A rigorous numerical method for the proof of Galaktionov-Svirshchevskii's conjecture

Mémoire

Abdoulaye Thiam

Maîtrise en mathématiques
Maître ès sciences (M.Sc.)

Québec, Canada

Résumé

La théorie des systèmes dynamiques étudie les phénomènes qui évoluent au cours du temps. Plus précisément, un système dynamique est donné par : un espace de phase dont les points correspondent à des états possibles du système étudié et une loi d'évolution décrivant l'infinitésimal (pour le cas continu) pas à pas (pour le cas discret) les changements des états du système. Le but de la théorie est de comprendre l'évolution dans le long terme. Dans ce travail, nous présentons une nouvelle méthode pour la résolution des systèmes linéaires avec preuve assistée par ordinateur dans le cadre de modèles linéaires réalistes. Après une introduction de quelques propriétés de la théorie des équations différentielles ordinaires, on introduit une méthode de calcul rigoureux pour trouver la solution périodique de la conjecture de Galaktionov-Svirshchevskii. On reformule le problème comme un problème à valeur initiale, puis on calcule la solution périodique dans le domaine positif et on déduit l'autre solution par symétrie. Notre résultat énonce une partie de la conjecture 3.2 dans le livre de Victor A. Galaktionov & Sergey R. Svirshchevskii *: Exact Solutions and Invariant Subspaces of Nonlinear Partial Differential Equations in Mechanics and Physics,* [Chapman & Hall/CRC, applied mathematics and nonlinear science series, (2007)].

Mots clés. Conjecture de Galaktionov-Svirshchevskii, Analyse d'intervalle, Théorème de contraction de Banach, Polynômes de rayons.

Abstract

The theory of dynamical systems studies phenomena which are evolving in time. More precisely, a dynamical system is given by the following data: a phase space whose points correspond to the possible states of the system under consideration and an evolution law describing the infinitesimal (for continuous time) or one-step (for discrete time) change in the state of the system. The goal of the theory is to understand the long term evolution of the system. In this work, we introduce a new method for solving piecewise linear systems with computer assisted proofs in the context of realistic linear models. After introducing some properties of the theory of ordinary differential equations, we provide a rigorous computational method for finding the periodic solution of Galaktionov-Svirshchevskii's conjecture. We reformulate the problem as an initial value problem, compute periodic solution in the positive domain and deduce the other solution by symmetry. Our result settles one part of the *Conjecture* 3.2 by Victor A. Galaktionov & Sergey R. Svirshchevskii*: Exact Solutions and Invariant Subspaces of Nonlinear Partial Differential Equations in Mechanics and Physics,* [Chapman & Hall/CRC, applied mathematics and nonlinear science series, (2007)].

Key words. Galaktionov-Svirshchevskii's conjecture, Interval analysis, Contraction mapping theorem, Radii polynomials.

Contents

Résumé iii

Abstract v

Contents vii

List of Figures ix

Acknowledgments xiii

Introduction 1

1 Interval Analysis in Matlab 3
- 1.1 Interval Arithmetic 4
- 1.2 Introduction to INTLAB 8

2 Basic notions of ordinary differential equations 13
- 2.1 Contraction mapping theorem 15
- 2.2 Existence and uniqueness of ODEs 18
- 2.3 Nonhomogeneous Linear Systems 27
- 2.4 Stability for Linear Systems 29

3 Rigorous computations for finite dimensional problems 31
- 3.1 A Newton-like operator 31
- 3.2 The radii polynomial approach in finite dimension 33

4 Galaktionov-Svirshchevskii's conjecture 41
- 4.1 Derivation of the Galaktionov-Svirshchevskii's conjecture 41
- 4.2 Set-up of the problem $\boldsymbol{f(x) = 0}$ 43
- 4.3 The radii polynomial approach for Galaktionov-Svirshchevskii's conjecture . 45
- 4.4 Validated numerics of Galaktionov-Svirshchevskii's conjecture 46

Conclusion 49

Bibliography 51

A numerical codes 53

List of Figures

0.1 The stable periodic behavior in (1). 1

1.1 Inclusion property of interval arithmetic. 7
1.2 Example of complex interval multiplication. 9

3.1 The radii polynomial $p(r) = r^2 - r + 0.130$, and $I = [0.16, 0.80]$, where the radii polynomial is strictly negative. 37

4.1 radii polynomials generated with ϕ. 47
4.2 graph of the period solution for Galaktionov-Svirshchevskii's conjecture. 47

to Jean-Christophe YOCCOZ,
who was awarded a Fields Medal
in 1994, for his work on
dynamical systems.

Acknowledgments

First I would like to thank my supervisor Professor Jean-Phillipe Lessard, whose generosity, energy and ideas have been essential during my two years of work with him. He gave me great opportunities and helpful financial support. Thank you Professor for your recommendation and your guidance during this project. He never stopped encouraging me.

Special thanks to Professor Jean-Christophe YOCCOZ at College de France (Paris), to his support, help and encouragement.

My Acknowledgements also go to all staff, lectures and students at the department of mathematics and statistic at the university of Laval, for giving me the opportunity to learn with them.

I would like to thanks the hospitality of Professor Robert Guenette, Sylvie Reny and Esther Payeur.

Lastly I offer my regards to all of those who supported me in any respect during the completion of this essay.

I dedicate this project to my defunct father Cheikh Thiam, my mother Ndack Tall, my sister Fatou Thiam, my uncle Vieux Tall and to all my family in Thies (Senegal).

Introduction

In general, the applied mathematician faces a highly nontrivial, perhaps impossible, task when trying to rigorously verify the hypotheses of general theorems for realistic models of physical systems. In fact, doing so might require the development of a new area of mathematics. Most often, we are left to face the realization that rigorous results can only be obtained for simplified models. [5, page 375]

We consider the autonomous fourth-order ODE

$$\phi^{(4)} + 10\phi^{(3)} + 35\phi^{(2)} + 50\phi' + 24\phi + sign(\phi) = 0, \tag{0.1}$$

where ϕ is the oscillatory component of the function $h(w) = w^4\phi(s)$, $s = \ln(w)$ and $h^{(4)} + sign(h) = 0$. Existence and uniqueness of periodic solutions are open questions. The Galaktionov-Svirshchevskii's conjecture states that (0.1) has a unique nontrivial periodic solution $\phi(s)$ which is asymptotically stable as $s \longrightarrow +\infty$. The Figure (0.1)[1] shows the stable periodic motion obtained for different initial data.

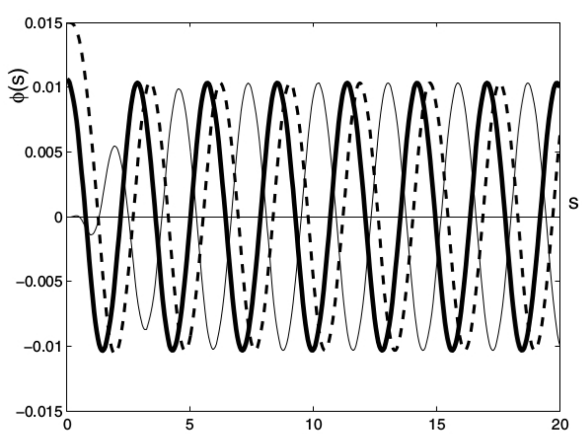

Figure 0.1: The stable periodic behavior in (1).

[1] Solutions and Invariant Subspaces of Nonlinear Partial Differential Equations in Mechanics and Physics, Chapman & Hall/CRC, applied mathematics and nonlinear science series, (2007), page 150.

In this project, a rigorous computational method to compute solutions of piecewise smooth systems is introduced. A general theory based on the radii polynomial approach is proposed to compute the periodic solution of Galaktionov-Svirshchevskii's conjecture. We present a form of the above mentioned "new area of mathematics". We introduce the theory of ordinary differential equation with computer assisted proofs in the context of realistic linear models. In the opening chapter, we introduce a method to compute with intervals of real numbers instead of real numbers. This is called interval analysis and we will use interval arithmetic and Intlab to obtain rigorous bounds for the exact solution. In chapter two, we introduce some classical results of ordinary differential equations and nonhomogeneous linear systems. The goal of chapter three is to develop a constructive method to prove the existence of zeros of finite dimensional maps and provide bounds on their location. This is done using the radii polynomial approach which is a variant of Newton's method. In chapter four, we introduce a rigorous computational method for finding the periodic solution of Galaktionov-Svirshchevskii's conjecture and validated numerics of this conjecture. We finish this work by giving some numerical codes in Matlab for our conjecture.

Chapter 1

Interval Analysis in Matlab

The concept of interval analysis is to compute with intervals of real numbers instead of real numbers. Sometimes floating point arithmetic has rounding errors and can produce inaccurate results. We use interval arithmetic to obtain rigorous bounds for the location of numerically computational solution. To do computations with some unknown parameters within a certain interval, it is recommended to use interval analysis. Algorithms may be implemented using interval arithmetic with uncertain parameters as intervals to produce an interval that bounds all possible results. If the lower and upper bounds of the interval can be rounded down and rounded up respectively then finite precision calculations can be performed using intervals, to give an enclosure of the exact solution. As it is not difficult to implement existing algorithms using intervals in place of real numbers, the result may be of no use if the interval obtained is too large. In this case, other computational methods must be considered or new ones developed in order to make the interval result as narrow as possible. In the 1950's, several people worked in this area but interval analysis began with a book *Interval Arithmetic. Prentice-Hall, Englewood Cliffs, NJ, USA, by Ramon Edgar Moore*[1] *in 1966.* It is an approach to bound rounding errors in a mathematical computation. This theory emerged considering the computation of the exact solution and the error as the single entity, *i.e* the interval. It is a powerful technique with many applications in mathematics, engineering, computer science, etc...

The goal of this chapter is to give some basic concepts of interval arithmetic using Intlab[2], which is a toolbox in Matlab for self-validating algorithms. Intlab supports real and complex intervals, vectors, full and sparse matrices, rigorous standard functions, multiple precision interval arithmetic and automatic differentiation. It can be implemented to all algorithms in Matlab and it is designed to be very fast.

[1]Ramon Edgar Moore, born in 1929 in California, is an American mathematician.
[2]source: http://www.ti3.tu-harburg.de/~ rump/intlab/index.html

1.1 Interval Arithmetic

As we are computing intervals instead of real numbers, they will be represented by brackets [.], used for intervals defined by an upper bound and a lower bound. With intervals defined by a radius and a midpoint, the brackets $< . >$ will be used.

1.1.1 Real Interval Arithmetic

We define a real interval $\underline{\overline{x}}$ as a nonempty set of reals numbers

$$\underline{\overline{x}} = [\underline{x}, \overline{x}] = \{x \in \mathbb{R} : \quad \underline{x} \leq x \leq \overline{x}\}.$$

$\underline{x}$ is the infimum, $\overline{x}$ the supremum and the set of all intervals over $\mathbb{R}$ is denoted by $\underline{\overline{\mathbb{R}}}$ where

$$\underline{\overline{\mathbb{R}}} = \{[\underline{x}, \overline{x}] : \underline{x}, \overline{x} \in \mathbb{R},\ \underline{x} \leq \overline{x}\}.$$

The midpoint and the radius of $\underline{\overline{x}}$ are denoted by

$$mid(\underline{\overline{x}}) = \tilde{x} = \frac{1}{2}(\underline{x} + \overline{x}), \quad rad(\underline{\overline{x}}) = \frac{1}{2}(\overline{x} - \underline{x}).$$

They can be used to define an interval $\underline{\overline{x}} \in \underline{\overline{\mathbb{R}}}$. In this case, if m is the midpoint and r the radius, this interval is represented by $< m,r >$. A point interval is an interval with zero radius, it contains a single point represented by

$$[x,x] \equiv x.$$

If the radius is greater than zero, then $\underline{\overline{x}}$ is a thick interval. The magnitude or absolute value and the mignitude of $\underline{\overline{x}}$ are defined as follows

$$|\underline{\overline{x}}| = mag(\underline{\overline{x}}) = max\{|\tilde{x}|, \quad \tilde{x} \in \underline{\overline{x}}\}, \quad mig(\underline{\overline{x}}) = min\{|\tilde{x}|, \quad \tilde{x} \in \underline{\overline{x}}\}.$$

We say that $\underline{\overline{x}} \subseteq \underline{\overline{y}}$ if, and only if $\underline{y} \leq \underline{x}$ and $\overline{y} \geq \overline{x}$. This inclusion is not an equivalence relation as it is not symmetric. The relation $\underline{\overline{x}} < \underline{\overline{y}}$ implies $\overline{x} < \underline{y}$.

Let $\underline{\overline{x}} = [\underline{x}, \overline{x}]$, $\underline{\overline{y}} = [\underline{y}, \overline{y}] \in \underline{\overline{\mathbb{R}}}$, the four elementary operations are defined by

$$\underline{\overline{x}}\ op\ \underline{\overline{y}} = \{x\ op\ y, \quad x \in \underline{\overline{x}}, \quad y \in \underline{\overline{y}}\} \qquad \text{for} \quad op \in \{+, -\ , \times\ , \div\}.$$

As the classical arithmetic operates on real numbers, interval arithmetic defines a set of operations on intervals. The base arithmetic operations are as follows.

Definition 1.

$$\underline{\overline{x}} + \underline{\overline{y}} = [\underline{x} + \underline{y}, \overline{x} + \overline{y}], \tag{1.1}$$

$$\underline{\overline{x}} - \underline{\overline{y}} = [\underline{x} - \overline{y}, \overline{x} - \underline{y}], \tag{1.2}$$

$$\underline{\overline{x}} \times \underline{\overline{y}} = [min(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}), max(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y})], \tag{1.3}$$

$$1/\underline{\overline{x}} = [1/\overline{x}, 1/\underline{x}], \quad \underline{x} > 0 \quad \text{or} \quad \overline{x} < 0, \tag{1.4}$$

$$\underline{\overline{x}} \div \underline{\overline{y}} = \underline{\overline{x}} \times 1/\underline{\overline{y}}. \tag{1.5}$$

1.1.2 Why interval arithmetic ?

There are many sources of errors by using numerical computations. Rounding, truncation and input errors are usually frequent. Let us give some examples to show how interval arithmetic is meant to keep track of them.

Example 1. (Rounding errors)
Let us consider the function $g(x) = 1 - x + \frac{x^2}{2}$, with $x = 0.531$ $i.e$ with 10^{-3} precision. If we evaluate this expression with classical arithmetic, we obtain the result $g(x) = 0.610$. And by computing this expression using interval arithmetic, we obtain

$$g(x) = 0.469 + \frac{0.531^2}{2} \in 0.469 + \tfrac{[0.281, 0.282]^2}{2}.$$

Thus $g(x) \in 0.469 + [0.140, 0.141] = [0.609, 0.610]$. This shows that the exact result is within the interval $[0.609, 0.610]$.

Example 2. (Truncation errors)
We are interested in the Taylor series of the exponential function $e^x = 1 + x + \frac{x^2}{2!}e^{\psi}$, with $\psi \in [0,x]$. If $x < 0$, $e^x \in 1 + x + \frac{x^2}{2!}[0,1]$. For $x = -0.531$, we obtain

$$e^{-0.531} \in 1 - 0.531 + \tfrac{(-0.531)^2}{2!}[0,1] = 0.469 + [0.140, 0.141][0,1] = [0.469, 0.610].$$

This shows how interval arithmetic keeps track of both, the truncation and the rounding errors.

Example 3. (Input errors)
Due to data uncertainly, suppose that $x \in [-0.532, -0.531]$. By using the previous expression, we obtain

$$\begin{aligned} e^x \in 1 + [-0.532, -0.531] + \frac{[-0.532, -0.531]^2}{2!}[0,1] &= [0.468, 0.470] + \frac{[0.280, 0.284]^2}{2}[0,1], \\ &= [0.468, 0.470] + [0, 0.142], \\ &= [0.468, 0.612]. \end{aligned}$$

This example shows how interval arithmetic can keep track of all error types simultaneously.

Remark 1. Note that division by zero is not defined for the elementary interval operations. We have to remove this restriction and apply what is called extended interval arithmetic. Let us finish this section by some important properties of inclusions, this theorem is labelled as the fundamental theorem of interval analysis.

Theorem 1.1. Moore's fundamental theorem
Suppose that the function $f(\underline{\overline{y}}_1, ..., \underline{\overline{y}}_n)$ defined an arithmetical expression (expression that results in a numeric value) with a finite number of intervals, $\underline{\overline{y}}_1, ...\underline{\overline{y}}_n \in \underline{\overline{\mathbb{R}}}$. Consider the four interval operations $(+, -, \times, \div)$. If

$$\underline{\overline{x}}_1 \subseteq \underline{\overline{y}}_1, ..., \underline{\overline{x}}_n \subseteq \underline{\overline{y}}_n$$

then

$$f(\underline{\overline{x}}_1, \underline{\overline{x}}_2, ..., \underline{\overline{x}}_n) \subseteq f(\underline{\overline{y}}_1, \underline{\overline{y}}_2, ..., \underline{\overline{y}}_n).$$

Proof 1.2. Suppose that $\underline{\overline{x}} \subseteq \underline{\overline{y}}$ and $\underline{\overline{u}} \subseteq \underline{\overline{v}}$. By the previous interval arithmetic operations, we have

$$\begin{aligned} \underline{\overline{x}} + \underline{\overline{u}} &\subseteq \underline{\overline{y}} + \underline{\overline{v}}, \\ \underline{\overline{x}} - \underline{\overline{u}} &\subseteq \underline{\overline{y}} - \underline{\overline{v}}, \\ \underline{\overline{x}} \times \underline{\overline{u}} &\subseteq \underline{\overline{y}} \times \underline{\overline{v}}, \\ \underline{\overline{x}} / \underline{\overline{u}} &\subseteq \underline{\overline{y}} / \underline{\overline{v}}. \end{aligned}$$

By the inclusion relation, we have

$$\underline{\overline{y}} \subseteq \underline{\overline{u}} \quad \text{and} \quad \underline{\overline{u}} \subseteq \underline{\overline{v}} \quad \Longrightarrow \underline{\overline{y}} \subseteq \underline{\overline{v}},$$

and by induction argument, the result follows.

Remark 2. By Moore's fundamental theorem of interval arithmetic, any function f defined by an arithmetical expression has a corresponding interval evaluation function F which is an inclusion function of f :

$$F(\underline{\overline{x}}) \supseteq f(\underline{\overline{x}}) = \{f(x) : x \in \underline{\overline{x}}\}.$$

It is an advantage in the sense that there is no restrictions to a particular class of functions that it can be applied to.

