Personal Research Project Report

Quadratization in the reachability problem for ODEs

Author: Yubo Cai

Institute: Laboratoire d'informatique de l'École polytechnique (LIX) - Team MAX (Joint Team with Centre national de la recherche scientifique CNRS)

Date: June 2nd, 2023

Supervisor: Gleb Pogudin

Contents

Abstract and Preface

I would like to thank the kind professor Gleb Pogudin for hosting me during this project in the MAX team at the Laboratoire d'informatique de l'École Polytechnique (LIX) and Centre national de la recherche scientifique (CNRS) the Spring of 2023. I am grateful for the guidance from my supervisor Prof. Gleb Pogudin on my research direction and the ideas I received when I encountered a bottleneck in my research. This research project was written as part of the Personal Research Project (LAB251 PRL) at École Polytechnique which I chose to explore Numerical ODE as my field of interest.

One of the fundamental problems for dynamical systems is the reachability problem. The problem is: given a dynamical system (defined by a system of differential equations) and some range of the initial conditions, find a (rigorous and guaranteed) bound for the state of the system at some other time. This is an important ingredient, for example, of many approaches to the verification of dynamical systems. The problem is very well studied for linear systems but is much less understood for the nonlinear case. Forets and Schilling [4] proposes to reduce the nonlinear case to the linear one by using Carleman linearization. The approach presented in the paper requires the system to have at most quadratic nonlinearities. On the other hand, Gleb Pogudin, the supervisor of the project, and his colleagues have recently designed an algorithm for transforming any system into such an at most quadratic form [1].

This project considers specific systems with high-degree nonlinearities and applies the composition of the algorithms described above to solve the reachability problem. The main challenge is to adjust the transformation at most quadratic systems in a way that would make the subsequent Carleman linearization as numerically stable as possible. In the project, we try to first explore 3 typical ODE systems with different characteristics, and then we prove that for any dissipative polynomial ODEs system, there exists a quadratization such that the quadratized system is dissipative as well. Based on this theorem, we later derive that if all the new variables introduced in quadratization can be quadratized by other variables, then the dissipativity of the resulting quadratic system can be ensured.

In Chapter 4, We reapply the conclusions of Chapter 3 to our three examples for verification. In subsequent work, we hope to extend the package of optimal quadratization to find an algorithm that will make the new quadratic system dissipative while introducing as few variables as possible.

Codes and Links

1. Github Repository for computation and verification of the report: link.
2. Qbee Package: link.
3. Code for Reachability Analysis: link.

Chapter 1 Introduction

In this chapter, we will introduce the concept of the **quadratization** problem and the approach to reachability based on **Carleman linearization** and **Taylor-model** which reduce the nonlinear case to the linear. At the same time, we will introduce the equation that satisfies the condition of stability properties and the associated error bound function based on this approach.

Since the Carleman linearization technique for weakly nonlinear systems relies on quadratic terms, which restricts the highest degree of the right-hand side to at most 2, we can generalize this theorem to higher-dimensional nonlinear ODE systems using the method of quadratization[1]. Therefore, we can establish a bridge between these two problems to investigate whether the quadratization algorithm can meet the stability requirements for the reachability of weakly nonlinear systems using Carleman linearization, and attempt to find the optimal quadratization approach.

1.1 Optimal quadratization of ODE system

The quadratization problem involves transforming a given system of ordinary differential equations (ODEs) with a polynomial right-hand side into a system in which the highest degree of the right-hand side is reduced to 2. Here is a more formal definition [1]:

Definition 1.1 (Quadratization of ODE)

Consider a system of ODEs

$$\begin{cases} x_1' = f_1(\bar{x}) \\ \dots \\ x_n' = f_n(\bar{x}) \end{cases} \tag{1}$$

where $\bar{x} = (x_1, \dots, x_n)$ *and* $f_1, \dots, f_n \in \mathbb{C}[\mathbf{x}]$. *Then a list of new variables*

$$y_1 = g_1(\bar{x}), \dots, y_m = g_m(\bar{x}) \tag{1.1}$$

is said to be a quadratization of (1) if there exist polynomials $h_1, \dots, h_{m+n} \in \mathbb{C}[\bar{x}, \bar{y}]$ *of degree at most two such that*

- $x_i' = h_i(\bar{x}, \bar{y})$ *for every* $1 \leqslant i \leqslant n$
- $y_j' = h_{j+n}(\bar{x}, \bar{y})$ *for every* $1 \leqslant j \leqslant m$

Here we call the number m *the* ***order of quadratization****. The* ***optimal quadratization*** *is a quadratization of smallest possible order.* ♣

Example 1.1 We illustrate the problem with a simple example of a scalar ODE:

$$x' = x^3 \tag{1.2}$$

Since on the right-hand side, the highest degree is 3, we can introduce a new variable of $y = x^n$ where $n = 2, 3$. We introduce a new variable $y := x^2$, then we can write

$$x' = xy \quad \text{and} \quad y' = 2xx' = 2x^4 = 2y^2$$

[1]This approach is feasible because it has been shown that all polynomial systems can be transformed into systems with the highest dimension of 2 [2].

Now, we can build a new ODEs system as:

$$\begin{cases} x' = xy \\ y' = 2y^2 \end{cases} \tag{1.3}$$

Definition 1.2 (Monomial quadratization)

*If all the polynomials $g_1, \dots, g_m$ are **monomials**, the quadratization is called a monomial quadratization. If a monomial quadratization of a system has the smallest possible order among all the monomial quadratizations of the system, it is called optimal monomial quadratization.* ♣

Remark Note that there can be multiple cases for the optimal quadratization of a system of ODEs. For example, both transformations of $y = x^2$ and $y = 3x^2$ are optimal quadratizations of equation (1.2).

Carothers' work has demonstrated that for all polynomial systems of ODEs , there exists a way to transform the system of ODEs into the highest quadratic term by introducing new variables[2] [2]. This means that quadratization can be performed for any ODEs system which is the theoretical cornerstone of this research work.

1.2 Reachability analysis of weakly nonlinear systems based on Carleman linearization

Thanks to the work of Forets and Schilling [4], we now have a new approach to reachbaility based on Carleman linearization. This approach involves transforming the nonlinear system into an infinite-dimensional linear system, which is then truncated to approximate the original system. Here we focus on the problem of reachability analysis of weekly nonlinear systems. Based on this paper, we have the following definitions and theorems.

