A study of the second-order dynamical systems for variational inequalities

by

Tumelo Ranoto

Supervisor:

Dr Chinedu Izuchukwu

Dissertation submitted in fulfilment of the requirements for the degree
of
Master of Science

The University of the Witwatersrand

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

School of Mathematics
University of the Witwatersrand, Johannesburg, South Africa.

May, 2025.

A study of the second-order dynamical systems for variational inequalities

by

Tumelo Ranoto

Declaration

Candidate's declarations:

I, Tumelo Ranoto, hereby certify that this dissertation submitted in partial fulfillment of the requirements for the award of Master of Science at the University of the Witwatersrand, is my work unless otherwise referenced or acknowledged. This work has not been submitted for any other qualification at any other academic institution.

Signed:

Date: 21-05-2025

Supervisor's declaration:

I, Chinedu Izuchukwu, hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Master of Science at the University of the Witwatersrand and that the candidate is qualified to submit this dissertation in an application for that degree.

Signed:

Date: 21-05-2025

Abstract

The purpose of this paper is to study the variational inequality problem through a second-order dynamical system. The dynamical system is a second-order dynamical system that involves the asymptotic vanishing of damping and Hessian-damping terms. Our approach primarily uses the asymptotic dynamics of the inertial system and the geometric characteristics of the damping factors. The asymptotic vanishing damping is directly linked to the method of the Nesterov accelerated gradient. The geometry of the Hessian damping term is in this form $\frac{d}{dt}(y(t)-w(t))$, where $y(t)$ is the generated trajectories by our proposed dynamical system, $w(t) = P_C(y(t)-Qy(t))$ and Q is our operator which is strongly monotone. The Hessian term assists with the neutralisation of the oscillation. This combination forms a damping force that assists with controlling the convergence speed of dynamical systems, stabilizing the system, and mostly ensuring we have a controlled velocity over time. As a main result, we establish the existence and uniqueness of the generated trajectories using the Cauchy-Lipschitz theorem. We give two numerical non-Hilbert examples. We construct a Lyapunov function to achieve weak convergence, assuming that the underlying operator is Lipschitz continuous and monotone. By considering the strong monotonicity of the operator, we establish exponential convergence. Furthermore, we given two numerical examples to show the applicability of our results. Finally, we conduct numerical experiments to show the performance of our dynamical system. The results presented in this paper build upon and significantly enhance recent findings in the existing literature.

Keywords: Variational inequality; Dynamical system; Continuous forward-backward-forward; Strongly monotone; Weak convergence; Exponential convergence

Acknowledgements

I express my heartfelt gratitude to God Almighty, whose boundless mercy and grace have been my foundation throughout this journey. His gift of life, good health, and unwavering strength has empowered me to see this program through to its completion. I extend my sincere thanks to my supervisor, Dr Izuchukwu, for entrusting me with this challenging and rewarding project. His insightful and unwavering patience has been invaluable to my growth as a researcher.

As there have been numerous direct and indirect contributors to this study, I would like to take this opportunity to express my gracious concern towards all these people who have helped in various ways throughout the development stages of this dissertation. I also express my appreciation to the University of the Witwatersrand (Wits) for providing a supportive academic environment and the resources necessary for the successful execution of this research. I would also like to extend my heartfelt gratitude to the Energy Mobility Education Trust (EMET) for their generous financial support toward this program. Their contribution and the opportunity they provided for me to pursue a research master's degree at the University of Witwatersrand (Wits) have been instrumental in making this achievement possible.

Finally, I want to express my sincerest appreciation and thanks to my family for supporting me endlessly all the time, and to my friends for their encouragement and the enjoyable time over the last years.

Dedication

This dissertation is devoted to God Almighty and my dear parents.

Table of Contents

Contributed papers from the dissertation

Tumelo Ranoto and Chinedu Izuchukwu, Second-order differential equation with Hessian-driven damping for variational inequality problem. Submitted to Computational Optimization and Applications.

Chapter 1

Introduction

1.1 Background of study of the variational inequality problem

Throughout this dissertation, we assume that C be a nonempty, closed, and convex subset of a real Hilbert space H endowed with the norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$. We denote by I the identity operator.

Let $Q : H \to H$ be an operator where $C \subseteq H$ then the *variational inequality problem* (VIP) for Q on C is to find $y^* \in C$ such that

$$\langle Qy^*, y - y^*\rangle \geq 0, \quad \forall y \in C. \tag{1.1.1}$$

Let Ω denote the set of solutions to (1.1.1). The earliest VIP was the Signorini problem, introduced by Signorini in 1959 [9, 52] and solved by Fichera in 1963 [42]. To address regularity issues in partial differential equations, Stampacchia [66] expanded on the Lax-Milgram theorem and referred to all such inequality-related problems as VIPs. VIP has been used to solve many mathematical problems, such as equilibrium problems and optimization problems; see, for example, [67, 40, 89, 78, 63, 91, 77, 43, 79, 80, 45, 58]. Moreover, the VIP serves as a unified formulation of several important mathematical problems. For example, when $C = \mathbb{R}^n$, then we have that VIP(C, Q) becomes the problem of finding a zero of the mapping Q. Many papers have focused extensively on solving the variational inequality problem when Q is pseudo-monotone. For instance, [28, 92, 75, 105]. Moreover, Phan Tu Vuong [100] specifically considers the case when Q is strongly pseudo-monotone with other authors see [97, 59, 81, 55, 90].

The projected gradient method is the simplest algorithmic method for solving VIPs when the operator $Q(y(t))$ is strongly (Pseudo) monotone or co-coercive. The method is given as follows:

$$y_{n+1} = P_C(y_n - \lambda Q(y_n)) \quad \forall n \geq 0, \tag{1.1.2}$$

where P_C represents the projection operator onto the closed convex set, and λ is a positive stepsize. We have that for $y_0 \in H$ the method will generate a sequence that converges to the solution set Ω. The method converges for the case when Q is a strongly pseudo-monotone operator [25]. Despite the simplicity of this method, the strong monotonicity requirement on $Q(y(t)$ may preclude the method from many applications, e.g, in situations where the underlying operator is (pseudo) monotone, not strongly (pseudo)monotone. To overcome this limitation, Korpelevich [33] introduced the extragradient method, which is defined as follows:

$$\begin{aligned} w_n &= P_C\left(y_n - \lambda Q(y_n)\right), \\ y_{n+1} &= P_C\left(y_n - \lambda Q(y_n)\right), \quad \forall n \geq 0. \end{aligned} \tag{1.1.3}$$

This method involves two evaluations of Q and two computations of the projection P_C at each iteration. This makes (1.1.3) computationally expensive. The method requires two projections at each iteration, which can be computationally expensive. To address this issue, Tseng [98] improved Algorithm (1.1.3) by reducing the number of projections P_C from two to one. The equation is given as follows:

$$\begin{cases} w(t) = P_C\big(y(t) - \lambda Q(y(t))\big), \\ \dot{y}(t) + y(t) - w(t) - \lambda\big[Q(y(t)) - Q(w(t))\big] = 0, \\ y(0) = y_0, \end{cases} \tag{1.1.4}$$

where $\lambda > 0$ and $x_0 \in H$. Here Q is the underlying operator that can be monotone or pseudomonotone. Radu I Botand Erno Robert Csetnek [25] studied the system when the operator is pseudomonotone and with the additional assumption that the function is weak to weak sequentially continuous.