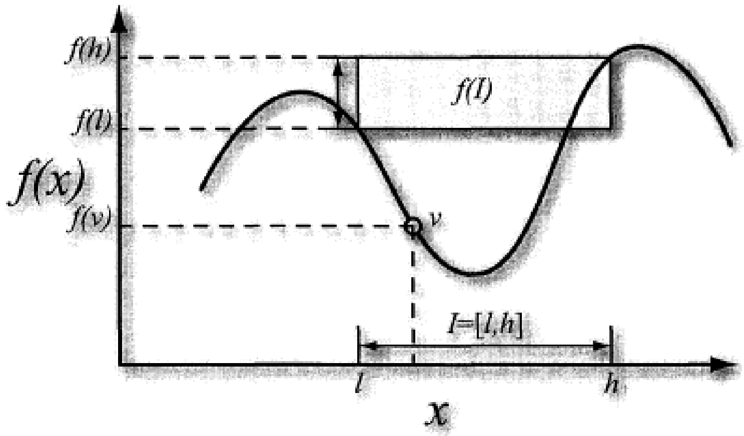

Figure 1.1: Inclusion property of interval arithmetic.

[3]

1.1.3 Interval Vectors and Matrices

Definition 2. Let $n \in \mathbb{N}$, an interval vector $\underline{\overline{X}} = \{\underline{\overline{x}}_1, \underline{\overline{x}}_2, ..., \underline{\overline{x}}_n\} \in \underline{\overline{\mathbb{R}}}^n$ is defined to be a vector with interval components, $\underline{\overline{x}}_1$, $\underline{\overline{x}}_2$, ...,$\underline{\overline{x}}_n \in \underline{\overline{\mathbb{R}}}$.

Definition 3. Let m,$n \in \mathbb{N}$, an $m \times n$ interval matrix $\underline{\overline{\mathcal{M}}} \in \underline{\overline{\mathbb{R}}}^{m\times n}$ is defined to be a matrix with interval components.

Remark 3. A point matrix or point vector has components all with zero radius, otherwise it is said to be a thick matrix or a thick vector.

The operations with interval vectors and interval matrices are carried out according to the operations on $\underline{\overline{\mathbb{R}}}$.

Definition 4. Let $\underline{\overline{X}} = \{\underline{\overline{x}}_1, \underline{\overline{x}}_2, ..., \underline{\overline{x}}_n\}$, $\underline{\overline{Y}} = \{\underline{\overline{y}}_1, \underline{\overline{y}}_2, ..., \underline{\overline{y}}_n\} \in \underline{\overline{\mathbb{R}}}^n$. We say that

$$\underline{\overline{X}} > \underline{\overline{Y}},\ if\ \underline{\overline{x}}_i > \underline{\overline{y}}_i \text{ for all } i.$$

As for vectors in $\mathbb{R}^n$, the infinity norm of an interval vector $\underline{\overline{X}} \in \underline{\overline{\mathbb{R}}}^n$ is given by

$$\|\underline{\overline{X}}\|_\infty = max\{\ |\underline{\overline{x}}_i|,\ i = 1, 2,...,n\}.$$

[3]Figure 1.1, Y. Hijazi, H. Hagen, C. Hansen, and Kenneth I. Joy, Why interval arithmetic is so useful subs.emis.de/LNI/Seminar/Seminar07/148.pdf, page 151.

1.1.4 Complex Interval Arithmetic

Definition 5. Let $\underline{\overline{x}}$, $\underline{\overline{y}} \in \underline{\overline{\mathbb{R}}}$ be real intervals. A rectangular complex interval is defined as

$$\underline{\overline{x}} + i\underline{\overline{y}} = \{x + iy, x \in \underline{\overline{x}}, y \in \underline{\overline{y}}\}.$$

It produces a rectangle of complex numbers in the complex plane with sides parallel to the coordinates axis. As for complex numbers in $\mathbb{C}$, complex interval operations are defined in the same way.

Example 4. Let $\underline{\overline{x}} = [1, 2] + i[1, 2]$ and $\underline{\overline{y}} = [3, 4] + i[3, 4]$, then

$$\underline{\overline{x}} \times \underline{\overline{y}} = [-5, 5] + i[6, 16].$$

Remark 4. The multiplication of rectangular complex intervals produces a rectangle in the complex plane. However if $\underline{\overline{x}} = \underline{\overline{x}}_1 + i\underline{\overline{x}}_2$ and $\underline{\overline{y}} = \underline{\overline{y}}_1 + i\underline{\overline{y}}_2$, the result

$$\underline{\overline{x}} \times \underline{\overline{y}} = \underline{\overline{x}}_1\underline{\overline{x}}_2 - \underline{\overline{y}}_1\underline{\overline{y}}_2 + i(\underline{\overline{x}}_1\underline{\overline{y}}_2 + \underline{\overline{x}}_2\underline{\overline{y}}_1)$$

produces a rectangle with actual range not in this shape. This implies that an overestimation of $\underline{\overline{x}} \times \underline{\overline{y}}$ is calculated. In the previous example, the product $\underline{\overline{x}} \times \underline{\overline{y}}$ is not rectangular but lies in the dashed lines for *Figure* 1.2[4].

Remark 5. (Outward Rounding) Note that if $\underline{x}$ and $\overline{x}$ are not machine numbers, the result $\underline{\overline{x}} = [\underline{x},\overline{x}]$ may not be representable on a machine. If $\underline{x}$ and $\overline{x}$ are not rounded, the rounded interval may not bound the original interval. We have to apply the *Outward Rounding*, *i.e.* $\underline{x}$ must be rounded downward and $\overline{x}$ rounded upward, to obtain $x \in \underline{\overline{x}}$. For floating point arithmetic, the standard ubiquitus IEEE has four rounding modes, nearest, round down, round up and round towards zero. Intlab uses the routine setround to change the rouding mode of the processor between nearest, round up and round down. This choice is used to create functions for input, output and arithmetic intervals.

Let us introduce some basic functions of Intlab in the next section.

1.2 Introduction to INTLAB

To initialize global variables and add the Intlab directories to the MATLAB seach path, type the command $\gg$ *startintlab*.

[4] source: Interval Analysis in MATLAB, G. I. Hargreaves, Numerical Analysis Report No. 416, Dec 2002.

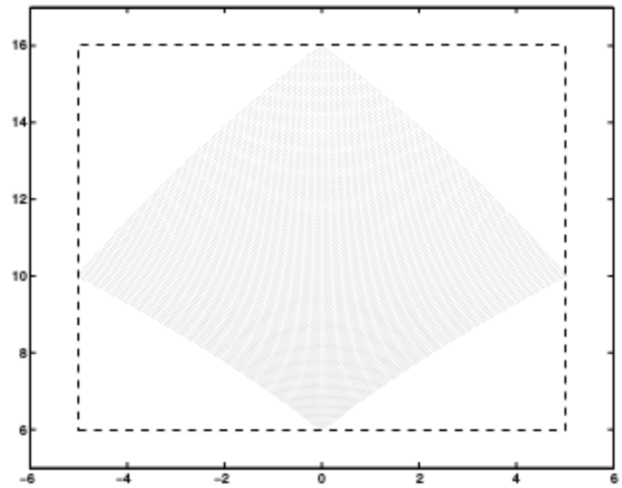

Figure 1.2: Example of complex interval multiplication.

Example 5. We can represent the interval $\underline{\overline{x}} = [1{,}1]$, using infimum and supremum as
$\gg \underline{\overline{x}} = infsup(-1{,}1)$.
We can also use the midpoint and the radius to represent the same interval as
$\gg \underline{\overline{x}} = midrad(0{,}1)$.
A circular region with midpoint $2 + i$ and radius 3 is represented by
$\gg x = midrad(2 + i{,}3)$.
Note that if we use the infimum and supremum to represent a rectangular region then it is stored with an overestimation as the smallest circular region enclosing it. The region with a infimum of $1 + i$ and a supremum of $2 + 2i$ is represented by
$\gg \underline{\overline{z}} = infsup(1 + i{,}2 + 2i)$.
The command
$\gg midrad(\underline{\overline{z}}) =< 1.50000000000000 + 1.50000000000000i, 0.70710678118655 >$
gives us the midpoint and the radius.

The function *intval* provides directly an interval variable. It can also be used to change variable to interval type. It gives us an interval and verified bound. We know that the value 0.1 cannot be expressed in binary floating point arithmetic, but we can use *intval* to provide a rigorous bound. In this case, this variable will not contains an interval including $\underline{\overline{x}} = 0.1$, since it is converted to binary format before being passed to *Intval*. So a thin interval is obtained, with radius $rad(\underline{\overline{x}}) = 0$.

We can use *intval* with a string argument to obtain rigorous bounds. For example, let

$\gg \underline{\overline{x}} = intval('0.1')$,
then we get
$\gg intval(\underline{\overline{x}}) = [0.09999999999999, 0.10000000000001]$.
This command use an INTLAB verified conversion to binary. Finally, $\underline{\overline{x}}$ contains 0.1 since the radius is
$\gg rad(\underline{\overline{x}}) = 1.387778780781446e^{-017}$.
The commands $inf(\underline{\overline{x}})$, $sup(\underline{\overline{x}})$, $mid(\underline{\overline{x}})$ and $rad(\underline{\overline{x}})$ give us the infimum, supremum, midpoint and radius of a interval $\underline{\overline{x}}$.Interval matrices and vectors are entered in a similar way by using the arguments being matrices or vectors with required size. In this example, we are going to show that the interval arithmetic operations are not distributive.

Example 6. Let us compute $I_1^2 I_2^3 / I_3^4$ and $e^{I_1}\sin(I_2 I_3)$, with

$$I_1 = [-1, 2],\ I_2 = [-10.987654321, 1.23456789] \text{ and } I_3 = [\pi, \pi + 0.1].$$

We have e^{I_1} = [0.3678,7.3890] and by the definition 1 we get

$$\begin{aligned} sin(I_2 I_3) = [&min(sin(I_2(1).I_3(1)), sin(I_2(1).I_3(2)), sin(I_2(2).I_3(1)), sin(I_2(2).I_3(2))),\\ &max(sin(I_2(1).I_3(1)), sin(I_2(1).I_3(2)), sin(I_2(2).I_3(1)), sin(I_2(2).I_3(2)))],\\ = [&min(sin(-34.5187), sin(-35.6175), sin(3.8785), sin(4.0019),\\ &max(sin(-34.5187), sin(-35.6175), sin(3.8785), sin(4.0019)],\\ = [&min(-0.5666, -0.5823, 0.067, 0.069), min(-0.5666, -0.5823, 0.067, 0.069)],\\ = [&-0.5823, 0.069]. \end{aligned}$$

then

$$e^{I_1} sin(I_2 I_3) = [0.3678,7.3890].[-0.5823,0.069] = [-0.2141, 0.5098].$$

On the other hand,

$$I_1^2 = [-1.000, 2.000].[-1.000,2.000] = [1.000,4.000],$$

$$\begin{aligned} I_2^3 = ([-10.9876, 1.2345])^3 &= [-10.9876, 1.2345].[-10.9876, 1.2345].[-10.9876, 1.2345],\\ &= [-1326.5038,1.8], \end{aligned}$$

$$I_3^4 = [\pi^4, (\pi + .1)^4] = [97.4091, 110.4164] \Rightarrow 1/I_3^4 = [0.0090,0.0102].$$

So we get,

$$\begin{aligned} I_1^2 I_2^3 / I_3^4 &= [1.000,4.000].[-1326.5038,1.8].[0.0090,0.0102], \\ &= [-11.9385,0.0734]. \end{aligned}$$

We have also,

$$I_1(I_2 + I_3) = [-1.000,2.000].[-7.8460, 4.4761] = [-15.6922,8.9524],$$

$I_1 I_2 = [-21.9753,10.9876]$ and $I_1 I_3 = [-3.2416,6.4832]$. This implies that

$$I_1 I_2 + I_1 I_3 = [-25.2169,17.4708].$$

We conclude that the interval arithmetic operations are not distributive.

Remark 6. To minimize errors, it is necessary to simplify operations for interval arithmetic. In the previous example, $I_1(I_2 + I_3)$ gives us the desired result.

Chapter 2

Basic notions of ordinary differential equations

In this chapter, we introduce some basic notions of ordinary differential equations (ODEs). More precisely, we provide results concerning existence, uniqueness, and continuity of solutions for ODEs (with respect to initial conditions).

More precisely, we establish the fundamental theorem of linear systems and provide results concerning existence, uniqueness, and continuity of solutions to ODEs (with respect the initial conditions). We introduce a study of linear systems of the form $\dot{x} = Ax$, where $x \in \mathbb{R}^n$ and A is an $n \times n$ matrix and solving the associated nonhomogeneous linear systems.

Definition 6. Let $\mathcal{X}$ be a set. A function $d : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}^+$ is called a metric if it satisfies the following properties

$$
\begin{aligned}
&d(x,y) = 0 \Leftrightarrow x = y, \quad \text{(positive definitness)},\\
&d(x,y) = d(y,x), \quad \text{for all} \quad x,y \in \mathcal{X} \quad \text{(symmetric)},\\
&d(x,z) \leq d(x,y) + d(y,z), \quad \text{for all} \quad x,y,z \in \mathcal{X} \quad \text{(triangle inequality)}.
\end{aligned}
$$

A paire $(\mathcal{X},d)$, where d is a metric is called a metric space.

Example 7. $\mathcal{X} = \mathcal{C}([0,1]) = \{f : f \quad \text{is continuous on } [0,1]\}$,

$$
\begin{aligned}
d_1(f,g) &= \int_0^1 |f(x) - g(x)| dx,\\
d_\infty(f,g) &= max_{x\in[0,1]} |f(x) - g(x)|,\\
d_2(f,g) &= \left(\int_0^1 |f(x) - g(x)|^2 dx \right)^{1/2}.
\end{aligned}
$$

$(\mathcal{X},d_1)$, $(\mathcal{X},d_2)$, and $(\mathcal{X},d_\infty)$ are a metric spaces.

Definition 7. Let $(\mathcal{X},d)$ be a metric space. A sequence (x_n) in $\mathcal{X}$ is called a Cauchy sequence if for any $\epsilon > 0$, there is an $n_\epsilon \in \mathbb{N}$ such that

$$d(x_m,x_n) < \epsilon, \quad \text{for any} \quad m, n \geq n_\epsilon.$$

Theorem 2.1. A convergent sequence in a metric space is a Cauchy sequence.

Proof 2.2. Suppose that (x_n) is a sequence which converges to x. Let $\epsilon > 0$, then there is an $N \in \mathbb{N}$ such that $d(x_n,x) < \epsilon/2$, for all $n \geq N$. Let m, $n \geq N$, then

$$d(x_m,x_n) \leq d(x_m,x) + d(x,x_n) < \epsilon/2 + \epsilon/2 = \epsilon,$$

hence (x_n) is a Cauchy sequence.

Note that the converse of this theorem is not true. For example, let $\mathcal{X} = (0,1]$, then the sequence $\left(\frac{1}{n}\right)$ is a Cauchy sequence but converges to 0, which is not in X.

Definition 8. A metric space $(\mathcal{X},d)$ is said to be complete if every Cauchy sequence in $\mathcal{X}$ converges (to a point in $\mathcal{X}$.)

Definition 9. A norm on a linear space $\mathcal{X}$ is a function $\|,\| : \mathcal{X} \longrightarrow \mathbb{R}$ with the following properties

$$\begin{aligned}
&\|x\| \geq 0, \quad \text{for all} \quad x \in \mathcal{X}, \quad \text{(nonnegative)},\\
&\|\lambda x\| = |\lambda|\|x\|, \quad \text{for all} \quad \lambda \in \mathbb{R} \quad \text{or} \quad \mathbb{C}, \quad x \in \mathcal{X}, \quad \text{(homogeneous)},\\
&\|x+y\| \leq \|x\| + \|y\|, \quad \text{for all} \quad x,y \in \mathcal{X}, \quad \text{(triangular inequality)},\\
&\|x\| = 0 \Longrightarrow x = 0, \quad \text{(strictly positive)}.
\end{aligned}$$

$(\mathcal{X},\|.\|)$ is a normed linear space.

Example 8. Let $f \in \mathcal{X} = \mathcal{C}([0,1])$, we define $\|f\|_\infty := sup\{\|f(t)\|, t \in [0,1]\}$. Then $(\mathcal{X},\|.\|_\infty)$ is a normed linear space.

Example 9. Let us show that $(\mathcal{X},\|.\|)_{\mathcal{C}(I)}$ is a complete metric space.

For the proof, let $(\alpha_n)_{n\geq 0}$ a Cauchy sequence in $\mathcal{X}$, then for $t \in I$, $(\alpha_n(t))_{n\geq 0}$ is a Cauchy sequence in $\overline{\mathcal{B}_\epsilon(x_0)} := \{x \in \mathbb{R}^n, \quad \|x - x_0\| \leq \epsilon\}$ and then converges. Let us denoted the limit by $\alpha(t)$. Let $\epsilon > 0$, then there exists $n_0 \in \mathbb{N}$ such that

$$\|\alpha_n - \alpha_m\|_{\mathcal{C}(I)} < \frac{\epsilon}{3}, \quad \text{for} \quad n,m \geq n_0.$$

For all $t \in I$, $n \geq n_0$,

$$\|\alpha(t) - \alpha_n(t)\| \leq \frac{\epsilon}{3}. \tag{2.1}$$

Let $t_0 \in I$, by the triangular inequality, we have

$$\|\alpha(t_0) - \alpha(t)\| \leq \|\alpha(t_0) - \alpha_{n_0}(t_0)\| + \|\alpha_{n_0}(t_0) - \alpha_{n_0}(t)\| + \|\alpha_{n_0}(t) - \alpha(t)\|, \tag{2.2}$$
$$\leq \frac{2\epsilon}{3} + \|\alpha_{n_0}(t_0) - \alpha_{n_0}(t)\|, \quad \text{by} \quad (2.1). \tag{2.3}$$

As α_{n_0} is bounded and using the previous inequality, it follows that α is bounded. α_{n_0} is continuous, then there exists $\delta > 0$ such that

$$\|\alpha(t_0) - \alpha(t)\| < \frac{\epsilon}{3}, \quad \text{for all} \quad t \in \overline{\mathcal{B}_\epsilon(\delta)}. \tag{2.4}$$

By (2.3), we get

$$\|\alpha(t_0) - \alpha(t)\| < \epsilon, \quad \text{for all} \quad t \in \overline{\mathcal{B}_\epsilon(\delta)}. \tag{2.5}$$

This shows that α in continuous at t_0. We have shown that $\alpha \in \mathcal{X}$ and (2.5) implies that $(\alpha_n)_{n\geq 0}$ converges to α.