Definition 1.3 (Kronecker product)

*For any pair of vectors $x \in \mathbb{R}^n, y \in \mathbb{R}^m$, their **Kronecker product** is $w = x \otimes y = (x_1y_1, \dots, x_1y_m, x_2y_1, \dots, x_ny_m)^T$, and the dimension is $\dim(w) = mn$. This product is not commutative. For matrices the definition is analogous: if $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$, then $A \otimes B \in \mathbb{R}^{mp\times nq}$ and*

$$A \otimes B := \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

*The **Kronecker power** $x^{\otimes i}$ of $x \in \mathbb{R}^n$ is defined as:*

$$x^{\otimes i} := \underbrace{x \otimes \dots \otimes x}_{i \text{ times}}$$

♣

Example 1.2 Assume we have $\dim(x) = 2$ where $x = [x_1, x_2]^T$. Then we have $x^{\otimes 2} = x \otimes x = [x_1^2, x_1x_2, x_2x_1, x_2^2]^T$

[2] Proof in Theorem 1 of the paper

Definition 1.4 (Represenation of quadratic polynomial DE with Kronecker product)

For a ***quadratic polynomial differential equation*** $x'(t) = f(x(t))$, *we can represent such a ODE as the following format*

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = F_1 x + F_2 x^{\otimes 2} \tag{1.4}$$

with initial condition $x(0) \in \mathbb{R}^n$ *and* $F_1 \in \mathbb{R}^{n\times n}$ *and* $F_2 \in \mathbb{R}^{n\times n^2}$. F_1 *and* F_2 *are independent of* t *and* F_1 *is associated with* ***linear*** *behavior of the dynamical system and* F_2 *is the* ***nonlinear*** *part.* ♣

Remark Here we focus on the weakly nonlinear system, therefore $\frac{\|F_2\|_2}{\|F_1\|_2}$ should be small since we want the linear part dominate the system of ODEs.

Example 1.3 Assume we have a nonlinear system of ODEs:

$$\begin{cases} x' = ax + by + cxy + dx^2 \\ y' = mx + ny + ky^2 \end{cases} \tag{1.5}$$

For system (1.5), we have the linear part matrix F_1 and nonlinear part matrix F_2

$$F_1 = \begin{array}{c c} & \begin{array}{cc} x & y \end{array} \\ \begin{array}{c} x' \\ y' \end{array} & \begin{pmatrix} a & b \\ m & n \end{pmatrix} \end{array} \quad \text{and} \quad F_2 = \begin{array}{c c} & \begin{array}{cccc} x^2 & xy & yx & y^2 \end{array} \\ \begin{array}{c} x' \\ y' \end{array} & \begin{pmatrix} d & c & 0 & 0 \\ 0 & 0 & 0 & k \end{pmatrix} \end{array} \tag{1.6}$$

Definition 1.5 (Weakly nonlinear system)

The system is said to be weakly nonlinear if the ratio

$$R := \frac{\|X_0\|\|F_2\|}{|\Re(\lambda_1)|} \tag{1.7}$$

satisfies $R < 1$. *Here* λ_1 *is the eigenvalue of* F_1 *with* ***the largest*** *real part, which means if* F_1 *has* n *eigenvalues, we have* $\Re(\lambda_n) \le \cdots \le \Re(\lambda_1)$. ♣

Remark The norm here can be generally any norm because all norms are equivalent up to a constant. Based on the code [3] here we use **Euclidean norm** and **operator norm** for computation [5].

Definition 1.6 (Dissipativity)

System (1.4) *is said to be dissipative if* $\Re(\lambda_1) < 0$ *(i.e., the real part of all eigenvalues is negative).*

$$\text{The conditions } \Re(\lambda_1) < 0 \text{ and } R < 1 \text{ ensure arbitrary-time convergence.} \tag{1.8}$$

♣

Theorem 1.1 (Error bound function)

Assuming that (1.4) *is weakly nonlinear and dissipative, the error bound associated with the linearized problem truncated at order* N *satisfies*

$$\|\eta_1(t)\| \le \varepsilon(t) := \|X_0\| R^N \left(1 - e^{\Re(\lambda_1)t}\right)^N,$$

with R *as defined in* (1.7). *This error bound holds for all* $t \ge 0$. ♡

[3] link to detailed code

Definition 1.7 (Euclidean norm and operator norm)

The vector 2-norm [6] $\|\cdot\|_2 : \mathbb{C}^m \to \mathbb{P}$ *defined for* $x \in \mathbb{C}^m$ *by*

$$\|x\|_2 = \sqrt{|x_0|^2 + \cdots + |x_{m-1}|^2} = \sqrt{\sum_{i=0}^{m-1} |x_i|^2}$$

Equivalently, it can be defined by

$$\|x\|_2 = \sqrt{x^H x}$$

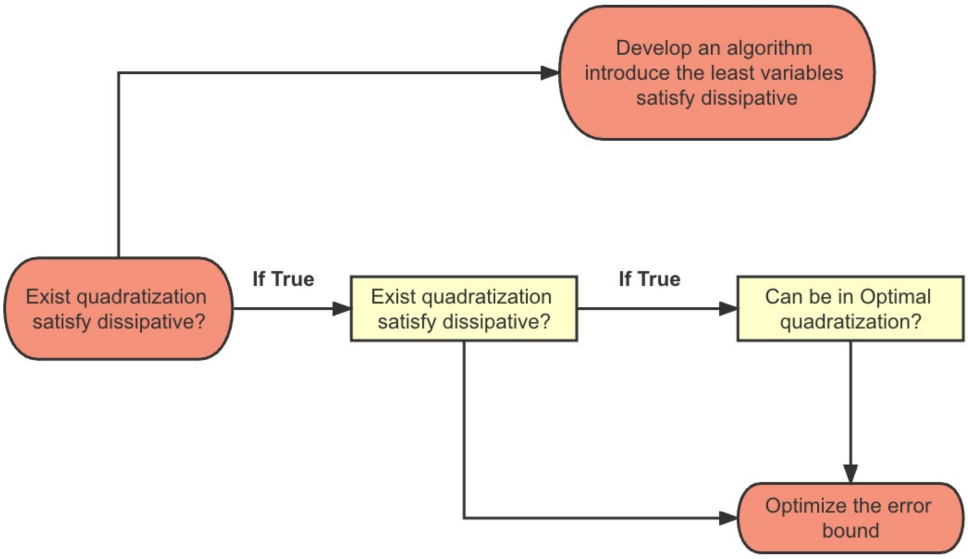
The operator norm of a matrix $A \in \mathbb{R}^{m,n}$ *is the square root of the largest eigenvalue of the symmetric matrix* $A^T A$. ♣

1.3 Connection and application between two problems and algorithms

For a high-dimensional nonlinear ODE system, our idea is to reduce the ODE system to a quadratic system through the optimal quadratization algorithm. However, there are multiple quadratization methods for an ODE system, and we hope to explore the impact of different quadratization methods on system reachability. We consider this problem based on five objectives for an ODE system:

1. Whether there exists a quadratization such that the resulting quadratic system is dissipative.
2. Whether there exists a quadratization such that the resulting quadratic system is weakly nonlinear.
3. Whether there exists an optimal quadratization that satisfies the conditions of being weakly nonlinear and dissipative.
4. If yes, we can find an optimal quadratization with a minimum error bound through optimization, which can be expressed by either analytical or numerical solutions.
5. Otherwise, we hope to develop an algorithm of optimal quadratization to achieve dissipativity

We aim to satisfy the first objective by finding an optimal quadratization for the system that meets the conditions of being weakly nonlinear and dissipative. After that, we will try to use optimization methods to find the possibility of achieving the minimum error bound. We will explore this problem by studying some typical and representative systems.



Chapter 2 Research of 3 examples of typical systems of ODEs

In this chapter, we research on three typical differential equations with different characteristics.