We now transition this algorithmic perspective into a dynamical system. One can interpret the iterative scheme as a discrete approximation of a continuous-time dynamical system. This perspective allows the behavior of solutions to be analyzed in terms of trajectories. The other traditional approach for solving variational inequalities is the dynamical system. Dynamical systems arise naturally in numerous applied and theoretical fields, including mechanics, financial forecasting, environmental applications, neuroscience, and brain modeling. In recent years, dynamical systems have been used to study the problems of variational inequality and optimization problems [44, 37, 36, 22, 38, 31, 25, 6, 7]. The continuous version of the gradient projection method is defined as follows:

$$\dot{y}(t) + y(t) = P_C\left(y(t) - \lambda Q(y(t))\right), \tag{1.1.5}$$

where $y(t)$ is the state at time t, $\lambda > 0$ is the step size, $Q(y(t))$ is the monotone operator, and $P_C(\cdot)$ is the projection operator that maps a point onto the set C. This method is useful for analysing how optimization algorithms behave over time in a continuous setting. In the pursuit of solving monotone VIPs, various first-order dynamical and algorithms have been proposed to see [88, 87, 61, 5, 20, 106, 2, 19, 46, 68, 70, 60, 106]. Additionally, there exist second-order dynamical systems corresponding to the first order, to improve the convergence rate; see [56, 23, 24, 57, 95, 3, 74, 101, 51]. The second-order method for the projected gradient is defined as follows:

$$\ddot{y}(t) + \alpha\dot{y}(t) + y(t) - P_C(y(t) - \epsilon Q(y(t))) = 0 \tag{1.1.6}$$

where $\epsilon > 0$ and Q is the underlying operators used to study the system. The convergence behavior of trajectories towards a solution of Q within the set C has been extensively analyzed and is well-known in the literature. Antipin [99] has proven the asymptotic and exponential stability for trajectories generated by (2.7.5). The trajectories generated by (2.7.5) converge faster than the first-order (2.6.1) counterparts for minimizing a smooth convex function. Phan Tu Vuong [100] extended the projection method proposed by Antipin [99] from solving optimization problems to variational inequality by replacing the gradient operator with a strong pseudo-monotone operator. The equation is given by:

$$\begin{cases} \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta(t)\left[y(t) - P_C\left(y(t) - \lambda Q y(t)\right)\right] = 0, \\ y(0) = y_0, \quad \dot{y}(0) = y_0. \end{cases} \tag{1.1.7}$$

However, the dynamical system given by (2.7.14) does not generate trajectories that converge to $\Omega \subset H$ for solving the VIP (1.1.1) when Q is monotone. As we want the rapid speed of convergence by introducing the acceleration term, we experience the challenge of oscillation, which causes the delay in achieving convergence. In that case, we introduce the Hessian-driven damping term to take care of that. In algorithms, damping helps regulate the momentum that propels the iterates forward. With too little damping, the iterates may overshoot, oscillate, or even diverge. With too much damping, the system will have slow convergence. We will modify to incorporate Hessian damping to address the oscillations. The method of Hessian damping is more effective than the standard damping. There have been many papers studying the system attached to the Hessian-driven damping term, see [7, 16, 14]. The following method, called the dynamical inertial Newtonian (DIN) method, introduces Hessian-based corrections to improve stability and convergence rates. The method is given as follows:

$$\ddot{y}(t) + \alpha\dot{y}(t) + \beta\nabla^2 Q(y(t))\dot{y}(t) + \nabla Q(y(t)) = 0, \tag{1.1.8}$$

where, $\beta > 0$ governs the contribution of the Hessian $\nabla^2 Q(y(t)$ and $\beta\nabla^2 Q(y(t))\dot{y}(t)$ serves to neutralise oscillations, providing a smoother and faster convergence, similar to Nesterov's accelerated gradient method.

In this dissertation, inspired by recent advances in second-order dynamical systems with variable damping, we investigate a novel approach incorporating a Hessian damping factor. Our contribution in this paper, is to propose the following system:

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta\frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Qy(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0 \end{cases} \tag{1.1.9}$$

where P_C is the projection operator onto a closed convex set C, $\alpha(t)$ is an asymptotically vanishing damping function, and Q is assumed to be a monotone or strongly monotone operator with Lipschitz continuity. Our primary contribution is the analysis of a second-order dynamical system associated with the forward-backward-forward method, achieved by introducing an acceleration term to the system proposed in [25]. Additionally, we incorporate a Hessian-driven factor to mitigate oscillations. We establish the existence and uniqueness of trajectories by employing the Cauchy–Lipschitz–Picard theorem under the assumption that the underlying operator is merely Lipschitz continuous. The second contribution is to introduce the Hessian-driven factor term to neutralize the oscillation. We will prove the existence and uniqueness of the trajectories using the Cauchy-Lipschitz

Picard theorem for the case when the underlying operator is only Lipschitz continuous. For the case when the underlying operator is monotone and Lipschitz continuous, we will achieve weak convergence of our trajectories to the solution set Ω. Furthermore, we will get the exponential convergence for the trajectories when the operator is Lipschitz continuous and strongly monotone.

If $\nabla Q(y)$ is the gradient of a differentiable real-valued function $Q : \mathbb{R}^n \to \mathbb{R}$,then solving the variational inequality problem (1.1.1) is the same as solving the optimization problem:

$$\min_{y \in C} Q(y). \tag{1.1.10}$$

In particular, y is a solution to $\mathrm{VI}(C, Q)$ if and only if y has no feasible descent direction, i.e., y satisfies the necessary optimality conditions for the problem.

1.2 Research motivation

The purpose of this research is to study variational inequality problems using second- order dynamical systems. The analysis of this system (1.1.9) is motivated by the observation that the inclusion of the acceleration term $\ddot{y}(t)$ often improves the convergence behavior of the trajectories. The motivation for this study can be traced to the heavy ball method of Polyak [84], where it was shown that the trajectories generated by the second-order dynamical systems converge faster than the first-order counterparts for minimizing a smooth convex function.

Phan Tu Vuong [100] utilised the Antipin equation proposed by Antipin [99] and replaced the gradient operator with a strongly pseudo-monotone operator. The equation is given by,

$$\begin{cases} \ddot{y}(t) + \alpha(t)\dot{y}(t) = -\beta(t)\left[y(t) - P_C\left(y(t) - \lambda Q y(t)\right)\right], \\ \quad y(0) = y_0, \quad \dot{y}(0) = y_0. \end{cases} \tag{1.2.1}$$

where λ is the positive stepsize, $\alpha(t)$ is the asymptotic vanishing damping and Q is strongly psuedomonotone. He obtained exponential convergence for (1.2.1), and the explicit discretization of (1.2.1) has been shown to converge linearly. However, the dynamical system given by Phan Tu Vuong [100] does not produce trajectories that converge to $\Omega \subseteq H$if Q is monotone.

Phan Tu Vuong [25] studied the first order forward-backward-forward method . He considered both the continuous and the discrete parts. For the continuous case, he achieved weak convergence, while for the discrete case, he achieved linear convergence where the underlying operator was pseudo-monotone. Radu I. Bot and Michael Sedlmayer [27] extend the work by developing a second-order dynamical system, specifically designed to approximate the solution of monotone inclusion problems and provide an asymptotic analysis of its trajectories. They have achieved weak convergence but for the case when the operator is a pseudo-monotone. Their results will be improved by incorporating the Hessian damping factor to solve the variational inequality problem.