2.1 Contraction mapping theorem

Definition 10. Let $(\mathcal{X}, d)$ denote a metric space and consider a function $T : \mathcal{X} \longrightarrow \mathcal{X}$. We say that $x \in \mathcal{X}$ is a fixed point of T if $T(x) = x$ and it is globally attracting if

$$\lim_{n\longrightarrow\infty} T^n(x) = x \quad \text{for all} \quad x \in \mathcal{X}.$$

Suppose that there exists $p \in \mathbb{N}^+$ such that $T^p(x) = x$, then we say that x is periodic and the integer p is called the period of x. Notice that x is periodic of period np, for all $n \in \mathbb{N}^+$

Example 10. The three fixed points of the function $f(x) = x^3$ are 0,1 and -1. The function $g(x) = x^2 - 1$ has two fixed point $\frac{1+\sqrt{5}}{2}$ (called golden number) and $\frac{1-\sqrt{5}}{2}$. 0 and -1 lie on the periodic orbit of period 2 because $g(0) = -1$ and $g(-1) = 0$, $g(-1) = 0$ and $g(0) = -1$.

Definition 11. A function $T : \mathcal{X} \longrightarrow \mathcal{X}$ is a contraction if there exists a real number $\lambda \in [0,1)$ such that

$$d(T(x),T(y)) \leq \lambda d(x,y), \quad \text{for all} \quad x,y \in \mathcal{X}.$$

λ is called a contraction constant.

The next theorem is fundamental. It shows that a contraction, viewed as a dynamical system, has a globally attracting fixed point.

Theorem 2.3. Let $(\mathcal{X},d)$ be a complete metric space. Assume that the function $T : \mathcal{X} \longrightarrow \mathcal{X}$ is a contraction with contraction constant λ. Then T has a unique fixed point $\bar{x} \in \mathcal{X}$. Furthermore, for any $x \in \mathcal{X}$,

$$d(T^n(x),\bar{x}) \leq \frac{\lambda^n}{1-\lambda}d(T(x),x).$$

Proof 2.4. Let us prove the uniqueness of the fixed point. Suppose that $T(x_1) = x_1$ and $T(x_2) = x_2$. As T is a contraction and x_1, x_2 fixed points, we have

$$d(T(x_1),T(x_2)) \leq \lambda d(x_1,x_2) \quad \text{and} \quad d(T(x_1),T(x_2)) = \lambda d(x_1,x_2).$$

Therefore $d(x_1,x_2) \leq \lambda d(x_1,x_2)$. If $x_1 \neq x_2$, then $d(x_1,x_2) \neq 0$ and this implies that $\lambda \geq 1$, a contradiction !

For the existence of the fixed point, let us consider the sequence of iterates $\{T^n(x)\}_{n=1}^{\infty}$, with $x \in \mathcal{X}$. By the contraction mapping, it follows that

$$d(T^{n+1}(x),T^n(x)) \leq \lambda d(T^n(x),T^{n-1}(x)) \leq ..... \leq \lambda^n d(T(x),x).$$

By applying the triangle inequality in definition 6 and using this result, it follows that

$$\begin{aligned}
d(T^{n+p}(x),T^n(x)) &\leq d(T^{n+p}(x),T^{n+p-1}(x)) + ... + d(T^{n+1}(x),T^n(x)) \\
&\leq (\lambda^{n+p-1} + ... + \lambda^n)d(T(x),x) \\
&\leq \lambda^n(1 + \lambda + ... + \lambda^{p-1})d(T(x),x) \\
&\leq \frac{\lambda^n}{1-\lambda}d(T(x),x).
\end{aligned}$$

Because $0 \leq \lambda < 1$, the sequence $\{\lambda^n\}_{n=1}^{\infty}$ converges to zero and $\{T^n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence. Since X is complete, this sequence converges to some point $\bar{x} \in \mathcal{X}$. Let us prove that $\bar{x}$ is a fixed point of the map T. Because the two sequences $\{T^{n+1}(x)\}_{n=0}^{\infty}$ and $\{T^n(x)\}_{n=1}^{\infty}$ are the same, we have that $\lim\limits_{n \longrightarrow \infty} T^{n+1}(x) = \bar{x}$. T is continuous by the contraction property and

$$d(T^{n+1}(x),T(\bar{x})) = d(T(T^n(x)),T(\bar{x})) \leq \lambda d(T^n(x),\bar{x}).$$

By the continuity of T, it implies that

$$\lim_{n \longrightarrow \infty} T^{n+1}(x) = \lim_{n \longrightarrow \infty} T(T^n(x)) = T(\bar{x}).$$

To obtain the inequality, pass to the limit as $p \longrightarrow \infty$ in

$$d(T^{n+p}(x),T^n(x)) \leq \frac{\lambda^n}{1-\lambda}d(T(x),x)$$

to obtain

$$d(T^n(x),\bar{x}) \leq \frac{\lambda^n}{1-\lambda} d(T(x),x).$$

Example 11. Let f be a continuous function on the interval $[-1,1]$ and consider the operator

$$Tf(x) = \sin(2\pi x) + \lambda \int_{-1}^{1} \frac{f(y)}{1+(x-y)^2} dy.$$

We are going to find conditions on λ such that T has a unique fixed point and give a best estimation for λ.

1) $Tf \in \mathcal{C}([-1,1])$.

Abstract: $x \longrightarrow \sin(2\pi x)$ is continuous. $x \longrightarrow \dfrac{1}{1+(x-y)^2}$ is continuous for all y, as integration is continuous, $x \longrightarrow Tf(x)$ is continuous.

$$\begin{aligned} \textit{Detailed}: \quad \frac{1}{1+(x-y)^2} - \frac{1}{1+(x+h-y)^2} &= \frac{(x+h-y)^2-(x-y)^2}{(1+(x-y)^2)(1+(x+h-y)^2)}, \\ &\leq (x+h-y)^2-(x-y)^2 \longrightarrow 0, \quad as \quad h \longrightarrow 0. \end{aligned}$$

This implies that

$$|Tf(x+h)-Tf(x)| \leq \|f\|_\infty \lambda \int_{-1}^{1} ((x+h-y)^2-(x-y)^2) dy \longrightarrow 0, \quad \text{as} \quad h \longrightarrow 0.$$

2) choose λ so that T is a contraction.

a) Rough estimate: Since $1+(x-y)^2 > 1$, we have

$$\begin{aligned} \|Tf-Tg\|_\infty &= \lambda \sup_{x\in[-1,1]} \Big| \int_{-1}^{1} \frac{f(y)-g(y)}{1+(x-y)^2} dy \Big|, \\ &\leq \lambda \sup_{x\in[-1,1]} \int_{-1}^{1} \big|f(y)-g(y)\big| dy, \\ &\leq \lambda \|f-g\|_\infty \int_{-1}^{1} dy, \\ &= 2\lambda \|f-g\|_\infty. \end{aligned}$$

Pick $\lambda < \frac{1}{2}$, then T is a contraction with respect to the norm on $\mathcal{C}([-1,1])$.

b) better estimate:

$$\begin{aligned} \|Tf-Tg\|_\infty &\leq \lambda \|f-g\|_\infty \sup_{x\in[-1,1]} \int_{-1}^{1} \frac{1}{1+(x-y)^2} dy, \\ &\leq \lambda \|f-g\|_\infty \sup_{x\in[-1,1]} (\arctan(x+1) - \arctan(x-1)), \\ &= \lambda \|f-g\|_\infty \cdot 2\arctan(1). \end{aligned}$$

Pick $\lambda < \dfrac{1}{2\arctan(1)} \approx 0.6366$ to get a contraction. Larger than in a).

For all $\lambda < \dfrac{1}{2\arctan(1)}$, the map T has a unique fixed point, *i.e.* there exists $f \in \mathcal{C}[-1,1]$ such that

$$Tf(x) = \sin(2\pi x) + \lambda \int_{-1}^{1} \frac{f(y)}{1+(x-y)^2} dy = f(x).$$

2.2 Existence and uniqueness of ODEs

In this section we will prove the basic existence and uniqueness theorem for ordinary differential equations.

Definition 12. We consider $f : U \longrightarrow \mathbb{R}^n$, be a continuous function defined on an open set $U \subset \mathbb{R}^n$. A differentiable function $\phi : I \longrightarrow U$, with I an open interval of $\mathbb{R}$ is the solution of the differential equation

$$\dot{x} := \frac{dx}{dt} = f(x)$$

if

$$\frac{\phi}{dt}(t) = f(\phi(t)), \quad \text{for all} \quad t \in I.$$

Next, we focus on solutions to the initial value problem (IVP), *i.e.*

$$\dot{x} = f(x), \quad x(t_0) = x_0,$$

given $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n$. We begin by starting a simple lemma, whose proof follows directly from differentiation and integration.

Lemma 2.5. Consider the initial value problem $\dot{x} = f(x)$ satisfying $x(t_0) = x_0$, with $t_0 \in I$, and suppose that $f(x)$ is continuous. Then ϕ is a solution of the IVP if and only if

$$\phi(t) = x_0 + \int_{t_0}^{t} \frac{d\phi}{ds}(s)ds = x_0 + \int_{t_0}^{t} f(\phi(s))ds.$$

Proof 2.6. If $x(t)$ is a continuous function on I that satisfies the previous integral equation, then

$$x(t_0) = x_0 \text{ and } \dot{x}(t) = \frac{d}{dt}\int_{t_0}^{t} f(x(s))ds = f(x(t)) \quad \forall t \in I.$$

By the fundamental theorem of calculus since $f(x(t)) \in \mathcal{C}(I)$, $x(t)$ is differentiable and satisfies the initial value problem, for all $t \in I$.
Conversely, if $x(t)$ is a solution of the initial value problem for all $t \in I$, then $x(t)$ is differentiable, hence continuous on I and $x(t) \in U$ for all $t \in I$, therefore

$$\dot{x}(t) = f(x(t)) \Longrightarrow x(t) = \int_0^t f(x(s))ds + c, \quad \text{for all} \quad t \in I.$$

Clearly $c = x(t_0) = x_0$. Thus $x(t)$ satisfies the integral equation for all $t \in I$.

2.2.1 Existence and uniqueness to solutions of ODEs

We consider two metric spaces (X, d_1) and (Y, d_2) and a function $f : X \longrightarrow Y$. f is *Lipschitz* if there exists a real constant $K \geq 0$ such that

$$d_2(f(x), f(y)) \leq C d_1(x, y), \quad \text{for all} \quad x, y \in X.$$

The smallest constant C satisfying this inequality is denoted by $Lip(f) := C$ and is called a Lipschitz constant.

Remark 7. f is locally *Lipschitz* if every point in $U \subset X$ has a neighborhood such that f restricted to that neighborhood is *Lipschitz*.

Theorem 2.7. Let $U \subset \mathbb{R}^n$ and open set and $f : U \longrightarrow \mathbb{R}^n$ be a *Lipschitz* function. Then there exists $l > 0$ and a solution $\phi : (t_0 - l, t_0 + l) \longrightarrow \mathbb{R}^n$ to the initial value problem $\dot{x} = f(x), \quad x(t_0) = x_0$ for any $t_0 \in \mathbb{R}$ and $x_0 \in U$. Furthermore, if ψ is another solution to the initial value, then $\phi(t) = \psi(t)$ on their common domain of definition.

Proof 2.8. Let us consider $\|.\|$ the norm on $\mathbb{R}^n$ and choose ϵ such that $\overline{\mathcal{B}_\epsilon(x_0)} \subset U$. By assumption f is *Lipschitz* and hence there exist positive constants C and M such that

$$\|f(x) - f(y)\| \leq C\|x - y\| \quad \text{and} \quad \|f(x)\| \leq M, \quad \text{for all} \quad x, y \in \overline{\mathcal{B}_\epsilon(x_0)}.$$

Using Lemma 2.5, it is sufficient to prove the existence and uniqueness of ϕ such that

$$\phi(t) = x_0 + \int_{t_0}^t \frac{d\phi}{ds}(s)ds = x_0 + \int_{t_0}^t f(\phi(s))ds. \tag{2.6}$$

The strategy for the proof is to define a function space $\mathcal{X}$ within which we expect to find a solution and define a contraction $T : \mathcal{X} \longrightarrow \mathcal{X}$. This implies that there exists a unique fixed point and we expect that this fixed point will be a solution of (2.6). Here the domain of the elements of the function space should be an interval $I \subset \mathbb{R}$ since solutions are functions of time.

Choose l such that

$$l < min\{\frac{\epsilon}{M}, \frac{1}{C}\} \quad set \quad I := (t_0 - l, t_0 + l) \quad \text{and define}$$

$$\mathcal{X} := \{\alpha : I \longrightarrow \overline{\mathcal{B}_\epsilon(x_0)}, \alpha \in \mathcal{C}^0, Lip(\alpha) < C\}.$$

We endow the space $\mathcal{X}$ with the $\mathcal{C} - norm$, $i.e$

$$\|\alpha\|_{\mathcal{C}^0(I)} := sup\{\|\alpha(t)\|, t \in I\}.$$

Then $(\mathcal{X}, \|.\|)_{\mathcal{C}(I)}$ is a complete metric space, see the next example. For $\alpha \in \mathcal{X}$, define

$$T(\alpha(t)) := x_0 + \int_{t_0}^{t} f(\alpha(s))ds.$$

Remark that $T(\alpha) = \alpha$ if and only if α satisfies (2.6). To apply the contraction mapping theorem, we need to prove that T is well defined and it is a contraction.

At the initial condition time t_0, we have

$$T(\alpha(t_0)) = x_0 + \int_{t_0}^{t_0} f(\alpha(s))ds = x_0.$$

Furthermore, given $t_1, t_2 \in I$,

$$\begin{aligned}
\Big\|T(\alpha(t_2)) - T(\alpha(t_1))\Big\| &= \Big\| \int_{0}^{t_2} f(\alpha(s))ds - \int_{0}^{t_1} f(\alpha(s))ds \Big\|, \\
&= \Big\| \int_{t_1}^{t_2} f(\alpha(s))ds \Big\|, \\
&\leq \Big| \int_{t_1}^{t_2} C ds \Big|, \\
&= C|t_2 - t_1|.
\end{aligned}$$

This implies that $Lip(T(\alpha)) \leq C$. Setting $t_1 = t$ and $t_2 = t_0$, we get

$$\|T(\alpha(t)) - x_0\| \leq C|t - t_0| \leq Cl < \epsilon, \quad \text{for all} \quad t \in I$$

and hence, $T(\alpha) : I \longrightarrow \overline{\mathcal{B}_\epsilon(x_0)}$ is well defined. For the contraction, let $\alpha_1,\alpha_2 \in \mathcal{X}$, then

$$\begin{aligned}
\|T(\alpha_1) - T(\alpha_2)\| &= \sup_{t\in I} \|T(\alpha_1)(t) - T(\alpha_2)(t)\|, \\
&= \sup_{t\in I} \left\| \int_{t_0}^{t} f(\alpha_1)(s)ds - \int_{t_0}^{t} f(\alpha_2)(s)ds \right\|, \\
&\le \sup_{t\in I} \int_{t_0}^{t} C\|\alpha_1(s) - \alpha_2(s)\|ds, \\
&\le C\|\alpha_1 - \alpha_2\|_{C^0} \sup_{t\in I} \int_{t_0}^{t} ds, \\
&\le Cl\|\alpha_1 - \alpha_2\|.
\end{aligned}$$

Since $Cl < 1$, then $T : \mathcal{X} \longrightarrow \mathcal{X}$ is a contraction and hence has a unique fixed point $\bar{\alpha}$ of $\dot{x} = f(x), \quad x(0) = x_0$ in $\mathcal{X}$ over the interval I.

Remark 8. Under the hypothesis of the Fundamental Existence-Uniqueness Theorem, if $x(t)$ is the solution of the initial value problem on an interval I, then the second derivative $\ddot{x}$ is continuous on I. For the proof, remark that

$$\ddot{x} = \frac{d}{dt}[f(x(t))] = Df[x(t)]\dot{x}(t) = Df[x(t)]f(x(t)) \in C(I).$$

By the chain rule, since $x(t) \in U$, $Df(x(t))$ and $f(x(t))$ are continuous for all $t \in I$.

Remark 9. Note that for this proof, we are limited to a small time interval of length $2l$, where $l < min\{\frac{\epsilon}{M}, \frac{1}{C}\}$. We have only proven uniqueness of solutions over the family of functions $Lip(\alpha) \le C$, as opposed to all differentiable functions. In the last line of the proof, we demonstrate that solutions to an IVP are *Lipschitz* continuous as a function of the initial value. First let us introduce the following fundamental result introduced by *Gronwall*[1]. Note that the applications to the theory of differential equations are later date. In the respect, we mention the name of the mathematician *Richard Bellman*.[2]

Lemma 2.9. (Gronwall´s inequality)
Let us consider a,α_1 and α_2 positive constants. Assume that $\phi(t)$ and $\psi(t)$ are positive continuous functions for $t_0 \le t \le t_0 + a$. If

$$\phi(t) \le \alpha_2 + \alpha_1 \int_{t_0}^{t} \psi(s)\phi(s)ds, \tag{2.7}$$

then

$$\phi(t) \le \alpha_2 e^{\alpha_1 \int_{t_0}^{t} \psi(s)ds}.$$

[1]Thomas Hakon Gronwall, 1877-1932, was a Swedish mathematician.
[2]Richard Ernest Bellman, 1920-1984, was an American applied mathematician.

Proof 2.10. From 2.7, we derive

$$\frac{\phi(t)}{\alpha_2 + \alpha_1 \int_{t_0}^{t} \psi(s)\phi(s)ds} < 1.$$

Multiplication with $\alpha_1\psi(t)$ and integration, yields

$$\int_{t_0}^{t} \frac{\alpha_1\psi(s)\phi(t)}{\alpha_2 + \alpha_1 \int_{t_0}^{t} \psi(s)\phi(s)ds} < \alpha_1 \int_{t_0}^{t} \psi(s)ds.$$

This implies that

$$ln(\alpha_1 \int_{t_0}^{t} \psi(s)\phi(s)ds + \alpha_2) - ln(\alpha_2) \leq \alpha_1 \int_{t_0}^{t} \psi(s)ds,$$

which produces

$$\alpha_1 \int_{t_0}^{t} \psi(s)\phi(s)ds + \alpha_2 \leq \alpha_2 e^{\alpha_1 \int_{t_0}^{t} \psi(s)ds}.$$

Applying the estimation (2.7) again, but now to the left hand side, yields the required inequality.

Remark 10. Remark that if $\alpha_2 = 0$, the estimation implies that $\phi(t) = 0, \quad t_0 \leq t \leq t_0 + a$.