1. $x' = -x + ax^3$
2. $x' = -x^3$
3. $x'' = kx + ax^3 + bx'$

In the first ODE, we have already a negative linear term -1 but for the second one we only have the nonliner term. We want to see whether we can quadratic it into a **weakly nonlinear** system. The third system is **Duffing equation** [8], We aim at investigating if there exists a general quadratization algorithm for more complex nonlinear ODE systems that enables the system to satisfy the conditions required for reachability analysis.

2.1 First system with known negative eigenvalue

Consider the system of ODEs

$$x' = -x + ax^3 \tag{2.1}$$

It is not difficult to see that the optimal quadratization is of order 2 (This also can be verified in **Qbee**[1]). Intuitively, we have quadratization by introducing a new variable $y = x^2$, then we got the following quadratic system:

$$\begin{cases} x' = -x + axy \\ y' = -2y + 2ay^2 \end{cases} \tag{2.2}$$

In this system, we can observe that all the eigenvalues of the linear part are negative, which meets the requirement of dissipativity. However, since a is a constant in the initial condition, we may not be able to satisfy the condition of weak nonlinearity. Hence, we need to change our quadratization in order to make the system weakly nonlinear.

2.1.1 Construction of the system

Since we know the optimal quadratization is of order 2, we can quadratize the system by introducing $y = \alpha x^2 - \beta x - \gamma$ where $\alpha, \beta, \gamma \in \mathbb{R}$. Here we take a new variable $w_0 = y = \alpha x^2$. We have our new system that:

$$\begin{cases} x' = -x + \frac{a}{\alpha}xy \\ y' = -2y + \frac{2a}{\alpha}y^2 \end{cases} \tag{2.3}$$

Remark We also tried with quadratization of $y = x^2 - mx$ and $y = x^2 - n$. However, the quadratization approach with $y = x^2 - mx$ is not stable with dissipativity but $y = x^2 - n$ works. Since after quadratization with $y = x^2 - n$, we have $x' = (an - 1)x + axy$ which should have same behavior with x' in (2.3) since both of them only have the term in x and xy and all the coefficients are changable, which means we can transform between $x' = (an - 1)x + axy$ and $x' = -x + \frac{a}{\alpha}xy$.

From previous analysis, we have the $F_1 \in \mathbb{R}^{n\times n}$ matrix for linear part:

$$F_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}.$$

[1] Qbee Package: https://github.com/AndreyBychkov/QBee/tree/master

And for the nonlinear part we have $F_2 \in \mathbb{R}^{n \times n^2}$ matrix:

$$F_2 = \begin{bmatrix} 0 & \frac{2a}{\alpha} & 0 & 0 \\ 0 & 0 & 0 & \frac{a}{\alpha} \end{bmatrix}.$$

For matrix F_1 we have the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$. Then we got $\Re(\lambda_1) = -1$ (the real part of λ_1). We can see that this new system is dissipative. We need to make the system is weakly nonlinear as well. Base on formula (1.7), we have

$$R := \frac{\|X_0\| \|F_2\|}{|\Re(\lambda_1)|}$$

Here we initialize the case that $X_0 = [x_0, y_0] = [x_0, \alpha x_0^2]$ where x_0 is the initial value of x and $y_0 = \alpha x_0^2$ and we set $x_0, \alpha, a > 0$. We have

$$\|X_0\| = \|[x_0, \alpha x_0^2]\| = \sqrt{x_0^2 + \alpha^2 x_0^4} = x_0 \sqrt{1 + \alpha^2 x_0^2} \quad \text{and} \quad \|F_2\| = 2\frac{a}{\alpha}$$

Then we can compute the value of R:

$$R := \frac{\|X_0\| \|F_2\|}{|\Re(\lambda_1)|} = x_0 \sqrt{1 + \alpha^2 x_0^2} \cdot 2\frac{a}{\alpha} < 1$$

Then we have the boundary condition that $\alpha > \frac{2ax_0}{\sqrt{1 - 4a^2 x_0^4}}$ and $ax_0^2 < \frac{1}{2}$, we can compute the error bound:

$$\begin{aligned} \|\eta_1(t)\| \leq \varepsilon(t) &:= \|X_0\| R^N \left(1 - e^{\Re(\lambda_1)t}\right)^N \\ &= \|X_0\|^{N+1} (2\frac{a}{\alpha})^N \left(1 - e^{-t}\right)^N \\ &= x_0^{N+1} (1 + \alpha^2 x_0^2)^{\frac{N+1}{2}} \frac{(2a)^N}{\alpha^N} \left(1 - e^{-t}\right)^N \end{aligned}$$

Since α is our parameter, then we only need to analyze the function $(1 + \alpha^2 x_0^2)^{\frac{N+1}{2}} \frac{1}{\alpha^N}$.

2.1.2 Optimization of the error bound function

We try to use optimization methods to find the α to make our error-bound function as small as possible, we set a function based on variable α:

$$f(\alpha) = \frac{(1 + \alpha^2 x_0^2)^{\frac{N+1}{2}}}{\alpha^N}$$

We first take the logarithm of the error equation, we have:

$$\log f(\alpha) = \frac{N+1}{2} \log(1 + \alpha^2 x_0^2) - N \log(\alpha)$$

We take the derivative of the logarithm of the error equation, we have:

$$\frac{d \log f(\alpha)}{d\alpha} = \frac{N+1}{2} \frac{2\alpha x_0^2}{1 + \alpha^2 x_0^2} - \frac{N}{\alpha} = \frac{(N+1) x_0^2 \alpha}{1 + \alpha^2 x_0^2} - \frac{N}{\alpha}$$

Since $\log$ is an increasing function, therefore, in order to find the local minimum of the error equation, we need to find the point where the derivative of the logarithm of the error equation is zero, i.e.

$$\frac{d \log f(\alpha)}{d\alpha} = \frac{(N+1) x_0^2 \alpha}{1 + \alpha^2 x_0^2} - \frac{N}{\alpha} = 0$$

Then we have:

$$\alpha = \sqrt{\frac{N}{x_0^2}}$$

Then, we find the optimal α for quadratization of the first system only depends on the order of Carleman linearization (N) and initial condition x_0. Since we have the condition that $ax_0^2 < \frac{1}{2}$ from the initial value, the optimal value of α exists since the boundary condition can be always satisfied with the optimal value.