Bot and Csetnek [62] approach this optimization problem from a continuous perspective using a second-order dynamical system that includes the asymptotic vanishing and Hessian damping, and also a penalty term corresponding to the constrained function. Constructing appropriate energy functionals, they proved weak convergence of the trajectories generated by this differential equation to a minimizer of the optimization problem. Furthermore, along the trajectories, they show the convergence of the objective function values. We will move from the optimization problem to solving a problem in variational inequality using a similar approach that includes the Hessian damping term.

1.3 Statement of problem

Problem Statement

Variational inequality problems (VIPs) are a class of mathematical problems that arise in numerous fields, including optimization, economics, game theory, and fluid mechanics. A variational inequality problem is defined as follows:

Given $Q : H \to H$ and $C \subseteq H$ then find a vector $y^* \in C$ such that:

$$\langle Q(y^*), y - y^* \rangle \geq 0, \quad \forall y \in C. \tag{1.3.1}$$

To address this problem, we propose to study the dynamical system (1.1.9) with a Hessian damping factor that leverages the properties of the mapping Q, specifically its monotonicity and Lipschitz continuity. The goal is to design a continuous-time trajectory that evolves toward a solution of the variational inequality, leveraging the energy dissipation inherent in second-order dynamics for faster convergence. We will use a dynamical system (1.1.9) to solve equation (1.1.1).

The advantages of including the Hessian damping is that it uses information from the Hessian matrix of Q to influence the system's movement. Allows the system to adjust its resistance based on the steepness in each direction. The approach can reduce oscillations around the optimal point more effectively than standard damping. Provides a mechanism for stabilizing the system as it approaches high-curvature regions near the solution.

1.4 Objectives

The main objectives of this dissertation are as follows:

1. We will review some known and useful methods and results on Variational Inequality Problems (VIPs) and Optimization.

2. We will introduce and study dynamical system(1.1.9) to generate trajectories that converge to the solution set Ω of (1.1.1).

3. We will achieve the existence and uniqueness of the trajectories generated by (1.1.9) when the Q is only Lipschitz continuous.

4. When Q is monotone and Lipschitz continuous, we will establish the weak convergence of the trajectories to the solution set Ω.

5. We will analyze the stability of the system and get the explicit bounds on the convergence rates.

6. Additionally, when we have that Q is strongly monotone and Lipschitz continuous, we will get the exponential convergence for (1.1.9).

7. We will give numerical examples to show the performance of the dynamical system (1.1.9).

1.5 Outline

The chapters in this dissertation are structured as follows:

The first chapter deals with the study of the background of variational inequalities. We summarise earlier work on the methods to solve variational inequality problems that are algorithmic and dynamical. We describe the aims and the problem statement for our paper.

The second chapter deals with the literature review. We outline the foundational elements of our study. We review essential definitions, concepts, theorems, lemmas, and propositions that serve as the building blocks for the subsequent results. Topics covered include monotonicity, Lipschitz continuity, and fundamental properties of the dynamical systems under consideration.

In the third chapter, we begin by noting that the underlying operator is only Lipschitz continuous. Using the Cauchy-Lipschitz-Picard theorem, we establish the existence and uniqueness of the trajectories generated by the dynamical system (3.1.1). Furthermore, we demonstrate that the acceleration term of the dynamical system is absolutely continuous, an essential property that ensures the integrability and smoothness of the system's evolution over time. We will also find the estimation of the third order.

In the fourth chapter, we modify the second-order forward-backward-forward dynamical system by including the Hessian-driven term. We will use the monotonicity and Lipschitz continuity of the operator Q. We will perform the asymptotic analysis and, using these properties, we prove the weak convergence of the proposed system under suitable assumptions.

In the fifth chapter, we will focus on the exponential convergence properties of the systems introduced in earlier chapters. We provide formal proof of exponential convergence under specific assumptions about the operator Q and the parameters of the system. In this chapter, we include a series of numerical experiments demonstrating the practical performance of the system. The experiments are presented with visualizations to emphasize the rate of convergence and the influence of parameter choices.

In the last chapter, we conclude the study by summarising the key contributions and insights derived from the research. We emphasize the novel aspects of our work, including the development of modified dynamical systems and the analysis of their convergence properties. Furthermore, we outline potential areas for future exploration, such as extending the systems to study the explicit discrete version of the system and generalizing our system by replacing the projection operator P_C with the resolvent operator to study the Monotone inclusion problem. Also, the interesting part is studying the variational inequality of the second kind.

Chapter 2

Literature Review

This chapter introduces the fundamental terms and concepts essential for the rest of our work. We also present key results and offer a comprehensive literature review of concepts pertinent to our study.

2.1 Preliminaries

2.1.1 Some basic concepts

Vector space

We will begin by studying vector spaces, laying the fundamental concepts that will be useful for our study. Once we have a solid understanding of vector spaces, we will include additional properties like inner product and more to be able to take our study to Hilbert space.

Definition 2.1.1. *[72] A vector space over a field $\mathbb{F}$ is a nonempty set denoted by $\mathbb{V}$ together with addition and scalar multiplication satisfying the following conditions:*

(i) $\tau + \mu = \mu + \tau$, for all $\mu, \tau \in \mathbb{V}$;

(ii) $\tau + (\mu + z) = (\tau + \mu) + z$, for all $\tau, \mu, z \in \mathbb{V}$;

(iii) $\alpha \times (\beta \times \tau) = (\alpha\beta) \times \tau$, for all $\alpha, \beta \in \mathbb{F}$ and $\tau \in \mathbb{V}$;

(iv) $\alpha \times (\tau + \mu) = \alpha \times \tau + \alpha \times \mu$, for all $\tau, \mu \in \mathbb{V}$ and $\alpha \in \mathbb{F}$;

(v) $(v + \theta) \times \tau = v \times \tau + \theta \times \tau$, for all $\theta, v \in \mathbb{F}$ and $\tau \in \mathbb{V}$;

(vi) there exists $1 \in \mathbb{F}$ such that $1 \times \tau = \tau$, $\forall \tau \in \mathbb{V}$.

Definition 2.1.2. *[32] Let $Q : \mathbb{V} \to \mathbb{W}$ be a function between two vector spaces where addition and scalar multiplication are preserved.*

(i) $Q(\tau + \mu) = Q(\tau) + Q(\mu)$, for all $\tau, \mu \in \mathbb{V}$

(ii) $Q(\beta\tau) = \beta Q(\tau)$, for all $\tau \in \mathbb{V}$

Definition 2.1.3. *Arithmetic Mean-Geometric Mean (AM-GM) Inequality states that for any nonnegative real numbers $a_1, a_2, \ldots, a_n$ then*

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}. \tag{2.1.1}$$

Lemma 2.1.4 (Young's Inequality)**.** *Let x and y be two positive real numbers such that*

$$\frac{1}{x} + \frac{1}{y} = 1. \tag{2.1.2}$$

Given that a and b are positive real number we will have the following,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq ab. \tag{2.1.3}$$

Equality holds if and only if $a^x = b^y$

Proof. Using the (2.1.3) and let the weight be given by $m_1 = \frac{1}{x}$ and $m_2 = \frac{1}{y}$. we have the following equation.

$$\frac{m_1 a^x + m_2 b^y}{m_1 + m_2} \geq ((a^x)^{m_1} (b^y)^{m_2})^{\frac{1}{m_1 + m_2}} . \tag{2.1.4}$$