We now use Gronwall's inequality to study the dependence of a solution on its initial condition.

Theorem 2.11. Consider the equation $\dot{x}(t) = f(x(t))$, $t \geq 0$ and f *Lipschitz* continuous with *Lipschitz* constant C. Consider the two initial value problems

$$\begin{gathered}\dot{x} = f(x(t)),\ x(0) = x_0 \in \mathbb{R}^n, \text{ solution } x_0(t) \text{ on interval } I,\\ \dot{x} = f(x(t)),\ x(0) = x_0 + \delta, \text{ solution } x_\epsilon(t) \text{ on interval } I, \text{ where } \delta \in \mathbb{R}^n.\end{gathered}$$

If $\|\delta\| \leq \epsilon$, (ϵ real, positive), we have

$$\|x_0(t) - x_\epsilon(t)\| \leq \epsilon e^{Ct} \text{ on the interval } I.$$

Proof 2.12. The solutions of the two initial value problems are given by

$$x_0(t) = x_0 + \int_0^t f(x_0(s))ds, \quad x_\epsilon(t) = x_0 + \delta + \int_0^t f(x_\epsilon(s))ds.$$

Substitution and applying the triangular inequality, produces

$$\|x_0(t) - x_\epsilon(t)\| \leq \|\delta\| + \int_0^t \|f(x_0(s)) - f(x_\epsilon(s))\| ds,$$

and using the *Lipschitz* condition,

$$\|x_0(t) - x_\epsilon(t)\| \leq \epsilon + C \int_0^t \|x_0(s) - x_\epsilon(s)\| ds.$$

Now by Gronwall's inequality, with $\alpha_1 = C$, $\psi(t) = 1$, $\phi(t) = \|x_0(t) - x_\epsilon(t)\|$ and $\alpha_2 = \epsilon$ give us the desired inequality.

Example 12. Let us consider the two differential equations

$$a) \quad \dot{x} = x^2, \quad x(0) = x_0 \quad \text{and} \quad b) \dot{x} = 3x^{2/3}, \quad x(0) = 0.$$

For $a)$ suppose that $x_0 > 0$. By the fundamental existence and uniqueness theorem, this IVP has a unique solution given by

$$\phi(t) = \frac{x_0}{1 - x_0 t}, \text{ defined for all } -\infty < t < 1/x_0.$$

For $b)$, we can see that f is not *Lipschitz* and in this case, it is possible to have nonunique solutions. Observe the two solutions for $b)$

$$\phi_1(t) \equiv 0, \quad \phi_2(t) = \begin{cases} t^3, & t \geq 0, \\ 0, & t \leq 0. \end{cases}$$

This implies that the both $\phi_1(t)$ and $\phi_2(t)$ equal zero but $\phi_1(t)$ and $\phi_2(t)$ are distinct on any interval about t_0.

2.2.2 Linear Systems

Definition 13. Let A be and $n \times n$ matrix, we defined the exponential matrix e^{At} by

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}, \quad \text{for all} \quad t \in \mathbb{R}.$$

Remark 11. Note that the series $\sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$ is absolutely and uniformly convergent for all $|t| \leq t_0$. This is proved by using the Weierstrass M-Test.

Let A be an $n \times n$ matrix and consider the initial value problem

$$\dot{x} = Ax, \quad x(0) = x_0 \in \mathbb{R}^n \tag{2.8}$$

We are going to establish that (2.8) has a unique solution defined by

$$x(t) = e^{At}x_0, \text{ for all } t \in \mathbb{R}.$$

Remark 12. This result is similar for the one dimension case $\dot{x} = ax, \quad x(0) = x_0 \in \mathbb{R}$, where the unique solution is given by $x(t) = x_0 e^{at}$.

To prove this theorem, we start by computing the derivative of the exponential function e^{At} by using the result that two convergent limit processes can be interchanged if one of them converges uniformly.

Lemma 2.13. Let A be an $n \times n$ matrix, then

$$\frac{d}{dt}e^{At} = Ae^{At}.$$

Proof 2.14. By definition, we have

$$\begin{aligned}
\frac{d}{dt}e^{At} &= \lim_{h \longrightarrow 0} \frac{e^{A(t+h)} - e^{A(t)}}{h}, \\
&= \lim_{h \longrightarrow 0} e^{At}\frac{e^{Ah} - I}{h}, \\
&= e^{At} \lim_{h \longrightarrow 0} \lim_{k \longrightarrow \infty} (A + \frac{A^2 h}{2!} + ... + \frac{A^k h^{k-1}}{k!}), \\
&= Ae^{At}.
\end{aligned}$$

Remark 13. The last equality follows by the Remark 11, since the series defining e^{Ah} converges uniformly for $|h| < 1$.

Theorem 2.15. (Fundamental Theorem for Linear Systems)

We consider A to be an $n \times n$ matrix, then the initial value problem (2.8) has a unique solution given by $x(t) = e^{At}x_0$, for all $t \in \mathbb{R}$.

Proof 2.16. By Lemma (2.13), if $x(t) = e^{At}x_0$ then we have

$$\dot{x}(t) = \frac{d}{dt}e^{At}x_0 = Ae^{At}x_0 = Ax(t), \quad \text{for all} \quad t \in \mathbb{R},$$

and $x(0) = Ix_0 = x_0$, this implies that $x(t) = e^{At}x_0$ is a solution. For uniqueness, let $x(t)$ be any solution of the initial value problem and set $z(t) = e^{-At}x(t)$, then we have

$$\begin{aligned}
\dot{z}(t) &= -Ae^{-At}x(t) + e^{-At}\dot{x}(t), \quad \text{for all} \quad t \in \mathbb{R}, \\
&= -Ae^{-At}x(t) + e^{-At}Ax(t), \quad \text{for all} \quad t \in \mathbb{R}, \\
&= -Ae^{-At}x(t) + Ae^{-At}x(t), \quad \text{for all} \quad t \in \mathbb{R}, \\
&= 0, \quad \text{for all} \quad t \in \mathbb{R}, \quad \text{since} \quad e^{-At} \quad \textit{and} \quad A \quad \textit{commute}.
\end{aligned}$$

Thus, $z(t)$ is a constant. For $t = 0$, we get $z(t) = x_0$ and therefore any solution of the initial value problem is given by $x(t) = e^{At}z(t) = e^{At}x_0$.

Example 13. Let A be an $n \times n$ matrix. Suppose that $T : E \longrightarrow E$, $x \longrightarrow Ax$, is a linear transformation of $\mathbb{R}^n$ that leaves a space $E \subset \mathbb{R}^n$ invariant. If $x(t)$ is the solution to the initial value problem,

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0 \in E,$$

then $x(t) \in E$ for all $t \in \mathbb{R}$.

For the proof, since $T(x) \in E$ for all $x \in E$ and $T(x) = Ax$ then $T(x_0) \in E$ for $x_0 \in E$. Because E is a linear subspace of $\mathbb{R}^n$, it follows that $tAx_0 \in E$. By induction, we have that $(\frac{t^k}{k!})A^k x_0 \in E$, for all $k \in \mathbb{N}$. Therefore

$$\sum_{k=0}^{N} \frac{A^k t^k x_0}{k!} \in E,$$

since E is a linear subspace of $\mathbb{R}^n$. Since a closed subspace of a complete metric space is complete, it follows that E is a complete normed linear space, $i.e$ every Cauchy sequence in E converges to a vector in E. Thus, for all $t \in \mathbb{R}$,

$$\lim_{N \longrightarrow \infty} \sum_{k=0}^{N} \frac{A^k t^k x_0}{k!} = e^{At}x_0 \in E.$$

Therefore, by the fundamental theorem of linear system,

$$x(t) = e^{At}x_0 \in E, \quad \text{for all} \quad t \in \mathbb{R}.$$

Remark 14. In practice, if A is an $n \times n$ matrix, we will use the algebraic technique of diagonalization to reduce the linear system $\dot{x} = Ax, \quad x(0) = x_0 \in \mathbb{R}^n$. Let us introduce this result from linear algebra in the case A has real and distinct eigenvalues.

Theorem 2.17. Let $\lambda_1,\lambda_2,....,\lambda_n$ be real and distinct eigenvalues of an $n \times n$ matrix A. If $\{v_1,v_2,....,v_n\}$ is a set of corresponding eigenvectors, then it forms a basis for $\mathbb{R}^n$. The matrix $P = \{v_1,v_2,....,v_n\}$ is invertible and

$$P^{-1}AP = diag\{\lambda_1,\lambda_2,.....,\lambda_n\}.$$

Remark 15. Let us define the linear transformation of coordinates $z = P^{-1}x$, where P is an invertible linear transformation defined in Theorem (2.17). We would like to reduce the system (2.8) to an uncoupled linear system. We get

$$x = Pz \quad \text{and} \quad \dot{z} = P^{-1}\dot{x} = P^{-1}Ax = P^{-1}APz.$$

By theorem (2.17), it follows that $\dot{z} = diag\{\lambda_1,\lambda_2,.....,\lambda_n\}z$, which is an uncoupled linear system. The solution is given by

$$z(t) = diag\{e^{\lambda_1 t},e^{\lambda_2 t},....,e^{\lambda_n t}\}z(0).$$

Since $x(t) = Pz(t)$ and $z(0) = P^{-1}x(0)$, it follows that the solution to the initial value problem (2.8) is given by $x(t) = PD(t)P^{-1}x(0)$, where $D(t)$ is the diagonal matrix

$$D(t) = diag\{e^{\lambda_1 t},e^{\lambda_2 t},....,e^{\lambda_n t}\}.$$

Corollary 2.18. By Theorem (2.17), the solution of the system (2.8) is given by

$$x(t) = PD(t)P^{-1}x(0), \quad D(t) = diag\{e^{\lambda_1 t},e^{\lambda_2 t},....,e^{\lambda_n t}\}.$$

Example 14. Solve the initial value problem

$$\dot{x} = Ax,\ x(0) = x_0 \qquad \text{where } A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = -1$. The associate eigenvectors are $v_1 = \{2,-2,1\}$, $v_2 = \{0,1,0\}$ and $v_3 = \{0,0,1\}$. So $P = \begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ and the solution is given by

$$x(t) = PD(t)P^{-1}x(0) = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} x(0),$$

$$= \frac{1}{2}\begin{pmatrix} 2e^t & 0 & 0 \\ 2e^{2t} - 2e^{-t} & 2e^{2t} & 0 \\ e^t - e^{-t} & 0 & 2e^{-t} \end{pmatrix} x(0).$$

2.3 Nonhomogeneous Linear Systems

In this section we consider the nonhomogeneous linear systems

$$\dot{x} = Ax(t) + b(t),$$

where A is an $n \times n$ matrix and $b(t)$ is a continuous vector valued function.

Definition 14. A fundamental matrix solution of $\dot{x} = Ax$ is an $n \times n$ matrix function ϕ (non singular) satisfying

$$\dot{\phi}(t) = A\phi(t), \quad \text{for all} \quad t \in \mathbb{R}. \tag{2.9}$$

Remark 16. By the Lemma 2.13, the matrix defined by $\phi(t) = e^{At}$ is a fundamental matrix solution which satisfying $\phi(0) = Id_n$. Furthermore, any fundamental matrix solution $\phi(t)$ of (2.9) is given by $\phi(t) = e^{At}K$, with K a nonsingular matrix. We deduce that, once we have found a fundamental matrix, we can easily solve the nonhomogeneous systems.

Theorem 2.19. Suppose that $\phi(t)$ is the fundamental matrix solution of $\dot{x} = Ax$, the solution of the nonhomogeneous system $\dot{x} = Ax(t) + b(t)$ with initial condition $x(0) = x_0$ is unique and is given by

$$x(t) = \phi(t)\phi^{-1}(0)x_0 + \int_0^t \phi(t)\phi^{-1}(\tau)b(\tau)d\tau \tag{2.10}$$

Proof 2.20. By differentiating the function (2.10) above, we have

$$\dot{x}(t) = \dot{\phi}(t)\phi^{-1}(0)x_0 + \phi(t)\phi^{-1}(t)b(t) + \int_0^t \dot{\phi}(t)\phi^{-1}(\tau)b(\tau)d\tau.$$

As $\phi(t)$ is the fundamental matrix solution of $\dot{x} = Ax$, it implies that

$$\begin{aligned}\dot{x}(t) &= A\left[\phi(t)\phi^{-1}(0)x_0 + \int_0^t \phi(t)\phi^{-1}(\tau)b(\tau)d\tau\right] + b(t)\\ &= Ax(t) + b(t).\end{aligned}$$

Remark 17. If $\phi(t) = e^{At}$, the solution of the nonhomogeneous linear system $\dot{x} = Ax(t)+b(t)$ is given by

$$x(t) = e^{At}x_0 + e^{At}\int_0^t e^{-A\tau}b(\tau)d\tau. \tag{2.11}$$

Example 15. Use the Theorem (2.19) to solve the nonhomogeneous linear system

$$\dot{x} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} t \\ 1 \end{pmatrix}, \quad x(0) = \begin{pmatrix} t \\ 1 \end{pmatrix}.$$

The eigenvalues are $\lambda_1 = 1$, $\lambda_2 = -1$, with associated eigenvectors $v_1 = (1{,}0)^T$ and $v_2 = (1, -2)^T$. The fundamental matrix solution $\Phi(t)$, with $\Phi(0) = I$ is given by

$$\phi(t) = e^{At} = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 & 1/2 \\ 0 & -1/2 \end{pmatrix} = \begin{pmatrix} e^t & (e^t - e^{-t})/2 \\ 0 & e^{-t} \end{pmatrix}.$$

Note that

$$\phi^{-}(t) = \begin{pmatrix} e^{-t} & -(e^t - e^{-t})/2 \\ 0 & e^t \end{pmatrix} = \phi(-t),$$

and the solution of the initial value problem is given by

$$\begin{aligned}x(t) &= \begin{pmatrix} e^t & (e^t - e^{-t})/2 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} e^t & (e^t - e^{-t})/2 \\ 0 & e^{-t} \end{pmatrix}\int_0^t \begin{pmatrix} -e^{-\tau} & (e^{-\tau} - e^{\tau})/2 \\ 0 & e^{\tau} \end{pmatrix}\begin{pmatrix} \tau \\ 1 \end{pmatrix} d\tau,\\ &= \begin{pmatrix} -t - 2 + \frac{5}{2}e^t + \frac{1}{2}e^{-t} \\ 1 - e^{-t} \end{pmatrix}.\end{aligned}$$

Remark 18. There are many interesting properties that can be study just by using the nature of the eigenvalues of the matrix A in (2.8). The stability condition is one such property.

2.4 Stability for Linear Systems

Definition 15. Recall the system (2.8), $\dot{x} = Ax, \quad x(0) = x_0 \in \mathbb{R}^n$. An equilibrium solution of (2.8) is the vector x_e such that $Ax_e = 0$. It is asymptotically stable if there exists $\delta > 0$ such that

$$\|x(t) - x_e\| \longrightarrow 0, \quad \text{as} \quad t \longrightarrow \infty, \quad \text{whenever} \quad \|x_0 - x_e\| \leq \delta. \tag{2.12}$$

If the equilibrium solution x_e est asymptotically stable, the system is called asymptotically stable, otherwise it is called unstable.

Remark 19. If the system is perturbed a little bit from the position of equilibrium, then by the asymptotic stability, this system will eventually return to that position after making small oscillations.

Theorem 2.21 (Stability theorem for the homogeneous Linear systems)**.** If all the eigenvalues of the matrix A have negative real part, the system is asymptotically stable. It is unstable if at least one eigenvalue has a positive real part.

Proof 2.22. Suppose that the matrix A is diagonalizable, there exists P such that $P^{-1}AP = D = diag(\lambda_1,...,\lambda_n)$. This implies that

$$e^{At} = Pe^{Dt}P^{-1} = Pdiag(\lambda_1,...,\lambda_n)P^{-1}.$$

If $\lambda_j = \alpha_j + i\beta_j, \quad j = 1,2,....,n$, then $e^{\lambda_j t} = e^{\alpha_j t}e^{i\beta_j t}$. And $e^{\lambda_j t}$ converges to zero, when $t \longrightarrow \infty$ if and only if $\lambda_j < 0$.

Theorem 2.23. Let A be an $n \times n$ matrix and consider the nonhomogeneous Linear system

$$\dot{x} = Ax(t) + b, \quad \text{where} \quad b \quad \text{is a constant.} \tag{2.13}$$

An equilibrium solution $\bar{x}$ of (2.13) is asymptotically stable if and only if all the eigenvalues of A have negative real parts. If there exists at least one eigenvalue of A with positive real part, this stable is unstable.

Proof 2.24. Consider (2.13) with b a constant vector. The stability of this system is given by the eigenvalues of the matrix A. This can be shown as follows using the previous theorem. Let $\bar{x}$ be an equilibrium solution of (2.13) and define $y(t) = x(t) - \bar{x}(t)$. Then

$$\dot{y}(t) = \dot{x}(t) - (\dot{\bar{x}})(t) = Ax(t) + b - A\bar{x}(t) - b = A(x(t) - \bar{x}(t)) = Ay(t).$$

Thus $x(t) \longrightarrow \bar{x}(t)$ if and only if $y(t) = 0$.

Example 16 (Richardson model)**.** Consider the following system of differential equations

$$\begin{cases} \dot{x}_1(t) = k_1 x_2 - \alpha_1 x_1 + g_1, \\ \dot{x}_2(t) = k_2 x_2 - \alpha_2 x_2 + g_2, \end{cases}$$ where g_1,g_2,α_1,α_2 are positive constants.

In matrix notation, this system can be written as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_1 \end{pmatrix} = \begin{pmatrix} -\alpha_1 & k_1 \\ k_2 & -\alpha_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_1 \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}.$$

The eigenvalues of the matrix A are

$$\lambda = \frac{-(\alpha_1 + \alpha_2) \pm \sqrt{(\alpha_1 - \alpha_2)^2 + 4k_1 k_2}}{2}.$$

Thus $\alpha_1\alpha_2 - k_1k_2 > 0$, both the eigenvalues of A are negative real parts and then the equilibrium solution $\bar{x}(t)$ is asymptotically stable. If $\alpha_1\alpha_2 - k_1k_2 < 0$, one of the eigenvalue has positive real part and the system is unstable.

Chapter 3

Rigorous computations for finite dimensional problems

The goal if this chapter is to provide a constructive approach for proving the existence of zeros of a function defined on a finite dimensional Banach space $\mathcal{X}$ (in our case $\mathbb{R}^n$ or $\mathbb{C}^n$). We are going to present an effective computational approach to the following two problems. Given a differentiable function $f : \mathcal{X} \longrightarrow \mathcal{Y}$ between Banach spaces, does there exist $x \in \mathcal{X}$ such that $f(x) = 0$ and, if so, provide an estimate of x.