2.1.3 Testing for Analytic Expression with Numerical Method

We need to verify our analytic solution and plot to see whether is true for our example in Julia code. We verify that our analytical solution is mathematically correct in numerical_optimi.ipynb. However, in the file quadra_reachability.jl we verify with different N and x_0 separately and see the plot of the error bound in reachability analysis. We choose the following cases

1. $N = 2$, $x_0 = 0.5$, with optimal choice $\alpha = \frac{\sqrt{N}}{x_0} = 2\sqrt{2}$
2. $N = 3$, $x_0 = 0.5$, with optimal choice $\alpha = \frac{\sqrt{N}}{x_0} = 2\sqrt{3}$

In order to see the result of the plot, you can find it in the file startup.ipynb and you may need to download **Julia** and some packages of reachability. You can see more introduction for initialization in the GitHub repository [2]. Here, we have our plot result:

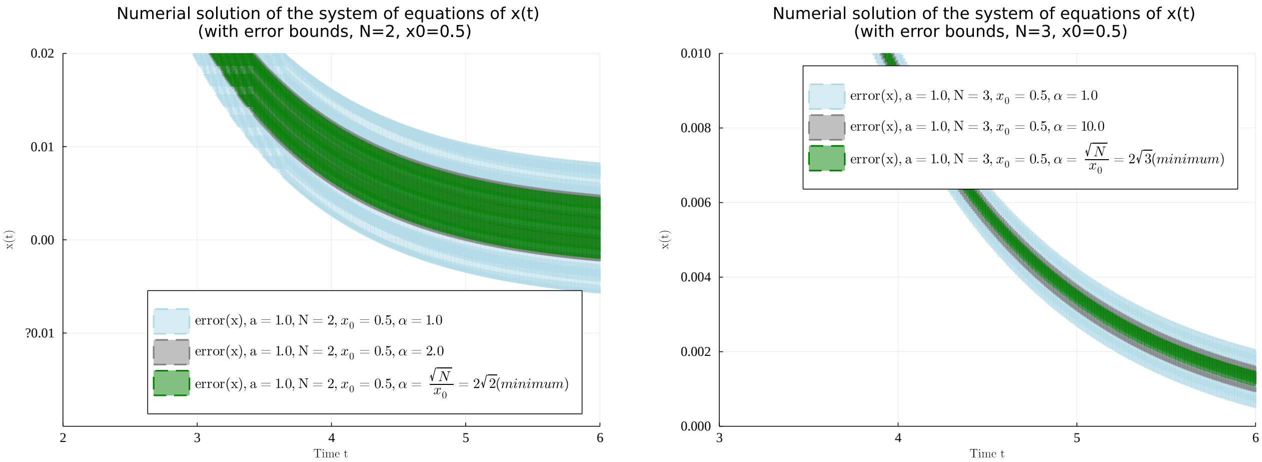


Figure 2.1: Plots with different N and x_0

We can clearly see from the graph that the green line, which is our optimal choice from the analytic solution, has the smallest error bound. This verifies our conclusion. Then we can conclude for our first model that the optimal quadratization is $y = x^2 + \frac{\sqrt{N}}{x_0}$. Therefore, we have the following system that

$$\begin{cases} x' = -x + \frac{ax_0}{\sqrt{N}} xy \\ y' = -2y + \frac{2ax_0}{\sqrt{N}} y^2 \end{cases}$$

2.2 Second system with no negative eigenvalue of the linear part

We have the second system of the following form:

$$x' = -x^3$$

Normally, we have the following quadratization, where $y = x^2$:

$$y' = -2xx' = -2x^4 = -2y^2$$

[2] https://github.com/yubocai-poly/Reachability-Problem-of-ODE/tree/main

Then we have the following system:

$$\begin{cases} y' = -2y^2 \\ x' = -xy \end{cases}$$

2.2.1 Construction of the system

However, we can find that **there is no linear part** in the system. Therefore, we try to add linear terms like ax or a into quadratization. After computation, we find that ax can't produce eigenvalues for matrix F_1. Therefore, we try with the term a. We can change $y = x^2 - a \Rightarrow a = x^2 - y$ (**this quadratization is not monomial**), then we have the following system:

$$\begin{cases} y' = -2y^2 \ - 2ax^2 - 2ay \\ x' = -xy - ax \end{cases}$$

Then we have the $F_1 \in \mathbb{R}^{n \times n}$ matrix for linear part:

$$F_1 = \begin{array}{c c} & \begin{array}{cc} x & y \end{array} \\ \begin{array}{c} x' \\ y' \end{array} & \begin{pmatrix} -a & 0 \\ 0 & -2a \end{pmatrix} \end{array} \tag{2.4}$$

And for the nonlinear part we have $F_2 \in \mathbb{R}^{n \times n^2}$ matrix:

$$F_2 = \begin{array}{c c} & \begin{array}{cccc} x^2 & xy & yx & y^2 \end{array} \\ \begin{array}{c} x' \\ y' \end{array} & \begin{pmatrix} 0 & -1 & 0 & 0 \\ -2a & 0 & 0 & -2 \end{pmatrix} \end{array} \tag{2.5}$$

For matrix F_1 we have the eigenvalues $\lambda_1 = -a$ and $\lambda_2 = -2a$. Then we got $\Re(\lambda_1) = -a$ (the real part of λ_1). Also, we need to compute the operator norm of F_2, we have

$$A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 4a^2 + 4 \end{bmatrix}.$$

Therefore, we got $\|F_2\|_2 = \sqrt{4 + 4a^2}$ and the weakly nonlinear condition that

$$\begin{aligned} R &:= \frac{\|X_0\| \|F_2\|}{|\Re(\lambda_1)|} \\ &= \frac{2\sqrt{a^2+1}\sqrt{x_0^4 + (1-2a)x_0^2 + a^2}}{a} \\ &= 2\sqrt{1 + \frac{1}{a^2}}\sqrt{(x_0^2 - a)^2 + x_0^2} \end{aligned}$$

However, we find that the weakly nonlinear condition can not be satisfied (here we apply the Broyden–Fletcher–Goldfarb–Shanno algorithm [3] to find the optimal value of R). From our optimization result, we find that $R \geq 2$ no matter what value chose for a and x_0. We can also see this in our plot [3]

Then, we can deduce that after quadratization for the ODE $x' = -x^3$, the boundary condition (R) can not be satisfied therefore we are not able to apply the Carleman linearization to the system.

[3] https://github.com/yubocai-poly/Reachability-Problem-of-ODE/blob/main/Reachability/evaluation/Nonlinear_example /nonlinear_sample_numerical.ipynb