Having that the following equation hold,

$$\frac{1}{x} + \frac{1}{y} = 1 \implies m_1 + m_2 = 1, \tag{2.1.5}$$

substituting this into (3.2.4) it gives the following equation ,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq (a^x)^{\frac{1}{x}} (b^y)^{\frac{1}{y}}. \tag{2.1.6}$$

Using the law of exponents then we have the following equation,

$$(a^x)^{\frac{1}{x}} = a, \quad (b^y)^{\frac{1}{y}} = b, \tag{2.1.7}$$

When we substitute equation (2.1.7) into equation (2.1.6) then we get,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq ab. \tag{2.1.8}$$

This is true if and only if $a^x = b^y$, therefore, Young's inequality holds. $\square$

Definition 2.1.5. *[32] A function $Q : P \to \mathbb{R}$ defined on a set P is said to be bounded if there exists a constant $K > 0$ such that*

$$|Q(\tau)| \leq K \quad \text{for all } \tau \in P. \tag{2.1.9}$$

Theorem 2.1.6. *[76] Let (P, ρ) be a metric space that is complete. Suppose $Q : P \to CB(P)$ is a multivalued mapping that satisfies the contraction property. Then, Q has a fixed point.*

2.1.2 Convex minimization problems

Definition 2.1.7 (Convex Set)**.** *[32] A subset C of a Hilbert space H is called convex if it satisfies the condition:*

$$\lambda u + (1-\lambda)v \in C, \quad \textit{for all } u, v \in C \textit{ and } \lambda \in [0,1]. \tag{2.1.10}$$

Remark 2.1.8. *The entire space H and the empty set $\emptyset$ are convex.*

Definition 2.1.9. *[18] A function Q is convex if*

$$Q(\beta a + (1-\beta)b) \leq \beta Q(a) + (1-\beta)Q(b) \tag{2.1.11}$$

holds for all $a, b \in H$ and for all $\beta \in [0,1]$.

Remark 2.1.10. *-Q concave implies Q is convex.*

Definition 2.1.11. *[18] A function Q is strictly convex if*

$$Q(\beta a + (1-\beta)b) < \beta Q(a) + (1-\beta)Q(b) \tag{2.1.12}$$

holds for all $a, b \in H$ and for all $\beta \in [0,1]$.

Definition 2.1.12. *[96] Let $Q : H \to \mathbb{R}$ and $C \subseteq H$. The constrained minimization problem is given by:*

$$\textit{minimize } Q(y) \quad \textit{with respect to } y \in C. \tag{2.1.13}$$

This means finding at least one local minimizer of the function Q on C, i.e., a point $y \in C$ such that

$$Q(y) \leq Q(y') \quad \textit{for all } y' \in C \setminus B(y, q)$$

for some $q > 0$, provided such a point exists.

If

$$Q(y) \leq Q(y') \quad \textit{for all } y' \in C,$$

then y is called a global minimizer of the function Q on C.

Definition 2.1.13. *[108] A continuous and convex function $Q : W \subset H \to \mathbb{R}$ attains its minimum at a point $y \in H$ if and only if $0 \in \partial Q(y)$.*

Definition 2.1.14. *[86] A function $Q : \mathbb{R}^n \to \mathbb{R}$ is strongly convex on $\mathbb{R}^n$ if $e > 0$ exists and is called modulus of strong convexity such that for any $a, u \in \mathbb{R}^n$, the following inequality holds:*

$$Q(u) \geq Q(a) + \langle \nabla Q(a), u - a \rangle + \frac{e}{2}\|u - a\|^2. \tag{2.1.14}$$

Theorem 2.1.15. *[32] If $C \subseteq H$ is convex, then both its closure $\overline{C}$ and its interior* $\operatorname{int}(C)$ *are convex subsets of H.*

Definition 2.1.16. *[32] If $C \subseteq H$ is a closed convex subset, then C is weakly closed.*

Example 2.1.17. *[18] Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an increasing function, let $\beta \in \mathbb{R}$, and define*

$$w : \mathbb{R} \to \mathbb{R} \cup \{+\infty\} \quad by \quad w(x) = \begin{cases} \int_\beta^x \varphi(t)\, dt & if\ x \geq \beta, \\ +\infty & otherwise. \end{cases}$$

Then, w is convex. Furthermore, if φ is strictly increasing, then w is strictly convex.

Theorem 2.1.18. *[32] Let $C \subseteq H$ be a convex subset and let $Q : H \to \mathbb{R}$ be a convex function. If y is a local minimizer of Q on C, then it is also a global minimizer. Furthermore, if Q is strictly convex, then the minimizer is uniquely determined. If Q is continuous and strongly convex and C is closed, then*

$$Argmin_{y \in C} Q(y) \neq \emptyset. \tag{2.1.15}$$

2.1.3 Differentiability

Definition 2.1.19. *[18] A scalar function $Q : \mathbb{R}^n \to \mathbb{R}$ it is differentiable at a point $z \in \mathbb{R}^n$ if there exists a vector $b \in \mathbb{R}^n$ such that for every $w \in \mathbb{R}^n$ and $\langle w, z \rangle = \sum_{i=1}^n w_i z_i$ is the standard dot product and $\|w\| = \sqrt{\langle w, w \rangle}$ is the Euclidean norm in $\mathbb{R}^n$, then the following holds:*

$$Q(w) = Q(z) + \langle b, w - z \rangle + o(\|w - z\|). \tag{2.1.16}$$

Example 2.1.20. *[18]*

$$\nabla Q(x) = \left(\frac{\partial Q}{\partial x_1}, \ldots, \frac{\partial Q}{\partial x_n} \right). \tag{2.1.17}$$

Definition 2.1.21. *[18] A function Q is differentiable on a set $\Omega \subset \mathbb{R}^n$ if it is differentiable at all points of Ω. A function Q is differentiable if it is differentiable on the entire space $\mathbb{R}^n$.*

Definition 2.1.22. *[18] A vector function $Q : \mathbb{R}^n \to \mathbb{R}^p$ It is differentiable at a specific point $\mu \in \mathbb{R}^n$,if there exists a matrix $A \in \mathbb{R}^{p \times n}$ such that for all $\tau \in \mathbb{R}^n$, the following holds:*

$$Q(\tau) = Q(\mu) + B(\tau - \mu) + o(\|\tau - \mu\|). \tag{2.1.18}$$

The matrix A is called the Jacobian matrix of Q at τ, we get that ,

$$Q(\mu) = Q(\tau) + \nabla Q(\tau)(\mu - \tau) + o(\|\mu - \tau\|). \tag{2.1.19}$$

the Jacobian provides a linear approximation of the function at τ.

Definition 2.1.23. *[32] An operator $Q : H_1 \to H_2$ is differentiable at a point $\tau \in H_1$ if there exists a bounded linear operator $N \in L(H_1, H_2)$ such that*

$$Q(\tau + \mu) = Q(\tau) + Nh + o(\|\mu\|). \tag{2.1.20}$$

The operator N is uniquely determined and is called a derivative of the operator Q given by $Q'\tau$ at the point τ.

Definition 2.1.24. *[32] A function $Q : H \to \mathbb{R}$ is Fréchet-differentiable, if it is differentiable at $\tau \in H$ and there exists $\mu \in H$ such that*

$$Q(\tau + h) = Q(\tau) + \langle \mu, h \rangle + o(\|h\|), \tag{2.1.21}$$

where $\lim_{h\to 0} \frac{o(\|h\|)}{\|h\|} = 0$. The element μ is a derivative of Q at τ and is denoted by $Q'(\tau)$.