3.1 A Newton-like operator

Let us start with a trivial proposition that sets the stage for our strategy for finding zeros of a function.

Proposition 3.1. We consider $X = \mathbb{R}^n$ or $\mathbb{C}^n$, and U ,$V \subset X$ be open sets. Let $f : U \longrightarrow V$. Suppose that $A : X \longrightarrow X$ is an invertible linear map and let defined by $T(x) := x - Af(x)$.

$$\text{If} \quad T(\bar{x}) = \bar{x} \quad \text{then} \quad f(\bar{x}) = 0.$$

Remark 20. The previous proposition is very important. It allows us to replace the problem or directly finding a root of f with that of proving the existence of a fixed point of T. If the operator T is a contraction, it provides existence and local uniqueness of a fixed point of T. Furthermore, it gives bounds on the location of the fixed point x as a function of the an initial guess (recall theorem 2.3). Therefore, the problems we are trying to prove are reduced to finding an invertible linear map A that makes T a contraction.

Example 17. In the one dimensional case, let $f \in \mathcal{C}(\mathbb{R},\mathbb{R})$, then the map $T : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $T(x) := x - \dfrac{f(x)}{f'(x)}$ is used iteratively in Newton´s method to find an approximate value to the root of f. More explicitly, given an initial value $x_0 \in \mathbb{R}$, then inductively we define $x_{k+1} = T(x_k)$. The hope is that $\lim\limits_{k \longrightarrow \infty} x_k$ exists and is the root. Assume that

$$\lim_{k \longrightarrow \infty} x_k = \bar{x}.$$

If T exists and is continuous at $\bar{x}$ then $f^{'}(\bar{x}) \neq 0$. Observe that $T(\bar{x}) = \bar{x}$ implies $f(\bar{x}) = 0$. With this in mind, and to see how this relates to the contraction mapping theorem assume $f(\bar{x}) = 0$ and $f^{'}(\bar{x}) \neq 0$. Let $|\epsilon|$ small, then

$$\begin{aligned}
\Big|T(\bar{x}+\epsilon) - T(\bar{x})\Big| &= \Big|\bar{x} + \epsilon - \frac{f(\bar{x}+\epsilon)}{f^{'}(\bar{x}+\epsilon)} - (\bar{x} - \frac{f(\bar{x})}{f^{'}(\bar{x})}\Big|, \\
\Big|T(\bar{x}+\epsilon) - T(\bar{x})\Big| &= \Big|\epsilon - \frac{f(\bar{x}+\epsilon)}{f^{'}(\bar{x}+\epsilon)}\Big|, \\
\Big|T(\bar{x}+\epsilon) - T(\bar{x})\Big| &\approx \Big|\epsilon - \frac{f(\bar{x}) + \epsilon f^{'}(\bar{x})}{f^{'}(\bar{x}+\epsilon)}\Big|, \\
\Big|T(\bar{x}+\epsilon) - T(\bar{x})\Big| &\approx |\epsilon|\Big|1 - \frac{f^{'}(\bar{x})}{f^{'}(\bar{x}+\epsilon)}\Big|.
\end{aligned}$$

Since $\epsilon = (\bar{x}+\epsilon) - \bar{x}$, then set $Lip(T) \approx \Big|1 - \frac{f^{'}(\bar{x})}{f^{'}(\bar{x}+\epsilon)}\Big|$, which is a contraction constant for Newton's near $\bar{x}$. It is an extremely strong contraction in a sufficiently small neighborhood of a nondegenerate zero of f. In the next section, this example is meant to motivate the idea of setting $A = A(x) := Df^{-1}(\bar{x})$.

Definition 16. Let $U \subset \mathbb{R}^n$ be an open set and $f : U \longrightarrow \mathbb{R}^n$ be a $\mathcal{C}^1$ map. Suppose that $Df(x)$ is invertible. The Newton operator is defined by

$$T(x) := x - Af(x), \tag{3.1}$$

where A is an approximation of $Df^{-1}(x)$. We write $A \approx Df^{-1}(x)$.

Remark 21. Note that A is not necessary equal $Df^{-1}(x)$. The approach of our method is as follows. Let $\bar{x}$ an initial guess for a zero of f, *i.e.* $f(\bar{x}) \approx 0$. In practice, we obtain the value of $\bar{x}$ using Newton's method. Since $\bar{x}$ is an approximate value of the exact solution, choose $A \approx Df^{-1}(\bar{x})$ and define the map $T(x) := x - Af(x)$. The question is now to prove that T is a contraction. In this case, we will be able to find the root of f.

Remark 22. If $f(\bar{x}) = 0$ and $Df^{-1}(\bar{x})$ is invertible, then in the small neighborhood of $\bar{x}$, the operator T is a contraction with small contraction constant. This is given by the same argument as the example (17). But we know that for an $n \times n$ matrix, the cost of computing the inverse is of order n^3. Thus for high dimensional problems we lost the regularity of the solution, because repeatedly computing the inverse is expensive.

3.2 The radii polynomial approach in finite dimension

The method of radii polynomials has been employed in mathematically rigorous computer assisted study of a wide variety of problems in differential equations and dynamical systems. It is an efficient tool for bounding the smallest and largest neighborhoods on wich a Newton-like operator associated with a nonlinear equation is a contraction mapping. The method has been introduced to study solutions of ordinary, partial, and delay differential equations, such has periodic orbits, equilibria and solutions of initial value problems (IVPs). In this section, we adapt the method of radii polynomials approach in finite dimensional problems.

In the following, we consider the sup norm on $\mathbb{R}$, *i.e.* given $x = (x_1,....,x_n) \in \mathbb{R}^n$, we define

$$\|x\|_\infty := max_{k=1,...,\infty}\{|x_k|\}.$$

The closed ball of radius r centred at x is denoted by

$$\overline{\mathcal{B}_r(x)} := \{z \in \mathbb{R}, \quad \|x - z\|_\infty \leq r\}.$$

Theorem 3.2. We consider $U \subset \mathbb{R}^n$ and let $T = (T_1,...,T_n) \in \mathcal{C}^1(U,\mathbb{R}^n)$, where $T_k : \mathbb{R}^n \longrightarrow \mathbb{R}$. Consider $\bar{x} \in U$ and suppose that $Y = (Y_1,...,Y_n) \in \mathbb{R}^n$, $Z(r) = (Z_1(r),...,Z_n(r)) \in \mathbb{R}^n$ provide the following bounds:

$$|T_k(\bar{x}) - \bar{x}_k| \leq Y_k \quad \text{and} \quad \sup_{b,c\in\overline{\mathcal{B}_r(0)}} |DT_k(\bar{x} + b)c| \leq Z_k(r). \tag{3.2}$$

If $\|Y + Z(r)\|_\infty < r$, then $T : \overline{\mathcal{B}_r(\bar{x})} \longrightarrow \overline{\mathcal{B}_r(\bar{x})}$ is a contraction mapping, with contraction constant

$$\kappa = \frac{\|Z(r)\|}{r} < 1.$$

In particular, there exists a unique $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})}$ such that $T(\widehat{x}) = \widehat{x}$.

Proof 3.3. By applying the Mean Value Theorem to T_k, for any $x,y \in \overline{\mathcal{B}_r(\bar{x})}$ and $k = 1,....,n$

$$T_k(x) - T_k(y) = DT_k(z)(x - y), \quad \text{for some} \quad z \in \{tx + (1-t)y, \quad t \in [0,1]\} \subset \overline{\mathcal{B}_r(\bar{x})}.$$

This implies that

$$\Big|T_k(x) - T_k(y)\Big| = \Big|DT_k(z)\frac{r(x-y)}{\|x-y\|_\infty}\Big|\frac{\|x-y\|_\infty}{r} \leq Z_k(r)\frac{\|x-y\|_\infty}{r}. \tag{3.3}$$

If $y = \bar{x}$ then $\dfrac{\|x-y\|_\infty}{r} = \dfrac{\|x-\bar{x}\|_\infty}{r} \leq 1$. This implies that

$$\Big|T_k(x) - T_k(\bar{x})\Big| \leq Z_k(r).$$

By the triangular inequality, one has that

$$\begin{aligned}\Big|T_k(x) - \bar{x}_k\Big| &\le \Big|T_k(x) - T_k(\bar{x})\Big| + \Big|T_k(\bar{x}) - \bar{x}_k\Big|,\\ &\le Z_k(r) + Y_k,\\ &\le \|Y + Z(r)\|_\infty,\\ &< r.\end{aligned}$$

That prove that $T(\overline{\mathcal{B}_r(\bar{x})}) \subset \overline{\mathcal{B}_r(\bar{x})}$. By (3.3), it follows that

$$\|T(x) - T(y)\|_\infty = \max_k\{|T_k(x) - T_k(y)|\} \le \|Z(r)\|_\infty \frac{\|x - y\|_\infty}{r}.$$

Since $\|Z(r)\|_\infty < \|Y + Z(r)\|_\infty < r$, T is a contraction on $\overline{\mathcal{B}_r(\bar{x})}$, with a contraction constant

$$\kappa := \frac{\|Z(r)\|}{r}.$$

Thus, by the contraction mapping theorem, T has a unique fixed point $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})}$, such that $T(\widehat{x}) = \widehat{x}$.

Remark 23. Remark that by the following splitting, $Z_k(r)$ is a polynomial.

$$\begin{aligned}DT(\bar{x} + b)c &= [I - A \cdot Df(\bar{x})]c - A[Df(\bar{x} + b) - Df(\bar{x})]c, \quad b,c \in \overline{\mathcal{B}_r(0)},\\ &= [I - A \cdot Df(\bar{x})]\tilde{c}r - A[Df(\bar{x} + \tilde{b}r) - Df(\bar{x})]\tilde{c}r, \quad \tilde{b},\tilde{c} \in \overline{\mathcal{B}_1(0)}.\end{aligned}$$

Definition 17. Given vectors Y and $Z(r)$ in $\mathbb{R}^n$ satisfying (3.2) and $T \in \mathcal{C}^1(U, \mathbb{R}^n)$, where $U \subset \mathbb{R}^n$ an open set, the associate radii polynomials $p_k(r), \quad k = 1,2,..,n$ are defined by

$$p_k(r) := Y_k + Z_k(r) - r.$$

Using this definition, we can modify the previous theorem in the best way.

Corollary 3.4. Let $f \in \mathcal{C}^1(U, \mathbb{R}^n)$, where $U \subset \mathbb{R}^n$ is an open set. Suppose that $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an invertible linear map and define

$$T(x) := x - Af(x).$$

Let Y, $Z(r)$ satisfying (3.2), $\bar{x} \in U$. Let $p_k(r)$, $k = 1,2,...,n$ be the radii polynomials. If there exists $r > 0$ such that $p_k(r) < 0$ for all $k = 1,2,...,n$, then there exists a unique $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})}$ such that $T(\widehat{x}) = \widehat{x}$, and hence, by Proposition 3.1, there exists a unique $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})}$ such that $f(\widehat{x}) = 0$.

Proof 3.5. By assumption, suppose that $p_k(r) < 0$, $k = 1,2,...,n$ for some $r > 0$. Hence

$$\|Y + Z(r)\|_\infty = \max_k\{|T_k(x) - T_k(y)|\} < r.$$

By theorem (3.2) there exists a unique $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})})$ such that $T(\widehat{x}) = \widehat{x}$ and hence, by Proposition (3.1) there exists a unique $\widehat{x} \in \overline{\mathcal{B}_r(\bar{x})})$ such that $f(\widehat{x}) = 0$.

Proposition 3.6. Let us consider the radii polynomials $p_k(r)$, $k = 1,2,...,n$ in Definition (16). Define

$$\mathcal{I} = \bigcap_{i=1}^{n} \mathcal{I}^{(i)} := \bigcap_{i=1}^{n} \{r > 0, \quad p_i(r) < 0\}.$$

If $\mathcal{I} \neq \emptyset$, then $\mathcal{I}$ is an open interval, an for any $r \in \mathcal{I}$, the ball $\mathcal{B}_{\bar{x}}(r)$ contains a unique solution $\tilde{x}$ such that $f(\tilde{x}) = 0$. Note that $\tilde{x}$ is that same solution for all $r \in \mathcal{I}$.

Proof 3.7. Assume that the highest degree of the polynomial nonlinearities of $f(x)$ is n. Fix $i \in \{1, ...,n\}$, then the coefficients of the radii polynomials will be of the form

$$p_i(r) = a_n^{(i)} r^n + a_{n-1}^{(i)} r^{n-1} + ... + a_1^{(i)} r - r + a_0^{(i)}.$$

with $a_j^{(i)} \geq 0$, (because $a_j^{(i)}$ are the coefficients given by the norms in Y and $Z(r)$). for all $j = 1,....,n$. Since $\mathcal{I} \neq 0$, then $a_1^{(i)} - 1 < 0$. Otherwise we would not be able to find $r > 0$ such that $p_i(r) < 0$. By Descarte's rule of signs and since $\mathcal{I} \neq 0$, each radii polynomial p_i has exactly two positive real zeros that we denote by $r_-^{(i)} < r_+^{(i)}$. Defining $\mathcal{I}^{(i)} = (r_-^{(i)}, r_+^{(i)})$, we obtain that $\mathcal{I} = \bigcap_{i=1}^{n} \mathcal{I}^{(i)}$. Let $\mathcal{I} := (r^-,r^+)$, this implies that $\mathcal{I}$ is an open interval.

Consider now $r \in \mathcal{I}$. Hence $p_i(r) < 0$, for all $i = 1,...,n$, and therefore

$$max_{i=1,...,n}\{Y_i + Z_i(r)\} = max_{i=1,...,n}(p_i(r) + r) < r.$$

The result follows from Corollary (3.4).

Remark 24. Proposition (3.6) demonstrates that the radii polynomials approach provides a strategy for obtaining bounds on the smallest ball $\mathcal{B}_{\bar{x}}(r^-)$ and the largest ball (given by $\mathcal{B}_{\bar{x}}(r^+)$) about the approximate solution on which the corresponding Newton-like operator is a contraction mapping.

Remark 25. Note that the existence of $r > 0$ such that $p_k(r) < 0$, $k = 1,2,...,n$, implies the existence of a range of values $\mathcal{I} = (r^-,r^+) \subset \mathbb{R}^+$ over which the inequalities are satisfied. Using the fact that $\tilde{x}$ is the unique zero of f in $\overline{\mathcal{B}_r(\bar{x})}$, for all $r \in \mathcal{I}$, the value r^- provides tight bounds for the location of the exact solution $\tilde{x}$ and r^+ provides information about the domain of isolation for $\tilde{x}$. This maximal interval is called the *existence interval* for the radii polynomials and it will be determined by using Intlab to obtain the desired bounds. We can conclude that if $\mathcal{X} \neq \emptyset$, then one can present an explicit domain in which there exists a unique root of f.

Let us consider two examples in one and two dimensions to demonstrate how the radii polynomials are used in practice.

Example 18. Application of radii polynomials for $f(x) = x^2 - x - 1$.
Let us consider a simplest nonlinear example $f(x) = x^2 - x - 1$. We want to prove the existence of a root and provide useful bounds on the value of the root. Let $\overline{x} = 1.5$ be an initial value. To define the radii polynomial, we first need to construct T or equivalently we need to choose A, with $T(x) = x - Af(x)$. Here A is chosen based on $Df(\overline{x})^{-1}$. We have

$$Df(\overline{x})^{-1} = \frac{1}{2.x_0 - 1} = \frac{1}{2} = 0.50, \quad \text{set} \quad A = 0.50,$$

then $T(x) = x - 0.50(x^2 - x - 1)$. We need to find Y and $Z(r)$ such that :

$$|T(\overline{x}) - \overline{x}| \le Y \quad \text{and} \quad \sup_{b,c \in \overline{\mathcal{B}_r(0)}} (|DT(\overline{x} + b)c|) \le Z(r).$$

We have

$$|T(\overline{x}) - \overline{x}| = |-0.50(\overline{x}^2 - \overline{x} - 1)| = 0.125,$$

so define $Y := 0.130$. To compute $Z(r)$, we can write

$$DT(x) = 1 - 0.50(2x - 1) = 1.50 - x.$$

Let $b = ru$ and $c = rv$ with $u,v \in \overline{\mathcal{B}_1(0)}$, then

$$|DT(\overline{x} + b)c| = |(1.50 - (\overline{x} + b))c| = |(1.5 - \overline{x} - b)c| = |(1.50 - 1.50 - ru)rv| = |-r^2uv|$$

and we get

$$\sup_{b,c \in \overline{\mathcal{B}_r(0)}} (|DT(\overline{x} + b)c|) = r^2 = \sup_{u,v \in \overline{\mathcal{B}_1(0)}} |r^2uv| = r^2.$$

Let $Z(r) = r^2$ then the associate Radii polynomial is

$$p(r) = Y + Z(r) - r = r^2 - r + 0.130.$$

Note that $p(r) < 0$ for all $r \in [0.16,0.80]$. We can conclude that there exists a unique $\tilde{x} \in \overline{\mathcal{B}_{1.5}(0.16)} = [1.34,1.66]$ such that $f(\tilde{x}) = 0$. Observe that the root $\frac{1+\sqrt{5}}{2} \approx 1.618 \in \overline{\mathcal{B}_{0.16}(1.5)}$. Using the quadratic formula we see that $p(r) < 0$ for all $r \in [0.16,0.80]$.

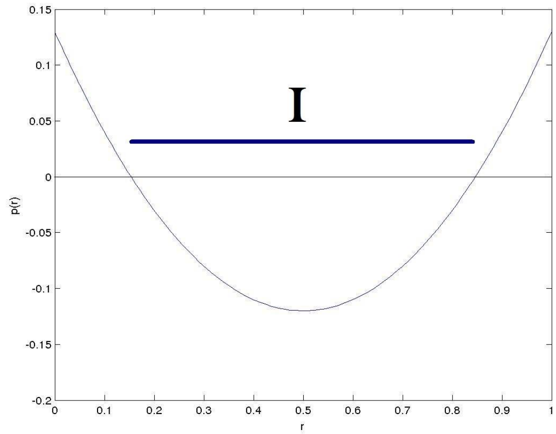

Figure 3.1: The radii polynomial $p(r) = r^2 - r + 0.130$, and $I = [0.16, 0.80]$, where the radii polynomial is strictly negative.