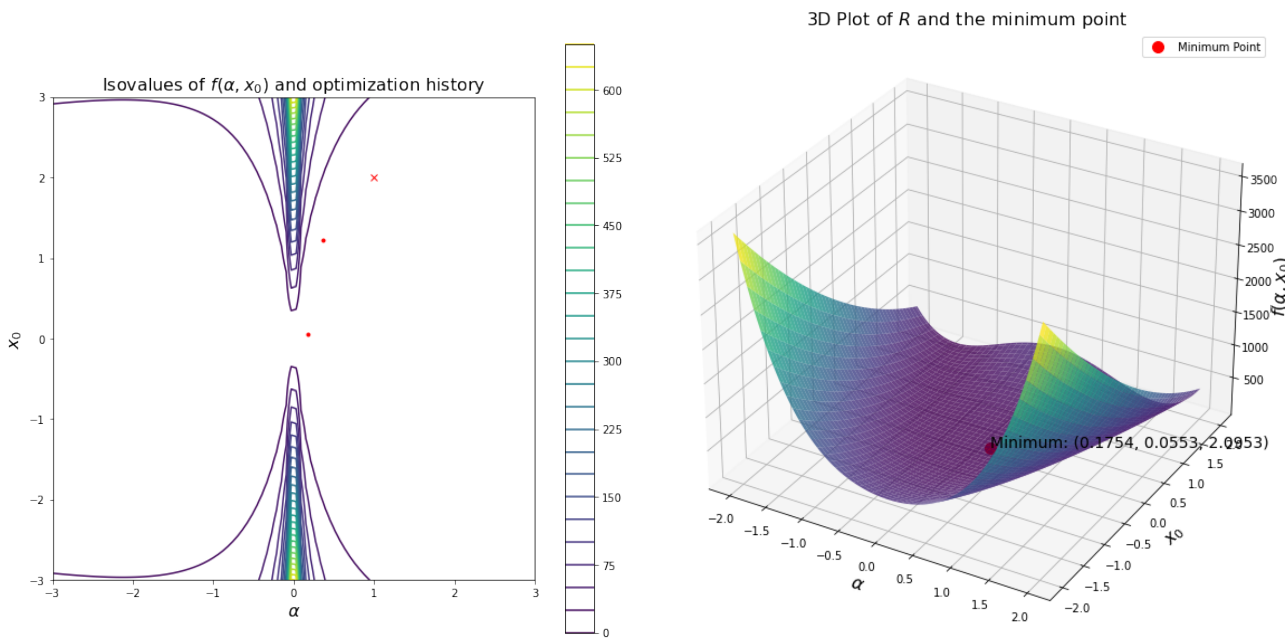


Figure 2.2: Plot for $R = 2\sqrt{1+\frac{1}{a^2}}\sqrt{(x_0^2-a)^2+x_0^2}$, we can see that the minimum of R is larger than 2

2.2.2 Computation for the case with quadratizaiton operator ax

We denote the quadratization method $y = x^2 - ax \Rightarrow ax = x^2 - y$, then we have the following system:

$$x' = -x^3 = -x(ax+y) = -ax^2 + xy = -a(ax+y) + xy = -a^2x - ay + xy$$

$$\begin{aligned} y' &= 2xx' = -2x^4 \\ &= -2(a^2x^2 + y^2 + 2axy) \quad \text{(Here we still not have linear part)} \\ &= -2(a^2y + a^3x + y^2 + 2axy) \end{aligned}$$

Then we can finalize the system as follows:

$$\begin{cases} y' = -2(a^2y + a^3x + y^2 + 2axy) \\ x' = -a^2x - ay + xy \end{cases}$$

Then we have the $F_1 \in \mathbb{R}^{n\times n}$ matrix for linear part:

$$\begin{bmatrix} -a^2 & -a \\ -2a^3 & -2a^2 \end{bmatrix}$$

However, we can still see that the first and second rows of the matrix F_1 are colinear, then we can't compute a negative eigenvalue. This is due to equation y', we don't have the linear part, which we need to substitute x^2 into $ax + y$. This transformation is the same as the transformation method in the equation of x' and causes linearity.

2.2.3 Analysis of the failure of Carleman linearization of the equation $x' = -x^3$

Since after we quadratic a system, the highest degree of the ODE is 2. Also, the operator of the quadratization is $[c, x^1, \cdots, x^n]$ where n is the highest degree of the variable that we introduced for quadratization. In this example, we don't have any linear part in ODE, therefore the nonlinear part dominates the behavior of ODE, and we can conjecture that if a nonlinear system has no linear part, we cannot generate a quadratic system with a linear part, no matter how we perform quadratization.

2.3 Third system of Duffing Equation

The **Duffing equation** is a non-linear second-order differential equation used to model certain damped and driven oscillators [8].

$$\ddot{x} + \delta\dot{x} + \alpha x + \beta x^3 = \gamma\cos(\omega t),$$

where the (unknown) function $x = x(t)$ is the displacement at time t, $\dot{x}$ is the first derivative of x with respect to time, i.e. **velocity**, and $\ddot{x}$ is the second time-derivative of x, i.e. **acceleration**. The numbers $\delta, \alpha, \beta, \gamma$ and ω are given constants. Here, we set $\gamma = 0$ which we don't consider the effect of the period function on the system. We have the following form:

$$x'' = kx + ax^3 + bx'$$

where $k < 0$ and $b < 0$. (we set $x_0 = x$ and $x_1 = x'$). Then we have the system of the ODEs:

$$\begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0^3 + bx_1 \end{cases}$$

Here we have the matrix to represent the linear part of the system:

$$F_1 = \begin{bmatrix} 0 & 1 \\ k & b \end{bmatrix}$$

From **Routh-Hurwitz Criterion** [7], we know F_1 has 2 negative eigenvalue $\lambda_1 = \frac{b}{2} + \sqrt{k + \frac{b2}{4}}, \lambda_2 = \frac{b}{2} - \sqrt{k + \frac{b2}{4}} < 0$. By introduce one new variable $w_0 = x_0^2$. Then we got the system of the ODEs:

$$\begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0w_0 + bx_1 \\ w_0' = 2x_0x_0' = 2x_0x_1 \\ \quad\;\; = -x_0^2 + x_0^2 + 2x_0x_1 \\ \quad\;\; = -w_0 + x_0^2 + 2x_0x_1 \end{cases}$$

For the new quadratic system, we have all 3 eigenvalues $\lambda_1 = \frac{b}{2} + \sqrt{k + \frac{b2}{4}}, \lambda_2 = \frac{b}{2} - \sqrt{k + \frac{b2}{4}}, \lambda_3 = -1 < 0$. Therefore, the new ODE system is dissipative and we have the matrix F_1 and F_2 that

$$F_1 = \begin{bmatrix} 0 & 1 & 0 \\ k & b & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad F_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Here we have the ratio R that

$$\begin{aligned} R &:= \frac{\|x_0\|\,\|F_2\|}{|\Re(\lambda_1)|} \\ &\leq \sqrt{x_0^2 + x_1^2 + x_0^4} \cdot \max(2, |a|) \end{aligned}$$

Here, we can see that there exist a quadratization method that makes the system dissipative. However, the initial condition and the coefficient constant decide whether the quadratization is weakly nonlinear or not.

Chapter 3 Expansion to the general case

3.1 Theorem of the arbitrary polynomial system of ODEs

Theorem 3.1

Consider a dissipative polynomial ODEs system. There exists a monomial quadratization such that the quadratized system is dissipative as well. ♡

Proof Let n and d denote the dimension and the degree of the system, respectively. Let the variables be $x_1, \dots, x_n$. Then the system is of the form:

$$\begin{cases} x_1' = \alpha_1^{(1)} x_1 + \cdots + \alpha_n^{(1)} x_n + \sum_{i_1+\cdots+i_n \le n} m_{i_1,\cdots,i_n}^{(1)} x_n^{i_1} \cdots x_n^{i_n} \\ \cdots \\ x_n' = \alpha_1^{(n)} x_1 + \cdots + \alpha_n^{(n)} x_n + \sum_{i_1+\cdots+i_n \le n} m_{i_1,\cdots,i_n}^{(n)} x_n^{i_1} \cdots x_n^{i_n} \end{cases} \tag{3.1}$$

where $m_{i_1,\cdots,i_n}^{(j)}$ means the coefficient of $x_1^{i_1} \cdots x_n^{i_n}$ in the j-th ODE and we assume that all eigenvalues have a negative real part in the linear part. We have the matrix $F_1 \in \mathbb{R}^{n \times n}$ that

$$F_1 = \begin{bmatrix} \alpha_1^{(1)} & \cdots & \alpha_n^{(1)} \\ \vdots & \ddots & \vdots \\ \alpha_1^{(n)} & \cdots & \alpha_n^{(n)} \end{bmatrix}$$