Definition 2.1.25. *[32] A function $Q : H \to \mathbb{R}$ is Gâteaux-differentiable at a point $y \in H$ if it has directional derivatives $Q'(y; s)$ for all $s \in H$ and*

$$Q'(y; s) = \langle g, s \rangle$$

holds for some $g \in H$. The element g is called a Gâteaux-differential of Q at y and is denoted by $Q'(y)$. The Gâteaux-differential of a Gâteaux-differentiable function is uniquely determined.

Theorem 2.1.26. *[86] If a function $Q : H \to \mathbb{R}$ is Fréchet-differentiable at $y \in H$, then it is Gâteaux-differentiable at y and*

$$Q'(y) = DQ(y), \text{ where } D \text{ is the derivative operator.}$$

2.1.4 Strongly monotone operator

Definition 2.1.27. *[73] Let $Q : C \to H$ be a mapping, where $C \subseteq H$, the mapping Q is said to satisfy the following properties:*

(i) L-Lipschitz if there exists $L > 0$ such that

$$\|Q(v) - Q(b)\| \leq L\|v - b\|, \quad \forall v, b \in H;$$

If $L = 1$, then Q is called nonexpansive, and when we have $L \in (0, 1)$, then Q is called a contraction.

(ii) τ-strongly monotone, if there exists $\tau > 0$ such that

$$\langle Q(v) - Q(b), v - b \rangle \geq \tau\|v - b\|^2, \quad \forall v, b \in H;$$

(iii) A mapping Q is said to be antimonotone on a set $C \subseteq \mathbb{R}^n$ if for all $v, b \in C$, it holds:

$$\langle Q(v) - Q(b), v - b \rangle \leq 0.$$

(iv) η-inverse strongly monotone , if there exists $\eta > 0$ such that

$$\langle Q(v) - Q(b), v - b \rangle \geq \eta\|Q(v) - Q(b)\|^2, \quad \forall v, b \in H;$$

if $\eta = 1$, then Q is called firmly nonexpansive.

(v) monotone, if

$$\langle Q(v) - Q(b), v - b \rangle \geq 0, \quad \forall v, b \in H;$$

(vi) pseudo-monotone, if

$$\langle Q(v), b - v \rangle \geq 0 \quad \Rightarrow \quad \langle Q(b), b - v \rangle \geq 0, \quad \forall v, b \in H;$$

(vii) an operator Q is said to be strongly pseudomonotone with modulus y on Ω if there exists $y > 0$ such that:

$$(Q(v))^\top (b - v) \geq 0 \implies (Q(b))^\top (b - v) \geq y\|v - b\|^2, \quad \textit{for all } v, b \in \Omega.$$

Remark 2.1.28. *We get that strongly monotone implies monotone and pseudo-monotone. Furthermore, strongly pseudo monotone implies pseudo monotone.*

Theorem 2.1.29. *Let Q be a continuous operator that is strongly monotone with modulus $\tau > 0$ on a nonempty, closed, and convex set C. Then, the variational inequality* $\mathrm{VI}(Q, C)$ *admits a unique solution, which coincides with the unique equilibrium point of the dynamical system* (5.1.1).

Example 2.1.30. *[18] Let $Q : H \to]-\infty, +\infty]$ be a proper function. Then, the operator ∂Q is monotone.*

Proof. Consider two points (w, u) and (a, v) in the graph of ∂Q, denoted as $\mathrm{gra}(\partial Q)$. We get the following inequalities that hold

$$\langle w - a, u \rangle + Q(a) \geq Q(w), \quad \text{and} \quad \langle a - w, v \rangle + Q(w) \geq Q(a).$$

Summing these two inequalities, we obtain

$$\langle w - a, u - v \rangle \geq 0. \tag{2.1.22}$$

Thus ∂Q is a monotone operator. □

2.2 Hilbert spaces

The notion of Hilbert spaces was introduced by David Hilbert (between 1862 and 1943) and is an extension of the concept of Euclidean spaces to infinite-dimensional spaces.

2.2.1 Properties of the inner product

Definition 2.2.1. *[72] Let H be a nonempty set. An inner product on H is a function $\langle \cdot, \cdot \rangle$ defined on $H \times H$ with values in $K = \mathbb{R}$ or $\mathbb{C}$ such that the following conditions hold:*

(i) $\langle \tau, \tau \rangle \geq 0$ for all $\tau \in H$ and $\langle \tau, \tau \rangle = 0 \Leftrightarrow \tau = 0$;

(ii) $\langle \tau, \mu \rangle = \overline{\langle \mu, \tau \rangle}$ for all $\tau, \mu \in H$ (the bar denotes the complex conjugate);

(iii) $\langle \alpha\tau, \mu \rangle = \alpha \langle \tau, \mu \rangle$;

(iv) $\langle \lambda u + \mu x, c\rangle = \lambda\langle u, c\rangle + \mu\langle x, c\rangle$ *for all* $x, u, c \in H$ *and* $\lambda, \mu \in K$.

The pair $(H, \langle\cdot,\cdot\rangle)$ *is called an inner product space. If* $K = \mathbb{R}$, *then (ii) becomes* $\langle a, b\rangle = \langle b, a\rangle$. *In this case,* $(H, \langle\cdot,\cdot\rangle)$ *is called a real inner product space.*

Remark 2.2.2. *A linear space* Y *with a norm defined on it, i.e.,* $(Y, \|\cdot\|)$, *is called a* normed linear space. *If* Y *is a normed linear space, the norm* $\|\cdot\|$ *always induces a metric* d *on* Y *given by* $d(a, c) = \|a - c\|$ *for each* $a, c \in Y$.

Definition 2.2.3. *[32] An inner product space* $(H, \langle\cdot,\cdot\rangle)$ *is considered complete if every Cauchy sequence in* H *converges to a limit within* H.

We have that a complete inner product space is referred to as a Hilbert space.

Proposition 2.2.4. *[72] (Cauchy-Schwarz Inequality) Let* $(H, \langle\cdot,\cdot\rangle)$ *be an inner product space. For arbitrary* $\tau, \mu \in H$,

$$|\langle \tau, \mu\rangle|^2 \le \langle \tau, \tau\rangle \cdot \langle \mu, \mu\rangle. \tag{2.2.1}$$

Lemma 2.2.5. *[72] The function* $\|\cdot\| : H \to [0, \infty)$, *defined by*

$$\|\tau\| = \sqrt{\langle \tau, \tau\rangle}, \tag{2.2.2}$$

is a norm on H.

Remark 2.2.6. *Following (2.2.1), the Cauchy-Schwarz inequality can generally be written as*

$$|\langle \tau, \mu\rangle| \le \|\tau\| \cdot \|\mu\|, \tag{2.2.3}$$

for arbitrary $\mu, \tau \in H$.