Example 19. (The radii polynomial approach for a two-dimensional example) Consider the FitzHugh-Nagumo equation

$$\dot{x} = f(x) = \begin{pmatrix} x_1(x_1 - a)(1 - x_1) - x_2 \\ \epsilon(x_1 - \gamma x_2) \end{pmatrix}, \tag{3.4}$$

at the parameter values $(a,\epsilon,\gamma) = (5, 1, 2)$.

An equilibrium solution $x = (x_1,x_2)$ is a solution of $f(x) = 0$, where $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by the rand-hand side of (3.4). let

$$Df(\bar{x}) = \begin{pmatrix} -3\bar{x}_1^2 + 2\bar{x}_1(1 + a) - a & -1, \\ \epsilon & -\epsilon\lambda \end{pmatrix}.$$

In this two-dimensional example, the explicit exact formula for the inverse of $Df^{-1}(\bar{x})$ is

$$Df^{-1}(\bar{x}) = \frac{1}{\epsilon - 2\lambda\epsilon\bar{x}_1(1 + a) - 3\lambda\epsilon\bar{x}_1} \begin{pmatrix} -\epsilon\lambda & 1 \\ -\epsilon & -3\bar{x}_1^2 + 2\bar{x}_1(1 + a) - a \end{pmatrix}.$$

Therefore, in this case, we set $A := Df^{-1}(\bar{x})$, and let $T(x) := x - Af(x)$. To apply the radii polynomial approach we compute the bounds Y and Z satisfying (3.2). First, realize that

$$\begin{aligned} T(\bar{x}) - \bar{x} &= -Af(\bar{x}) = -Df^{-1}(\bar{x})f(\bar{x}), \\ &= \frac{1}{\epsilon - 2\lambda\epsilon\bar{x}_1(1 + a) - 3\lambda\epsilon\bar{x}_1} \begin{pmatrix} -\epsilon\lambda & 1 \\ -\epsilon & -3\bar{x}_1^2 + 2\bar{x}_1(1 + a) - a \end{pmatrix} \begin{pmatrix} \bar{x}_1(\bar{x}_1 - a)(1 - \bar{x}_1) - \bar{x}_2 \\ \epsilon(\bar{x}_1 - \gamma\bar{x}_2) \end{pmatrix}. \end{aligned}$$

Using this expression, we can compute (for instance with interval arithmetic) a bound Y_1,Y_2 such that

$$|[T(\bar{x}) - \bar{x}]_k| \leq Y_k, \quad \text{for} \quad k = 1{,}2.$$

The next step is to compute bounds Z_1,Z_2 such that

$$\sup_{b,c \in \overline{\mathcal{B}_r(0)}} |DT_k(\bar{x}+b)c| \leq Z_k(r), \quad \text{for} \quad k = 1{,}2.$$

Let us consider the following splitting

$$DT(\bar{x}+b)c = [I - A \cdot Df(\bar{x})]c - A[Df(\bar{x}+b) - Df(\bar{x})]c.$$

We have

$$Df(\bar{x}+b) - Df(\bar{x}) = \begin{pmatrix} -6\bar{x}_1 b_1 - 3b_1^2 + b_1(2+2a) & 0 \\ 0 & 0 \end{pmatrix}.$$

Let $b = ur$ and $c = vr$ with $\|u\| \leq 1$ and $\|v\| \leq 1$. This implies

$$\begin{aligned}
DT(\bar{x}+b)c &= \underbrace{[I - ADf(\bar{x})]}_{\epsilon} c - A \begin{pmatrix} -6\bar{x}_1 b_1 - 3b_1^2 + b_1(2+2a) & 0 \\ 0 & 0 \end{pmatrix} c, \\
DT(\bar{x}+b)c &= \epsilon \begin{pmatrix} 1 \\ 1 \end{pmatrix} r - A \begin{pmatrix} -6\bar{x}_1 b_1 - 3b_1^2 + b_1(2+2a) & 0 \\ 0 & 0 \end{pmatrix} c, \\
|DT(\bar{x}+b)c| &\leq \epsilon \begin{pmatrix} 1 \\ 1 \end{pmatrix} r + |A| \left| \begin{pmatrix} -6\bar{x}_1 b_1 - 3b_1^2 + b_1(2+2a) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 r \\ v_2 r \end{pmatrix} \right|, \\
|DT(\bar{x}+b)c| &\leq \epsilon \begin{pmatrix} 1 \\ 1 \end{pmatrix} r + |A| \left| \begin{pmatrix} -6\bar{x}_1 u_1 v_1 r^2 - 3u_1^2 v_1 r^3 + u_1 v_1 r^2 (2+2a) \\ 0 \end{pmatrix} \right|, \\
|DT(\bar{x}+b)c| &\leq \epsilon \begin{pmatrix} 1 \\ 1 \end{pmatrix} r + |A| \left| \begin{pmatrix} (6\bar{x}_1 + 2 + 2a) r^2 + 3r^3 \\ 0 \end{pmatrix} \right|,
\end{aligned}$$

where the inequalities are considered component-wise. Let

$$Z_2(r) := |A| \begin{pmatrix} 6\bar{x} + 2 + 2a \\ 0 \end{pmatrix} r^2 \quad \text{and} \quad Z_3(r) := |A| \begin{pmatrix} 3 \\ 0 \end{pmatrix} r^3.$$

The radii polynomials $p_1(r)$ and $p_2(r)$ are defined by

$$\begin{pmatrix} p_1(r) \\ p_2(r) \end{pmatrix} = Z_3(r) + Z_2(r) + (\epsilon - 1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} r + Y.$$

It is important to realize that the definition of the radii polynomials is the same for any approximate solution $\bar{x} \in \mathbb{R}^2$. This means that once we have derived explicit formulas of the radii polynomials, we can use them for any $\bar{x}$ that we obtained. We fixed $(a,\epsilon,\gamma) = (5,1,2)$, and using Newton's method, we computed the approximate solution

$$\bar{x} = \begin{pmatrix} 1.1292 \\ 0.5646 \end{pmatrix}.$$

For each $i = 1,2$, we could compute the interval $\mathcal{I}^{(i)}$ on which each radii polynomial is negative. Hence,

$$\begin{aligned} \mathcal{I} &= \bigcap_{k=1}^{2} \{r > 0 : p_k(r) < 0\}, \\ &= \mathcal{I}^{(1)} \cap \mathcal{I}^{(2)}. \end{aligned}$$

We conclude that for all $r \in \mathcal{I}$, $B_r(\bar{x})$ contains a unique solution $\hat{x}$ such that $f(\hat{x}) = 0$.

```
>> int_Radii(x,2)

ans =
0.71004      3.0237    -0.99999  2.8693e-05
0.35502      7.5119    -0.99999  1.4347e-05

The roots are

ans =
  -4.5667e+00
   3.0837e-01
   2.8696e-05
ans =
  -2.1291e+01
   1.3229e-01
   1.4348e-05
intval ans =
[  2.8695e-005,  1.3229e-001]
```

The first vector column is the coefficients of $Z_3(r)$, the second the coefficients of $Z_2(r)$, the third the coefficients of $Z_1(r)$ and the last column represents the coefficients of Y. The codes associated to this example can be found in annexe.

Chapter 4

Galaktionov-Svirshchevskii's conjecture

4.1 Derivation of the Galaktionov-Svirshchevskii's conjecture

The Galaktionov-Svirshchevskii's conjecture originally comes from the study of the Kuramoto-Sivashinsky (KS) equation, which is given by the semi-linear fourth order parabolic PDE

$$u_t = -u_{yyyy} - u_{yy} + (u_y)^2.$$

This model of equation was originally introduced to describe flame front propagation in turbulent flows of gaseous combustible mixtures, but also have many applications in physics, including 2D turbulence.

In [6], the authors Galaktionov and Svirshchevskii considered a modified KS equation of the form

$$u_t = -u_{yyyy} - u_{yy} + (1-\lambda)(u_y)^2 + \lambda(u_{yy})^2,$$

where $\lambda \in [0,1]$ is a constant which describes dynamical properties of hyper-cooled melt. They modified yet again the equation and introduced an absorption term, in order to describe the extinction phenomena, that is

$$u_t = -u_{yyyy} + u_{yy}^2 - 1.$$

Here the presence of the non Lipschitz absorption term -1 implies that the problem is rather consistent for the Cauchy problem, as we can impose the zero contact angle at each interface. Finally, in order to face the question of the maximal regularity of the solutions at the interfaces, they considered the signed version of this PDE for solutions of changing signs, that is

$$u_t = -u_{yyyy} + u_{yy}^2 - sign(u), \quad (y,t) \in \mathbb{R} \times \mathbb{R}^+ \tag{4.1}$$

We would like to study the equation (4.1) by considering the travelling wave solutions. We detect the oscillatory component for the travelling waves by using the ansatz

$$u(y,t) = h(w), \quad w = y - \lambda t, \;\; \lambda \in \mathbb{R}. \tag{4.2}$$

Here, $\lambda \neq 0$ is the travelling wave speed and this implies that $u(w) = h(w)$ is just a stationary solution of (4.1).

Suppose that at $w = 0$, $h(w)$ has the interface with the trivial extension $h \equiv 0$ for $w < 0$. For $w > 0$, we can exhibit "the maximal regularity" for this PDE, and by substituting in (4.1), we get

$$-\lambda h' = -h^{(4)} + (h'')^2 - sign(h), \quad w > 0, \quad h(0) = 0. \tag{4.3}$$

Remark 26. To exhibit the "maximal regularity", the expression is like

$$u(y,t) \sim (y - s(t))^4 \phi(ln(y - s(t))), \text{ as } y \longrightarrow s^+(t).$$

$s(t)$ is the interface. The strong absorption term is involved in the oscillatory behavior. The leading change sign asymptotic are governed by the terms $h^{(4)}$ and $sign(h)$ in (4.3).

By neglecting the non-stationary, $\lambda-$dependent term for small h implies that $\lambda h' = 0$ and $h'' = 0$, so we keep two leading terms and get

$$h^{(4)} + sign(h) = 0. \tag{4.4}$$

Set $h(w) = w^4 \phi(s)$, where $s = \ln(w)$. We call $\phi(s)$ the *oscillatory component* for h. By plugging it in (4.4), we obtain

$$\begin{aligned}
h^{'}(w) &= 4w^3\phi(ln(w)) + w^3\phi^{'}(ln(w)),\\
h^{''}(w) &= 12w^2\phi(ln(w)) + 7w^2\phi^{'}(ln(w)) + w^2\phi^{''}(ln(w)),\\
h^{'''}(w) &= 24w\phi(ln(w)) + 26w\phi^{'}(ln(w)) + 9w\phi^{''}(ln(w)) + w\phi^{'''}(ln(w)),\\
h^{(4)}(w) &= 24\phi(ln(w)) + 50\phi^{'}(ln(w)) + 35\phi^{''}(ln(w)) + 10\phi^{'''}(ln(w)) + \phi^{(4)}(ln(w)).
\end{aligned}$$

Hence, ϕ satisfies the autonomous fourth-order ODE

$$\phi^{(4)} + 10\phi^{(3)} + 35\phi^{(2)} + 50\phi' + 24\phi + sign(\phi) = 0, \quad sign(\phi) = \begin{cases} +1, & \phi \geq 0, \\ -1, & \phi < 0 \end{cases} \tag{4.5}$$

Conjecture 4.1 (Galaktionov-Svirshchevskii [6])**.** The ODE (4.5) has a unique nontrivial periodic solution $\phi(s)$ which is asymptotically stable as $s \longrightarrow +\infty$.

The rest of the present work is dedicated to demonstrate one part of the Galaktionov-Svirshchevskii's conjecture, that is that ODE (4.5) has a locally unique nontrivial periodic solution. To demonstrate this, we recast the problem of looking for a periodic solution of (4.5) as a problem of the form $f(x) = 0$ posed on a finite dimensional Banach space. Setting up the problem $f(x) = 0$ requires some basic notions of ODEs, as introduced in Chapter 2. Then, we combine the theory of the radii polynomial approach as introduced in Chapter 3 together with the theory of interval arithmetic as introduced in Chapter 1, to prove the result.

4.2 Set-up of the problem $f(x) = 0$

Consider a time rescaling factor $L > 0$, and let

$$\Psi(t) := \phi(tL).$$

Hence, we can see that if ϕ is a solution of (4.5), then $-\phi$ is also a solution. We use this symmetry to set-up the problem. Let

$$\begin{aligned}
\phi_1 &:= \phi \\
\phi_2 &:= \phi' \\
\phi_3 &:= \phi'' \\
\phi_4 &:= \phi^{(3)}.
\end{aligned}$$

This implies that

$$\begin{aligned}
\begin{pmatrix} \phi_1' \\ \phi_2' \\ \phi_3' \\ \phi_4' \end{pmatrix} &= \begin{pmatrix} \phi_2 \\ \phi_3 \\ \phi_4 \\ -10\phi_4 - 35\phi_3 - 50\phi_2 - 24\phi_1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -sign(\phi_1) \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -24 & -50 & -35 & -10 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -sign(\phi_1) \end{pmatrix} = A\phi + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ -sign(\phi_1) \end{pmatrix}}_{b},
\end{aligned}$$

where

$$\phi := \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix}, \quad A := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -24 & -50 & -35 & -10 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -sign(\phi_1) \end{pmatrix}.$$

We obtain that solving (4.5) is equivalent to solve

$$\begin{cases} \phi' = g_1(\phi) := A\phi + b, & \phi \geq 0, \\ \phi' = g_2(\phi) := A\phi - b, & \phi < 0. \end{cases}$$

Before continuing our discussions, let us introduce a lemma which help us to solve the problem in the positive domain and deduce the solution in the negative domain.

Lemma 4.2. Let $\phi : [0,L] \longrightarrow \mathbb{R}^4$, $\phi'(t) = A\phi(t) + b$ with $\phi(0) = \phi_0 = -\phi(L)$. Then $\varphi(t) := -\phi(t)$ satisfies $\varphi'(t) = A\varphi(t) - b$ and $\varphi(0) = -\varphi(L)$.

Proof 4.3. We know from the general theory of ordinary differential equations, as introduced in Chapter 2, that the general solution of $\phi(t)$ is given by

$$\begin{aligned}\phi(t) &= e^{At}\phi_0 + \int_0^t e^{A(t-s)}b(ds,\\ -\varphi(t) &= -e^{A(t)}\varphi(0) + \int_0^t e^{(t-s)}bds,\\ \varphi(t) &= e^{A(t)}\varphi_0 - \int_0^t e^{(t-s)}bds.\end{aligned}$$

This implies that $\varphi'(t) = A\varphi(t) - b$ and $\varphi(0) = -\phi(0) = \phi(L) = -\varphi(L)$. $\square$

By the previous lemma, it is sufficient to consider the problem in the upper domain $t \in [0,L]$. Let $x = (L, a_1,a_2,a_3) \in \mathbb{R}^4$, where L is the period and $\phi_0 = (0,a_1,a_2,a_3)$, and consider the nonlinear model

$$\begin{cases} f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4, \\ x \longrightarrow f(x) = \phi_0 + \phi(L), \quad \text{where} \quad x = (L,a_1,a_2,a_3) \quad \text{solution on the positive domain.}\end{cases}$$

More explicitly,

$$f(x) := \begin{pmatrix} 0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} + e^{AL}\begin{pmatrix} 0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} + \int_0^L e^{A(L-s)}b(s)ds. \tag{4.6}$$

The problem $f(x) = 0$ implies that $\phi_0 + \phi(L) = 0$. So we get,

$$\phi_0 + e^{AL}\phi_0 + \int_0^L e^{(A-L)}bds = 0,$$

$$\text{where } A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -24 & -50 & -35 & -10 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

It is a non homogeneous linear systems, so we turn on chapter two to solve this equation. The eigenvalues and associated eigenvectors are

$$\lambda_1 = -4, \quad \lambda_2 = -1, \quad \lambda_3 = -2 \quad \text{and} \quad \lambda_4 = -3,$$

with associated eigenvectors

$$v_1 = \begin{pmatrix} -1/64 \\ 1/10 \\ -1/4 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1/8 \\ 1/4 \\ -1/2 \\ 1 \end{pmatrix} \quad \text{and} \quad v_4 = \begin{pmatrix} -1/27 \\ 1/9 \\ -1/3 \\ 1 \end{pmatrix}.$$

Let

$$P = \begin{pmatrix} -1/64 & -1 & -1/8 & -1/27 \\ 1/10 & 1 & 1/4 & 1/9 \\ -1/4 & -1 & -1/2 & -1/3 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

The fundamental matrix solution $\Phi(t) = e^{At}$, with $\Phi(0) = I$ is given by

$$e^{At} = P^{-1} diag(e^{\lambda_1 t}, e^{\lambda_2 t}, e^{\lambda_3 t}, e^{\lambda_4 t})P$$

$$= \begin{pmatrix} 64 & 352/3 & 64 & 32/3 \\ -4 & -13/3 & -3/2 & -1/6 \\ 4/8 & 76 & 32 & 4 \\ -108 & -189 & -189/2 & -27/2 \end{pmatrix} \begin{pmatrix} e^{-4t} & 0 & 0 & 0 \\ 0 & e^{-t} & 0 & 0 \\ 0 & 0 & e^{-2t} & 0 \\ 0 & 0 & 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} -1/64 & -1 & -1/8 & -1/27 \\ 1/10 & 1 & 1/4 & 1/9 \\ -1/4 & -1 & -1/2 & -1/3 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

We now use Newton's method to obtain an approximation of the solution

$$\bar{x} = \begin{pmatrix} 1.4183 \\ 0.0225 \\ 0.0084 \\ -0.0488 \end{pmatrix}.$$

The graph of the periodic orbit can be found in Figure 4.2.

4.3 The radii polynomial approach for Galaktionov-Svirshchevskii's conjecture

Let us consider the problem $f(x) = 0$ as given by (4.6). To use the radii polynomial approach as introduced in Chapter 3, we start to compute the bound $Y \in \mathbb{R}^4$ such that $|Df^{-1}(\bar{x})f(\bar{x})| \leq Y$. For the bounds Z_i, let us consider $h : [0,1] \longrightarrow \mathbb{R}^4$ by $h(s) = D_x f(\bar{x} + sb)c$. We have

$$h(1) - h(0) = D_x f(\bar{x} + b)c - D_x f(\bar{x}).$$

For all $k \in \{1,....,4\}$, there exists $s_k \in [0,1]$ such that

$$(D_x f_k(\bar{x}+b) - D_x f_k(\bar{x}))c = h_k(1) - h_k(0) = h_k'(s_k) = D_x^2 f_k(\bar{x}+s_k b)(b,c).$$

Let $\tilde{b}$ and $\tilde{c}$ in $\overline{B}_1(0)$ such that $b = \tilde{b}r$ and $c = \tilde{c}r$. It implies

$$(D_x f_k(\bar{x}+b) - D_x f_k(\bar{x}))c \leq D_x^2 f_k(\bar{x}+s_k b)(\tilde{b},\tilde{c})r^2.$$

Let $r^* = 10^{-4}$ an a-priori upper bound for the left point of the existence interval of the radii polynomials. We will have to show a-posteriori that $r < r^*$. Denote by $b^* = [-r^*,r^*]^4$ a vector in $\mathbb{R}^4$ with entries given by the interval $[-r^*,r^*]$.