In order to quadratize the polynomial system, we can introduce new variables in the form (**here we use monomial quadratization**)

$$v_{i_1,\dots,i_n} = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \quad \text{where } \sum_{j=1}^{n} i_j \le d. \tag{3.2}$$

According to the chain rule, we got

$$v'_{i_1,\dots,i_n} = \sum_{j=0}^{n} i_j v_{i_1,\dots,i_j-1,\dots,i_n} x_j' \tag{3.3}$$

By plugging in our newly introduced variables, we can quadratize (3.1) with the format

$$\begin{cases} x_1' = \alpha_1^{(1)} x_1 + \cdots + \alpha_n^{(1)} x_n + \text{sum of quadratic terms} \\ \cdots \\ x_n' = \alpha_1^{(n)} x_1 + \cdots + \alpha_n^{(n)} x_n + \text{sum of quadratic terms} \end{cases}$$

Next, we will show that for each new variable v, the derivative v' can be written as

$$v'_{i_1,\dots,i_n} = -\alpha v_{i_1,\dots,i_n} + \text{sum of quadratic terms} \quad \text{where } \alpha > 0.$$

Since we have the degree of $v_{i_1,\dots,i_n} \le n$, the highest degree of the right-hand side is less or equal to $2n - 1$. Therefore, each term on the right-hand of equation (3.3) can be written as a product of two variables. The other thing we need to consider is how to produce the term $-\alpha v_{i_1,\dots,i_n}$. From equation (3.3), we got

$$\begin{aligned}
v'_{i_1,\dots,i_n} &= \sum_{j=0}^{n} i_j v_{i_1,\dots,i_j-1,\dots,i_n} x'_j \\
&= \underbrace{-\alpha v_{i_1,\dots,i_n}}_{\text{keep it linear}} + \underbrace{\alpha v_{i_1,\dots,i_n} + \sum_{j=0}^{n} i_j v_{i_1,\dots,i_j-1,\dots,i_n} x'_j}_{\text{change it into quadratic form}} \\
&= -\alpha v_{i_1,\dots,i_n} + \text{sum of quadratic terms}
\end{aligned}$$

By applying this algorithm to all new variables, we have the matrix F'_1 of the linear part in the new system as follows:

$$F'_1 = \begin{bmatrix}
\alpha_1^{(1)} & \cdots & \alpha_n^{(1)} & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & & \vdots \\
\alpha_1^{(n)} & \cdots & \alpha_n^{(n)} & 0 & \cdots & 0 \\
0 & \cdots & 0 & -\beta_1 & \cdots & 0 \\
\vdots & & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & -\beta_l
\end{bmatrix}$$

It's easy to find that all eigenvalues of F'_1 have a negative real part. Therefore, we finish the proof of the theorem.

Example 3.1 Consider a system of ODEs

$$\begin{cases} x' = -3x + 1.9y + x^3 \\ y' = -1.9x + y + y^3 \end{cases}$$

In this example, our main problem is that there are no negative coefficients in y of the ODE y'. We can change y' into $y' = -1.9x - y + y^3 + 2y$. Then, we have our two basic functions that

$$\begin{cases} x' = -3x + 1.9y + xv_{2,0} \\ y' = -1.9x + y + yv_{0,2} \end{cases}$$

Here, we introduced new variables $v_{2,0}, v_{3,0}, v_{0,2}, v_{0,3}$, then we apply the algorithm above then we get the system that

$$\begin{cases}
x' = \boxed{-3x + 1.9y} + xv_{2,0} \\
y' = \boxed{-1.9x + y} + yv_{0,2} \\
v'_{2,0} = 2xx' = 2x(-3x + 1.9y + v_{3,0}) = \boxed{-6v_{2,0}} + 3.8xy + 2xv_{3,0} \\
v'_{3,0} = 3x^2x' = 3x^2(-3x + 1.9y + v_{3,0}) = \boxed{-9v_{3,0}} + 5.7v_{2,0}y + 3v_{2,0}v_{3,0} \\
v'_{0,2} = 2yy' = 2y(-1.9x - y + 2y + v_{0,3}) = \boxed{-2v_{0,2}} - 3.8xy + 4y^2 + 2yv_{0,3} \\
v'_{0,3} = 3y^2y' = 3y^2(-1.9x - y + 2y + v_{0,3}) = \boxed{-3v_{0,3}} - 5.8xv_{0,2} + 6yv_{0,2} + 3v_{0,2}v_{0,3}
\end{cases}$$

Then, we can build our matrix for linear part F_1:

$$F_1 = \begin{bmatrix}
-3 & 1.9 & 0 & 0 & 0 & 0 \\
-1.9 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -6 & 0 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0 \\
0 & 0 & 0 & 0 & -2 & 0 \\
0 & 0 & 0 & 0 & 0 & -3
\end{bmatrix}$$

Theorem 3.2

*The smallest highest degree of new variables we introduced in **Theorem 3.1** is* $d-1$.

Proof Suppose we have a $n \times d$ polynomial system of ODEs with the highest degree of d. The introduced variables are in this format:

$$v_{i_1,\ldots,i_n} = x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$$

According to the chain rule, we got

$$v'_{i_1,\ldots,i_n} = \sum_{j=0}^{n} i_j v_{i_1,\ldots,i_j-1,\ldots,i_n} x'_j$$

We assume $v^*_{i_1,\ldots,i_n}$ be the introduced variable with highest degree and $\deg(v^*_{i_1,\ldots,i_n}) = d^*$. Then we got $\deg(v^{*'}_{i_1,\ldots,i_n}) = d^* + d - 1$. Since all the terms in $v^{*'}_{i_1,\ldots,i_n}$ is able to quadratic, therefore we got $2d^* \geq d^* + d - 1 \Rightarrow d^* \geq d - 1$. Then we finish the proof.

Theorem 3.3

Let $v_1, \ldots, v_N$ be the new variables that we introduced in a monomial quadratization. If for each i, v_i is quadratic in $v_1, \ldots, v_N, x_1, \ldots, x_n$, then we can get all negative eigenvalues of the linear part in the ODEs system.