2.2.2 Properties of Hilbert spaces

Hilbert spaces have several important geometric properties due to their inner product structure. One such property is that the nearest point map in a real Hilbert space H *onto a closed convex set* C *is Lipschitz continuous with a constant of 1.*

Definition 2.2.7. *[32] The following inequality holds for all* $\tau, \mu \in H$:

$$|\|\tau\| - \|\mu\|| \le \|\tau + \mu\| \le \|\tau\| + \|\mu\|. \tag{2.2.4}$$

Furthermore,

(i) $\|\tau + \mu\| = \|\tau\| + \|\mu\|$

(ii) $\|\tau - \mu\| = |\|\tau\| - \|\mu\|$

Proposition 2.2.8. *[32] Let* H *be a real Hilbert space. For each* $\tau, \mu \in H$, *we have*

$$\|\tau - \mu\|^2 = \|\tau\|^2 + \|\mu\|^2 - 2\langle \tau, \mu\rangle, \tag{2.2.5}$$
$$\|\tau + \mu\|^2 \le \|\tau\|^2 + 2\langle \mu, \tau + \mu\rangle. \tag{2.2.6}$$

Remark 2.2.9. *It follows from Proposition (2.2.8)*

$$\|\tau+\mu\|^2 = \|\tau\|^2 + \|\mu\|^2 + 2\langle \tau, \mu\rangle \quad \forall \tau, \mu \in H. \tag{2.2.7}$$

Proposition 2.2.10. *[18] For $\tau, \mu, k, u \in H$, we have*

$$2\langle \tau-\mu, k-u\rangle = \|\tau-u\|^2 + \|\mu-k\|^2 - \|\tau-k\|^2 - \|\mu-u\|^2. \tag{2.2.8}$$

Definition 2.2.11 (Parallelogram Law)**.** *[54] For all vectors $\tau, \mu \in H$, the following identity holds*

$$\|\tau+\mu\|^2 + \|\tau-\mu\|^2 = 2(\|\tau\|^2 + \|\mu\|^2). \tag{2.2.9}$$

Theorem 2.2.12. *[32] Let H be Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ defined by the relation*

$$\langle \tau, \mu\rangle = \frac{1}{2}\left(\|\tau\|^2 + \|\mu\|^2 - \|\tau-\mu\|^2\right), \tag{2.2.10}$$

if and only if the parallelogram law holds.

Lemma 2.2.13. *[18] Let $k, \tau, \mu \in H$. Then the following hold:*

(i)

$$\|k+\mu\|^2 = \|k\|^2 + 2\langle k, \mu\rangle + \|\mu\|^2.$$

(ii) Polarization Identity

$$4\langle \mu, k\rangle = \|\mu+k\|^2 - \|\mu-k\|^2.$$

(iii) Apollonius's Identity

$$\|\mu-k\|^2 = 2\|\tau-\mu\|^2 + 2\|\tau-k\|^2 - 4\left\|\tau - \frac{\mu+k}{2}\right\|^2.$$

2.2.3 Example of Hilbert space

(i) The space $\mathbb{R}^n$ is a Hilbert space having the following inner product ,

$$\langle a, c\rangle = a_1c_1 + \cdots + a_nc_n = \sum_{i=1}^{n} a_ic_i,$$

where $a = (a_1, a_2, \ldots, a_n)$ and $c = (c_1, c_2, \ldots, c_n)$ are elements of $\mathbb{R}^n$.

(ii) The space $L^2(\mathbb{R})$ is the space of real-valued functions such that

$$\int_{\mathbb{R}} |Q(x)|^2\, dx < \infty,$$

and it is a real Hilbert space, with an inner product

$$\langle Q, M\rangle = \int_{\mathbb{R}} Q(x)M(x)\, dx.$$

with

$$\|Q\| = \left(\int_a^c Q^2(x)\, dx\right)^{\frac{1}{2}},$$

where M and Q are real valued functions .

(iii) This is an infinite-dimensional generalization of (ii). Let

$$\ell^2(H) = \{(Q_1, Q_2, \dots) : Q_n \in H, \forall n \in \mathbb{N}, \text{ and } \sum_{n=1}^{\infty} \|Q_n\|_H^2 < +\infty\}.$$

For $Q = (Q_1, Q_2, \dots)$ and $M = (M_1, M_2, \dots)$ in $\ell^2(H)$, set

$$\langle Q, M\rangle_{\ell^2(H)} = \sum_{n=1}^{\infty} \langle Q_n, M_n\rangle_H.$$

2.2.4 Convergence in Hilbert space

Definition 2.2.14. *[32] A sequence $\{y_k\}_{k=0}^{\infty}$ of elements of a Hilbert space H converges weakly to $y \in H$ if for any $a \in H$, the sequence $\{\langle a, y_k\rangle\}_{k=0}^{\infty}$ converges to $\langle a, y\rangle$. We call the point y a weak limit of the sequence $\{y_k\}_{k=0}^{\infty}$ and write $y_k \xrightarrow{*} y$. If a subsequence $\{y_{n_k}\}_{k=0}^{\infty}$ of $\{y_k\}_{k=0}^{\infty}$ converges weakly to y, then y is called a weak cluster point of the sequence $\{y_k\}_{k=0}^{\infty}$. We say that $\{y_k\}_{k=0}^{\infty}$ converges (strongly) to y if $\lim_{k\to\infty} y_k = y = 0$. Weakly convergent sequences have the following properties:*

(i) A weakly convergent sequence $\{y_k\}_{k=0}^{\infty} \subseteq H$ has exactly one weak limit.

(ii) A weakly convergent sequence $\{y_k\}_{k=0}^{\infty} \subseteq H$ is bounded.

(iii) A bounded sequence $\{y_k\}_{k=0}^{\infty} \subseteq H$ includes a weakly convergent subsequence.

(iv) If a sequence $\{y_k\}_{k=0}^{\infty} \subseteq H$ is bounded and has exactly one weak cluster point $y \in H$, then $y_k \xrightarrow{*} y$.

(v) If a sequence $\{y_k\}_{k=0}^{\infty}$ converges to $y \in H$, then it converges weakly to $x \in H$.

(vi) A weakly convergent sequence $\{y_k\}_{k=0}^{\infty}$ of a finite dimensional Hilbert space H is convergent.

Theorem 2.2.15. *Let $\beta > 0$. A point $y \in C$ is a solution of the variational inequality (1.1.1) if and only if $y \in Fix_{P_C}(Id - Q)$.*

Theorem 2.2.16. *(Banach, 1922) Let B be a complete metric space and $Q : B \to B$ be a contraction mapping. Then Q has a unique fixed point $b \in B$. Moreover, for any $b_0 \in B$, the sequence $\{Q^k b_0\}_{k=0}^{\infty}$ converges to b at a geometric rate.*

Lemma 2.2.17. *[32] If $y_k \rightharpoonup y \in H$, then $\liminf_k \|y_k\| \geq \|y\|$.*

Proof. Let $y_k \rightharpoonup y$. $y = 0$ is the trival case . Now assume $y \neq 0$. By the Cauchy–Schwarz inequality, we have

$$\liminf_k \|y_k\|^2 \geq \liminf_k \langle y, y_k\rangle = \|y\|^2,$$

which implies $\liminf_k \|y_k\| \geq \|y\|$. □

Note that a sequence which is weakly convergent does not necessarily converge strongly.

Example 2.2.18. *Let $H = \ell^2$ and $y_k = e_k = (e_{k1}, e_{k2}, \dots)$, where*

$$e_{ki} = \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

Then $\|y_k\| = 1$, although the sequence $\{y_k\}_{k=0}^{\infty}$ does not contain a convergent subsequence because

$$\|y_k - y_l\| = \sqrt{2}, \quad \text{for all } k \neq l,$$

i.e., $\{y_k\}_{k=0}^{\infty}$ is not a Cauchy sequence. However, note that $y_k \rightharpoonup 0$.

Under additional assumptions, weak convergence implies strong convergence.

