0\}$  quand  $\varepsilon \rightarrow 0$ . Mais la non-linéarité de  $G_\varepsilon$  est la même que celle de  $G$  au voisinage de 0 (pour  $\varepsilon > 0$  fixé).

### Propriétés.

**Lemme 3.3.4**  $G_\varepsilon \in C^1$ ;  $G_\varepsilon = G$  au voisinage de 0. De plus, on a:

$$i) \sup_{x \in \mathbb{R}^2} \|G_\varepsilon(x)\| = \sup_{\|x\| \leq \varepsilon} \|G(x)\| \rightarrow 0, \text{ quand } \varepsilon \rightarrow 0.$$

ii)  $\exists C \geq 0$  tel que  $\|DG_\varepsilon(x)\| \leq C\varepsilon$ , pour tout  $\varepsilon > 0$ ,  $\varepsilon$  assez petit, pour tout  $x \in \mathbb{R}^2$ .

$$iii) \sup_{x \in \mathbb{R}^2} \|DG_\varepsilon(x)\| \rightarrow 0 \text{ quand } \varepsilon \rightarrow 0.$$

**Preuve.** Si  $\varepsilon \rightarrow 0$ ,  $x \rightarrow 0$  et comme  $G(0) = 0$ , alors

$$\sup_{\|x\| \leq \varepsilon} \|G(x)\| \rightarrow G(0) = 0, \text{ ceci pour } i).$$

Pour *ii*), on a

$$DG_\varepsilon(x) = \frac{1}{\varepsilon} D(\chi_\varepsilon * G)(x) + \chi_\varepsilon(x) DG(x),$$

où  $*$  représente le produit de convolution.

Ce qui donne avec  $\|\chi_\varepsilon(x) DG(x)\| \leq \sup_{\|x\| \leq \varepsilon} \|DG(x)\|$ ,

$$\|DG_\varepsilon(x)\| \leq \frac{1}{\varepsilon} \left\| D\chi\left(\frac{x}{\varepsilon}\right) \right\| \|G(x)\| + \sup_{\|x\| \leq \varepsilon} \|DG(x)\|.$$

Et comme  $\sup_{\|x\| \leq \varepsilon} \|DG(x)\| \rightarrow 0$  quand  $\varepsilon \rightarrow 0$ , (car  $DG(0) = 0$ ) et  $\|G(x)\| = O(\varepsilon^2)$  alors,

$$\|DG_\varepsilon(x)\| \leq C\varepsilon$$

pour tout  $x \in \mathbb{R}^2$ , et tout  $\varepsilon > 0$ ,  $\varepsilon$  assez petit.

Enfin, les hypothèses faites sur  $G$  entraînent que la constante de Lipschitz de  $G$  sur  $B(0, \varepsilon)$ , tend vers 0 quand  $\varepsilon \rightarrow 0$ , on en déduit *iii*).  $\square$

Dans la suite, on supposera  $\varepsilon > 0$  fixé et on omettra de préciser l'indice  $\varepsilon$ . On considérera donc l'équation (3.26) où la fonction  $G$  vérifie les conditions suivantes:

$G \in C^1$ ,

$G(0) = 0$ ,  $DG(0) = 0$ ,

$\|G\|$  est aussi petit que nécessaire,  $DG(x)$  est uniformément bornée.

### Relation de récurrence.

On applique la procédure de dérivation optimale, c'est à dire, on fixe  $x_0 \neq 0$  et on définit dans  $M_2(\mathbb{R})$  l'application  $\varphi : A \rightarrow \varphi(A)$  telle que

$$\begin{aligned}\varphi(A) &= \left[ \int_0^{+\infty} F(e^{tA}x_0) [e^{tA}x_0]^T dt \right] [\Gamma(A)]^{-1} \\ &= M + \left[ \int_0^{+\infty} G(e^{tA}x_0) [e^{tA}x_0]^T dt \right] [\Gamma(A)]^{-1}\end{aligned}\quad (3.29)$$

avec

$$\Gamma(A) = \int_0^{+\infty} [e^{tA}x_0] [e^{tA}x_0]^T dt. \quad (3.30)$$

On pose

$$r(A) = \left[ \int_0^{+\infty} G(e^{tA}x_0) [e^{tA}x_0]^T dt \right] [\Gamma(A)]^{-1},$$

et on construit la suite d'approximations  $A_{j+1} = \varphi(A_j)$ , il vient que

$$A_{j+1} = M + r(A_j). \quad (3.31)$$

### Procédure de calcul.

La matrice initiale étant  $A_0 = DF(x_0)$ , on suppose que  $G$  a été choisi tel que le spectre  $\sigma(A_0)$  soit très proche du spectre  $\sigma(M)$ . Par conséquent, on peut considérer que les valeurs propres de  $A_0$  sont complexes conjuguées. On note  $H = G_\varepsilon$  et on choisit  $x_0$  assez voisin de 0 pour que  $H = G$  dans un voisinage de  $x_0$ .

$e^{tA_0}x_0 \rightarrow 0$  exponentiellement quand  $t \rightarrow +\infty$  et donc la forme quadratique

associée à  $\Gamma(A_0)$  telle que

$$v^T \Gamma(A_0) v > 0, \text{ pour tout } v \neq 0$$

et par conséquent  $\Gamma(A_0)$  est définie positive.

### Etude de la trace.

Supposons que l'on ait construit  $A_1, \dots, A_j$  vérifiant la propriété suivante:  $\sigma(A_i)$  est constitué de deux valeurs propres complexes conjuguées à partie réelle négative, alors  $A_{j+1} = M + r(A_j)$  est bien définie.  $\sigma(A_{j+1}) \subset \mathcal{C} \setminus \mathbb{R}$  tant que  $r(A_j)$  reste de norme assez petite.

On a

$$tr(A_{j+1}) = tr \left( \left[ \int_0^{+\infty} G(e^{tA_j} x_0) (e^{tA_j} x_0)^T dt \right] [\Gamma(A_j)]^{-1} \right). \quad (3.32)$$

En utilisant la propriété  $tr(AB) = tr(BA)$  on obtient

$$\begin{aligned} tr(A_{j+1}) &= \int_0^{+\infty} tr \left( [\Gamma(A_j)]^{-1} \left[ G(e^{tA_j} x_0) (e^{tA_j} x_0)^T \right] \right) dt \\ &= \int_0^{+\infty} tr \left( \left( [\Gamma(A_j)]^{-1} G(e^{tA_j} x_0) \right) (e^{tA_j} x_0)^T \right) dt. \end{aligned} \quad (3.33)$$

Sachant que

$$\left( [\Gamma(A_j)]^{-1} G(e^{tA_j} x_0) \right) = C$$

est un vecteur colonne et

$$\left( (e^{tA_j} x_0)^T \right) = L$$

est un vecteur ligne, on a

$$tr(CL) = tr(LC) = LC.$$

Ceci permet d'écrire

$$\begin{aligned}
tr(A_{j+1}) &= \int_0^{+\infty} (e^{tA_j} x_0)^T [\Gamma(A_j)]^{-1} G(e^{tA_j} x_0) dt \\
&= \int_0^{+\infty} (e^{tA_j} x_0)^T [\Gamma(A_j)]^{-1} \chi_\varepsilon \left( \frac{e^{tA_j} x_0}{\varepsilon} \right) (-\Phi(\|e^{tA_j} x_0\|) (e^{tA_j} x_0)) dt \\
&= - \int_0^{+\infty} \Phi(\|e^{tA_j} x_0\|) (e^{tA_j} x_0)^T [\Gamma(A_j)]^{-1} \chi_\varepsilon \left( \frac{e^{tA_j} x_0}{\varepsilon} \right) (e^{tA_j} x_0) dt.
\end{aligned} \tag{3.34}$$

Les termes sous l'intégrale étant positifs, alors  $tr(A_{j+1}) < 0$ .

Sachant que le spectre  $\sigma(A_j)$  possède deux valeurs propres complexes conjuguées, on peut donc écrire  $e^{tA_j}$  avec  $\delta_j = tr(A_j)$ , sous la forme

$$e^{tA_j} = e^{-\delta_j t} \begin{bmatrix} \cos \beta_j t & -\sin \beta_j t \\ \sin \beta_j t & \cos \beta_j t \end{bmatrix} = e^{-\delta_j t} R_j(t), \tag{3.35}$$

et avec  $[\gamma] = [(R_j(t) x_0)^T [\Gamma(A_j)]^{-1} (R_j(t) x_0)]$ , la relation (3.34) s'écrit

$$tr(A_{j+1}) = - \int_0^{+\infty} \Phi(e^{-\delta_j t} \|x_0\|) (e^{-2\delta_j t}) \chi_\varepsilon \left( \frac{e^{-\delta_j t} R_j(t) x_0}{\varepsilon} \right) [\gamma] dt. \tag{3.36}$$

### Existence de la dérivée optimale.

**Etude de  $\|[\Gamma(A)]^{-1}\|$ .** En utilisant la relation (3.35) pour écrire  $e^{tA}$ , on a

$$\begin{aligned}
v^T \Gamma(A_0) v &= \int_0^{+\infty} [v^T e^{-\delta t} R(t) x_0]^2 dt \\
&= \int_0^{+\infty} e^{-2\delta t} [v^T R(t) x_0]^2 dt \\
&= \int_0^{+\infty} e^{-2\delta t} [(R(t) v)^T x_0]^2 dt.
\end{aligned} \tag{3.37}$$

On choisit un cône  $(C)$  convexe d'angle  $\kappa < \frac{\pi}{4}$ , symétrique autour de  $x_0$ . Si  $w \in (C)$ , on a

$$|wx_0| \geq \|x_0\| \|w\| \cos \kappa. \tag{3.38}$$

Si  $R(t)v \in (C)$  quand  $t \in [t_1, t_2]$  ou  $t \in [t_1 + \frac{2k\pi}{\beta}, t_2 + \frac{2k\pi}{\beta}]$ , alors

$$\begin{aligned} \int_0^{+\infty} e^{-2\delta t} [(R(t)v)x_0]^2 dt &\geq \sum_{k=0}^{\infty} \int_{t_1 + \frac{2k\pi}{\beta}}^{t_2 + \frac{2k\pi}{\beta}} e^{-2\delta t} (\cos \kappa)^2 \|v\|^2 \|x_0\|^2 dt \\ &\geq \left( \sum_{k=0}^{\infty} e^{-\frac{4\delta\pi}{\beta}k} \right) \left( \frac{e^{-2\delta t_1} - e^{-2\delta t_2}}{2\delta} \right) (\cos \kappa)^2 \|v\|^2 \|x_0\|^2. \end{aligned} \quad (3.39)$$

Ceci permet d'avoir

$$\int_0^{+\infty} e^{-2\delta t} [(R(t)v)x_0]^2 dt \geq \frac{1}{1 - e^{-\frac{4\delta\pi}{\beta}}} \left( \frac{e^{-2\delta t_1} - e^{-2\delta t_2}}{2\delta} \right) (\cos \kappa)^2 \|v\|^2 \|x_0\|^2. \quad (3.40)$$

$(1 - e^{-\frac{4\delta\pi}{\beta}})$  et  $(e^{-2\delta t_1} - e^{-2\delta t_2})$  étant de l'ordre de  $\delta$ , on aura

$$\int_0^{+\infty} e^{-2\delta t} [(R(t)v)x_0]^2 dt \geq \frac{1}{2\delta} (\cos \kappa)^2 \|v\|^2 \|x_0\|^2 = \frac{C_1}{\delta} \|v\|^2 \|x_0\|^2. \quad (3.41)$$

Avec  $\lambda_{\min} = \min \{ \lambda : \lambda \in \sigma(\Gamma(A)) \}$ , il vient que

$$\lambda_{\min} \geq \frac{C_1}{\delta} \|x_0\|^2, \quad (3.42)$$

ce qui implique  $\forall \lambda \in \sigma([\Gamma(A)]^{-1})$ ,  $\lambda < \frac{1}{\lambda_{\min}}$  et

$$\lambda([\Gamma(A)]^{-1}) \leq \frac{\delta}{C_1 \|x_0\|^2}. \quad (3.43)$$

Ce qui donne

$$v^T [\Gamma(A)]^{-1} v \leq \frac{\delta}{C_1 \|x_0\|^2} \|v\|^2. \quad (3.44)$$

D'où, l'on en déduit qu'avec la matrice  $[\Gamma(A)]^{-1}$  symétrique positive,

$$\|[\Gamma(A)]^{-1}\| \leq \frac{\delta}{C_1 \|x_0\|^2}. \quad (3.45)$$

**Etude de  $r(A)$ .**

**Lemme 3.3.5** *Sous les hypothèses sur la fonction  $G$ , et pour tout  $A \in \mathcal{M}_2(\mathbb{R})$ , vérifiant  $\|A - M\| \leq \rho$ , on a*

$$\|r(A)\| \leq C_4 \|x_0\|. \quad (3.46)$$

**Preuve.** De la relation (3.35) on déduit

$$\|e^{tA}\| \leq e^{-\delta t}. \quad (3.47)$$

De (3.45) et de l'estimation (3.47), on déduit la majoration

$$\begin{aligned} \|r(A)\| &\leq \frac{\delta}{C_1 \|x_0\|^2} \int_0^{+\infty} \Phi(e^{-\delta t} \|x_0\|) e^{-2\delta t} \|x_0\|^2 dt \\ &\leq \frac{\delta}{C_1} \int_0^{+\infty} C_2 \|x_0\| e^{-3\delta t} dt = \frac{C_2 \delta \|x_0\|}{3C_1 \delta} \\ &\leq C_4 \|x_0\|, \end{aligned} \quad (3.48)$$

avec  $C_4 = \frac{C_2}{3C_1}$ .

**Proposition 3.3.6** *Sous les hypothèses des lemmes (3.3.4) et (3.3.5), il existe  $\rho > 0$  tel que l'équation*

$$A = M + r(A) \quad (3.49)$$

*admette au moins une solution dans  $B(M, \rho)$ .*

**Preuve.** Si on impose  $C_4 \|x_0\| < \rho$ , on a

$$\|r(A)\| < \rho. \quad (3.50)$$

Alors, l'application  $\varphi : A \longrightarrow M + r(A)$ , définie de  $B(M, \rho)$  dans  $\mathcal{M}_2(\mathbb{R})$  est en fait à valeurs dans  $B(M, \rho)$ .

Cette application est continue. La continuité de  $\varphi$  est conséquence de la continuité

de chacune des applications  $A \longrightarrow (\Gamma(A))^{-1}$  et  $A \longrightarrow \int_0^{+\infty} G(x) [e^{tA}x_0]^T dt$ .

Pour la première application, il suffit de montrer la continuité de l'application qui à  $A \longrightarrow \Gamma(A)$ , puisque nous savons que  $\Gamma(A)$  est inversible uniformément (dans un voisinage de  $A = 0$ ). Or,

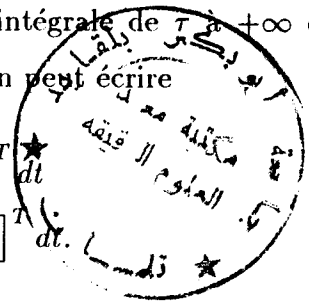
$$\Gamma(A) = \int_0^{+\infty} [x][x]^T dt = \int_0^\tau [x][x]^T dt + \int_\tau^{+\infty} [x][x]^T dt$$

l'intégrale de 0 à  $\tau$  dépend continument de  $A$ , et l'intégrale de  $\tau$  à  $+\infty$  est uniformément petite quand  $\tau$  est assez grand. De la même manière pour la deuxième application

$$r(A) = \int_0^{+\infty} G(x) [e^{tA}x_0]^T dt = \int_0^\tau G(x) [e^{tA}x_0]^T dt + \int_\tau^{+\infty} G(x) [e^{tA}x_0]^T dt.$$

L'intégrale de 0 à  $\tau$  dépend continument de  $A$ , et l'intégrale de  $\tau$  à  $+\infty$  est uniformément petite quand  $\tau$  est assez grand. D'où l'on peut écrire

$$\begin{aligned} \Gamma(A) &= \lim_{\tau \rightarrow +\infty} \text{unif} \int_0^\tau (e^{tA}x_0) (e^{tA}x_0)^T dt \\ r(A) &= \lim_{\tau \rightarrow +\infty} \text{unif} \int_0^\tau G(e^{tA}x_0) [e^{tA}x_0]^T dt. \end{aligned}$$



$\Gamma(A)$  et  $r(A)$  sont limite uniforme d'une suite de fonctions continues et sont donc continues.