Proof Assume we have a quadratic ODEs system with all introduced variables $v_1, \ldots, v_N$ and for each i, v_i is quadratic in $v_1, \ldots, v_N, x_1, \ldots, x_n$. Since in monomial quadratization all the new variables we introduced are monomial. For any introduced variables can be represented by $v_{i_1,\ldots,i_n} = x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$. From (3.3), we got

$$\begin{aligned} v'_{i_1,\ldots,i_n} &= \sum_{j=0}^{n} i_j v_{i_1,\ldots,i_j-1,\ldots,i_n} x'_j \\ &= \sum_{j=0}^{n} i_j v_{i_1,\ldots,i_j-1,\ldots,i_n} \Big(\alpha_1^{(j)} x_1 + \cdots + \alpha_j^{(j)} x_j + \sum_{i_1+\cdots+i_j \leq j} m^{(j)}_{i_1,\cdots,i_j} x_j^{i_1} \cdots x_j^{i_j}\Big) \end{aligned}$$

Since we have $\forall j \in [0, n]$ that $\deg(v_{i_1,\ldots,i_j-1,\ldots,i_n}) \geq 1$, we know that for all the terms in v'_i, the degree of them is larger or equal to 2 which means no linear term in contained in v'_i. Then for each i, we got

$$\begin{aligned} v'_i &= \text{sum of nonlinear terms} \\ &= \text{sum of quadratic terms} \\ &= \underbrace{-\alpha v_i}_{\text{keep it in linear}} + \underbrace{\alpha v_i}_{\text{change it into quadratic form}} + \text{sum of quadratic terms} \\ &= -\alpha v_i + \text{sum of quadratic terms} \end{aligned}$$

Since we assumed that each i, v_i is quadratic which means all v_i can be represented by the product of the other two terms in $v_1, \ldots, v_N, x_1, \ldots, x_n$. **QED**

Chapter 4 Application of the algorithm in the previous example

The main difference between **Chapter 2** and **Chapter 3** is our quadratization is based on the optimal method and the algorithms we proposed don't have such requirements. Therefore, we can see in our third example that the optimal quadratization is unstable with dissipativity and weakly nonlinearity. We would like to try to apply the algorithm to those three examples and compare it with the previous quadratization.

4.1 Computation with the algorithm

Example 4.1 We have our first example that

$$x' = -x + ax^3$$

By introducing new variable $v_2 = x^2$, we got

$$v_2' = 2xx' = 2x(-x + ax^3) = -2x^2 + 2ax^4 = -2v_2 + 2av_2^2$$

Therefore, we got the system

$$\begin{cases} x' = -x + ax^3 \\ v_2' = -2v_2 + 2av_2^2 \end{cases}$$

with the linear part matrix

$$F_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$$

Example 4.2 For the second example

$$x' = -x^3$$

we are not able to quadratic the function with dissipativity and weakly nonlinearity since there is no linear part in the original equation.

Example 4.3 For the Duffing system

$$x'' = kx + ax^3 + bx'$$

Since it's a second-order ode, we need to introduce a term x' in order to transfer the ode into a polynomial system that

$$\begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0^3 + bx_1 \end{cases}$$

Since we have $k < 0$ and $b < 0$ (previous condition), the system has a linear part with negative eigenvalues. We introduce a new variable $v_{2,0} = x_0^2$ then we got $v_{2,0}' = 2x_0x_0' = 2x_0x_1$. Then we have our new quadratic system that

$$\begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0^3 + bx_1 \\ v_{2,0} = x_0^2 \end{cases}$$

Here we use this example to explore the effect of the negative linear part on the nonlinear part. We have

$$\begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0^3 + bx_1 \\ v_{2,0} = 2x_0x_1 \end{cases} \Rightarrow \begin{cases} x_0' = x_1 \\ x_1' = kx_0 + ax_0^3 + bx_1 \\ v_{2,0} = -cv_{2,0} + cx_0^2 + 2x_0x_1 \end{cases}$$

We have the $F_1 \in \mathbb{R}^{3\times3}$ matrix and $F_2 \in \mathbb{R}^{3\times9}$ matrix that

$$F_1 = \begin{bmatrix} 0 & 1 & 0 \\ k & b & 0 \\ 0 & 0 & -c \end{bmatrix} \quad \text{and} \quad F_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ c & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then we try to analyse the weakly nonlinearity of this system, we got

$$R := \frac{\|X_0\|\|F_2\|}{|\Re(\lambda_1)|}$$

Here $|\Re(\lambda_1)| = \left|\max(-c, \frac{b+\sqrt{b^2+4k}}{2})\right|$ and $\|F_2\| = \max(|a|, \sqrt{c^2+4})$. Then we got

$$R := \frac{\|X_0\| \max(|a|, \sqrt{c^2+4})}{\left|\max(-c, \frac{b+\sqrt{b^2+4k}}{2})\right|}$$

Case 1. We assume $c > -\frac{b+\sqrt{b^2+4k}}{2}$, then we got $R := \frac{\|X_0\| \max(|a|, \sqrt{c^2+4})}{-\frac{b+\sqrt{b^2+4k}}{2}}$. As $c \to -\frac{b+\sqrt{b^2+4k}}{2}^+$, the value of R is decreasing.

Case 2. We assume $0 < c < -\frac{b+\sqrt{b^2+4k}}{2}$, then we have

$$R := \frac{\|X_0\| \max(|a|, \sqrt{c^2+4})}{-c} = \|X_0\| \max(\left|\frac{a}{c}\right|, \sqrt{1+\frac{4}{c^2}})$$

Since c is positive, we can see that as $c \to -\frac{b+\sqrt{b^2+4k}}{2}^-$, the value of R is decreasing. This implies that when we have $c = -\frac{b+\sqrt{b^2+4k}}{2}$ which is the largest eigenvalue of the F_1 matrix before quadratization, we have the minimal value of R. We can plot the graph to show the value of R to verify our conclusion:

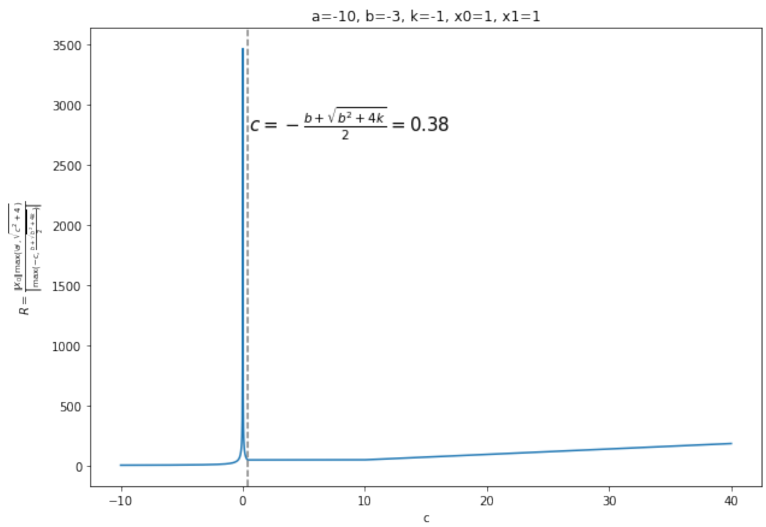


Figure 4.1: The graph of $a = -10, b = -3, k = -1, x_0 = 1, x_1 = 1, -\frac{b+\sqrt{b^2+4k}}{2} = 0.38$

This attempt gives us some insight into setting the value of the linear part of the newly introduced variables. We can set all the values of linear part of the newly introduced variables equal to the $\Re(\lambda_1)$ of the original F_1. However, it's just a conjecture in this simple instance, and we will explore this further during the summer.

Chapter 5 Future working direction

In this report, we focus on quadratization in the reachability problem for ODEs, which we study through three typical ODE systems and extend to the general case and prove some theorems. However, there is still very much we can go furture about this subject.