Let $x^* = \bar{x} + b^* \in \mathbb{R}^4$, with its $k-th$ entry given by the interval $[\bar{x}_k - r^*, \bar{x}_k + r^*]$. Denote by $\delta = [-1,1]^4$ a vector in $\mathbb{R}^4$, whose entries are given by the interval $[-1,1]$, then for each b, r in $\overline{B_r(0)}$, we get

$$Df^{-1}(\bar{x})|(D_x f(\bar{x}+b) - D_x f(\bar{x}))c| \leq |AD_x^2 f(\bar{X})(\delta,\delta)r^2|.$$

Using interval arithmetic, compute $Z^{(2)} \in \mathbb{R}^5$, such that

$$|Df^{-1}(\bar{x})D_x^2 f(\bar{X})(\delta,\delta)r^2| \leq Z^{(2)}.$$

Using the previous bounds, define the radii polynomials

$$p_k(r) = Z_k^{(2)} r^2 - r + Y_k, \quad k = 1,...,4.$$

4.4 Validated numerics of Galaktionov-Svirshchevskii's conjecture

We now used the Z and Y bounds of Section 4.3 to compute the the radii polynomials defined by

$$p_k(r) = Z_k^{(2)} r^2 - r + Y_k, \quad k = 1,...,4,$$

hence, we obtained the following result

```
>> conj_Radii(x, 1e-4)
P =
   7.4544e+00  -1.0000e+00   4.9428e-15
   1.3096e+01  -1.0000e+00   1.3486e-16
   2.6877e+01  -1.0000e+00   2.8171e-16
   6.2463e+01  -1.0000e+00   5.1074e-16
ans =
   4.9923e-15   1.0001e-04
```

Figure 4.1: radii polynomials generated with ϕ.

The first column represent the vector $(Z_i^{(2)})_{i=1}^4$, the second vector $(-1)_{i=1}^4$ and the third $(Y_i)_{i=1}^4$.

Theorem 4.4. For every $r \in \mathcal{I} = [4.9923e^{-15}, 1.0001e^{-4}]$, there exists a unique $\hat{x} \in \mathcal{B}_{\bar{x}}(r)$ such that $f(\hat{x}) = 0$, with f given in (4.6). That corresponds to a periodic orbit of the Galaktionov-Svirshchevskii's conjecture with period $L = 1.4183$.

Proof 4.5. The follows from an application of Proposition (3.6) of Chapter 3.

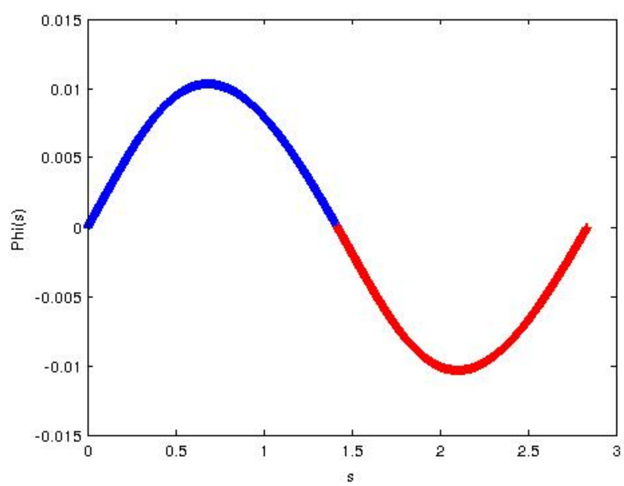

Figure 4.2: graph of the period solution for Galaktionov-Svirshchevskii's conjecture.

After defining $F\_conj(x)$ and $DF\_conj(x)$, we use Newton's method to find $\bar{x}$, which is an approximation of the exact solution $\tilde{x}$. We now define the operator $T(x) := x - Af(x)$, where $A = Df(\bar{x})^{-1}$. Then we can define the radii polynomial $p_i(r), \quad r = 1,2$, by the function $conj\_Radii(x, r\_star)$. We use this function to compute $\mathcal{I} = \{r > 0, \quad p_i(r) < 0\} \neq \emptyset$. The codes associated to this work can be found in annexe.

Conclusion

Judicious use of interval arithmetic, combined with careful pen and paper estimates, leads to effective strategies for computer assisted analysis of nonlinear operator equations. In this work, we provide a rigorous computational method for finding the periodic solution as conjectured by Galaktionov and Svirshchevskii in its conjecture, which was an open problem prior to the present work. By the Radii polynomial approach and interval arithmetic, we can provide a strategy for obtaining bounds about the approximate solution on which the corresponding Newton-like operator is a contraction mapping. Note that in this work, we did not prove the stability for the solution of the Galaktionov-Svirshchevskii's conjecture which required other concepts in dynamical systems.

This method has been introduced to study solutions of ordinary, partial, and delay differential equations, such has periodic orbits, equilibria and solutions of initial value problems (IVPs). It is a powerful technique in mathematics but for infinite dimensional Banach spaces, an explicit representation of the inverse of the derivative is not possible. In this case, other tools will be used.

Bibliography

[1] M. Gameiro, J.-P. Lessard, J. Mireles James and K. Mischaikow, *Ordinary Differential Equations: A Constructive Approach*, in preparation, 2015.

[2] G. I. Hargreaves, *Interval Analysis in MATLAB*, Numerical Analysis Report No. 416, December 2002.

[3] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems-Springer Berlin*, Springer-Verlag Berlin Heidelberg 1996.

[4] M. Gameiro, J.-P. Lessard and Y. Ricaud, *Rigorous numerics for piecewise-smooth systems: a functional analytic approach based on Chebyshev series*, To appear in Journal of Computational and Applied Mathematics, 2015.

[5] C. Chicone, *Ordinary differential equations with applications*, (Texts in applied mathematics, 34), Springer (2006).

[6] Victor A Galaktionov, Sergey R Svirshchevskii, *Exact Solutions and Invariant Subspaces of Nonlinear Partial Differential Equations in Mechanics and Physics*, Chapman & Hall/CRC, applied mathematics and nonlinear science series, (2007).

[7] C. Robinson, *Dynamical Systems, Stability, Symbolic Dynamics, and Chaos*, Studies in Advanced Mathematics CRC Press, 1994.

[8] L. Perko, *Differential Equations and Dynamical Systems*, Springer-Verlag, New-York, 1998.

[9] Y. Hijazi, H. Hagen, C. Hansen, and Kenneth I. Joy, *Why interval arithmetic is so useful*, subs.emis.de/LNI/Seminar/Seminar07/148.pdf.

[10] A. Hungria, J.-P. Lessard and J.D. Mireles James, *Rigorous numerics for analytic solutions of differential equations: the radii polynomial approach*, To appear in Mathematics of Computation, 2015.

[11] William R Holmes, May A. Mata, Leah E. Keshet, *Local Perturbation Analysis in MatCont*, User's Guide & Tutorials, 0ctober 2014.

[12] B. N. Datta, *Numerical Linear Algebra and Applications*, second edition, SIAM, 2010.