$\varphi$  applique la boule fermée dans elle même et  $\varphi$  est continue. D'après le théorème de Brouwer [13], on conclut à l'existence d'un point fixe.  $\square$

### Unicité de la dérivée optimale.

**Proposition 3.3.7** *Sous les mêmes hypothèses sur  $M$  et sous les hypothèses du lemme (3.3.4) et (3.3.5), pour tout  $\rho > 0$ , l'application  $\varphi$  est lipschitzienne de rapport strictement inférieur à 1 de la boule  $B(M, \rho)$  dans elle même. Dans ce cas, l'équation (3.49) a une et une seule solution dans  $B(M, \rho)$ .*



**Preuve:** Pour montrer que  $A$  est unique, il suffit de vérifier que l'application  $r$  est lipschitzienne de rapport strictement plus petit que 1. Pour cela, considérons l'application qui à la matrice  $B$  fait correspondre  $r(B)$  tel que

$$r(B) = \left[ \int_0^{+\infty} G(e^{Bt}x_0) [e^{Bt}x_0]^T dt \right] [\Gamma(B)]^{-1}. \quad (3.51)$$

On va estimer la différence

$$\begin{aligned} r(B_2) - r(B_1) = & \left[ \int_0^{+\infty} G(e^{B_2 t}x_0) [e^{B_2 t}x_0]^T dt - \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] [\Gamma(B_2)]^{-1} \\ & + \left[ \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] \left[ [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right]. \end{aligned} \quad (3.52)$$

Pour l'étude du caractère lipschitzien dans la boule  $B(M, \rho)$ , on pose

$$\begin{aligned} U = & \left[ \int_0^{+\infty} G(e^{B_2 t}x_0) [e^{B_2 t}x_0]^T dt - \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] [\Gamma(B_2)]^{-1} \end{aligned} \quad (3.53)$$

et

$$R = \left[ \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right] \left[ [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right]. \quad (3.54)$$

**Estimation du terme  $R$ .** on a

$$\|R\| \leq \left\| \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right\| \left\| [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right\| \quad (3.55)$$

avec

$$\begin{aligned} \left\| \int_0^{+\infty} G(e^{B_1 t}x_0) [e^{B_1 t}x_0]^T dt \right\| &= \left\| \int_0^{+\infty} \chi_\varepsilon(x) \Phi(\|x\|) [e^{B_1 t}x_0] [e^{B_1 t}x_0]^T dt \right\| \\ &\leq \int_0^{+\infty} k \|x_0\| e^{-3\delta t} \|x_0\|^2 dt = k \frac{\|x_0\|^3}{3\delta}. \end{aligned} \quad (3.56)$$

et

$$\left\| [\Gamma(B_2)]^{-1} - [\Gamma(B_1)]^{-1} \right\| = \left\| [\Gamma(B_2)]^{-1} [\Gamma(B_1) - \Gamma(B_2)] [\Gamma(B_1)]^{-1} \right\|. \quad (3.57)$$

En posant

$$S = [\Gamma(B_1)] - [\Gamma(B_2)],$$

on obtient

$$\left\| [\Gamma(B_2)]^{-1} [[\Gamma(B_1)] - [\Gamma(B_2)]] [\Gamma(B_1)]^{-1} \right\| \leq \frac{\delta}{C_1 \|x_0\|^2} [S] \frac{\delta}{C_1 \|x_0\|^2}. \quad (3.58)$$

En utilisant l'estimation déjà faite au chapitre 2, et en particulier la relation (2.45), il vient

$$\begin{aligned} \|\Gamma(B_1) - \Gamma(B_2)\| &\leq \|B_2 - B_1\| \frac{C'^2}{(1 - C''\rho)^2} \left[ \int_0^{+\infty} e^{-\frac{\eta}{2}t} \|x_0 x_0^T\| \|e^{B_2 t}\| dt \right] \\ &\quad + \|B_2 - B_1\| \frac{C'^2}{(1 - C''\rho)^2} \left[ \int_0^{+\infty} e^{-\frac{\eta}{2}t} \|x_0 x_0^T\| \|e^{B_1 t}\| dt \right]. \end{aligned}$$

On obtient avec  $\|e^{B_2 t}\| \leq e^{-\delta t}$ , et  $\|e^{B_1 t}\| \leq e^{-\delta t}$  (voir relation 3.47),

$$\|\Gamma(B_1) - \Gamma(B_2)\| \leq 2 \|B_2 - B_1\| \frac{C'^2}{(1 - C''\rho)^2} \|x_0\|^2 \frac{2}{(\eta + 2\delta)}, \quad (3.59)$$

d'où

$$\|\Gamma(B_1) - \Gamma(B_2)\| \leq C_4 \|B_2 - B_1\|. \quad (3.60)$$

Ceci implique pour la relation (3.57)

$$\begin{aligned} &\left\| \Gamma(B_2)^{-1} - \Gamma(B_1)^{-1} \right\| \\ &\leq \frac{\delta^2}{C_1^2 \|x_0\|^4} \frac{2C'^2}{(1 - C''\rho)^2} \|x_0\|^2 \frac{2}{(\eta + 2\delta)} \|B_2 - B_1\|. \end{aligned} \quad (3.61)$$

En utilisant l'estimation (3.61), il vient

$$\|R\| \leq \frac{\delta^2}{C_1^2 \|x_0\|^4} \frac{C'^2}{(1 - C''\rho)^2} \|x_0\|^2 \frac{2}{(\eta + 2\delta)} k \frac{\|x_0\|^3}{3\delta} \|B_2 - B_1\| \quad (3.62)$$

et

$$\|R\| \leq \frac{2}{3} C \|x_0\| \frac{\delta}{(\eta + 2\delta)} \|B_2 - B_1\|. \quad (3.63)$$

Avec  $\eta$  fixé,  $\delta$  pouvant tendre vers 0

$$\|R\| \leq \frac{2\delta}{3} C \|x_0\| \|B_2 - B_1\|, \quad (3.64)$$

$C = \frac{k}{C_1^2 (\eta + 2\delta)} \frac{C'^2}{(1 - C''\rho)^2}$ ,  $C_5 = \frac{k}{C_1^2 (\eta + 2\delta)} \frac{C'^2}{(1 - C''\rho)^2} \frac{2\delta}{3} \|x_0\|$  ne dépendent que de  $\delta$ ,  $\rho$  et  $x_0$ .

Finalement

$$\|R\| \leq C_5 \|B_2 - B_1\|.$$

**Estimation du terme  $U$ .**

En posant  $\Psi(x) = \chi_\varepsilon(x) \Phi(\|x\|)$ ,  $U$  s'écrit

$$U = \int_0^{+\infty} [\Psi(x_1) - \Psi(x_2)] [e^{B_1 t} x_0] [e^{B_1 t} x_0]^T dt [\Gamma(B_2)]^{-1} \\ + \int_0^{+\infty} \Psi(x_2) \left[ [e^{B_1 t} x_0] [e^{B_1 t} x_0]^T - [e^{B_2 t} x_0] [e^{B_2 t} x_0]^T \right] dt \left[ [\Gamma(B_2)]^{-1} \right]. \quad (3.65)$$

L'estimation de  $U$  donne

$$\|U\| \\ \leq \int_0^{+\infty} \left\| [\Psi(x_1) - \Psi(x_2)] [e^{B_1 t} x_0] [e^{B_1 t} x_0]^T dt \right\| \left\| [\Gamma(B_2)]^{-1} \right\| \\ + \left\| \int_0^{+\infty} \Psi(x_2) \left[ [e^{B_1 t} x_0] [e^{B_1 t} x_0]^T - [e^{B_2 t} x_0] [e^{B_2 t} x_0]^T \right] dt \right\| \left\| [\Gamma(B_2)]^{-1} \right\|. \quad (3.66)$$

De la relation

$$|\Psi(x_1) - \Psi(x_2)| \leq \sup_{x \in R^2} |D\Psi(x)| \|e^{B_1 t} x_0 - e^{B_2 t} x_0\| \quad (3.67)$$

où

$$D\Psi(x) = \frac{1}{\varepsilon} D(\chi_\varepsilon * \Phi)(x) + \chi_\varepsilon(x) D\Phi(x)$$

et

$$|D\Psi(x)| \leq \frac{1}{\varepsilon} \left| D\chi\left(\frac{x}{\varepsilon}\right) \right| |\Phi(x)| \leq k_1,$$

il vient que  $|D\Psi(x)|$  est bornée, et

$$\begin{aligned} \|U\| &\leq \int_0^{+\infty} k_1 \|e^{B_1 t} x_0 - e^{B_2 t} x_0\| e^{-2\delta t} \|x_0\|^2 dt \|\Gamma(B_2)^{-1}\| \\ &+ \int_0^{+\infty} k_2 \|x\| \left\| [e^{B_1 t} x_0] [e^{B_1 t} x_0]^T - [e^{B_2 t} x_0] [e^{B_2 t} x_0]^T \right\| dt \|\Gamma(B_2)^{-1}\|. \end{aligned} \quad (3.68)$$

Finalement,

$$\begin{aligned} \|U\| &\leq \int_0^{+\infty} k_1 \|B_2 - B_1\| e^{-\frac{\eta}{2}t} \frac{(C')^2}{(1 - C''\rho)^2} e^{-2\delta t} \|x_0\|^3 dt \frac{\delta}{C_1 \|x_0\|^2} \\ &+ \int_0^{+\infty} 2k_2 \|B_2 - B_1\| e^{-\frac{\eta}{2}t} \frac{(C')^2}{(1 - C''\rho)^2} e^{-2\delta t} dt \|x_0\|^3 \frac{\delta}{C_1 \|x_0\|^2} \\ &\leq (k_1 + 2k_2) \left( \frac{(C')^2 \|x_0\|}{C_1 (1 - C''\rho)^2} \right) \left( \frac{2\delta}{\eta + 2\delta} \right) \|B_2 - B_1\| \\ &\leq 2\delta K \|x_0\| \|B_2 - B_1\|. \end{aligned} \quad (3.69)$$

avec  $\eta$  fixé,  $\delta$  pouvant tendre vers 0,

$$K = (k_1 + 2k_2) \left( \frac{(C')^2}{C_1 (1 - C''\rho)^2 (\eta + 2\delta)} \right) \quad (3.70)$$

et

$$C_6 = 2K\delta \|x_0\|. \quad (3.71)$$

$C_6$  ne dépend que de  $\rho, \delta$  et  $x_0$ , d'où

$$\|U\| \leq C_6 \|B_2 - B_1\|. \quad (3.72)$$

Finalement

$$\|r(B_2) - r(B_1)\| \leq (C_5 + C_6) \|B_2 - B_1\|. \quad (3.73)$$

En prenant  $(C_5 + C_6) < 1$ , et  $C_4 \|x_0\| < \rho$ , on obtient que l'application  $r$  est lipschitzienne, de rapport strictement  $< 1$  de la boule  $B(M, \rho)$  dans elle-même. On conclut que  $\hat{A}$  est unique.

**Théorème 3.3.8** *Sous les hypothèses des propositions (3.3.6) et (3.3.7), on a*

$$\varphi : B(M, \rho) \longrightarrow B(M, \rho),$$

*$\varphi$  est une contraction stricte.*

*Alors, une condition suffisante pour que la suite  $A_j$  converge vers l'unique point fixe dans la boule est que  $A_0 \in B(M, \rho)$ .*

**Preuve.** On a

$$A_0 \in B(M, \rho) \iff [A_0 - M] = DG(x_0) \in B(0, \rho). \quad (3.74)$$

D'après les hypothèses sur la fonction  $G$ , pour tout  $x$  et pour tout  $\varepsilon > 0$ ,  $\varepsilon$  petit

$$\|DG(x)\| \leq C\varepsilon. \quad (3.75)$$

Donc, si l'on impose la condition

$$\|DG(x_0)\| < C' < \rho, \quad (3.76)$$

alors,  $A_0 \in B(M, \rho)$ .

Dans ce cas, la suite

$$\begin{cases} A_{j+1} = M + r(A_j) \\ A_0 \in B(M, \rho) \end{cases} \quad (3.77)$$

converge vers l'unique point fixe  $\tilde{A}$  solution de l'équation (3.49).  $\square$

**Conclusion:**

Le problème traité a permis de vérifier que les conditions sous lesquelles la conjecture a été formulée peuvent être satisfaites, *i.e* : la convergence de la procédure.

### 3.4 Commentaires.

La réponse à la question de la relation entre les propriétés qualitatives de l'équation non-linéaire et la dérivée optimale peut se résumer aux résultats suivants:

La dérivée optimale est d'ordre deux au moins en fonction de la donnée initiale, en général elle est de l'ordre de la non-linéarité.

Dans le cas scalaire, la dérivée optimale est une fonction de Lyapunov pour l'équation non-linéaire.

Dans le cas vectoriel, l'étude faite porte sur la perturbation d'une matrice dont les valeurs propres sont imaginaires. On a montré l'existence d'un point fixe stable, ce qui nous a permis d'avoir des renseignements sur la possibilité offerte par la dérivée optimale pour l'étude de la stabilité de ce type d'équations.

**QUATRIEME CHAPITRE**  
**APPROXIMATION OPTIMALE.**

# Chapitre 4

## Méthode d'approximation.

### 4.1 Introduction.

Dans le premier chapitre, nous avons introduit une procédure qui permet de calculer une certaine approximation définie comme la dérivée optimale d'une équation différentielle ordinaire non-linéaire. Cette approximation globale par rapport à des courbes démarrant du point initial  $x_0$  et allant vers 0 quand  $t \rightarrow +\infty$ , a été conçue initialement pour permettre d'associer une équation linéaire à une E.D.O non-linéaire au voisinage d'un point d'équilibre, en particulier quand la non-linéarité n'est pas assez régulière autour de ce point.

Le but de ce chapitre est de construire une méthode d'approximation <sup>1</sup> de la solution d'une équation différentielle ordinaire non-linéaire basée sur la notion de dérivée optimale, c'est à dire résoudre numériquement le problème avec valeur initiale du type (4.1). On calcule une approximation de la solution à chaque pas en remplaçant l'équation non-linéaire par la dérivée optimale correspondante dans l'intervalle considéré. Cette méthode sera appelée "approximation optimale".

---

<sup>1</sup>Ce sujet a fait l'objet d'une proposition d'article actuellement soumis à "Computers And Mathematics With Applications".



## 4.2 Idée du problème.

En fait, nous nous intéressons à la résolution numérique du problème avec valeur initiale de la forme

$$\begin{cases} \frac{dx}{dt} = F(x) \\ x(0) = x_0 \end{cases} \quad (4.1)$$

pour  $t \in [0, T]$ ,  $x \in \mathbb{R}^n$ .  $F$  est définie dans un certain ouvert  $\Omega$  à valeurs dans  $\mathbb{R}^n$ .

Nous considérons une subdivision  $t_0 = 0 < t_1 < \dots < t_i < t_{i+1} < \dots < t_n = T$  de l'intervalle  $[0, T]$ , et on pose  $\tau_{i+1} = t_{i+1} - t_i$ .

Plusieurs méthodes de résolution de ce problème existent (Euler, Runge-Kutta, ...), l'idée étant d'approcher la fonction non-linéaire à l'aide de ses dérivées successives.

Ce que nous proposons ici [25], c'est de remplacer la fonction  $F$  non-linéaire par une application linéaire dans le sens de la dérivée optimale dans chacun des intervalles de la subdivision. Ceci permet le calcul d'une approximation  $\tilde{x}_i$  de la solution en tout point  $t_i$  de la subdivision. Une approximation  $\tilde{x}$  de la solution est déduite par interpolation linéaire entre  $t_i$  et  $t_{i+1}$ . Cette approximation n'utilise pas de valeurs de la solution de l'équation non-linéaire et donc, elle est basée uniquement sur le calcul de la valeur de la solution obtenue en utilisant l'approximation linéaire optimale.

## 4.3 Préliminaires.

### 4.3.1 Erreur sur les conditions initiales.

Nous allons supposer que l'on commet une erreur sur les conditions initiales c'est à dire, qu'on démarre avec  $\bar{x}(0) = \bar{x}_0$  une approximation de  $x(0) = x_0$ . Cela veut

dire, qu'en réalité, c'est un problème de la forme

$$\begin{cases} \frac{d\bar{x}}{dt} = F(\bar{x}) \\ \bar{x}(0) = \bar{x}_0, \end{cases} \quad (4.2)$$

que l'on va résoudre. Par la suite nous calculerons l'erreur introduite par cette formulation.

Pour cela, nous allons supposer que

H1)  $\|F(x)\| \leq M, M > 0$ .

H2)  $F$  est continue, Lipschitzienne de constante de Lipschitz  $\gamma$ .

H3) Il existe  $M_2 < +\infty$  tel que  $\|F(y+x) - F(x) - DF(x)y\| \leq M_2 \|y\|^2$ , pour tout  $x, y \in \mathbb{R}^n$ .

### 4.3.2 Changement de Variable et de fonction.

Nous commençons par faire un changement de variable et de fonction en centrant  $F$  autour de  $\bar{x}_0$ . Nous posons

$$\begin{cases} \bar{x} = \bar{y} + \bar{x}_0 \\ G(\bar{y}) = F(\bar{x}_0 + \bar{y}) - F(\bar{x}_0) \\ b = F(\bar{x}_0). \end{cases} \quad (4.3)$$

L'équation (4.2) devient

$$\begin{cases} \frac{d\bar{y}}{dt} = G(\bar{y}) + b, \\ \bar{y}(0) = 0. \end{cases} \quad (4.4)$$

Notons que la solution du système (4.4) vérifie la relation

$$\bar{x}(t) = \bar{y}(t) + \bar{x}_0, \quad (4.5)$$

où  $\bar{x}(t)$  représente la solution du problème (4.2) et  $\bar{y}(t)$  la solution du système (4.4). En appliquant la procédure de dérivation optimale au système (4.4), la fonction  $G$  est remplacée par une application linéaire  $\tilde{A}$ . Ainsi, nous obtenons l'équation qui définit le problème optimal

$$\begin{cases} \frac{du}{dt} = \tilde{A}u + b, \\ u(0) = 0, \end{cases} \quad (4.6)$$

et  $u(t)$  est sa solution.

#### 4.4 Procédure de la dérivée optimale dans l'intervalle $[\alpha, \beta]$ .

Soit  $[\alpha, \beta]$  un intervalle quelconque de temps réel,  $x \in \mathbb{R}^n$ , la fonction  $G(\bar{y})$  peut s'écrire

$$G(\bar{y}) = F(\bar{y} + x) - F(x). \quad (4.7)$$

Nous allons maintenant rappeler brièvement la procédure suivie pour permettre le calcul de la dérivée optimale de  $G$ .

On minimise la fonctionnelle

$$J(A) = \int_{\alpha}^{\beta} \|G(\bar{y}) - A\bar{y}(t)\|^2 dt \quad (4.8)$$

le long d'une solution donnée. Ce qui donne

$$A = \left( \int_{\alpha}^{\beta} [G(\bar{y}(t))] [\bar{y}(t)]^T dt \right) \left( \int_{\alpha}^{\beta} [\bar{y}(t)] [\bar{y}(t)]^T dt \right)^{-1}. \quad (4.9)$$

Ce calcul est alors utilisé itérativement sur l'intervalle  $[\alpha, \beta]$  comme suit.

**Premier pas:**

**Calcul de la matrice initiale  $A^0$ .**  $A^0$  est la matrice initiale: généralement, on

choisit pour  $A^0$  la matrice jacobienne de  $G$  en un point.

**Deuxième pas:**

Calcul de  $A^1$  à partir de la solution de l'équation

$$\begin{cases} \frac{d\bar{y}}{dt} = (A^0)\bar{y} + b \\ \bar{y}(0) = 0 \end{cases} \quad (4.10)$$

en minimisant la fonctionnelle

$$J(A) = \int_{\alpha}^{\beta} \|G(\bar{y}(t)) - A\bar{y}(t)\|^2 dt, \quad (4.11)$$

où  $\bar{y}(t)$  est la solution de l'équation (4.10).  $A^1$  est déterminée de manière unique par la formule (4.9).