Lemma 2.2.19. *[32] If $y_k \rightharpoonup y \in H$ and $\|y_k\| \to \|y\|$, then $y_k \to y$ strongly.*

Proof. Let

A Hilbert space has an important property, expressed in the following lemma.

Lemma 2.2.20. *[32] If $x_k \rightharpoonup y \in H$, then for any $y_0 \in H$, $y_0 \neq y$, the following inequality holds:*

$$\liminf_k \|x_k - y_0\| > \liminf_k \|x_k - y\|. \tag{2.2.11}$$

Proof. Let $x_k \rightharpoonup y$ and $y_0 \neq y$. Denote $\delta = \|y - y_0\|^2$. Since a weakly convergent sequence is bounded, both lower limits in (2.2.11) are finite. The following give the properties of the inner product

$$\|x_k - y_0\|^2 = \|x_k - y + y - y_0\|^2 = \|x_k - y\|^2 + \|y - y_0\|^2 + 2\langle x_k - y, y - y_0\rangle.$$

By assumption, $\langle x_k - y, y - y_0\rangle \to 0$, so

$$\liminf_k \|x_k - y_0\|^2 = \liminf_k \|x_k - y\|^2 + \delta > \liminf_k \|x_k - y\|^2,$$

which implies

$$\liminf_k \|x_k - y_0\| > \liminf_k \|x_k - y\|.$$ □

Lemma 2.2.21. *Let H be the Hilbert space and $(y_n)_{n\in\mathbb{N}}$ be a bounded sequence in the space . Then $(x_n)_{n\in\mathbb{N}}$ possesses a weakly convergent subsequence.*

Definition 2.2.22. *[25] The minty type of the variational inequality problem (1.1.1) is given below by finding $y^* \in C$ such that,*

$$\langle Q(y, y - y^*\rangle \geq 0 \quad \forall y \in C \tag{2.2.12}$$

Remark 2.2.23. *For any $\lambda > 0$, y^* is a solution of VI(Q, C)(1.1.1) if and only if $y^* = P_C(y^* - \lambda Q y^*)$.*

2.3 Lebesgue spaces

2.3.1 Properties of Lebesgue spaces

Definition 2.3.1. *[102] For $x \in H$ the space $L^2(C)$ consists of all measurable functions $Q : C \to H$ such that ,*

$$\|Q\|_{L^2(C)} = \left(\int_C |Q(x)|^2 \, dx\right)^{1/2} < \infty. \tag{2.3.1}$$

Definition 2.3.2. *Given $p \in [1, \infty)$ and $x \in H$, the space $L^p(C)$ consists of all measurable functions $Q : C \to H$ such that,*

$$L_p(\mu) = \{Q : Q \text{ measurable}, \|Q\|_p < \infty\}, \tag{2.3.2}$$

where

$$\|Q\|_p = \left(\int |Q(x)|^p \, d\mu(x)\right)^{1/p}. \tag{2.3.3}$$

Definition 2.3.3. *[102] The space $L^\infty(C)$ consists of all measurable functions $Q : C \to H$ such that*

$$\|Q\|_{L^\infty(C)} := \operatorname{ess\,sup}_{x \in C} |Q(x)| < \infty, \tag{2.3.4}$$

where ess sup *denotes the essential supremum.*

2.3.2 Properties of locally Lebesgue space

Definition 2.3.4. *[102] The space $L^1_{loc}(\mathbb{R}^n)$ consists of all locally integrable functions on $\mathbb{R}^n$. A function $Q : \mathbb{R}^n \to \mathbb{R}$ belongs to $L^1_{loc}(\mathbb{R}^n)$ if Q is measurable and integrable over every compact subset of $\mathbb{R}^n$. this means*

$$Q \in L^1_{loc}(\mathbb{R}^n) \iff \int_K |Q(x)| \, dx < \infty \quad \text{for every compact set } K \subset \mathbb{R}^n. \tag{2.3.5}$$

Definition 2.3.5. *Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, and let $1 \le p < \infty$. The locally Lebesgue space $L^p_{loc}(\Omega)$ consists of measurable functions $Q : \Omega \to \mathbb{R}$ such that for every compact subset $K \subset \Omega$,*

$$\int_K |Q(x)|^p \, dx < \infty.$$

Equivalently,

$$L^p_{loc}(\Omega) = \{Q : \forall K \subset \Omega, \ K \text{ compact}, \ Q \in L^p(K)\}.$$

2.3.3 Examples

Example 2.3.6. *Let*

$$g(x) = \frac{1}{|x|} \tag{2.3.6}$$

for $x \neq 0$.

While $g \notin L^1(\mathbb{R})$ *globally (it diverges near* $x = 0$*), for any compact set* $K \subset \mathbb{R} \setminus \{0\}$*, the integral:*

$$\int_K |g(x)|\, dx < \infty. \tag{2.3.7}$$

Thus,

$$g \in L^1_{loc}(\mathbb{R}).$$

Example 2.3.7. *Let*

$$h(x) = e^{-x^2}$$

for $x \in \mathbb{R}$. *Then:*

$$\int_{-\infty}^{\infty} |h(x)|^2\, dx = \int_{-\infty}^{\infty} e^{-2x^2}\, dx = \sqrt{\frac{\pi}{2}} < \infty.$$

Hence,

$$h \in L^2(\mathbb{R}).$$

2.4 Metric projection

2.4.1 Non-expansive operator

Definition 2.4.1. *[48] Let* C *be a nonempty subset of* H*, and let* $Q : C \to H$*. Then* Q *is*

(i) firmly nonexpansive if

$$\forall \tau, \mu \in C, \quad \|Q\tau - Q\mu\|^2 + \|(Id - Q)\tau - (Id - Q)\mu\|^2 \leq \|\tau - \mu\|^2; \tag{2.4.1}$$

(ii) nonexpansive if it is Lipschitz continuous with having a comnstant that is equal to 1,

$$\forall \tau, \mu \in C, \quad \|Q\tau - Q\mu\| \leq \|\tau - \mu\|; \tag{2.4.2}$$

(iii) quasinonexpansive if

$$\forall \tau \in C, \forall \mu \in \operatorname{Fix} Q, \quad \|Q\tau - \mu\| \leq \|\tau - \mu\|; \tag{2.4.3}$$

(iv) strictly quasinonexpansive if

$$\forall \tau \in C \setminus \operatorname{Fix} Q, \forall \mu \in \operatorname{Fix} Q, \quad \|Q\tau - \mu\| < \|\tau - \mu\|. \tag{2.4.4}$$

Proposition 2.4.2. *Let* C *be a nonempty closed convex subset of* H*. Then the projector* P_C *is firmly nonexpansive.*

2.4.2 Existence and uniqueness

Theorem 2.4.3. *[32] For $C \subseteq H$ and any $x \in H$, then there exists a unique metric projection $P_C(x) \in C$.*

Theorem 2.4.4. *For $C \subset H$ and any $x \in H$, then there exists a metric projection P_Cx and it is uniquely determined.*

Proposition 2.4.5. *[32]. For $C \subseteq H$ and let Π_C be the orthogonal projection onto C. For all $x, b \in H$ and all $z \in C$, the following hold: $\|\Pi_C(x) - \Pi_C(b)\|^2 \leq \|x - b\|^2 - \|(x - \Pi_C(x)) - (b - \Pi_C(b))\|^2$;*

Definition 2.4.6. *For all $a \in H$ and $b \in C$ and $P_C : H \to C$ represent the projection operator , it holds*

$$\langle a - P_C(a), b - P_C(a)\rangle \leq 0. \tag{2.4.5}$$

2.4.3 Characterization of the metric projection

Theorem 2.4.7 (Characterization Theorem)**.** *[32] Let $y \in H$, $C \subset H$ be a convex subset, and $b \in C$. Then the following statements are equivalent,*

(i) $y = P_C(x)$,

(ii) $\langle y - b, c - b\rangle \geq 0$ *for all* $c \in C$.