Appendix A

numerical codes

Codes associated to the FitzHugh-Nagumo equation

```
function [y] = int_myfunction(ix, ig)

% a,eps and g are parameters.
a = intval(5);
eps = intval(1);

y=[ix(1)*(ix(1)-a)*(1- ix(1))-ix(2); eps*( ix(1)-ig*ix(2))];

end
```

```
% This function computes de Jacobian matrix for FitzHugh-Nagumo
    equation.

function [dy] = int_function_df(ix,ig)

% ix is the initial guess of the root
% ig is a parameter in the Lorenz equation
a = intval(5);
eps = intval(1);
%The Jacobian matrix
dy(1,1) = -3*ix(1)^2 + ix(1)*(2+2*a)-a;
dy(1,2) = intval(-1);
dy(2,1) = eps;
```

```
dy(2,2)=-eps*ig;

end
```

```
function I = int_Radii(x0,g)
% x is the initial point and g is a parameter;
ig=intval(g);
%x0=intval(newton(x,20,g));
x11=x0(1,:);
int_f=int_myfunction(x0,ig);
A=inv(int_function_df(x0,ig));
A=intval(A);
Y=abs(A*int_f);
num2str(mid(Y));
fprintf('\n');
eps=abs(intval(eye(2))-A*(int_function_df(x0,ig)));
Z1=eps*(intval([1;1]));
Z2=abs(A)*(intval([6*x11;12]));
Z3=abs(A)*(intval([3;0]));
p_int=[Z3,Z2,Z1-intval([1;1]),Y];
p=sup(p_int);
num2str(p)
fprintf('\n');
fprintf('The roots are \n\n')
for i=1:2
    roots(p(i,:))
end
for i=1:2
    a(i,:) = sort (roots(p(i,:)));
end
if norm(imag(a))== 0
    for i=1:2
        I1= infsup(a(1,2),a(1,3));
        I2= infsup(a(2,2),a(2,3));
        I=intersect(I1,I2);
        radii_min=inf(I)*1.1;
%         evaluate_p_at_rad_min = radii_min^3*p(:,1)+radii_min^2*p
```

```
    (:,2)+ radii_min*p(:,3)+p(:,4);
     end
else
     I=infsup(-1,-1);
%      evaluate_p_at_rad_min = 1;
end

% if max(sup(evaluate_p_at_rad_min))<=0
%
%      fprintf('\n\n');
%      display(['Success! The radius is = ',num2str(radii_min)])
%      fprintf('\n\n\n')
% else
%      display('No radius ')
% end

end
```

```
function [F] = F_conj(x)

x is the vector in R^4, L is the period and Phi0 the initial
    condition.

% x = [L;a]; a = (a1,a2,a3); phi0 = (0,a1,a2,a3);
L=x(1);
a=x(2:4);

%%%%%%%%%%%%%%%%%%
%%% Definition of F %%%
%%%%%%%%%%%%%%%%%%

Homogeneous part of F.

t_hom1= (-(11/6)*exp(-4*L)+7*exp(-3*L)-(19/2)*exp(-2*L)+(13/3)*exp
    (-L))*a(1)+(-exp(-4*L)+(7/2)*exp(-3*L)-4*exp(-2*L)+(3/2)*exp(-L)
    )*a(2)+(-(1/6)*exp(-4*L)+(1/2)*exp(-3*L)-(1/2)*exp(-2*L)+(1/6)*
    exp(-L))*a(3);
t_hom2=((22/3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(13/3)*exp(-L))*
    a(1)+(4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L))*a
    (2)+((2/3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(1/6)*exp(-L))*a
    (3);
t_hom3=(-(88/3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(13/3)*exp(-L))
    *a(1)+(-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*exp(-L)
    )*a(2)+(-(8/3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(1/6)*exp(-
    L))*a(3);
t_hom4=((352/3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(13/3)*exp(-L)
    )*a(1)+(64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*exp(-L
    ))*a(2)+((32/3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(1/6)*exp
    (-L))*a(3);


t_hom=[t_hom1;t_hom2;t_hom3;t_hom4];

Inhomogeneous part of F.
```

```
t_inhom1=   -(1/24)*exp(-4*L)+(1/6)*exp(-3*L)-(1/4)*exp(-2*L)+(1/6)*
    exp(-L)-1/24;
t_inhom2=   (1/6)*exp(-4*L)-(1/2)*exp(-3*L)+(1/2)*exp(-2*L)-(1/6)*
    exp(-L);
t_inhom3=  -(2/3)*exp(-4*L)+(3/2)*exp(-3*L)-exp(-2*L)+(1/6)*exp(-L);
t_inhom4=  (8/3)*exp(-4*L)-(9/2)*exp(-3*L)+2*exp(-2*L)-(1/6)*exp(-L)
    ;

t_inhom=[t_inhom1;t_inhom2;t_inhom3;t_inhom4];

F= t_hom + t_inhom  +[0;a];

end
```

```
function  [DF] = DF_conj(x)

%%%% x = [L;a]; a = (a(1),a(2),a(3)); Phi0 = (0,a(1),a(2),a(3));

L=x(1);
a=x(2:4);
DF=zeros(4);

We give the definition for each component of DF.

DF(1,1)= ((22/3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(13/3)*exp(-L)
    )*a(1)+(4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L))*
    a(2)+((2/3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(1/6)*exp(-L))*a
    (3)+(1/6)*exp(-4*L)-(1/2)*exp(-3*L)+(1/2)*exp(-2*L)-(1/6)*exp(-L
    );
DF(1,2)=-(11/6)*exp(-4*L)+7*exp(-3*L)-(19/2)*exp(-2*L)+(13/3)*exp(-
    L);
DF(1,3)=-exp(-4*L)+(7/2)*exp(-3*L)-4*exp(-2*L)+(3/2)*exp(-L);
DF(1,4)=-(1/6)*exp(-4*L)+(1/2)*exp(-3*L)-(1/2)*exp(-2*L)+(1/6)*exp
    (-L);

DF(2,1)= (-(88/3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(13/3)*exp(-L
    ))*a(1)+(-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*exp(-
    L))*a(2)+(-(8/3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(1/6)*exp
    (-L))*a(3)-(2/3)*exp(-4*L)+(3/2)*exp(-3*L)-exp(-2*L)+(1/6)*exp(-
```

```
    L);
DF(2,2)=(22/3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(13/3)*exp(-L)
    +1;
DF(2,3)=4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L);
DF(2,4)=(2/3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(1/6)*exp(-L);

DF(3,1)= ((352/3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(13/3)*exp(-
    L))*a(1)+(64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*exp
    (-L))*a(2)+((32/3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(1/6)*
    exp(-L))*a(3)+(8/3)*exp(-4*L)-(9/2)*exp(-3*L)+2*exp(-2*L)-(1/6)*
    exp(-L);
DF(3,2)=-(88/3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(13/3)*exp(-L);
DF(3,3)=-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*exp(-L)
    +1;
DF(3,4)=-(8/3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(1/6)*exp(-L);


DF(4,1)= (-(1408/3)*exp(-4*L)+567*exp(-3*L)-152*exp(-2*L)+(13/3)*
    exp(-L))*a(1)+(-256*exp(-4*L)+(567/2)*exp(-3*L)-64*exp(-2*L)
    +(3/2)*exp(-L))*a(2)+(-(128/3)*exp(-4*L)+(81/2)*exp(-3*L)-8*exp
    (-2*L)+(1/6)*exp(-L))*a(3)-(32/3)*exp(-4*L)+(27/2)*exp(-3*L)-4*
    exp(-2*L)+(1/6)*exp(-L);
DF(4,2)=(352/3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(13/3)*exp(-L)
    ;
DF(4,3)=64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*exp(-L);
DF(4,4)=(32/3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(1/6)*exp(-L)
    +1;

end
```

```
function [I,success] = conj_Radii(x, r_star)

% x = [L;a]; a = (a(1),a(2),a(3)); phi0 = (0,a(1),a(2),a(3));

x=intval(x);
L=x(1);
a=x(2:4);
```

```
Because the radius of intval(2/3), intval(1/6), etc... are
    different to zero, we replace
intval(2/3), intval(1/6), etc... by intval(2)/intval(3), intval(1)/
    intval(6), etc....

i1=intval(1);
i2=intval(2);
i3=intval(3);
i6=intval(6);
i8=intval(8);
i11=intval(11);
i13=intval(13);
i22=intval(22);
i24=intval(24);
i32=intval(32);
i88=intval(88);
i128=intval(128);
i352=intval(352);
i512=intval(512);
i1408=intval(1408);
i5632=intval(5632);

%%%%%%%%%%%%%%%%%%
%%% Definition of F %%%
%%%%%%%%%%%%%%%%%%

We define each component of the homogeneous part of F.

t_hom1= (-(i11/i6)*exp(-4*L)+7*exp(-3*L)-(19/2)*exp(-2*L)+(i13/i3)*
    exp(-L))*a(1)+(-exp(-4*L)+(7/2)*exp(-3*L)-4*exp(-2*L)+(3/2)*exp
    (-L))*a(2)+(-(i1/i6)*exp(-4*L)+(1/2)*exp(-3*L)-(1/2)*exp(-2*L)+(
    i1/i6)*exp(-L))*a(3);
t_hom2=((i22/i3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(i13/i3)*exp(-
    L))*a(1)+(4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L)
    )*a(2)+((i2/i3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(i1/i6)*exp
    (-L))*a(3);
t_hom3=(-(i88/i3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(i13/i3)*exp
    (-L))*a(1)+(-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*
    exp(-L))*a(2)+(-(i8/i3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(
```

```
    i1/i6)*exp(-L))*a(3);
t_hom4=((i352/i3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(i13/i3)*exp
    (-L))*a(1)+(64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*
    exp(-L))*a(2)+((i32/i3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(
    i1/i6)*exp(-L))*a(3);

t_hom=[t_hom1;t_hom2;t_hom3;t_hom4];

Inhomogeneous part of F.

t_inhom1=  -(i1/i24)*exp(-4*L)+(i1/i6)*exp(-3*L)-(1/4)*exp(-2*L)+(
    i1/i6)*exp(-L)-i1/i24;
t_inhom2=  (i1/i6)*exp(-4*L)-(1/2)*exp(-3*L)+(1/2)*exp(-2*L)-(i1/i6
    )*exp(-L);
t_inhom3= -(i2/i3)*exp(-4*L)+(3/2)*exp(-3*L)-exp(-2*L)+(i1/i6)*exp
    (-L);
t_inhom4= (i8/i3)*exp(-4*L)-(9/2)*exp(-3*L)+2*exp(-2*L)-(i1/i6)*exp
    (-L);

t_inhom=[t_inhom1;t_inhom2;t_inhom3;t_inhom4];

F= t_hom + t_inhom +[0;a];

%%%%%%%%%%%%%%%%%%%%%%
%%% Definition of invDF %%%
%%%%%%%%%%%%%%%%%%%%%%

We define the inverse of DF.

DF11= ((i22/i3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(i13/i3)*exp(-L
    ))*a(1)+(4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L))
    *a(2)+((i2/i3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(i1/i6)*exp(-
    L))*a(3)+(i1/i6)*exp(-4*L)-(1/2)*exp(-3*L)+(1/2)*exp(-2*L)-(i1/
    i6)*exp(-L);
DF12=-(i11/i6)*exp(-4*L)+7*exp(-3*L)-(19/2)*exp(-2*L)+(i13/i3)*exp
    (-L);
DF13=-exp(-4*L)+(7/2)*exp(-3*L)-4*exp(-2*L)+(3/2)*exp(-L);
DF14=-(i1/i6)*exp(-4*L)+(1/2)*exp(-3*L)-(1/2)*exp(-2*L)+(i1/i6)*exp
    (-L);
```

```
DF21= (-(i88/i3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(i13/i3)*exp(-
    L))*a(1)+(-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*exp
    (-L))*a(2)+(-(i8/i3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(i1/
    i6)*exp(-L))*a(3)-(i2/i3)*exp(-4*L)+(3/2)*exp(-3*L)-exp(-2*L)+(
    i1/i6)*exp(-L);
DF22=(i22/i3)*exp(-4*L)-21*exp(-3*L)+19*exp(-2*L)-(i13/i3)*exp(-L)
    +1;
DF23=4*exp(-4*L)-(21/2)*exp(-3*L)+8*exp(-2*L)-(3/2)*exp(-L);
DF24=(i2/i3)*exp(-4*L)-(3/2)*exp(-3*L)+exp(-2*L)-(i1/i6)*exp(-L);

DF31= ((i352/i3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(i13/i3)*exp
    (-L))*a(1)+(64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*
    exp(-L))*a(2)+((i32/i3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(
    i1/i6)*exp(-L))*a(3)+(i8/i3)*exp(-4*L)-(9/2)*exp(-3*L)+2*exp(-2*
    L)-(i1/i6)*exp(-L);
DF32=-(i88/i3)*exp(-4*L)+63*exp(-3*L)-38*exp(-2*L)+(i13/i3)*exp(-L)
    ;
DF33=-16*exp(-4*L)+(63/2)*exp(-3*L)-16*exp(-2*L)+(3/2)*exp(-L)+1;
DF34=-(i8/i3)*exp(-4*L)+(9/2)*exp(-3*L)-2*exp(-2*L)+(i1/i6)*exp(-L)
    ;

DF41= (-(i1408/i3)*exp(-4*L)+567*exp(-3*L)-152*exp(-2*L)+(i13/i3)*
    exp(-L))*a(1)+(-256*exp(-4*L)+(567/2)*exp(-3*L)-64*exp(-2*L)
    +(3/2)*exp(-L))*a(2)+(-(i128/i3)*exp(-4*L)+(81/2)*exp(-3*L)-8*
    exp(-2*L)+(i1/i6)*exp(-L))*a(3)-(i32/i3)*exp(-4*L)+(27/2)*exp
    (-3*L)-4*exp(-2*L)+(i1/i6)*exp(-L);
DF42=(i352/i3)*exp(-4*L)-189*exp(-3*L)+76*exp(-2*L)-(i13/i3)*exp(-L
    );
DF43=64*exp(-4*L)-(189/2)*exp(-3*L)+32*exp(-2*L)-(3/2)*exp(-L);
DF44=(i32/i3)*exp(-4*L)-(27/2)*exp(-3*L)+4*exp(-2*L)-(i1/i6)*exp(-L
    )+1;

invDF = inv([DF11 DF12 DF13 DF14; DF21 DF22 DF23 DF24; DF31 DF32
    DF33 DF34; DF41 DF42 DF43 DF44]);

%%%%%%%%%%%%%%%% Definition of Y

A = invDF;
```

```
Y=abs(A*F); We define the bound Y.

%%%%%%%%%%%%%%% Definition Z2(r)

L=mid(L);
a=mid(a);
X_ast = [ infsup(L-r_star,L+r_star);
          infsup(a(1)-r_star,a(1)+r_star);
          infsup(a(2)-r_star,a(2)+r_star);
          infsup(a(3)-r_star,a(3)+r_star)  ];

delta=[infsup(-1,1);infsup(-1,1);infsup(-1,1);infsup(-1,1)];

We define D^2(F), which we will used in Z^{2}(r). See page 36.

DDF1 = (i1/i6)*delta(1)*delta(1)*exp(-X_ast(1))-delta(1)*delta(1)*
    exp(-2*X_ast(1))+4*delta(3)*delta(1)*exp(-4*X_ast(1))-(i2/i3)*
    delta(1)*delta(1)*exp(-4*X_ast(1))+(3/2)*delta(1)*delta(1)*exp
    (-3*X_ast(1))+(i2/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))-(3/2)*
    delta(4)*delta(1)*exp(-3*X_ast(1))+delta(4)*delta(1)*exp(-2*
    X_ast(1))+(i2/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))-(3/2)*delta
    (4)*delta(1)*exp(-3*X_ast(1))+(i13/i3)*X_ast(2)*delta(1)*delta
    (1)*exp(-X_ast(1))+(3/2)*X_ast(3)*delta(1)*delta(1)*exp(-X_ast
    (1))+(i1/i6)*X_ast(4)*delta(1)*delta(1)*exp(-X_ast(1))-(i13/i3)*
    delta(2)*delta(1)*exp(-X_ast(1))-(i13/i3)*delta(2)*delta(1)*exp
    (-X_ast(1))-(3/2)*delta(3)*delta(1)*exp(-X_ast(1))-(3/2)*delta
    (3)*delta(1)*exp(-X_ast(1))-(i1/i6)*delta(4)*delta(1)*exp(-X_ast
    (1))-(i1/i6)*delta(4)*delta(1)*exp(-X_ast(1))+delta(4)*delta(1)*
    exp(-2*X_ast(1))+8*delta(3)*delta(1)*exp(-2*X_ast(1))-(i88/i3)*
    X_ast(2)*delta(1)*delta(1)*exp(-4*X_ast(1))+63*X_ast(2)*delta(1)
    *delta(1)*exp(-3*X_ast(1))-38*X_ast(2)*delta(1)*delta(1)*exp(-2*
    X_ast(1))-16*X_ast(3)*delta(1)*delta(1)*exp(-4*X_ast(1))+(63/2)*
    X_ast(3)*delta(1)*delta(1)*exp(-3*X_ast(1))-16*X_ast(3)*delta(1)
    *delta(1)*exp(-2*X_ast(1))-(i8/i3)*X_ast(4)*delta(1)*delta(1)*
    exp(-4*X_ast(1))+(9/2)*X_ast(4)*delta(1)*delta(1)*exp(-3*X_ast
    (1))-2*X_ast(4)*delta(1)*delta(1)*exp(-2*X_ast(1))+(i22/i3)*
    delta(2)*delta(1)*exp(-4*X_ast(1))-21*delta(2)*delta(1)*exp(-3*
    X_ast(1))+19*delta(2)*delta(1)*exp(-2*X_ast(1))+(i22/i3)*delta
```

```
(2)*delta(1)*exp(-4*X_ast(1))-21*delta(2)*delta(1)*exp(-3*X_ast
(1))+19*delta(2)*delta(1)*exp(-2*X_ast(1))-(21/2)*delta(3)*delta
(1)*exp(-3*X_ast(1))+4*delta(3)*delta(1)*exp(-4*X_ast(1))-(21/2)
*delta(3)*delta(1)*exp(-3*X_ast(1))+8*delta(3)*delta(1)*exp(-2*
X_ast(1));

DDF2 = -(i1/i6)*delta(1)*delta(1)*exp(-X_ast(1))+2*delta(1)*delta
(1)*exp(-2*X_ast(1))-16*delta(3)*delta(1)*exp(-4*X_ast(1))+(i8/
i3)*delta(1)*delta(1)*exp(-4*X_ast(1))-(9/2)*delta(1)*delta(1)*
exp(-3*X_ast(1))-(i8/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))
+(9/2)*delta(4)*delta(1)*exp(-3*X_ast(1))-2*delta(4)*delta(1)*
exp(-2*X_ast(1))-(i8/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))
+(9/2)*delta(4)*delta(1)*exp(-3*X_ast(1))-(i13/i3)*X_ast(2)*
delta(1)*delta(1)*exp(-X_ast(1))-(3/2)*X_ast(3)*delta(1)*delta
(1)*exp(-X_ast(1))-(i1/i6)*X_ast(4)*delta(1)*delta(1)*exp(-X_ast
(1))+(i13/i3)*delta(2)*delta(1)*exp(-X_ast(1))+(i13/i3)*delta(2)
*delta(1)*exp(-X_ast(1))+(3/2)*delta(3)*delta(1)*exp(-X_ast(1))
+(3/2)*delta(3)*delta(1)*exp(-X_ast(1))+(i1/i6)*delta(4)*delta
(1)*exp(-X_ast(1))+(i1/i6)*delta(4)*delta(1)*exp(-X_ast(1))-2*
delta(4)*delta(1)*exp(-2*X_ast(1))-16*delta(3)*delta(1)*exp(-2*
X_ast(1))+(i352/i3)*X_ast(2)*delta(1)*delta(1)*exp(-4*X_ast(1))
-189*X_ast(2)*delta(1)*delta(1)*exp(-3*X_ast(1))+76*X_ast(2)*
delta(1)*delta(1)*exp(-2*X_ast(1))+64*X_ast(3)*delta(1)*delta(1)
*exp(-4*X_ast(1))-(189/2)*X_ast(3)*delta(1)*delta(1)*exp(-3*
X_ast(1))+32*X_ast(3)*delta(1)*delta(1)*exp(-2*X_ast(1))+(i32/i3
)*X_ast(4)*delta(1)*delta(1)*exp(-4*X_ast(1))-(27/2)*X_ast(4)*
delta(1)*delta(1)*exp(-3*X_ast(1))+4*X_ast(4)*delta(1)*delta(1)*
exp(-2*X_ast(1))-(i88/i3)*delta(2)*delta(1)*exp(-4*X_ast(1))+63*
delta(2)*delta(1)*exp(-3*X_ast(1))-38*delta(2)*delta(1)*exp(-2*
X_ast(1))-(i88/i3)*delta(2)*delta(1)*exp(-4*X_ast(1))+63*delta
(2)*delta(1)*exp(-3*X_ast(1))-38*delta(2)*delta(1)*exp(-2*X_ast
(1))+(63/2)*delta(3)*delta(1)*exp(-3*X_ast(1))-16*delta(3)*delta
(1)*exp(-4*X_ast(1))+(63/2)*delta(3)*delta(1)*exp(-3*X_ast(1))
-16*delta(3)*delta(1)*exp(-2*X_ast(1));

DDF3 = (i1/i6)*delta(1)*delta(1)*exp(-X_ast(1))-4*delta(1)*delta(1)
*exp(-2*X_ast(1))+64*delta(3)*delta(1)*exp(-4*X_ast(1))-(i32/i3)
*delta(1)*delta(1)*exp(-4*X_ast(1))+(27/2)*delta(1)*delta(1)*exp
(-3*X_ast(1))+(i32/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))-(27/2)
```

```
    *delta(4)*delta(1)*exp(-3*X_ast(1))+4*delta(4)*delta(1)*exp(-2*
    X_ast(1))+(i32/i3)*delta(4)*delta(1)*exp(-4*X_ast(1))-(27/2)*
    delta(4)*delta(1)*exp(-3*X_ast(1))+(i13/i3)*X_ast(2)*delta(1)*
    delta(1)*exp(-X_ast(1))+(3/2)*X_ast(3)*delta(1)*delta(1)*exp(-
    X_ast(1))+(i1/i6)*X_ast(4)*delta(1)*delta(1)*exp(-X_ast(1))-(i13
    /i3)*delta(2)*delta(1)*exp(-X_ast(1))-(i13/i3)*delta(2)*delta(1)
    *exp(-X_ast(1))-(3/2)*delta(3)*delta(1)*exp(-X_ast(1))-(3/2)*
    delta(3)*delta(1)*exp(-X_ast(1))-(i1/i6)*delta(4)*delta(1)*exp(-
    X_ast(1))-(i1/i6)*delta(4)*delta(1)*exp(-X_ast(1))+4*delta(4)*
    delta(1)*exp(-2*X_ast(1))+32*delta(3)*delta(1)*exp(-2*X_ast(1))
    -(i1408/i3)*X_ast(2)*delta(1)*delta(1)*exp(-4*X_ast(1))+567*
    X_ast(2)*delta(1)*delta(1)*exp(-3*X_ast(1))-152*X_ast(2)*delta
    (1)*delta(1)*exp(-2*X_ast(1))-256*X_ast(3)*delta(1)*delta(1)*exp
    (-4*X_ast(1))+(567/2)*X_ast(3)*delta(1)*delta(1)*exp(-3*X_ast(1)
    )-64*X_ast(3)*delta(1)*delta(1)*exp(-2*X_ast(1))-(i128/i3)*X_ast
    (4)*delta(1)*delta(1)*exp(-4*X_ast(1))+(81/2)*X_ast(4)*delta(1)*
    delta(1)*exp(-3*X_ast(1))-8*X_ast(4)*delta(1)*delta(1)*exp(-2*
    X_ast(1))+(i352/i3)*delta(2)*delta(1)*exp(-4*X_ast(1))-189*delta
    (2)*delta(1)*exp(-3*X_ast(1))+76*delta(2)*delta(1)*exp(-2*X_ast
    (1))+(i352/i3)*delta(2)*delta(1)*exp(-4*X_ast(1))-189*delta(2)*
    delta(1)*exp(-3*X_ast(1))+76*delta(2)*delta(1)*exp(-2*X_ast(1))
    -(189/2)*delta(3)*delta(1)*exp(-3*X_ast(1))+64*delta(3)*delta(1)
    *exp(-4*X_ast(1))-(189/2)*delta(3)*delta(1)*exp(-3*X_ast(1))+32*
    delta(3)*delta(1)*exp(-2*X_ast(1));

DDF4 = -(i1/i6)*delta(1)*delta(1)*exp(-X_ast(1))+8*delta(1)*delta
    (1)*exp(-2*X_ast(1))-256*delta(3)*delta(1)*exp(-4*X_ast(1))+(
    i128/i3)*delta(1)*delta(1)*exp(-4*X_ast(1))-(81/2)*delta(1)*
    delta(1)*exp(-3*X_ast(1))-(i128/i3)*delta(4)*delta(1)*exp(-4*
    X_ast(1))+(81/2)*delta(4)*delta(1)*exp(-3*X_ast(1))-8*delta(4)*
    delta(1)*exp(-2*X_ast(1))-(i128/i3)*delta(4)*delta(1)*exp(-4*
    X_ast(1))+(81/2)*delta(4)*delta(1)*exp(-3*X_ast(1))-(i13/i3)*
    X_ast(2)*delta(1)*delta(1)*exp(-X_ast(1))-(3/2)*X_ast(3)*delta
    (1)*delta(1)*exp(-X_ast(1))-(i1/i6)*X_ast(4)*delta(1)*delta(1)*
    exp(-X_ast(1))+(i13/i3)*delta(2)*delta(1)*exp(-X_ast(1))+(i13/3)
    *delta(2)*delta(1)*exp(-X_ast(1))+(3/2)*delta(3)*delta(1)*exp(-
    X_ast(1))+(3/2)*delta(3)*delta(1)*exp(-X_ast(1))+(i1/i6)*delta
    (4)*delta(1)*exp(-X_ast(1))+(i1/i6)*delta(4)*delta(1)*exp(-X_ast
    (1))-8*delta(4)*delta(1)*exp(-2*X_ast(1))-64*delta(3)*delta(1)*
```

```
    exp(-2*X_ast(1))+(i5632/i3)*X_ast(2)*delta(1)*delta(1)*exp(-4*
    X_ast(1))-1701*X_ast(2)*delta(1)*delta(1)*exp(-3*X_ast(1))+304*
    X_ast(2)*delta(1)*delta(1)*exp(-2*X_ast(1))+1024*X_ast(3)*delta
    (1)*delta(1)*exp(-4*X_ast(1))-(1701/2)*X_ast(3)*delta(1)*delta
    (1)*exp(-3*X_ast(1))+128*X_ast(3)*delta(1)*delta(1)*exp(-2*X_ast
    (1))+(i512/i3)*X_ast(4)*delta(1)*delta(1)*exp(-4*X_ast(1))
    -(243/2)*X_ast(4)*delta(1)*delta(1)*exp(-3*X_ast(1))+16*X_ast(4)
    *delta(1)*delta(1)*exp(-2*X_ast(1))-(i1408/i3)*delta(2)*delta(1)
    *exp(-4*X_ast(1))+567*delta(2)*delta(1)*exp(-3*X_ast(1))-152*
    delta(2)*delta(1)*exp(-2*X_ast(1))-(i1408/i3)*delta(2)*delta(1)*
    exp(-4*X_ast(1))+567*delta(2)*delta(1)*exp(-3*X_ast(1))-152*
    delta(2)*delta(1)*exp(-2*X_ast(1))+(567/2)*delta(3)*delta(1)*exp
    (-3*X_ast(1))-256*delta(3)*delta(1)*exp(-4*X_ast(1))+(567/2)*
    delta(3)*delta(1)*exp(-3*X_ast(1))-64*delta(3)*delta(1)*exp(-2*
    X_ast(1));

DDF = [DDF1;DDF2;DDF3;DDF4];

We define the bound Z^2(r).

Z2 = abs(DDF);


%%%%%%%%%%%%% Definition of the Radii Polynomial

P=[sup(Z2)  -ones(4,1)  sup(Y)];

r1=sort(roots(P(1,:)));
r2=sort(roots(P(2,:)));
r3=sort(roots(P(3,:)));
r4=sort(roots(P(4,:)));

I=[max(r1(1),max(r2(1),max(r3(1),r4(1))))  min(r1(2), min(r2(2),min
    (r3(2),r4(2))))];

I=[1.01*I(1)  0.99*I(2)];

success=1;
```

```
% We make sure that the radii polynomials are negative

Y=intval(sup(Y)); Z2=intval(sup(Z2));

r_minus=intval(I(1));
r_plus=intval(I(2));

p1_minus=Z2(1)*r_minus^2-r_minus+Y(1);
p1_plus=Z2(1)*r_plus^2-r_plus+Y(1);
p2_minus=Z2(2)*r_minus^2-r_minus+Y(2);
p2_plus=Z2(2)*r_plus^2-r_plus+Y(2);

if sup(p1_minus)>=0
    success=0;
    display('p1_minus>0')
end

if sup(p1_plus)>=0
    success=0;
    display('p1_plus>0')
end

if sup(p2_minus)>=0
    success=0;
    display('p2_minus>0')
end

if sup(p2_plus)>=0
    success=0;
    display('p2_plus>0')
end

if r_plus>=r_star
   I(2)=r_star;
end

end
```