**Troisième pas:**

En supposant que les matrices  $A^1, \dots, A^{(j-1)}$  ont été calculées, alors pour calculer  $A^{(j)}$  à partir de  $A^{(j-1)}$ , on doit résoudre

$$\begin{cases} \frac{d\bar{y}}{dt} = A^{(j-1)}\bar{y} + b \\ \bar{y}(0) = 0. \end{cases} \quad (4.12)$$

Si  $\bar{y}_j(t)$  est solution de l'équation (4.12), la minimisation de la fonctionnelle

$$J_j(A) = \int_{\alpha}^{\beta} \|G(\bar{y}_j(t)) - A\bar{y}_j(t)\|^2 dt \quad (4.13)$$

donne  $A^{(j)}$ .

En effet, nous avons la relation suivante entre  $A^{(j-1)}$  et  $A^{(j)}$

$$A^{(j)} = \left( \int_{\alpha}^{\beta} [G(\bar{y}_j(t))] [\bar{y}_j(t)]^T dt \right) \left( \int_{\alpha}^{\beta} [\bar{y}_j(t)] [\bar{y}_j(t)]^T dt \right)^{-1}, \quad (4.14)$$

où

$$G(\bar{y}_j) = F(\bar{y}_j + x) - F(x), \quad (4.15)$$

et

$$\begin{aligned} \bar{y}_j(\beta) &= \int_{\alpha}^{\beta} \exp[(\beta - s) A^{(j-1)}] (b) ds, \\ b &= F(x). \end{aligned} \quad (4.16)$$

La limite de cette suite représente la matrice optimale  $\tilde{A}(x, \theta)$ , relative au problème démarrant en  $x_0$  et intégré sur l'intervalle  $[\alpha, \beta]$  de longueur  $\theta$ .

La résolution du système linéaire obtenu

$$\begin{cases} \frac{du}{dt} = \tilde{A}(x, \theta) u + b \\ u(0) = 0, \end{cases} \quad (4.17)$$

par rapport à  $u$ , donne

$$\begin{aligned} u(\beta) &= \int_{\alpha}^{\beta} \exp[(\beta - s) \tilde{A}(x, \theta)] (b) ds \\ b &= F(x). \end{aligned} \quad (4.18)$$

Ce qui permet de calculer une approximation  $u(\beta)$ , de  $\bar{y}(t)$  au point  $\beta$ , et

$$\tilde{y}_{\beta} = u(\beta). \quad (4.19)$$

## 4.5 Procédure d'approximation optimale dans l'intervalle $[0, T]$ .

Nous allons maintenant construire une approximation optimale sur l'ensemble de l'intervalle  $[0, T]$ , en utilisant la dérivée optimale sur chacun des sous-intervalles  $[\alpha, \beta]$  constituant l'intervalle  $[0, T]$ . Pour le moment, nous considérons une subdivision arbitraire  $[t_i, t_{i+1}]$ ,  $i = 0, \dots, n-1$ ,  $t_0 = 0$ ,  $t_n = T$ . Nous notons  $\tau_{i+1} = t_{i+1} - t_i$ .

Dans cette subdivision, la fonction non-linéaire  $G$  peut s'écrire

$$G(\bar{y}) = F(\bar{y} + \bar{x}(t_i)) - F(\bar{x}(t_i)). \quad (4.20)$$

Notons qu'il n'existe pas une fonction unique  $G$ , mais toutes les fonctions obtenues en centrant  $F$  autour des points  $\bar{x}(t_i)$  sur chaque subdivision  $[t_i, t_{i+1}]$  de l'intervalle  $[0, T]$ . Nous commençons avec la valeur initiale  $\bar{x}(t_0) = \bar{x}_0 = \tilde{x}_0$ .

#### 4.5.1 Algorithme.

**Premier pas:**

La solution du problème de la dérivée optimale sur l'intervalle  $[t_0, t_1]$  à partir de  $\tilde{x}_0 = \bar{x}_0$ , et en démarrant avec  $A_0 = DF(x_0)$ , nous permet de calculer  $\tilde{A}_1(\tilde{x}_0, \tau_1)$ .

La solution du système linéaire correspondant sur l'intervalle  $[t_0, t_1]$

$$\begin{cases} \frac{du}{dt} = (\tilde{A}_1(\tilde{x}_0, \tau_1)) u + b_0 \\ u(0) = 0 \\ b_0 = F(\tilde{x}_0), \end{cases} \quad (4.21)$$

donne une approximation  $\tilde{y}_1 = u(t_1)$ , de  $\bar{y}_1(t_1)$  et

$$\tilde{x}_1 = \tilde{y}_1 + \tilde{x}_0. \quad (4.22)$$

**Deuxième pas:**

La solution du problème de la dérivée optimale sur l'intervalle  $[t_1, t_2]$  à partir de  $\tilde{x}(t_1) = \tilde{x}_1$ , et en démarrant avec  $\tilde{A}_1(\tilde{x}_0, \tau_1)$  nous permet de calculer  $\tilde{A}_2(\tilde{x}_1, \tau_2)$ .

Le système linéaire correspondant peut s'écrire

$$\begin{cases} \frac{du}{dt} = \hat{A}_2(\tilde{x}_1, \tau_2) u + b_1 \\ u(0) = 0 \\ b_1 = F(\tilde{x}_1) \end{cases} \quad (4.23)$$

dont la solution donne une approximation  $\tilde{y}_2 = u(t_2)$ , de  $\bar{y}(t_2)$  et par conséquent

$$\tilde{x}_2 = \tilde{y}_2 + \tilde{x}_1. \quad (4.24)$$

**Troisième pas:**

En supposant que  $\tilde{x}_1, \dots, \tilde{x}_i$ , ont été calculés, pour calculer  $\tilde{x}_{i+1}$  à partir de  $\tilde{x}_i$ , nous devons résoudre premièrement le problème de la dérivée optimale sur l'intervalle  $[t_i, t_{i+1}]$ .

Nous obtenons la matrice optimale correspondante  $\tilde{A}_{i+1}(\tilde{x}_i, \tau_{i+1})$  qui définit une équation linéaire de la forme

$$\begin{cases} \frac{du}{dt} = \tilde{A}_{i+1}(\tilde{x}_i, \tau_{i+1}) u + b_i \\ u(0) = 0 \\ b_i = F(\tilde{x}_i) \end{cases} \quad (4.25)$$

dont la solution sur l'intervalle considéré s'écrit

$$u(t_{i+1}) = \int_{t_i}^{t_{i+1}} \exp\left[(t_{i+1} - s) \left(\tilde{A}_{i+1}(\tilde{x}_i, \tau_{i+1})\right)\right] (b_i) ds. \quad (4.26)$$

Dans ce cas, une approximation  $\tilde{y}_{i+1} = u(t_{i+1})$ , de  $\bar{y}(t_{i+1})$  est calculée et par conséquent, le schéma général de l'approximation optimale de la solution du

système (4.2) peut s'écrire

$$\begin{cases} \tilde{x}_{i+1} = \tilde{y}_{i+1} + \tilde{x}_i \\ \tilde{x}_0 = \tilde{x}_0 \end{cases}, 0 \leq i \leq n-1. \quad (4.27)$$

Finalement, la procédure d'approximation optimale permet de construire une fonction  $\tilde{x}(t)$  en recollant bout à bout les approximations trouvées sur chaque intervalle  $[t_i, t_{i+1}]$ .  $\tilde{x}(t)$  est défini par

$$\tilde{x}(t) = u_{i+1}(t) + \tilde{x}_i, \text{ pour } t_i \leq t \leq t_{i+1}, 0 \leq i \leq n-1. \quad (4.28)$$

$u$  est solution de l'équation (4.25),  $\tilde{x}$  est l'approximation optimale de la solution  $x$  sur  $[0, T]$ .

**Remarque 4.5.1** Notons que  $\tilde{x}$  et  $\tilde{x}$  sont en fait fonction de la subdivision  $\tau = (\tau_{i+1})$ . Cette dépendance peut-être caractérisée par la notation  $\tilde{x}^{(\tau)}, \tilde{x}^{(\tau)}, |\tau| = \sup_i \tau_{i+1}$ .

## 4.6 Estimation de l'erreur.

Dans cette section, nous allons estimer l'erreur induite par la transformation du problème (4.1) en problème (4.2), et l'erreur entre la solution du problème (4.2) et l'approximation optimale. Nous calculons cette erreur dans l'intervalle  $[t_i, t_{i+1}]$  où  $\hat{A}_{i+1}(\tilde{x}_i, \tau_{i+1}) = \tilde{A}_{i+1}$  représente la matrice optimale. Nous allons prouver que sous certaines conditions la solution donnée par l'approximation optimale converge vers la solution du système non-linéaire initial dans  $L^1(0, T)$ .

Premièrement, nous considérons le cas où la fonction est dissipative dans un ouvert contenant la trajectoire de la solution désirée. C'est à dire, nous supposons que en choisissant la norme euclidienne canonique dans  $\mathbb{R}^n$ , et en notant  $\langle \cdot, \cdot \rangle$ , le produit scalaire correspondant



il existe  $R > 0$ ,  $R > \|x_0\| + MT$ ,  $\alpha > 0$  tel que

$$\langle F(x) - F(y), x - y \rangle \leq -\alpha \|x - y\|^2, \quad (4.29)$$

pour tout  $x, y \in B_R = B(0, R)$ .

Les hypothèses de base sur  $F$  nous permettent d'affirmer que la solution désirée reste à l'intérieur d'une sphère de centre 0, et de rayon  $\|x_0\| + MT$ , tant que  $t \leq T$ .

**Lemme 4.6.1** *La matrice donnée par la relation (4.9) est bornée dans l'intervalle  $[t_i, t_{i+1}]$ .*

**Preuve:** En effet,

$$\|A\| = \left\| \left( \int_{t_i}^{t_{i+1}} [G(\bar{y}(t))] [\bar{y}(t)]^T dt \right) \left( \int_{t_i}^{t_{i+1}} [\bar{y}(t)] [\bar{y}(t)]^T dt \right)^{-1} \right\| \leq \left\| \int_{t_i}^{t_{i+1}} [G(\bar{y}(t))] [\bar{y}(t)]^T dt \right\| \left\| \left( \int_{t_i}^{t_{i+1}} [\bar{y}(t)] [\bar{y}(t)]^T dt \right)^{-1} \right\| \quad (4.30)$$

au vu de H1) et H2) dans la section 4.3, nous avons

$$\int_{t_i}^{t_{i+1}} \|G(\bar{y}(t))\| \left\| [\bar{y}(t)]^T \right\| dt \left( \int_{t_i}^{t_{i+1}} \|\bar{y}(t)\| \left\| [\bar{y}(t)]^T \right\| dt \right)^{-1} \leq \left( \int_{t_i}^{t_{i+1}} M \|\bar{y}(t)\|^2 dt \right) \left( \int_{t_i}^{t_{i+1}} \|\bar{y}(t)\|^2 dt \right)^{-1}. \quad (4.31)$$

Finalement

$$\|A\| \leq M \frac{\int_{t_i}^{t_{i+1}} \|\bar{y}(t)\|^2 dt}{\int_{t_i}^{t_{i+1}} \|\bar{y}(t)\|^2 dt} \leq M. \quad \square \quad (4.32)$$

**Proposition 4.6.2** *Sous les hypothèses du (4.6.1), sous les hypothèses H1), H2) H3) sur  $F$  énoncées à la section 4.3, et si  $F$  est  $\alpha$ -dissipative pour un  $\alpha > 0$ , alors  $\tilde{x}^{(\tau)}$  converge vers  $\bar{x}^{(\tau)}$  dans  $L^1(0, T)$ , quand le pas de la subdivision tend*

vers zéro, c'est à dire

$$\lim_{|\tau| \rightarrow 0} \int_0^T \left\| \bar{x}^{(\tau)}(s) - \tilde{x}^{(\tau)}(s) \right\| ds = 0.$$

**Preuve:** Nous allons évaluer premièrement, l'erreur introduite par l'approximation optimale sur l'intervalle  $[t_i, t_{i+1}]$ . Nous avons

$$\begin{aligned} \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\|^2 &= \int_{t_i}^{t_{i+1}} \frac{d}{ds} \left( \left\| \bar{x}^{(\tau)}(s) - \tilde{x}^{(\tau)}(s) \right\|^2 \right) ds \\ &= 2 \int_{t_i}^{t_{i+1}} \left\langle \bar{x}^{(\tau)}(s) - \tilde{x}^{(\tau)}(s), \dot{\bar{x}}(s) - \dot{\tilde{x}}(s) \right\rangle ds. \end{aligned} \quad (4.33)$$

En notant entre  $t_i \leq s \leq t_{i+1}$

$$\begin{aligned} \bar{x}^{(\tau)}(s) &= v_{i+1}(s) + \tilde{x}_i^{(\tau)} \\ \tilde{x}^{(\tau)}(s) &= u_{i+1}(s) + \tilde{x}_i^{(\tau)} \end{aligned} \quad (4.34)$$

et

$$\begin{aligned} \dot{v}_{i+1}(s) &= G(v_{i+1}) + b_i \\ \dot{v}_{i+1}(s) &= F\left(v_{i+1} + \tilde{x}_i^{(\tau)}\right) \\ \dot{u}_{i+1}(s) &= \left( \tilde{A}_{i+1} \left( \tilde{x}_i^{(\tau)}, \tau_{i+1} \right) \right) u_{i+1} + b_i, \end{aligned} \quad (4.35)$$

nous obtenons

$$\left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\|^2 = 2 \int_{t_i}^{t_{i+1}} \left\langle v_{i+1}(s) - u_{i+1}(s), \dot{v}_{i+1}(s) - \dot{u}_{i+1}(s) \right\rangle ds. \quad (4.36)$$

Maintenant

$$\begin{aligned} \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\|^2 &= 2 \int_{t_i}^{t_{i+1}} \left\langle v_{i+1} - u_{i+1}, G(v_{i+1}(s)) - \left( \tilde{A}_{i+1} \right) u_{i+1}(s) \right\rangle ds \\ &= 2 \int_{t_i}^{t_{i+1}} \left\langle v_{i+1} - u_{i+1}, G(v_{i+1}(s)) - G(u_{i+1}(s)) \right\rangle ds \\ &\quad + 2 \int_{t_i}^{t_{i+1}} \left\langle v_{i+1} - u_{i+1}, G(u_{i+1}(s)) - \left( \tilde{A}_{i+1} \right) u_{i+1}(s) \right\rangle ds. \end{aligned} \quad (4.37)$$

La première integrale est  $\leq 0$ , puisque  $G$  est dissipative, on peut l'éliminer dans

le calcul

$$\left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\|^2 \leq 2 \int_{t_i}^{t_{i+1}} \langle v_{i+1} - u_{i+1}, G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s) \rangle ds. \quad (4.38)$$

En utilisant l'inégalité de Cauchy-Schwartz, on a

$$\begin{aligned} & \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\|^2 \leq \\ & 2 \sup_{t_i \leq s \leq t_{i+1}} \|v_{i+1} - u_{i+1}\| \left( \sqrt{\tau_{i+1}} \left( \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s)\|^2 ds \right)^{\frac{1}{2}} \right), \end{aligned} \quad (4.39)$$

et avec

$$\begin{aligned} & \sup_{t_i \leq s \leq t_{i+1}} \|v_{i+1} - u_{i+1}\| \leq \\ & 2 \left( \sqrt{\tau_{i+1}} \left( \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s)\|^2 ds \right)^{\frac{1}{2}} \right), \end{aligned} \quad (4.40)$$

on obtient

$$\left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\| \leq 2 \sqrt{\tau_{i+1}} \left( \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s)\|^2 ds \right)^{\frac{1}{2}}. \quad (4.41)$$

Dans l'intervalle  $[t_i, t_{i+1}]$ , en utilisant la propriété de minimisation de  $\tilde{A}_{i+1}$  par rapport à  $G$ , [22], il vient

$$\begin{aligned} & \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s)\|^2 ds \\ & = \inf_{\forall A \in M_n(\mathbb{R}), \operatorname{Re}(\sigma) \in ]-\infty, 0[} \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - Au_{i+1}(s)\|^2 ds. \end{aligned} \quad (4.42)$$

En particulier pour  $A = DG(0)$ , nous avons

$$\int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - (\tilde{A}_{i+1}) u_{i+1}(s)\|^2 ds \leq \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - DG(0) u_{i+1}(s)\|^2 ds, \quad (4.43)$$

et

$$\begin{aligned} \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\| &\leq 2\sqrt{\tau_{i+1}} \left( \int_{t_i}^{t_{i+1}} \|G(u_{i+1}(s)) - DG(0)u_{i+1}(s)\|^2 ds \right)^{\frac{1}{2}} \\ &\leq 2\sqrt{\tau_{i+1}}\sqrt{\tau_{i+1}} \sup_{t_i \leq s \leq t_{i+1}} \|G(u_{i+1}(s)) - DG(0)u_{i+1}(s)\|. \end{aligned} \quad (4.44)$$

Au vu de H3) section 4.3, il vient que

$$\|G(u_{i+1}(s)) - DG(0)u_{i+1}(s)\| \leq M_2 \|u_{i+1}(s)\|^2, \quad (4.45)$$

et

$$\left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1}) \right\| \leq 2M_2(\tau_{i+1}) \sup_{t_i \leq s \leq t_{i+1}} [\|u_{i+1}(s)\|^2]. \quad (4.46)$$

En appliquant le lemme de Gronwall à l'équation linéaire définissant l'approximation optimale dans l'intervalle  $[t_i, t_{i+1}]$ , on obtient

$$\begin{aligned} \frac{d}{dt} \|u_{i+1}(s)\| &\leq \|\tilde{A}_{i+1}\| \|u_{i+1}(s)\| + \|b_i\| \\ \|u_{i+1}(0)\| &= 0 \\ b_i &= F\left(\tilde{x}_i^{(\tau)}\right) \end{aligned} \quad (4.47)$$

et

$$\sup_{t_i \leq s \leq t_{i+1}} \|u_{i+1}(s)\| \leq \|b_i\| \left( \frac{e^{\|\tilde{A}_{i+1}\|(t_{i+1}-t_i)} - 1}{\|\tilde{A}_{i+1}\|} \right). \quad (4.48)$$

Au vu du lemme (4.6.1), et avec  $\tau_{i+1}$  suffisamment petit, on obtient

$$\sup_{t_i \leq s \leq t_{i+1}} \|u_{i+1}(s)\| \leq 2(\|b_i\|)\tau_{i+1}. \quad (4.49)$$

Ce qui donne que

$$\sup_{t_i \leq s \leq t_{i+1}} \|u_{i+1}(s)\|^2 \leq 4(\|b_i\|)^2(\tau_{i+1})^2. \quad (4.50)$$

Finalement, en notant

$$k = 8 \sup_{0 \leq i \leq n} \|b_i\|^2 M_2 = 8 \sup_{\|x\| \leq \|x_0\| + MT} \|F'(x)\|^2 M_2, \quad (4.51)$$

on obtient

$$\left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}_{i+1}^{(\tau)} \right\| \leq k (\tau_{i+1})^3. \quad (4.52)$$

En multipliant par  $\tau_{i+1}$ , il vient que

$$\tau_{i+1} \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}_{i+1}^{(\tau)} \right\| \leq k (\tau_{i+1})^4. \quad (4.53)$$

Sur tout l'intervalle  $[0, T]$ , l'erreur peut s'écrire

$$\sum_{i=0}^n \tau_{i+1} \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}_{i+1}^{(\tau)} \right\| \leq k \sum_{i=0}^n (\tau_{i+1})^4. \quad (4.54)$$

Avec  $\sum_{i=0}^n (\tau_{i+1})^4 \leq T \left( \sup_i \tau_{i+1} \right)^3 = T |\tau|^3$ , on a

$$\sum_{i=0}^n \left\| \bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}_{i+1}^{(\tau)} \right\| \leq k |\tau|^3 T. \quad (4.55)$$

En passant à la limite

$$\lim_{|\tau| \rightarrow 0} \int_0^T \left\| \bar{x}^{(\tau)}(s) - \tilde{x}^{(\tau)}(s) \right\| ds = 0. \quad \square \quad (4.56)$$

On en conclut que la solution  $\tilde{x}^{(\tau)}$  calculée à partir de la procédure d'approximation optimale converge vers la solution  $\bar{x}^{(\tau)}$  du problème (4.2) dans  $L^1(0, T)$ , quand  $|\tau| \rightarrow 0$ .