Further research direction 1. In the quadratization, we need to introduce a large amount of new variables which requires huge computations. Our future work will be based on the optimal monomial quadratizaiton algorithm to develop an optimisation algorithm through theorem 3.3 to make the system dissipative while introducing as few new variables as possible.

Further research direction 2. Although we have shown that all dissipative systems can be transformed into dissipative quadratic systems, we have not yet investigated how to satisfy weakly nonlinearity as far as possible. We will be discussing with the Reachability Analysis of Carlemen linearization group to see how we can work on this direction.

Bibliography

[1] Andrey Bychkov and Gleb Pogudin. *Optimal monomial quadratization for ODE systems*. 2021. arXiv: 2103.08013 [cs.SC].

[2] D. C. Carothers et al. "SOME PROPERTIES OF SOLUTIONS TO POLYNOMIAL SYSTEMS OF DIFFERENTIAL EQUATIONS". In: 2005.

[3] John E Dennis Jr and Robert B Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM, 1996.

[4] Marcelo Forets and Christian Schilling. "Reachability of Weakly Nonlinear Systems Using Carleman Linearization". In: *Lecture Notes in Computer Science*. Springer International Publishing, 2021, pp. 85–99. DOI: 10.1007/978-3-030-89716-1_6. URL: https://doi.org/10.1007%2F978-3-030-89716-1_6.

[5] Marcelo Forets and Christian Schilling. "Reachability of Weakly Nonlinear Systems Using Carleman Linearization". In: *RP*. Vol. 13035. LNCS. Springer, 2021, pp. 85–99. DOI: 10.1007/978-3-030-89716-1_6. URL: https://doi.org/10.1007/978-3-030-89716-1%5C_6.

[6] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[7] Adolf Hurwitz et al. "On the conditions under which an equation has only roots with negative real parts". In: *Selected papers on mathematical trends in control theory* 65 (1964), pp. 273–284.

[8] Ivana Kovacic and Michael J Brennan. *The Duffing equation: nonlinear oscillators and their behaviour*. John Wiley & Sons, 2011.

Appendix A Mathematical Tools

This appendix covers some of the basic mathematics used in this research. We briefly discuss the properties of summation operators and study the properties of linear and some nonlinear equations. This chapter covers the basic knowledge of ordinary differential equations, linear algebra, and calculus.

A.1 Linear Algebra

A.1.1 Eigenvectors and eigenvalues

A pair $(\lambda, V) \in \mathbb{K} \times \mathbb{K}^N \backslash \{0_{\mathrm{K}^N}\}$ is a pair of eigenvalue-eigenvector of a matrix $A \in \mathbb{K}^{N\times N}$ if and only if

$$AV = \lambda V$$

The **spectrum** of A is the set of all eigenvalues

$$Sp(A) = \left\{\lambda \in \mathbb{K} \quad \text{s.t.} \quad \exists 0_{\mathbb{K}^N} \neq V \in \mathbb{K}^N, \quad AV = \lambda V\right\}$$

A.1.2 characteristic polynomial

The **characteristic polynomial** of A is the polynomial p_A defined by

$$p_A(x) = \det(xI - A)$$

The eigenvalues of A fit with the zeros of p_A

$$Sp(A) = Z\left(p_A\right)$$

A.1.3 Norms in Matrix Analysis

The l_1-**norm** is defined as following:

$$\|A\|_1 = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$$

(the maximum absolute column sum). Put simply we sum the absolute values down each column and then take the biggest answer.

The **infinity norm** is defined as following:

$$\|A\|_\infty = \max_{1\leq i\leq n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

(the maximum absolute row sum). Put simply we sum the absolute values along each row and then take the biggest answer.

The **Euclidean norm** is defined as following:

$$\|A\|_E = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2}$$

(the square root of the sum of all the squares). This is similar to ordinary "Pythagorean" length where the size of a vector is found by taking the square root of the sum of the squares of all the elements.

The **operator norm** of a matrix $A \in \mathbb{R}^{m,n}$ is the square root of the largest eigenvalue of the symmetric matrix $A^T A$.

A.1.4 Positive definite matrix

An $n \times n$ symmetric real matrix M is said to be positive-definite if $\mathbf{x}^\top M \mathbf{x} > 0$ for all non-zero $\mathbf{x}$ in $\mathbb{R}^n$. Formally,

$$M \text{ positive-definite } \iff \mathbf{x}^\top M \mathbf{x} > 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$$

An $n \times n$ symmetric real matrix M is said to be positive-semidefinite or non-negative-definite if $\mathbf{x}^\top M \mathbf{x} \geq 0$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Formally,

$$M \text{ positive semi-definite } \iff \quad \mathrm{x}^\top M \mathrm{x} \geq 0 \text{ for all } \mathrm{x} \in \mathbb{R}^n$$

A.2 Calculus and Optimization

A.2.1 Partial Derivatives

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. The partial derivative with respect to x_i is

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}$$

In practice, $\frac{\partial f}{\partial x_i}$ is computed by differentiating f w.r.t x_i, supposing that the other coordinates are constant.

A.2.2 Gradient

Let $\mathbf{w} = [w_1, \ldots, w_d]^\top$ (note: a column vector). The gradient of a function E (w) wrt $\mathbf{w}$ is

$$\mathbf{g} = \nabla_{\mathbf{w}} E(\mathbf{w}) = \left[\frac{\partial}{\partial w_1} E(\mathbf{w}), \ldots, \frac{\partial}{\partial w_d} E(\mathbf{w})\right]^\top$$

i.e., a column vector of d partial derivatives $\mathbf{g} = [g_1, \ldots, g_d]$ where

$$\begin{aligned} g_1 &= [\nabla_{\mathbf{w}} E(\mathbf{w})]_1 = \frac{\partial}{\partial w_1} E(\mathbf{w}) \\ \vdots &= \vdots \\ g_d &= [\nabla_{\mathbf{w}} E(\mathbf{w})]_d = \frac{\partial}{\partial w_d} E(\mathbf{w}) \end{aligned}$$

A.2.3 Jacobian matrix

Suppose $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ is a function such that each of its first-order partial derivatives exist on $\mathbf{R}^n$. This function takes a point $\mathbf{x} \in \mathbf{R}^n$ as input and produces the vector $\mathbf{f}(\mathbf{x}) \in \mathbf{R}^m$ as output. Then the Jacobian matrix of $\mathbf{f}$ is defined to be an $m \times n$ matrix, denoted by $\mathbf{J}$, whose (i, j) th entry is $\mathbf{J}_{ij} = \frac{\partial f_i}{\partial x_j}$, or explicitly

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

where ∇f_i is the covector (row vector) of the gradient of the i component.

A.2.4 Hessian matrix

The Hessian matrix contains all combinations of two successive partial derivatives: $\mathcal{D}^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\ldots,n}$.

A.2.5 Gradient direction

A direction $d \in \mathbb{R}^n$ is called a descent direction for f at x if $\nabla f(x) \cdot d < 0$

A.2.6 Generic gradient descent

Initialization: Choose a starting point x_0 and set $i = 0$
Step i :

- compute $f(x_i)$ and $\nabla f(x_i)$
- choose a step size t and set

$$x_{i+1} = x_i - t \nabla f(x_i)$$