Maintenant, nous allons évaluer l'erreur entre les problèmes (4.1) et (4.2).

**Lemme 4.6.3** *Sous les hypothèses H1), H2) section 4.3, sur la fonction  $F$ , et si*

$F$  est  $\alpha$ -dissipative pour un  $\alpha > 0$ , alors,

$$\|x_{i+1} - \bar{x}^{(\tau)}(t_{i+1})\| \leq e^{-\alpha\tau_{i+1}} \|x_i - \tilde{x}_i^{(\tau)}\|.$$

**Preuve:** Nous avons

$$\begin{aligned} 2\langle x - \bar{x}, \dot{x} - \dot{\bar{x}} \rangle &= 2\langle x(t) - \bar{x}(t), \dot{x}(t) - \dot{\bar{x}}(t) \rangle \\ &= 2\langle x(t) - \bar{x}(t), F(x(t)) - F(\bar{x}(t)) \rangle. \end{aligned} \quad (4.57)$$

Le caractère dissipatif de  $F$  permet d'écrire

$$2\langle x - \bar{x}, \dot{x} - \dot{\bar{x}} \rangle \leq -\alpha \|x(t) - \bar{x}(t)\|^2, \quad (4.58)$$

et

$$\frac{d}{dt} \|x(t) - \bar{x}(t)\| \leq -\alpha \|x(t) - \bar{x}(t)\|. \quad (4.59)$$

En intégrant entre  $t_i$  et  $t$ , pour  $t_i \leq t \leq t_{i+1}$

$$\|x(t) - \bar{x}(t)\| \leq e^{-\alpha(t-t_i)} \|x(t_i) - \bar{x}^{(\tau)}(t_i)\|, \quad (4.60)$$

et en posant  $x(t_i) = x_i$  et  $\bar{x}^{(\tau)}(t_i) = \tilde{x}_i^{(\tau)}$ , on obtient

$$\|x(t) - \bar{x}(t)\| \leq e^{-\alpha(t-t_i)} \|x_i - \tilde{x}_i^{(\tau)}\|. \quad (4.61)$$

La relation (4.61) peut s'écrire

$$\|x(t_{i+1}) - \bar{x}^{(\tau)}(t_{i+1})\| \leq e^{-\alpha(t_{i+1}-t_i)} \|x_i - \tilde{x}_i^{(\tau)}\| \quad (4.62)$$

et

$$\|x_{i+1} - \bar{x}^{(\tau)}(t_{i+1})\| \leq e^{-\alpha\tau_{i+1}} \|x_i - \tilde{x}_i^{(\tau)}\|. \quad \square \quad (4.63)$$

**Proposition 4.6.4** *Sous les mêmes hypothèses sur la fonction  $F$  et sous les*

hypothèses du lemme (4.6.3), et la proposition (4.6.2), alors,  $\tilde{x}^{(\tau)}$  converge vers  $x$  dans  $L^1(0, T)$ , quand le pas de la subdivision tend vers zéro, c'est à dire

$$\lim_{|\tau| \rightarrow 0} \int_0^T \|x(s) - \tilde{x}^{(\tau)}(s)\| ds = 0.$$

**Preuve:** L'erreur globale peut s'écrire

$$\|x(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1})\| \leq \|x(t_{i+1}) - \bar{x}^{(\tau)}(t_{i+1})\| + \|\bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}^{(\tau)}(t_{i+1})\| \quad (4.64)$$

et

$$\|x_{i+1} - \tilde{x}_{i+1}^{(\tau)}\| \leq \|x_{i+1} - \bar{x}^{(\tau)}(t_{i+1})\| + \|\bar{x}^{(\tau)}(t_{i+1}) - \tilde{x}_{i+1}^{(\tau)}\|. \quad (4.65)$$

Au vu des relations (4.52) et (4.63),

$$\begin{aligned} \|x_{i+1} - \tilde{x}_{i+1}^{(\tau)}\| &\leq e^{-\alpha\tau_{i+1}} \|x_i - \tilde{x}_i^{(\tau)}\| + k(\tau_{i+1})^3 \\ &\leq e^{-\alpha\tau_{i+1}} \left( e^{-\alpha\tau_i} \|x_{i-1} - \tilde{x}_{i-1}^{(\tau)}\| + k(\tau_i)^3 \right) + k(\tau_{i+1})^3 \\ &\leq e^{-\alpha(\tau_{i+1}+\tau_i)} \|x_{i-1} - \tilde{x}_{i-1}^{(\tau)}\| + ke^{-\alpha\tau_{i+1}}(\tau_i)^3 + k(\tau_{i+1})^3 \\ &\leq e^{-\alpha(\tau_{i+1}+\tau_i+\tau_{i-1})} \|x_{i-2} - \tilde{x}_{i-2}^{(\tau)}\| + ke^{-\alpha(\tau_{i+1}+\tau_i)}(\tau_{i-1})^3 + ke^{-\alpha\tau_{i+1}}(\tau_i)^3 + k(\tau_{i+1})^3 \end{aligned} \quad (4.66)$$

finalement, pour l'ensemble des pas, on a

$$\begin{aligned} \|x_{i+1} - \tilde{x}_{i+1}^{(\tau)}\| &\leq e^{-\alpha(\tau_{i+1}+\tau_i+\tau_{i-1}+\dots+\tau_1)} \|x_0 - \tilde{x}_0\| \\ &\quad + k \sum_{j=1}^i e^{-\alpha(\tau_{i+1}+\tau_i+\dots+\tau_{i+2-j})} (\tau_{i+1-j})^3 + k(\tau_{i+1})^3. \end{aligned} \quad (4.67)$$

Avec  $(\tau_{i+1})^3 \leq |\tau|^2 \tau_{i+1}$  et

$$\sum_{j=1}^i e^{-\alpha(\tau_{i+1}+\tau_i+\dots+\tau_{i+2-j})} (\tau_{i+1-j})^3 \leq \sum_{j=1}^i (\tau_{i+1-j})^3 \leq |\tau|^2 \sum_{j=1}^i \tau_{i+1-j}, \quad (4.68)$$

il vient que

$$\begin{aligned} \left\| x_{i+1} - \tilde{x}_{i+1}^{(\tau)} \right\| &\leq e^{-\alpha(\tau_{i+1} + \tau_i + \tau_{i-1} + \dots + \tau_1)} \left\| x_0 - \tilde{x}_0 \right\| \\ &\quad + k |\tau|^2 \left( \tau_{i+1} + \sum_{j=1}^i \tau_{i+1-j} \right) \\ &\leq e^{-\alpha(\tau_{i+1} + \tau_i + \tau_{i-1} + \dots + \tau_1)} \left\| x_0 - \tilde{x}_0 \right\| + k |\tau|^2 T. \end{aligned} \quad (4.69)$$

En multipliant par  $\tau_i$  et dans tout l'intervalle  $[0, T]$ , l'erreur peut s'écrire

$$\begin{aligned} \sum_{i=0}^n \tau_{i+1} \left\| x_{i+1} - \tilde{x}_{i+1}^{(\tau)} \right\| &\leq \sum_{i=0}^{n-1} (\tau_{i+1}) e^{-\alpha(\tau_{i+1} + \tau_i + \tau_{i-1} + \dots + \tau_1)} \left\| x_0 - \tilde{x}_0 \right\| + k \sum_{i=0}^{n-1} (\tau_{i+1}) |\tau|^2 T \\ &\leq T e^{-\alpha T} \left\| x_0 - \tilde{x}_0 \right\| + k |\tau|^3 T. \end{aligned} \quad (4.70)$$

En passant à la limite, on obtient

$$\lim_{|\tau| \rightarrow 0} \int_0^T \left\| x(s) - \tilde{x}^{(\tau)}(s) \right\| ds \leq T e^{-\alpha T} \left\| x_0 - \tilde{x}_0 \right\|. \quad (4.71)$$

L'erreur globale est majorée par l'erreur de départ qu'on commet sur les conditions initiales, en considérant que  $\tilde{x}_0$  est une approximation de  $x(0) = x_0$ . Si on suppose cette erreur négligeable, ce qui est le cas en général, alors

$$\lim_{|\tau| \rightarrow 0} \int_0^T \left\| x(s) - \tilde{x}^{(\tau)}(s) \right\| ds = 0. \quad \square \quad (4.72)$$

Dans ce cas, la solution  $\tilde{x}^{(\tau)}$  donnée par la procédure d'approximation optimale converge vers la solution  $x$  du problème théorique dans  $L^1(0, T)$ , quand  $|\tau| \rightarrow 0$ .

#### 4.6.1 Cas où la fonction $F$ n'est pas dissipative

Les calculs ont été fait en supposant que la fonction  $F$  est dissipative. Dans le cas lipschitzien, en général, on peut toujours s'y ramener en faisant le changement de variable suivant

$$x(t) = e^{\lambda t} z(t). \quad (4.73)$$



L'équation initiale s'écrit alors

$$\frac{dx}{dt} = \lambda e^{\lambda t} z(t) + e^{\lambda t} \frac{dz(t)}{dt} = F(e^{\lambda t} z(t)) \quad (4.74)$$

d'où

$$\begin{aligned} \frac{dz}{dt} &= e^{-\lambda t} F(e^{\lambda t} z(t)) - \lambda z(t) \\ &= H(t, z(t)). \end{aligned} \quad (4.75)$$

On obtient ainsi, une nouvelle fonction  $H(t, z)$  dépendant du temps. Pour  $\lambda$  assez grand, celle-ci est dissipative en  $y$  uniformément par rapport au temps.

Dans le contexte où nous travaillons, nous subdivisons l'intervalle  $[0, T]$  en un ensemble d'intervalles  $[t_i, t_{i+1}]$  que nous allons considérer assez petits pour que l'on puisse prendre les coefficients comme constants. Dans ce qui suit, nous allons évaluer l'erreur due à cette approximation.

Il vient que

$$\begin{aligned} \|H(t, z) - H(t_i, z)\| &\leq |e^{-\lambda t} - e^{-\lambda t_i}| \|F(e^{\lambda t} z)\| + e^{-\lambda t} \|F(e^{\lambda t} z) - F(e^{\lambda t_i} z)\| \\ &\leq \lambda |t - t_i| M_0 + M_1 \lambda |t - t_i| \|z\| \end{aligned} \quad (4.76)$$

avec  $\lambda > M_1$ . En notant  $\delta = |t - t_i|$ , on obtient

$$\|H(t, z) - H(t_i, z)\| \leq \lambda \delta (M_0 + M_1 \|z\|). \quad (4.77)$$

Avec  $\delta$  assez petit, on peut écrire

$$H(t, z) = H(t_i, z) = H(z). \quad (4.78)$$

C'est à dire que moyennant l'erreur ci-dessus, la fonction  $H$  peut être considérée comme indépendante explicitement du temps dans chaque intervalle  $[t_i, t_{i+1}]$ . Ce qui permet d'utiliser  $H(t_i, z)$  comme une approximation de  $H(t, z)$  pour  $t \in [t_i, t_{i+1}]$ .

C'est à dire que, quand on considère la fonction  $H$  indépendante du temps dans l'intervalle  $[t_i, t_{i+1}]$ , nous faisons une erreur majorée par la relation (4.77).

## 4.7 Commentaires.

La méthode présentée dans ce chapitre est basée essentiellement sur l'utilisation de la dérivée optimale. Elle permet de remplacer sur chaque intervalle de la subdivision choisie, le problème non-linéaire par un problème linéaire. L'approche proposée est fondamentalement différente des méthodes existantes (Euler, RK4, ...), car celles-ci essayent d'approcher la fonction non-linéaire par ses dérivées successives.

C'est une méthode d'ordre 3 en fonction du pas de discrétisation  $\tau_i$ . De ce point de vue, l'approximation optimale est d'ordre supérieur à celui d'Euler, et d'ordre inférieur à celui de Runge-kutta 4, et donc se situe entre les deux.

**CINQUIEME CHAPITRE  
MISE EN OEUVRE  
NUMERIQUE.**

# Chapitre 5

## Mise en oeuvre numérique.

### 5.1 Introduction.

L'objet de ce chapitre est essentiellement la mise en oeuvre de la procédure de la dérivée optimale ainsi que celle de la méthode d'approximation optimale. Nous présentons les différents organigrammes qui interviennent dans l'élaboration des programmes. Précisons que la résolution des équations différentielles a été faite à l'aide de la méthode de Runge-Kutta d'ordre 4. Le logiciel utilisé pour tracer les courbes est grapher. Le langage de programmation utilisé est le Fortran. Enfin, nous proposons un certain nombre d'exemples d'équations différentielles ordinaires non-linéaires, pour lesquelles on applique les méthodes introduites.

### 5.2 Procédure de la dérivée optimale.

#### 5.2.1 Schéma général.

La procédure de calcul est basée sur le schéma présenté au chapitre 1.

Entrées  $[x_0, A_0]$

**Niveau (I) :** Calcul de  $A_1$  à partir de  $A_0$

$$A_1 = \left[ \int_0^{+\infty} [F(e^{A_0 t} x_0)] [e^{A_0 t} x_0]^T dt \right] \left[ \int_0^{+\infty} [e^{A_0 t} x_0] [e^{A_0 t} x_0]^T dt \right]^{-1} \quad (5.1)$$

**Niveau (II) :** Calcul de  $A_{(j)}$  à partir de  $A_{(j-1)}$

$$A_{(j)} = \left[ \int_0^{+\infty} [F(e^{A_{(j-1)} t} x_0)] [e^{A_{(j-1)} t} x_0]^T dt \right] \left[ \int_0^{+\infty} [e^{A_{(j-1)} t} x_0] [e^{A_{(j-1)} t} x_0]^T dt \right]^{-1} \quad (5.2)$$

**Niveau (III) :** Calcul de

$$\|A_{(j)} - A_{(j-1)}\|. \quad (5.3)$$

**Niveau (IV) :** Si

$$\|A_{(j)} - A_{(j-1)}\| < \varepsilon, \quad (5.4)$$

où  $\varepsilon$  représente le niveau d'approximation désiré, alors  $\tilde{A} = A_{(j)}$ .  $\tilde{A}$  est la dérivée optimale de  $F$  en  $x_0$ . Sinon  $A_{(j-1)} = A_{(j)}$  et on va au **Niveau (II)**.

La formule générale de calcul de la suite  $[A_j]$  peut s'écrire

$$[A_j][\Gamma] = \int_0^{+\infty} [F(x)] [x]^T dt = [FX] \quad (5.5)$$

avec

$$[FX] = \begin{bmatrix} \int_0^{+\infty} x_1 f_1(x) dt & \dots & \int_0^{+\infty} x_n f_1(x) dt \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \int_0^{+\infty} x_1 f_n(x) dt & \dots & \int_0^{+\infty} x_n f_n(x) dt \end{bmatrix} \quad (5.6)$$

et

$$[\Gamma] = \begin{bmatrix} \int_0^{+\infty} x_1^2 dt & \dots & \int_0^{+\infty} x_n x_1 dt \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \int_0^{+\infty} x_1 x_n dt & \dots & \int_0^{+\infty} x_n^2 dt \end{bmatrix} \quad (5.7)$$

Les éléments des matrices  $[\Gamma]$ ,  $[FX]$  sont sous forme d'intégrales. Pour calculer toutes ces intégrales, il nous faut connaître l'expression des  $x(t)$ , solutions des différentes équations. Techniquement, on doit faire appel à la procédure de calcul de l'intégrale un nombre de fois égal au nombre d'intégrales composant les matrices, c'est à dire  $2n^2$  fois. Afin d'éviter ce problème dans le programme, on va les écrire sous forme de série.

### 5.2.2 Organigramme.

Les principales étapes suivies par le programme pour résoudre le problème de la dérivée optimale sont:

1. Lecture des données du problème d'approximation  $[A_0, x_0, \varepsilon]$ .
2. Initialisation des paramètres de calcul des matrices  $[\Gamma]$ ,  $[FX]$ .
3. Initialisation des matrices  $[\Gamma]$ ,  $[FX]$ .
4. Résolution de l'équation linéaire

$$\begin{cases} \frac{dx}{dt} = A_{(j)}x(t) \\ x(0) = x_0 \end{cases}$$

et calcul des matrices  $[\Gamma]$ ,  $[FX]$ .

5. Calcul de la matrice  $A_{(j)}$ , en résolvant le système

$$A_{(j)}[\Gamma] = [FX].$$

6. Test d'arrêt pour le niveau d'approximation désirée

$$\|A_{(j)} - A_{(j-1)}\| < \varepsilon.$$

7. Calcul des valeurs propres de  $A_{(j)}$ , pour le suivi de la stabilité.

8. Dans le cas où le niveau d'approximation  $\varepsilon$ -désirée est atteint, alors

$$\hat{A} = A_{(j)}.$$

Sinon, on réinitialise les données et on recommence en 2.

### Commentaires.

L'essentiel de ce programme revient en fait à la résolution du système linéaire caractérisant l'équation linéaire optimale. Alors, trois questions se posent en voulant résoudre ce type de problème sur ordinateur:

- 1) Ce système est-il non dégénéré ?
- 2) Le programme utilisé est-il capable d'en fournir une solution ?
- 3) Cette solution obtenue est-elle représentative de la solution algébrique ?

Les réponses à ces trois questions sont fournies par l'utilisation des programmes suivants <sup>1</sup>:

"CNSL" -Contrôle numérique d'un système linéaire- permettant l'étude du conditionnement numérique et l'évaluation de l'incertitude sur les inconnues. Lui-même fait appel à plusieurs sous-programmes dont les principaux sont  
-"DNSL", pour la détection de la dégénérescence numérique d'un système linéaire,  
-"IXNL", pour la détermination de l'incertitude sur les inconnues du système linéaire, chacun des vecteurs-solutions est contrôlé par le critère des résidus normés et amélioré si nécessaire.  
-"RSLCA", pour la résolution d'un système linéaire avec contrôle par le résidu

---

<sup>1</sup>Les organigrammes et les programmes sont exposés dans la référence [8]

et améliorations éventuelles.

- "DETG", pour le calcul du déterminant par la méthode de Gauss.

- "HASARD", permettant d'obtenir une valeur numérique tirée au hasard entre deux bornes données.

b) La norme matricielle utilisée pour le test d'arrêt est

$$\|A\|_1 = \max_j \sum_i |a_{i,j}|.$$

### 5.2.3 Applications.

On présente tous les exemples d'applications relatifs à la procédure de la dérivée optimale. Ceux-ci vont nous permettre d'illustrer par des courbes ou de comparer par un calcul d'erreur relative, les méthodes proposées. On verra l'alternative offerte par l'association d'une équation linéaire optimale à une équation différentielle ordinaire non-linéaire, et particulièrement dans le cas où celle-ci n'admet pas de linéarisée classique à l'origine, ou que sa dérivée est nulle.

Le calcul de l'erreur relative se fera en utilisant la relation suivante

$$E_r = \frac{\|x(t) - \tilde{y}(t)\|}{\|x(t)\|}, \quad (5.8)$$

$x(t)$  représente la solution du système non-linéaire,

$\tilde{y}(t)$  représente la solution du système obtenu à partir de l'approximation.

#### Cas où la fonction n'est pas différentiable en 0.

C'est par exemple le cas où  $F$  n'est pas régulière, et cela peut arriver quand celle-ci est seulement localement lipschitzienne. Nous allons considérer deux exemples de systèmes comprenant une fonction de type valeur absolue et donc non-différentiable en 0.



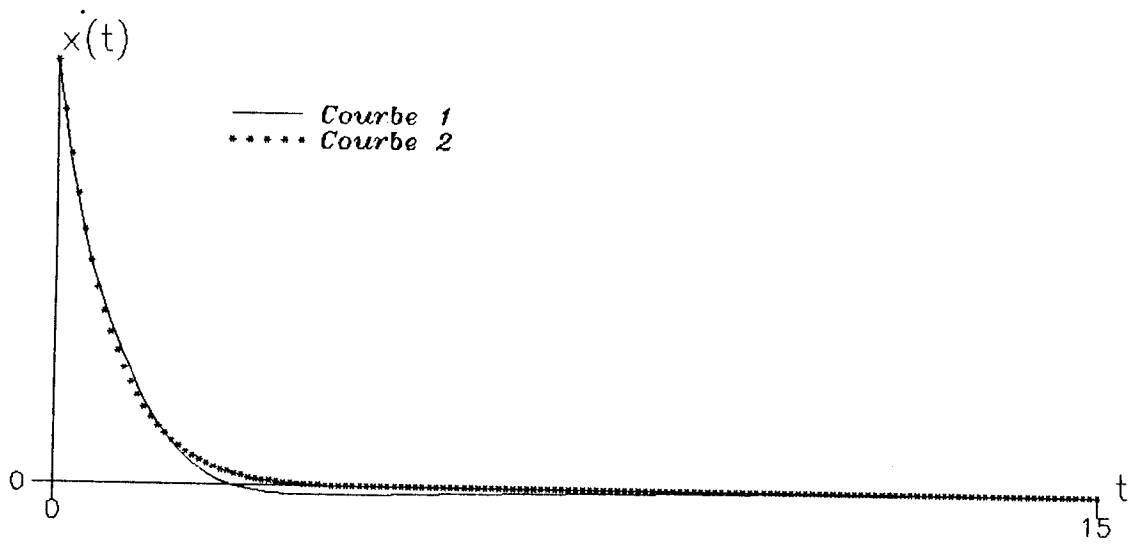


Figure 5.1

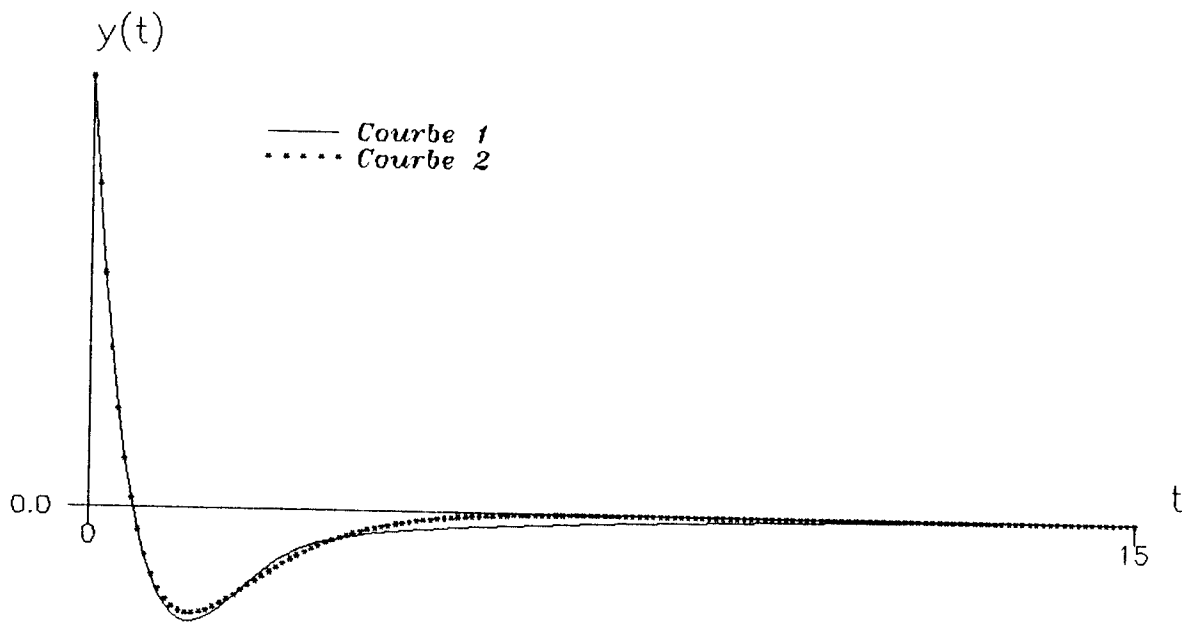


Figure 5.2

**Exemple 5.2.1**

$$\begin{cases} \frac{dx}{dt} = -x + \alpha \cos\left(\frac{\pi}{2} \exp(|y|)\right) \\ \frac{dy}{dt} = -y + \alpha \cos\left(\frac{\pi}{2} \exp(|x|)\right) \end{cases}, (x_0, y_0) = (.5, .25), |\alpha| < 1. \quad (5.9)$$

Alors, pour  $\alpha = 0.5$ , nous avons

$$DF(x_0, y_0) = \begin{bmatrix} -1 & -0.9097 \\ -0.6799 & -1 \end{bmatrix}, (x_0, y_0) = (.5, .25). \quad (5.10)$$

Après 10 itérations la procédure de calcul donne ( $\varepsilon = 10^{-6}$ )

$$\tilde{A} = \begin{bmatrix} -1.2925 & 0.1232 \\ -0.8522 & -1.0391 \end{bmatrix}, (x_0, y_0) = (.5, .25). \quad (5.11)$$

Les courbes des figures (5.1) and (5.2) représentent le graphe des composantes respectives  $(x(t), y(t))$  de la solution des systèmes (5.9) et (5.11) en fonction du temps.

**Exemple 5.2.2**

$$\begin{cases} \frac{dx}{dt} = -x + \alpha \sin(|y|) \\ \frac{dy}{dt} = -y + \alpha \sin(|x|) \end{cases}, (x_0, y_0) = (1, .5), |\alpha| < 1. \quad (5.12)$$

Pour  $\alpha = 0.5$ , la matrice jacobienne de  $F$  en  $x_0$  s'écrit

$$DF(x_0, y_0) = \begin{bmatrix} -1 & 0.4387 \\ 0.2701 & -1 \end{bmatrix}, (x_0, y_0) = (1, .5). \quad (5.13)$$

Après 5 itérations, la procédure de calcul donne ( $\varepsilon = 10^{-6}$ )

$$\tilde{A} = \begin{bmatrix} -1.0207 & 0.5172 \\ 0.3502 & -0.8336 \end{bmatrix}, (x_0, y_0) = (1, .5) \quad (5.14)$$

Les courbes des figures (5.3) et (5.4) représentent le graphe des composantes respectives  $x(t), y(t)$  de la solution des systèmes (5.12) et (5.14) en fonction du temps. Le Tableau I représente les valeurs de la solution à des instants  $t \in [0, T]$  de l'équation non-linéaire et de son approximation au sens de la dérivée optimale, ainsi que l'erreur relative.

t	Xnl(t)	Xlin(t)	Ynl(t)	Ylin(t)	Er
0	.1000000E+01	.1000000E+01	.5000000E+00	.5000000E+00	0
1	.5008458E+00	.5012365E+00	.3794262E+00	.379014E+00	.0009
2	.2781739E+00	.2783531E+00	.2502546E+00	.2505841E+00	.001
3	.1628160E+00	.1630481E+00	.1565178E+00	.1580888E+00	.007
4	.9741862E-01	.9789484E-01	.9600762E-01	.9793934E-01	.014
5	.5878825E-01	.5941472E-01	.5847295E-01	.6022978E-01	.022
6	.3559002E-01	.3622664E-01	.3551962E-01	.3692688E-01	.030
7	.2157151E-01	.2213122E-01	.2155580E-01	.2261112E-01	.039
8	.1308045E-01	.1353121E-01	.1307694E-01	.1383805E-01	.047
9	.7932950E-02	.8275914E-02	.7932167E-02	.8467011E-02	.056
10	.4811411E-02	.5062412E-02	.4811237E-02	.5180184E-02	.065
11	.2918231E-02	.3096885E-02	.2918192E-02	.3169155E-02	.074
12	.1769989E-02	.1894538E-02	.1769980E-02	.1938808E-02	.083
13	.1073550E-02	.1159008E-02	.1073549E-02	.1186105E-02	.093
14	.6511409E-03	.7090407E-03	.6511404E-03	.7256214E-03	.10
15	.3949368E-03	.4337673E-03	.3949367E-03	.4439117E-03	.11

Tableau I

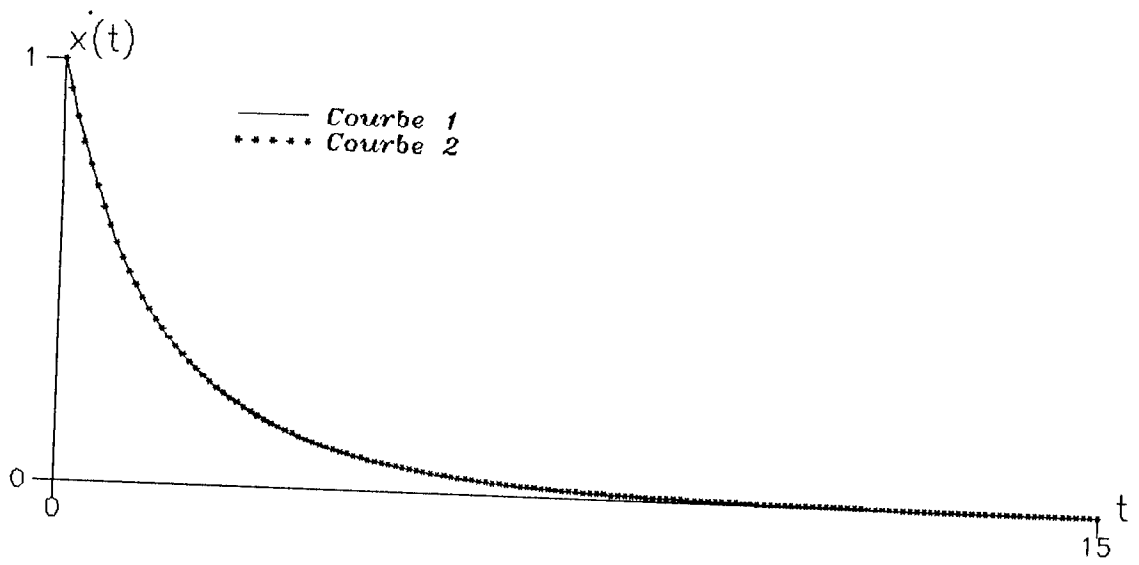


Figure 5.3

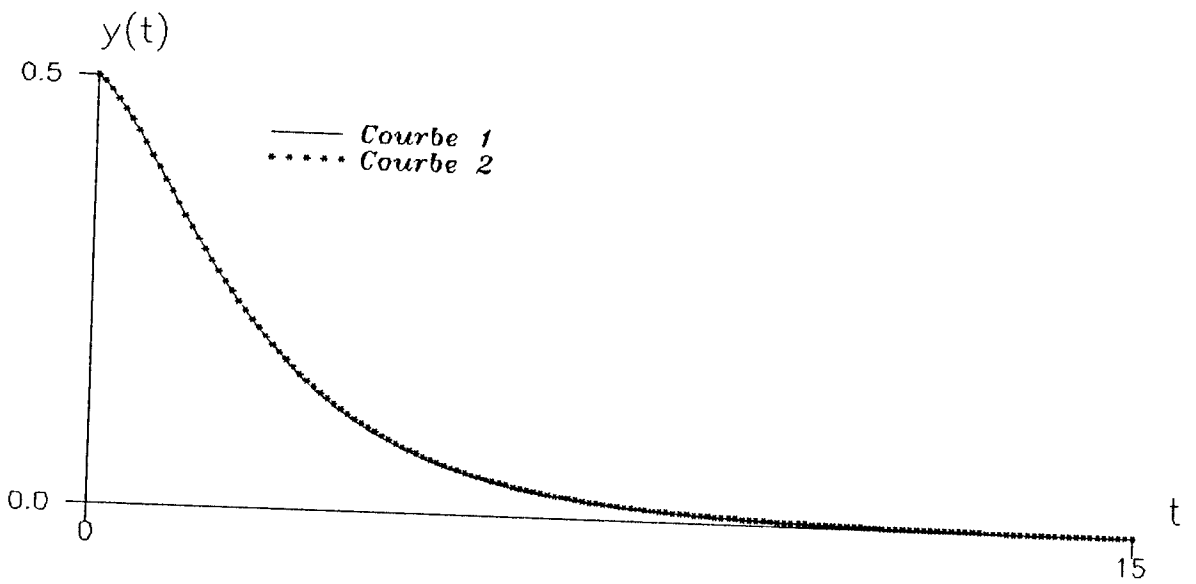


Figure 5.4

On remarque que la solution donnée par la dérivée optimale est du même ordre de grandeur que celle donnée par l'équation non-linéaire. De plus l'erreur relative atteint au temps  $t = T = 15s$ , 11% et reste donc petite.

**Cas où la différentielle en 0 est nulle.**

Dans cet exemple, nous allons considérer un système pour lequel  $DF(0) = 0$ .

**Exemple 5.2.3**

$$\begin{cases} \frac{dx}{dt} = -\frac{x^4}{x^2 + y^2} \\ \frac{dy}{dt} = -\frac{y^4}{x^2 + y^2} \end{cases}, (x_0, y_0) = (1, .5). \quad (5.15)$$

Avec  $DF(x)$  calculée en  $(x_0, y_0)$

$$DF(x_0, y_0) = \begin{bmatrix} -1.12 & 0.64 \\ 0.16 & -0.52 \end{bmatrix}, (x_0, y_0) = (1, .5) \quad (5.16)$$

et après sept itérations, la procédure de calcul donne ( $\varepsilon = 10^{-6}$ )

$$\tilde{A} = \begin{bmatrix} -1.1030 & 0.6127 \\ 0.3085 & -0.7844 \end{bmatrix}, (x_0, y_0) = (1, .5). \quad (5.17)$$

**Remarque 5.2.4** Dans ce cas l'approximation par dérivation optimale est d'ordre 3 en fonction de la donnée initiale, car  $DF(0) = 0$ , et dans ce cas,  $\|F(x) - D^2F(0)x^2\| = O(\|x\|^3)$ . En utilisant cette relation dans la formule donnant l'ordre au chapitre 3, on obtient

$$\|x(t) - \tilde{y}(t)\| \leq O(\|x_0\|^3)$$

où  $x(t)$  représente la solution du système définie par la relation (5.15), et  $\tilde{y}(t)$  la solution du système définie par la relation (5.17).

Cas où le comportement des solutions de l'équation linéarisée diffère de celui de l'équation non-linéaire.

**Exemple 5.2.5** *Considérons le système suivant*

$$\begin{cases} \frac{dx}{dt} = -x - \frac{2y}{\ln(x^2 + y^2)} \\ \frac{dy}{dt} = -y + \frac{2x}{\ln(x^2 + y^2)} \end{cases}, (x_0, y_0) = (0, .5) \quad (5.18)$$

dans le disque unité  $\{(x, y) \in \mathbb{R}^2; x^2 + y^2 < 1\}$ .

La linéarisation de  $F$  en  $(x_0, y_0) = (0, .5)$  donne

$$DF(x_0, y_0) = \begin{bmatrix} -1 & 3.524 \\ -1.4426 & -1 \end{bmatrix}, (x_0, y_0) = (0, .5). \quad (5.19)$$

Après 10 itérations, la procédure de calcul donne ( $\varepsilon = 10^{-6}$ )

$$\tilde{A} = \begin{bmatrix} -1.4934 & 1.2489 \\ -0.5213 & -1.1254 \end{bmatrix}, (x_0, y_0) = (0, .5). \quad (5.20)$$

Les courbes des figures (5.5) and (5.6) représentent le graphe des composantes respectives  $(x(t), y(t))$  de la solution des systèmes (5.18) et (5.20) en fonction du temps. Le Tableau II représente les valeurs de la solution à des instants  $t \in [0, T]$  de l'équation non-linéaire, de son approximation au sens de la dérivée optimale, ainsi que l'erreur relative. On remarque que la solution donnée par la dérivée optimale obéit à la même dynamique que celle donnée par l'équation non-linéaire même si au temps  $t = T = 15s$ , l'erreur relative est de l'ordre de 100%.

t	Xnl(t)	Xlin(t)	Ynl(t)	Ylin(t)	Er
0	.0000000E+00	.0000000E+00	.5000000E+00	.5000000E+00	0
1	.1432933E+00	.1517739E+00	.1153286E+00	.1177924E+00	.048
2	.6613008 E-01	.5793614E-01	.1434192E-01	.8519892E-02	.148
3	.2476355E-01	.1105308E-01	-.2539680E-02	-.5333531E-02	.562
4	.8627967E-02	-.3666260E-05	-.3069739E-02	-.2656959E-02	.946
5	.2898245E-02	-.8070500E-03	-.1717583E-02	-.6254748E-03	1.14
6	.9504736E-03	-.3078052E-03	-.7953907E-03	-.4509668E-04	1.18
7	.3056629E-03	-.5867217E-04	-.3383072E-03	.2837578E-04	1.06
8	.9643728E-04	.3895649E-07	-.1372356E-03	.1411885E-04	.932
9	.2978096E-04	.4291440E-05	-.5404247E-04	.3321252E-05	1.01
10	.8954184E-05	.1635317E-05	-.2085933E-04	.2386980E-06	.983
11	.2595794E-05	.3114442E-06	-.7937176E-05	-.1509661E-06	.971
12	.7123896E-06	-.3104541E-09	-.2988374E-05	-.7502629E-07	.971
13	.1780326E-06	-.2281944E-07	-.1116058E-05	-.1763572E-07	.988
14	.3641637E-07	-.8688157E-08	-.4141681E-06	-.1263416E-08	.999
15	.3343012E-08	-.1653209E-08	-.1529153E-06	.8031747E-09	1.00

Tableau II

**Comparaison entre la linéarisation par dérivation au sens de Fréchet et l'approximation optimale.**

C'est le cas où la linéarisation au sens de Fréchet existe. On calcule l'approximation par la procédure de la dérivée optimale, et on fait une comparaison.

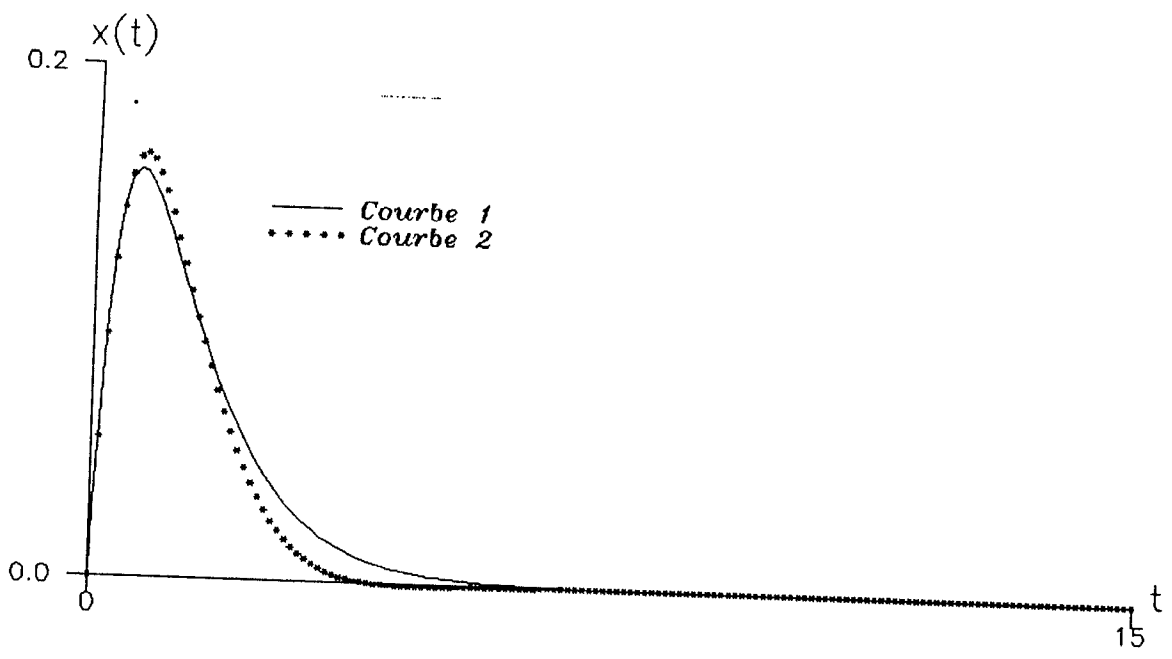


Figure 5.5

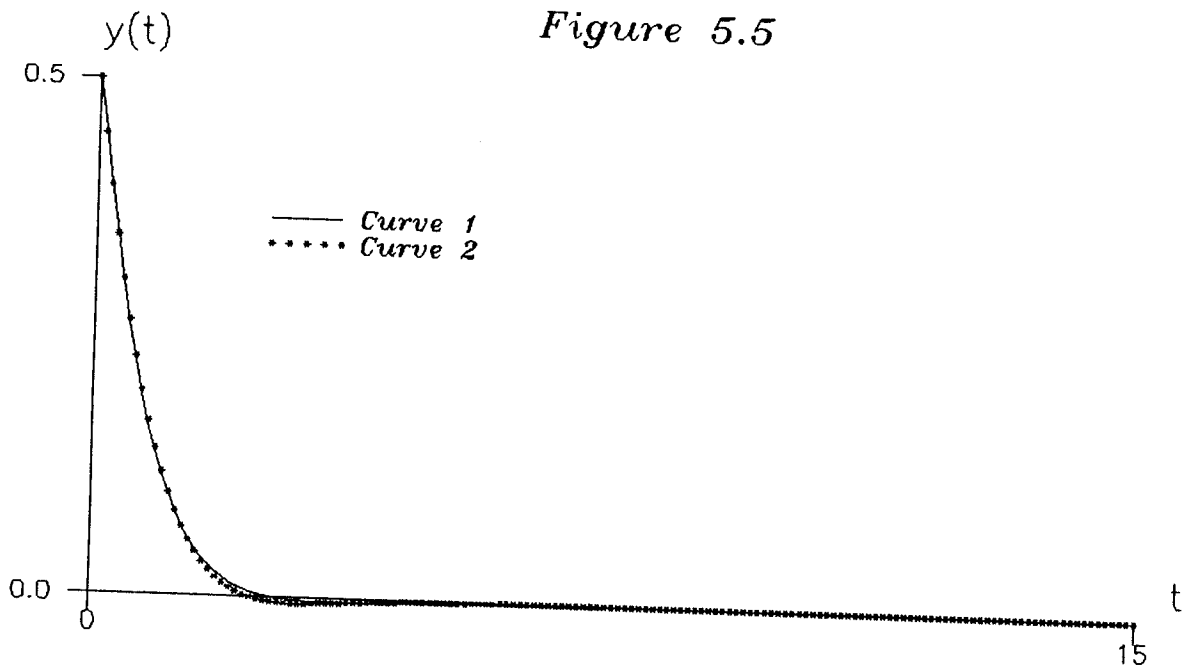


Figure 5.6



**Exemple 5.2.6** *Considérons le système*

$$\begin{cases} \frac{dx}{dt} = -7.10^2 x - 2.10^3 x^2 - 2.10^5 y \\ \frac{dy}{dt} = 2.10^3 x - 2.10^5 y \end{cases}, (x_0, y_0) = (5, 0). \quad (5.21)$$

*Le calcul de la dérivée de Fréchet en 0 donne*

$$DF(0, 0) = \begin{bmatrix} -7.10^2 & -2.10^5 \\ 2.10^3 & -2.10^5 \end{bmatrix}, (x_0, y_0) = (5, 0) \quad (5.22)$$

*et la dérivée optimale s'écrit*

$$\tilde{A} = \begin{bmatrix} -13.2710^3 & 4.3110^5 \\ 2.10^3 & -2.10^5 \end{bmatrix}, (x_0, y_0) = (5, 0). \quad (5.23)$$

Dans les figures (5.7) et (5.8), la courbe 1 correspond à la solution du système non-linéaire (5.21), la courbe 2 à la solution de la dérivée optimale (5.23), et la courbe 3 à la solution correspondante au système (5.22) obtenu par dérivation au sens de Fréchet en 0. La différence entre les deux approximations apparaît très bien dans les figures (5.7) et (5.8). On peut voir sur les courbes que l'approximation donnée par la dérivée optimale est meilleure que celle donnée par la dérivation au sens de Fréchet de l'équation non-linéaire, tant que la solution  $x(t)$  reste suffisamment loin de l'origine.

#### **Commentaire général.**

Les exemples exposés montrent que l'approximation obtenue par dérivation optimale donne des résultats satisfaisants par rapport aux résultats exacts tout en respectant la dynamique du problème initial. Ceci est confirmé par le calcul de l'erreur relative qui n'excède pas 100%, ce taux élevé correspond à la portion de la solution qui est très proche de l'origine, où en fait, la meilleure approximation est certainement celle donnée par la linéarisation classique, quand elle existe.

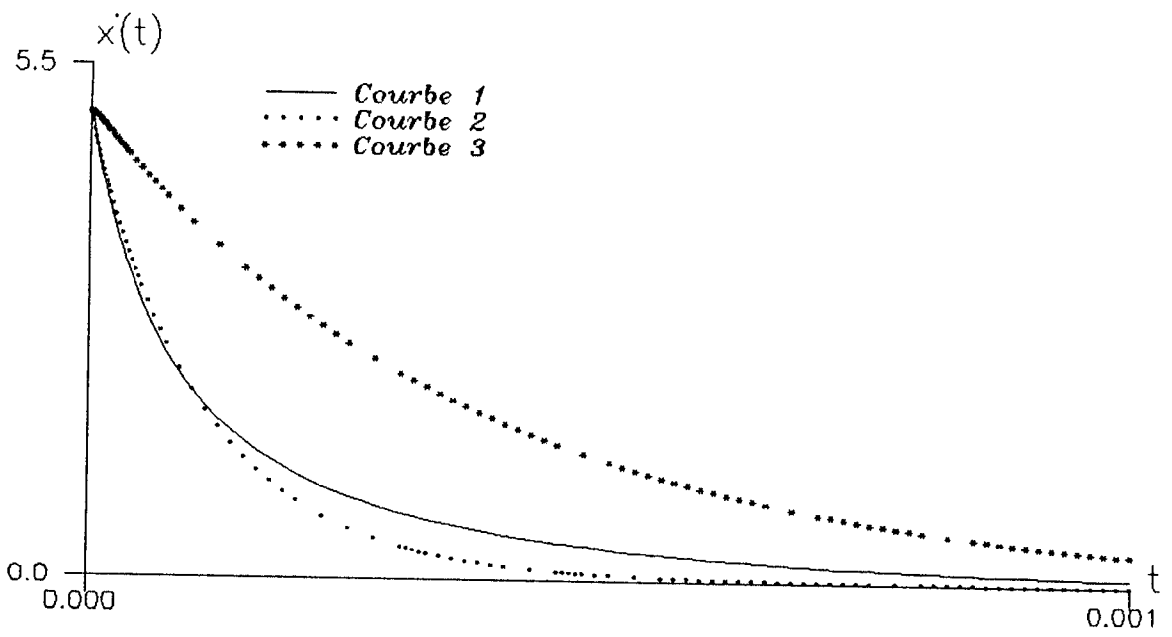


Figure 5.7

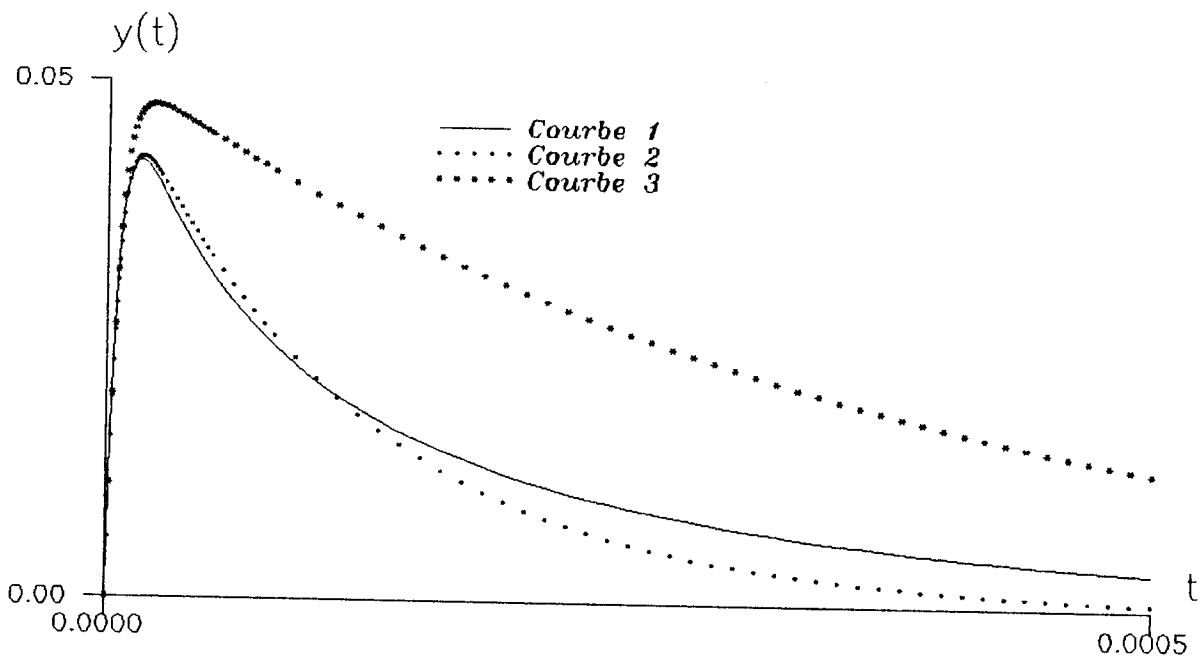
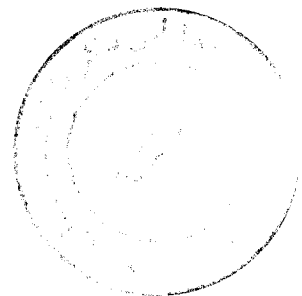


Figure 5.8



Notons que tant que la solution reste autour de la donnée initiale, l'erreur relative est petite.

En général, on peut définir la qualité de l'approximation en comparant les résultats obtenus à partir d'une donnée initiale  $x_0$ , tant que la solution reste en dehors de la boule de centre 0, de rayon  $\alpha \|x_0\|$ ,  $0 < \alpha < 1$ , aussi petit que possible. Ceci permet d'affirmer, que dans le cas d'ordre deux par exemple, l'approximation est bonne tant que la solution reste en dehors de la boule, c'est à dire dans une couronne autour du point d'équilibre.

### 5.3 Approximation et stabilité.

On présentera deux types d'exemples qui, on va le voir, vérifient la conjecture (3.3.3) énoncée au chapitre 3. En appliquant la procédure de dérivation optimale aux exemples qui vont suivre, on linéarise en réalité la perturbation qui devient dépendante de la donnée initiale  $x_0$ . On choisit alors  $x_0$  très voisin de 0, c'est à dire qu'on perturbe légèrement la dérivée au point 0.

Le premier type concerne une certaine classe de fonctions pour laquelle on a fait une étude d'existence, d'unicité et de convergence de la dérivée optimale. Il s'agit, en fait, de la perturbation non-linéaire de la forme

$$G(x) = -\Phi(\|x\|)x,$$

d'une matrice dont les valeurs propres sont imaginaires.

On a montré l'existence d'un point fixe stable pour la procédure, et par conséquent, il y a convergence effective vers la matrice optimale cherchée.

Le deuxième type concerne le cas où un système linéarisé par la méthode classique admet une valeur propre double. L'approximation par la linéarisation classique répond au problème de la stabilité, mais pas à la géométrie. C'est le cas particulier où, les courbes intégrales d'un champ de vecteurs quelconque, au voisinage d'un

point singulier ne ressemblent pas toujours à celles du champs linéaire qui lui est associé.

**Cas où la linéarisée présente un centre.**

Ici, on se place dans le cadre du chapitre 3.

**Exemple 5.3.1**

$$\begin{cases} \frac{dx}{dt} = -y - x\sqrt{x^2 + y^2} \\ \frac{dy}{dt} = x - y\sqrt{x^2 + y^2} \end{cases}, x(0) = (0, .01). \quad (5.24)$$

Le calcul de  $DF(x_0, y_0)$  donne

$$DF(x_0, y_0) = \begin{bmatrix} -1.10^{-2} & -1 \\ 1 & -2.10^{-2} \end{bmatrix}. \quad (5.25)$$

La procédure de calcul de la dérivée optimale donne après 13 itérations ( $\varepsilon = 10^{-10}$ )

$$\tilde{A} = \begin{bmatrix} -0.764510^{-2} & -1 \\ 1 & -0.765710^{-2} \end{bmatrix} \quad (5.26)$$

dont les valeurs propres sont

$$\begin{cases} \tilde{\lambda}_1 = -0.765110^{-2} - i \\ \tilde{\lambda}_2 = -0.765110^{-2} + i \end{cases}$$

$Re(\tilde{\lambda}_1) < 0$  et  $Re(\tilde{\lambda}_2) < 0$ . Donc, la dérivée optimale permet de conclure à la stabilité asymptotique et fait apparaître l'origine comme un foyer. L'équation (5.24) se prête à un calcul direct explicite quand on utilise les coordonnées polaires, et

$$\begin{cases} r' = -r^2 \\ \theta' = 1. \end{cases} \quad (5.27)$$

La solution passant par  $(r_0, \theta_0)$ , où  $r_0 \neq 0$  à  $t = 0$  est donnée par

$$\begin{cases} \rho(t) = \left(t + \frac{1}{r_0}\right)^{-1} \\ \omega(t) = t + \theta_0 \end{cases} \quad (5.28)$$

et alors  $\rho(t) \rightarrow 0$  et  $\omega(t) \rightarrow +\infty$ , quand  $t \rightarrow +\infty$ . Donc, l'origine est un foyer stable pour l'équation non-linéaire. Ceci est en accord avec la conclusion donnée par la dérivée optimale, et vérifie la conjecture (3.3.3) au chapitre 3.

**Exemple 5.3.2**

$$\begin{cases} \frac{dx}{dt} = -y - x(x^2 + y^2) \\ \frac{dy}{dt} = x - y(x^2 + y^2) \end{cases} ; (x_0, y_0) = (0, .1). \quad (5.29)$$

Avec  $DF(x, y)$  calculée en  $(x_0, y_0) = (0, .1)$ ,

$$DF(x_0, y_0) = \begin{bmatrix} -1.10^{-2} & -1 \\ 1 & -3.10^{-2} \end{bmatrix}, \quad (5.30)$$

le calcul de la dérivée optimale donne après 13 itérations, ( $\varepsilon = 10^{-10}$ )

$$\tilde{A} = \begin{bmatrix} -0.554010^{-2} & -1 \\ 1 & -0.554710^{-2} \end{bmatrix}, \quad (5.31)$$

dont les valeurs propres sont

$$\begin{cases} \tilde{\lambda}_1 = -0.5544 + i \\ \tilde{\lambda}_2 = -0.5544 - i. \end{cases}$$

$\operatorname{Re}(\tilde{\lambda}_1) < 0$  et  $\operatorname{Re}(\tilde{\lambda}_2) < 0$ . La dérivée optimale est asymptotiquement stable, et fait apparaître l'origine comme un foyer. Alors, l'origine est asymptotiquement stable. L'équation (5.29), se prête à un calcul direct en utilisant les coordonnées

polaires et

$$\begin{cases} r' = -r^3 \\ \theta' = 1. \end{cases} \quad (5.32)$$

La solution passant par  $(r_0, \theta_0)$ , où  $r_0 \neq 0$  à  $t = 0$  est donnée par

$$\begin{cases} \rho^2(t) = 2 \left( t + \frac{1}{2r_0^2} \right)^{-1} \\ \omega(t) = t + \theta_0 \end{cases} \quad (5.33)$$

on voit bien que  $\rho(t) \rightarrow 0$  et  $\omega(t) \rightarrow +\infty$  quand  $t \rightarrow +\infty$ . Donc, l'origine est un foyer stable pour l'équation non-linéaire. Ceci est en accord avec la conclusion de la dérivée optimale, et vérifie la conjecture (3.3.3) énoncée au chapitre 3.

### Exemple 5.3.3

$$\begin{cases} \frac{dx}{dt} = -y + x(x^2 + y^2) \\ \frac{dy}{dt} = x + y(x^2 + y^2) \end{cases} ; (x_0, y_0) = (0, 1). \quad (5.34)$$

Dans ce cas, la convergence de l'intégrale donnée par la relation (1.3) se pose. Pour le résoudre dans le cas de dimension 2 par exemple, on calcule l'intégrale entre 0 et  $-\infty$ , c'est à dire

$$G(A) = \int_0^{-\infty} \|F(x(t)) - Ax(t)\|^2 dt. \quad (5.35)$$

La procédure de dérivée optimale donne

$$\hat{A} = \begin{bmatrix} 0.554010^{-2} & -1 \\ 1 & 0.554710^{-2} \end{bmatrix}, \quad (5.36)$$

dont les valeurs propres sont

$$\begin{cases} \tilde{\lambda}_1 = 0.5544 + i \\ \tilde{\lambda}_2 = 0.5544 - i. \end{cases}$$

$\operatorname{Re}(\tilde{\lambda}_1) > 0$  and  $\operatorname{Re}(\tilde{\lambda}_2) > 0$ . La dérivée optimale est instable en faisant apparaître l'origine comme un foyer. Alors, l'origine est asymptotiquement instable. L'équation (5.34) se prête à un calcul direct en utilisant les coordonnées polaires, et

$$\begin{cases} r' = r^3 \\ \theta' = 1. \end{cases} \quad (5.37)$$

La solution passant par  $(r_0, \theta_0)$ , où  $r_0 \neq 0$  à  $t = 0$  est donnée par

$$\begin{cases} \rho^2(t) = 2 \left( \frac{1}{2r_0^2} - t \right)^{-1} \\ \omega(t) = t + \theta_0 \end{cases} \quad (5.38)$$

on voit bien que  $\rho(t) \rightarrow \infty$  et  $\omega(t) \rightarrow +\infty$  quand  $t = 1/2r_0^2$ . Donc l'origine est un foyer instable pour l'équation non-linéaire. Ceci est en accord avec la conclusion donnée par la dérivée optimale, tout en vérifiant la conjecture (3.3.3) du chapitre 3.

#### Exemple 5.3.4

$$\begin{cases} \frac{dx}{dt} = -y + x(x^2 + y^2) \sin\left(\frac{\pi}{(x^2 + y^2)^{\frac{1}{2}}}\right) \\ \frac{dy}{dt} = x + y(x^2 + y^2) \sin\left(\frac{\pi}{(x^2 + y^2)^{\frac{1}{2}}}\right) \end{cases}, x(0) = (0, .01) \quad (5.39)$$

La partie non-linéaire admet une dérivée première continue partout et par conséquent, il existe une solution unique pour tout point  $(x, y) \neq 0$  à  $t = 0$ . Les équations en

coordonnées polaires du système s'écrivent

$$\begin{aligned} r' &= r^3 \sin \frac{\pi}{r} \\ \theta' &= 1, \end{aligned} \quad (5.40)$$

les cercles  $r = 1/n$ ,  $n = 1, 2, \dots$  sont des orbites périodiques représentées par les solutions  $\rho(t) = 1/n$ ,  $\theta(t) = t + \theta_0$ ,  $\theta_0$  est constante.

$$\begin{aligned} r' &> 0 & r &> 1 \\ r' &< 0 & \frac{1}{2m} < r < \frac{1}{2m-1} \\ r' &> 0 & \frac{1}{2m+1} < r < \frac{1}{2m} \end{aligned} \quad (5.41)$$

$r = 1/n$  peut être périodique et toute solution non-périodique doit rester complètement à l'intérieur d'une des régions définies par la relation (5.41). Puisque  $\rho$  et  $\omega$  sont monotones quand  $t \rightarrow +\infty$ , ( $\theta \rightarrow +\infty$ ) ces orbites non-périodiques doivent tendre vers les cercles  $r = 1/n$  quand  $t \rightarrow \pm\infty$ , ou bien dans le cas  $r > 1$ ,  $\rho \rightarrow +\infty$ . Donc, l'origine est un centre pour l'équation non-linéaire. La procédure de calcul de la dérivée optimale avec  $DF(x_0, y_0) = DF(0, .01)$  telle que

$$DF(x_0, y_0) = \begin{bmatrix} -0.158510^{-4} & -1 \\ 1 & -0.320410^{-1} \end{bmatrix}, \quad (5.42)$$

donne en intégrant sur  $\mathbb{R}^+$  si la trace est positive, et sur  $\mathbb{R}^-$  si la trace est négative, la série de matrices suivantes

$$A_1 = \begin{bmatrix} -0.755510^{-5} & -1 \\ 1 & 0.575510^{-6} \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0.366510^{-4} & -1 \\ 1 & 0.366910^{-4} \end{bmatrix}$$



$$A_3 = \begin{bmatrix} 0.479010^{-5} & -1 \\ 1 & 0.471610^{-5} \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 0.634110^{-4} & -1 \\ 1 & 0.634310^{-4} \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 0.288910^{-5} & -1 \\ 1 & 0.291210^{-5} \end{bmatrix}$$

$$A_6 = \begin{bmatrix} 0.595910^{-4} & -1 \\ 1 & 0.595810^{-4} \end{bmatrix}$$

$$A_7 = \begin{bmatrix} 0.352810^{-5} & -1 \\ 1 & 0.350710^{-5} \end{bmatrix}$$

$$A_8 = \begin{bmatrix} 0.652510^{-4} & -1 \\ 1 & 0.642410^{-4} \end{bmatrix}$$

$$A_9 = \begin{bmatrix} 0.136310^{-5} & -1 \\ 1 & 0.137010^{-5} \end{bmatrix}$$

*Si on continue le calcul, on remarque que la suite de matrices a la forme suivante*

$$A_j = \begin{bmatrix} \alpha \cdot 10^{-5} & -1 \\ 1 & \alpha \cdot 10^{-5} \end{bmatrix} \quad (5.43)$$

*avec  $\alpha$  variable. Donc il n'y a pas de convergence, et cela veut dire que l'on ne peut pas conclure, ce qui n'est pas en contradiction avec la conjecture (3.3.3).*

### **Cas où la linéarisée présente un noeud.**

Cet exemple concerne le cas où la linéarisée classique en  $(0, 0)$  présente un noeud, alors que l'analyse du système non-linéaire montre que c'est un foyer. Mais ici, en général le problème est lié aux conditions sur la non-linéarité. Ce qui est très

bien illustré par l'exemple suivant.

**Exemple 5.3.5** *Considérons le système différentiel suivant*

$$\begin{cases} \frac{dx}{dt} = -x - \frac{2y}{\ln(x^2 + y^2)} \\ \frac{dy}{dt} = -y + \frac{2x}{\ln(x^2 + y^2)} \end{cases}, (x_0, y_0) = (0, 0.5) \quad (5.44)$$

sur le disque unité ouvert  $x^2 + y^2 < 1$ . Observons que  $\frac{2y}{\ln(x^2 + y^2)}$  se prolonge en une fonction de classe  $C^1$  au voisinage de  $(0, 0)$ . Elle admet en effet une limite égale à 0 à l'origine, ainsi que ses dérivées partielles

$$\frac{-4xy}{(x^2 + y^2)(\ln(x^2 + y^2))^2}, \frac{2}{\ln(x^2 + y^2)} - \frac{4y^2}{(x^2 + y^2)\ln(x^2 + y^2)^2}. \quad (5.45)$$

Il en est de même pour le terme  $\frac{2x}{\ln(x^2 + y^2)}$ . L'origine est donc un point singulier.

La linéarisée au 1er ordre

$$\begin{cases} \frac{dx}{dt} = -x \\ \frac{dy}{dt} = -y \end{cases} \quad (5.46)$$

admet pour valeurs propres  $\lambda_1 = \lambda_2 = -1$ . Le système (5.46) présente donc un noeud propre stable, comme le montre la courbe 1 de la figure (5.9). La linéarisée  $DF(x)$  calculée en  $x_0$  s'écrit

$$DF(x_0, y_0) = \begin{bmatrix} -1 & 3.524 \\ -1.4426 & -1 \end{bmatrix}, (x_0, y_0) = (0, 0.5) \quad (5.47)$$

La procédure de calcul après la (8<sup>ième</sup>) itération donne

$$\tilde{A} = \begin{bmatrix} -1.4934 & 1.2489 \\ -0.5213 & -1.1254 \end{bmatrix}, (x_0, y_0) = (0, 0.5) \quad (5.48)$$

dont les valeurs propres sont

$$\begin{cases} \tilde{\lambda}_1 = -1.3094 - i0.7856 \\ \tilde{\lambda}_2 = -1.3094 + i0.7856. \end{cases}$$

$Re(\tilde{\lambda}_1), Re(\tilde{\lambda}_2) < 0$ , la dérivée optimale fait apparaître l'origine comme un foyer stable, et l'origine est asymptotiquement stable comme le montre la courbe 2 de la figure (5.9). Maintenant on va résoudre le problème initial non-linéaire en utilisant les coordonnées polaires  $(r, \theta)$ , il vient

$$\begin{cases} \frac{dr}{dt} = -\frac{x^2 + y^2}{r} = -r \\ \frac{d\theta}{dt} = \frac{1}{x^2 + y^2} \frac{2x^2 + 2y^2}{\ln(x^2 + y^2)} = \frac{1}{\ln r} \end{cases} \quad (5.49)$$

La solution du problème de Cauchy est donnée par

$$\begin{cases} r = r_0 e^{-t}, r_0 < 1 \\ d\theta = \frac{dt}{\ln(r_0) - t}, \theta = \theta_0 - \ln\left(1 - \frac{t}{\ln r_0}\right). \end{cases} \quad (5.50)$$

Pour une condition initiale  $(r_0, \theta_0)$  en  $t = 0$ , la solution est définie sur  $[\ln r_0, +\infty[$ .  $\lim_{t \rightarrow +\infty} r(t) = 0$ ;  $\lim_{t \rightarrow +\infty} \theta(t) = -\infty$ . On a donc une spirale qui converge vers 0. La courbe s'enroule en spirale à l'intérieur du cercle  $r = 1$  quand  $t \rightarrow \ln(r_0) + 0$ , comme le montre la courbe 3 de la figure (5.9).

**Cas où la non-linéarité n'est pas régulière dans le voisinage du point stationnaire.**

L'exemple (5.2.1) illustre bien le cas où la fonction non-linéaire n'est pas régulière au voisinage du point stationnaire. En utilisant la dérivée au sens de Fréchet, il n'est pas possible de faire une étude sur la stabilité en  $(0, 0)$ . La méthode qu'on propose donne le même comportement que celui du système non-linéaire au voisinage du point stationnaire.

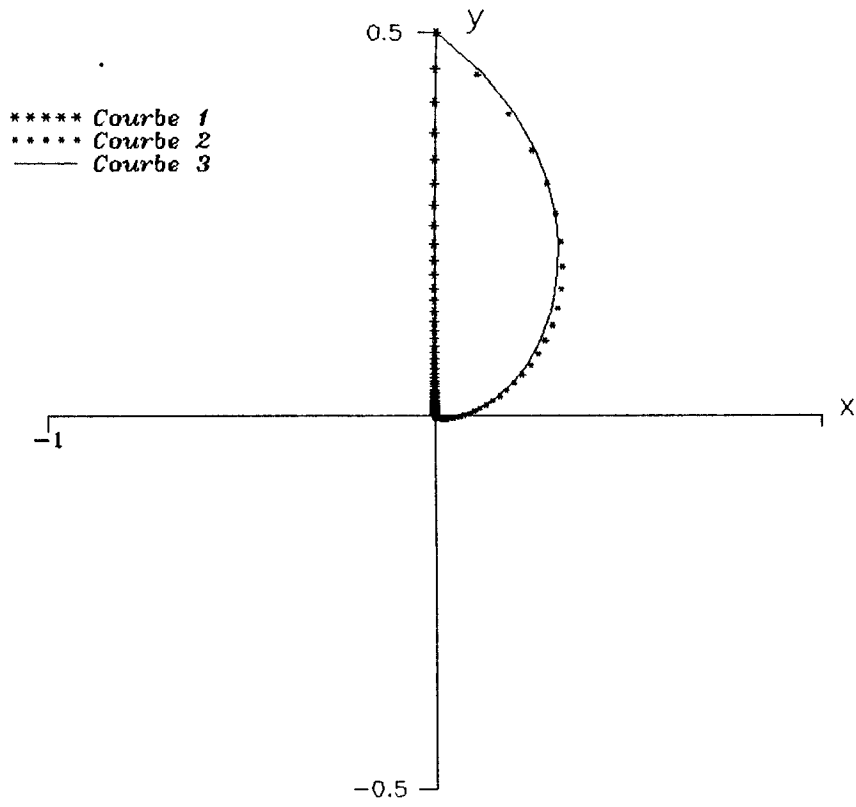


Figure 5.9

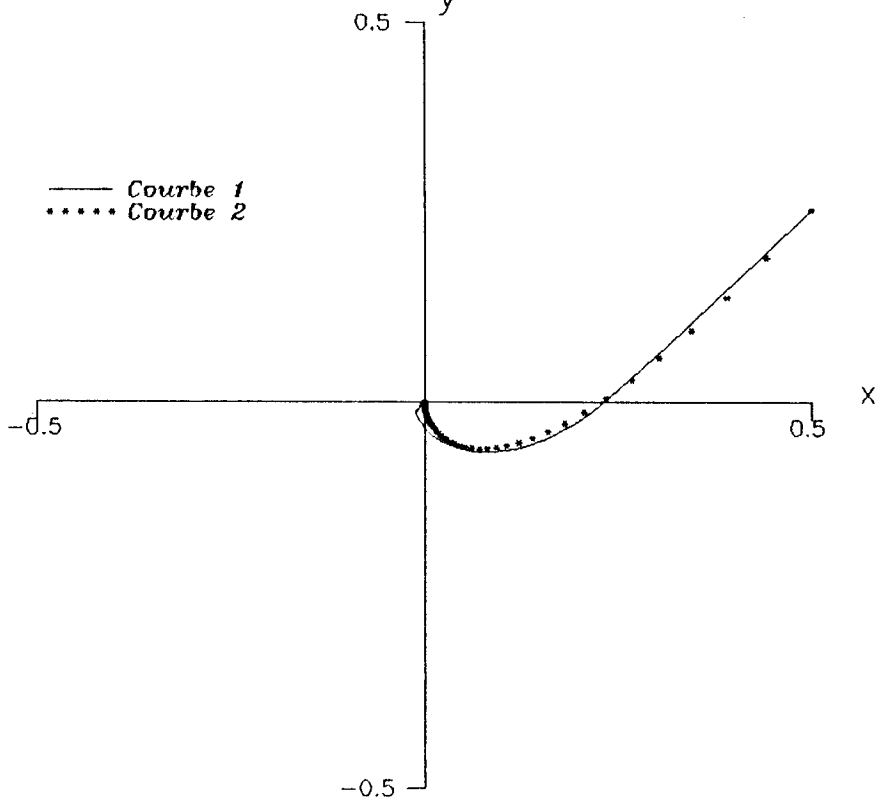


Figure 5.10

La figure (5.10) représente la solution  $(x(t), y(t))$ , dans l'espace des phases, des systèmes (5.9) et (5.11).

### Commentaires.

Dans l'ensemble des exemples qu'on vient de présenter, la conjecture sur la propriété de la stabilité a été vérifiée.

### Cas du centre:

Dans le cas où la perturbation  $\Phi > 0$ , on a intégré sur  $\mathbb{R}^+$ , dans le cas  $\Phi < 0$ , on a intégré sur  $\mathbb{R}^-$ . Ce qui nous a permis, conformément au résultat théorique de voir que la procédure converge vers la matrice optimale. Dans le cas où la perturbation  $\Phi$  change de signe, la procédure ne permet pas de conclure, car il n'y a pas de convergence. Le fait que la procédure ne converge pas est un résultat en lui même dans le sens où on est pas induit en erreur par un résultat contraire.

### Cas du noeud:

L'exemple montre que pour la vérification (les attracteurs du système (3.25) vont dans les attracteurs du système (3.24)) dans le cas d'un noeud, il est nécessaire de restreindre les fonctions  $g_1, g_2$  [1].

On montre que si

$$g_1 = o(r^{1+\epsilon}), g_2 = o(r^{1+\epsilon}), r \rightarrow 0, \text{ pour } \epsilon > 0$$

alors, si l'origine est un noeud propre pour le système linéaire (3.25), il l'est aussi pour le système non-linéaire (3.24).

En général, une condition nécessaire est donnée par le théorème <sup>2</sup> suivant.

**Théorème 5.3.6** [1] Avec  $g_1, g_2$  dans le système non-linéaire (3.24) satisfaisant les inégalités

$$\|g_i(x_1, x_2)\| \leq \Psi(\|x\|), i = 1, 2 \quad (5.51)$$

---

<sup>2</sup>La démonstration de ce théorème se trouve dans la référence [1]

$f(r)$  est une fonction continue, définie sur  $0 \leq r \leq r_0$  et pour  $r \rightarrow 0$

(5.52)

otenu par la procédure de l'approximation optimale.

Avec les données  $t_0 = 0, T = 10, step = 1.10^{-1}, \varepsilon = 10^{-4}$ , on obtient

t	Xnl(t)	Xlin1(t)	Xlin2(t)
0	0.0000000	0.0000000E+00	0.0000000E+00
1	0.1432933E+00	0.1517739E+00	0.1432945 E+00
2	0.6613008 E-01	0.5793614E-01	0.6613086E-01
3	0.2476355E-01	0.1105308E-01	0.2476381E-01
4	0.8627967E-02	-0.3666260E-05	0.8628033E-02
5	0.2898245E-02	-0.8070500E-03	0.2898258E-02
6	0.9504736E-03	-0.3078052E-03	0.9504751E-02
7	0.3056629E-03	-0.5867217E-04	0.3056622E-03
8	0.9643728E-04	0.3895649E-07	0.9643653E-04
9	0.2978096E-04	0.4291440E-05	0.2978051E-04
10	0.8954184E-05	0.1635317E-05	0.8953958E-05

Tableau I

t	Ynl(t)	Ylin1(t)	Ylin2(t)
0	.5000000E+00	.5000000E+00	.5000000E+00
1	0.1153286E+00	0.1177924E+00	0.1153296E+00
2	0.1434192E-01	0.8519892E-02	0.1434179E-01
3	-0.2539680E-02	-0.5333531E-02	-0.2539879E-02
4	-0.3069739E-02	-0.2656959E-02	-0.3069846E-02
5	-0.1717583E-02	-0.6254748E-03	-0.1717630E-02
6	-0.7953907E-03	-0.4509668E-04	-0.7954108E-03
7	-0.3383072E-03	0.2837578E-04	-0.3383153E-03
8	-0.1372356E-03	0.1411885E-04	-0.1372386E-03
9	-0.5404247E-04	0.3321252E-05	-0.5404361E-04
10	-0.2085933E-04	0.2386980E-06	-0.2085973E-04

Tableau II

Dans le tableau III, nous donnons pour comparaison. 1) l'erreur relative  $Er1$  entre la solution supposée exacte et obtenue à partir de la résolution du système non-linéaire (5.44) et la solution calculée à partir du système linéaire optimal (5.48). 2) l'erreur relative  $Er2$ , entre la même solution exacte, et la solution donnée par la procédure d'approximation optimale.

t	Er1	Er2
0	0.	0.
1	0.048	0.8E-05
2	0.148	1.1E-05
3	0.562	1.3E-05
4	0.946	1.3E-05
5	1.14	1.4E-05
6	1.18	1.6E-05
7	1.06	1.7E-05
8	0.932	1.8E-05
9	1.01	1.9E-05
10	0.983	2.E-05

Table III

Dans les figures (5.11) et (5.12), la courbe 1, représente la solution de l'équation non-linéaire, la courbe 2 la solution donnée par l'approximation optimale.

#### 5.4.1 Commentaires.

Cette méthode permet la résolution numérique du problème à valeur initiale. L'exemple (5.3.5), montre que la méthode donne des résultats satisfaisants, dans le sens de l'approximation de la solution, d'abord par rapport à la solution exacte, ensuite par rapport à la solution de la dérivée optimale. Ceci est confirmé par le calcul de l'erreur relative, qui permet de voir que la procédure d'approximation optimale est meilleure, au sens de l'erreur relative, que la dérivée optimale globale.



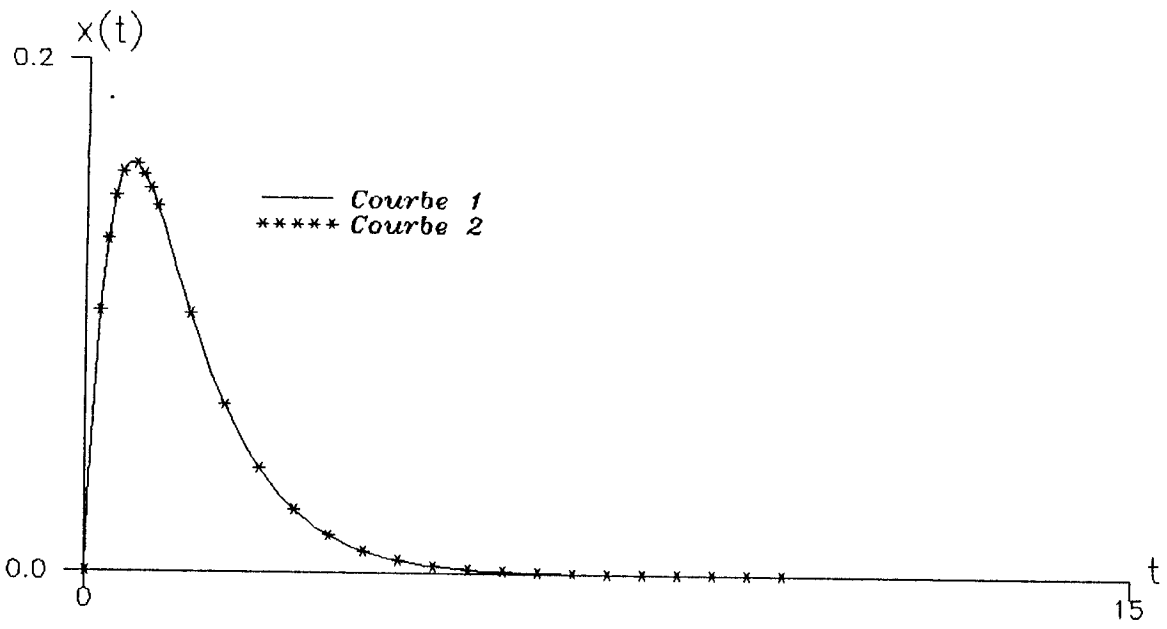


Figure 5.11

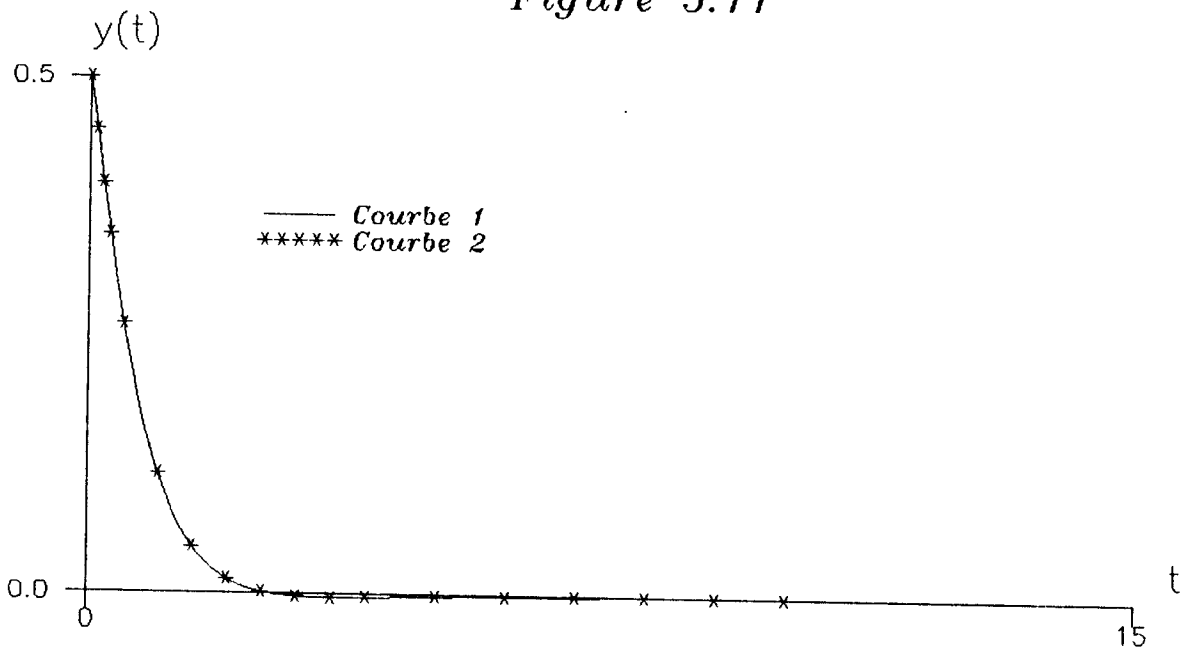


Figure 5.12

L'approche proposée est fondamentalement différente des méthodes existantes, dans le sens où on calcule une approximation de la solution à chaque pas en remplaçant l'équation non-linéaire par la dérivée optimale correspondante dans l'intervalle considéré. L'erreur de la méthode est d'ordre 3 en fonction du pas de discrétisation  $\tau_i$ .

En faisant une comparaison entre les erreurs relatives  $Er_1$  et  $Er_2$ , on remarque une nette amélioration de la précision par rapport à l'approximation donnée par la dérivée optimale. En plus la précision sur la solution obtenue en appliquant la méthode de Runge-Kutta d'ordre 4, à l'équation non-linéaire et en même temps au problème approché par une famille d'équations linéaires optimales permet de conclure à la presque équivalence des problèmes non-linéaires et linéaire optimale.

**CONCLUSIONS GENERALES.**

# CONCLUSIONS GENERALES

Dans cette thèse, nous avons présenté une contribution à l'étude des équations différentielles ordinaires non-linéaires en dimension finie par l'analyse numérique. Les principaux résultats de notre travail sont les suivants:

**1-**Nous avons introduit une procédure d'approximation d'une E.D.O non-linéaire, définie comme la "dérivée optimale", (**def. 1.4.1**).

**2-**Nous avons précisé l'ordre de l'approximation par dérivation optimale (**lem. 3.2.1, prop. 3.2.2**).

**3-**Nous avons, en particulier, prouvé l'existence, l'unicité, et la convergence de la dérivée optimale dans le cas de la perturbation non-linéaire:

a) d'une matrice asymptotiquement stable (**chap. 2**).

b) d'une matrice dont les valeurs propres sont imaginaires (centre) (**chap. 3**).

**4-**Dans le cas scalaire, on a vu que l'approximation converge, et que le résultat représente une fonction de Lyapunov pour l'équation non-linéaire scalaire (**prop. 3.3.1**).

On a montré aussi que si la fonction non-linéaire définissant l'équation est continue et que sa dérivée existe en 0, alors  $\lim_{x_0 \rightarrow 0} \tilde{a}(x_0) = f'(0)$ , (**lem.1.6.2**).

On a vu que même si  $f'(0)$  n'existe pas, l'approximation existe.

**5-**Dans le cas vectoriel, une étude sur le lien entre la dérivée optimale et la stabilité asymptotique a été proposée (**chap. 3**).

**6-**Mise au point d'une méthode d'approximation pour la résolution du problème avec valeur initiale pour une E.D.O non-linéaire en dimension finie, basée sur la dérivation optimale (**chap.4**).

Par ailleurs, les résultats de cette thèse et les autres résultats connus dans le cadre des E.D.O non-linéaires montrent que le problème de l'étude du comportement ainsi que la synthèse de la stabilité dans le cas en particulier de la non régularité de la fonction non-linéaire définissant l'équation sont loin d'être résolus. Ce qui justifie à notre avis, la mise au point et l'utilisation d'autres techniques, méthodes où calculs pour le résoudre.

Enfin, le travail entrepris dans le cadre de cette thèse a ouvert une direction de recherche qui mérite à notre avis, d'être approfondie, en particulier:

- 1-L'approfondissement de la relation entre la stabilité asymptotique et la dérivée optimale,
- 2-L'étude dans le cas critique, c'est à dire le cas d'une valeur propre nulle, ou deux valeurs propres imaginaires.
- 3-L'étude d'une équation dépendant d'un paramètre, dans le cas où l'équation change de stabilité (stable, instable).
- 4-L'étude de l'approximation dans le cas des systèmes de la forme  $F(x, u)$ .
- 5-L'existence, l'unicité et la convergence vers la matrice optimale dans le cas général.

## Références

- [1] **E. A. Coddington and N. Levinson**, (1955), *Theory of ordinary differential equations*, Mc Graw-Hill.
- [2] **W. Ralston & Wilf**, (1960), *Mathematical methods for digital computers*, Wiley, New York, pp.110-120 .
- [3] **R. E. Kalman, E. J. E. Bertran**, (1960), *Control system analysis and design via second method of Liapunov*, journal of basic engineering, 371-393.
- [4] **J. Hale**, (1969), *Ordinary differential equations*, Springer Verlag.
- [5] **N. Rouche, J. Mawhin**, (1973), *Equations différentielles ordinaires, Tome I, Théorie générale*, Masson et C<sup>ie</sup>.
- [6] **N. Rouche, J. Mawhin**, (1973), *Equations différentielles ordinaires, Tome II, Stabilité et solutions périodiques*, Masson et C<sup>ie</sup>.
- [7] **B. Vujanovic**, (1973), *Application of the optimal linearization method to the heat transfer problem. International*, Int. Heat mass transfer, 16, pp. 1111-1117.
- [8] **M. La Porte et J. Vignes**, (1974) *Algorithmes numériques, Analyse et mise en oeuvre*, Ed. Technip.
- [9] **H. Reinhart**, (1975), *Equations différentielles, fondements et applications*. Gauthier-Villars.

- [10] **L. Pontriaguine:** (1975) *Equations différentielles ordinaires*, collection Mir.
- [11] **B. Vujanovic, B. Baclic,** (1975) *Application of Gauss principle of least constraint to the non-linear heat transfer problem*, Int. Heat. Mass Transfer, vol.13, pp. 721-730. Pergamon Press, Oxford.
- [12] **W. Rudin,** (1975), *Analyse complexe et réelle*, Masson et C<sup>ie</sup>.
- [13] **L. E. J. Bouwer,** (1912), *Über Abbildungen von Mannigfaltigkeiten*, Math. Ann. 71, 97-115.
- [14] **A. Jordan et al,** (1987), *Optimal linearization method applied to the resolution of state equation*, RAIRO-APII, 21, 175-185.
- [15] **A. Jordan et al,** (1987), *Optimal linearization of non-linear state equation*, RAIRO-APII, 21, 263-271.
- [16] **M. Sibony et J.CI. Mardon,** (1988), *Analyse numérique II, Approximations et équations différentielles*, Hermann, Editeurs des sciences et arts.
- [17] **J. P. Demailly,** (1991), *Analyse numérique et équations différentielles*, Presses universitaires de Grenoble.
- [18] **T. Benouaz,** *Application of the optimal linearization for study stability problem*, Third International Colloquium on Numerical Analysis, Plovdiv, Bulgaria (13-17 august 1994).
- [19] **T. Benouaz et O. Arino,** (1994), *Existence, unicité et convergence de l'approximation au sens des moindres carrés d'une équation différentielle ordinaire non-linéaire*, 94/14 Université de Pau, CNRS URA 1204.
- [20] **T. Benouaz and O. Arino,** (1995), *Determination of the stability of a nonlinear ordinary differential equation by least square approximation. Computational procedure*, Appl. Math. and Comp sci., N<sup>o</sup>.1, Vol. 5, pp 33-48.

- [21] **T. Benouaz et O. Arino**, (1995) *Relation entre l'approximation optimale et la stabilité asymptotique*, Publications de l' U.A, CNRS 1204, N<sup>o</sup>.95/10.
- [22] **T. Benouaz and O. Arino**, (1996), *Least square approximation of a nonlinear ordinary differential equation*, Computers Math. Applic. , Vol. 31, N<sup>o</sup>8, pp. 69-84.
- [23] **T. Benouaz**, *Least square approximation of a nonlinear ordinary differential equation: The scalar case*, à paraitre dans the "Proceeding of the Fourth International Colloquium on Numerical Analysis", Plovdiv, Bulgaria (13-17 august 1995).
- [24] **T. Benouaz**, *Approximation of a nonlinear ordinary differential equation by an optimal procedure*, à paraitre dans the "Proceeding of the 2nd International Conférence on Differential Equations in Marrakech", (16-20 june1995).
- [25] **T. Benouaz and O. Arino**, *Optimal approximation of the initial value problem*, (Submitted).