Definition 2.4.8. *For $C \subset H$ then the metric projection onto C, also known as the nearest point projection, is the mapping $P_C : H \to C$ that assigns to each $x \in H$ a unique point $P_Cx \in C$ such that*

$$\|x - P_Cx\| = \min\{\|x - a\| : a \in C\}. \tag{2.4.6}$$

Lemma 2.4.9. *Let $\tau, w, \mu \in H$. The following conditions are equivalent:*

(i) $\langle \tau - w, \tau - w\rangle = 0$,

(ii) $\langle \tau - \mu, \tau - \mu\rangle = \|\tau - \mu\|^2$,

(iii) $\|\tau - w\|^2 = \|\tau - \mu\|^2 + \|\mu - w\|^2$,

(iv) $\langle \tau - \mu, \tau - w\rangle = 0$.

Lemma 2.4.10. *For $C \subset H$ and $P_C : H \to H$ be a metric projection onto C. Then for any $x \in H$, the following conditions are equivalent:*

(i) $y = P_Cx$,

(ii) $y \in C$ *and* $x - y \in N_C(y)$, *where* $N_C(y)$ *denotes the normal cone to* C *at* y.

Proposition 2.4.11. *[34]. For $C \subseteq H$ and $Q : C \to C$ be an inverse strongly monotone operator. If x^* is the solution set of the VIP (1.1.1), then for any $\mu > 0$, we have that $G\left(P_C(I - \mu Q)\right) = VIP(C, Q)$.*

Proof. Let $\mu > 0$ and $z \in C$, then from the characterization of the metric projection, we have that

$$\begin{aligned} w \in G(P_C(I-\mu Q)) &\iff w = P_C(I-\mu Q)w, \\ &\iff \langle w-(I-\mu Q)w, z-w\rangle \geq 0, \\ &\iff \langle \mu Q(w), z-w\rangle \geq 0, \\ &\iff \langle Q(w), z-w\rangle \geq 0, \\ &\iff w \in VIP(C,Q). \end{aligned}$$

Hence, $G\left(P_K(I-\mu Q)\right) = VIP(C,Q)$. □

2.4.4 Examples of metric projection

Example 2.4.12. *We now present several examples of metric projections in Hilbert spaces, highlighting their properties and applications.*

1. *Let $C = \{z \in H : \|z - z_0\| \leq \rho\}$, which represents a closed ball in H centered at $z_0 \in H$ with radius $\rho > 0$. The metric projection $P_C : H \to C$ assigns to each $z \in H$ the closest point in C. It is given explicitly by:*

$$P_C z = \begin{cases} z_0 + \rho\frac{(z-z_0)}{\|z-z_0\|}, & \text{if } \|z-z_0\| > \rho, \\ z, & \text{if } \|z-z_0\| \leq \rho. \end{cases}$$

   This means that if the point z lies outside the ball C, the projection $P_C z$ is the point on the boundary of C in the direction of z. On the other hand, if z is already inside C, the projection simply returns z itself.

2. *Let $C = \{x \in H : \|x\| \leq r\}$ be a closed ball centered at the origin with radius $r > 0$. Then, the metric projection onto C is:*

$$P_C x = \begin{cases} r\frac{x}{\|x\|}, & \text{if } \|x\| > r, \\ x, & \text{if } \|x\| \leq r. \end{cases}$$

   This is another example of the metric projection onto a closed ball, where the projection scales the vector x to ensure that its norm does not exceed r.

3. *Suppose that $H \neq \{0\}$ and let $\alpha \in]0,1]$. Let Q_1 and Q_2 be the soft and hard thresholders, respectively, defined for every $\tau \in H$ by*

$$Q_1\tau = \begin{cases} \left(1 - \frac{1}{\|\tau\|}\right)\tau, & \text{if } \|\tau\| > 1; \\ 0, & \text{if } \|\tau\| \leq 1. \end{cases} \tag{2.4.7}$$

   and

$$Q_2\tau = \begin{cases} \alpha\tau, & \text{if } \|\tau\| > 1; \\ 0, & \text{if } \|\tau\| \leq 1. \end{cases} \tag{2.4.8}$$

Then Q_1 is firmly nonexpansive. Moreover, for $\alpha < 1$, Q_2 is quasinonexpansive but not nonexpansive, and for $\alpha = 1$, Q_2 is not quasinonexpansive.

Finally, the operator Q_3 defined for every $\tau \in H$ by

$$Q_3\tau = \begin{cases} \left(1 - \frac{2}{\|\tau\|}\right)\tau, & \text{if } \|\tau\| > 1; \\ -\tau, & \text{if } \|\tau\| \leq 1. \end{cases} \tag{2.4.9}$$

is nonexpansive but not firmly nonexpansive.

In Chapters 3 ,4 and 5 we will observe that the operator $P_C(I - \lambda Q)y(t)$ plays a key role in proving our main theorems.

Theorem 2.4.13. *[32] Let $G : D \to D$ be a continuous mapping on a closed set $D \subset \mathbb{R}^n$. Then G has at least one fixed point.*

2.5 Variational inequality problem

The following theorem helps us explore whether a unique solution exists for the general variational inequality.

Theorem 2.5.1. *[18] Let u and v be two distinct solutions to the variational inequality (1.1.1). If the condition*

$$\langle Q(u) - Q(v), u - v\rangle \geq 0, \tag{2.5.1}$$

is satisfied, where Q is the underlying operator then the solution to the variational inequality (1.1.1) in H is unique.

Proof. Let $u, v \in H$ be two distinct solutions of the variational inequality (1.1.1), meaning:

$$\langle Q(u), z - u\rangle \geq 0 \quad \text{for all } z \in H,$$

$$\langle Q(v), w - v\rangle \geq 0 \quad \text{for all } w \in H.$$

Now, choose $z = v$ and $w = u$. Adding these two inequalities gives:

$$\langle Q(u) - Q(v), u - v\rangle \leq 0.$$

Since condition (2.5.1) holds, we conclude that it is impossible for $u \neq v$. Thus, the solution to the variational inequality is unique. The condition (2.5.1) defines a strictly monotone mapping, ensuring the uniqueness of the solution. □

2.5.1 Examples

Example 2.5.2. *(a) Consider the following variational inequality problem (VIP), where the function Q and the set C are defined as:*

$$Q(y) = \begin{bmatrix} 22y_1 - 2y_2 + 6y_3 - 4 \\ 2y_2 - 2y_1 \\ 2y_3 + 6y_1 \end{bmatrix}, \quad C = \left\{ y \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} y_1 - y_2 \geq 1, \\ -3y_1 - y_3 \geq -4, \\ 2y_1 + 2y_2 + y_3 = 0, \\ l \leq y \leq h \end{array} \right\},$$

where $l = (-6, -6, -6)^T$ *and* $h = (6, 6, 6)^T$*. This VIP is linear, and after solving the system, we find that the problem has a unique solution:*

$$y^* = (2, 1, -6)^T \quad \textit{(refer to [107])}.$$

This example demonstrates how a linear VIP with specific constraints can be solved to yield a unique solution.

(b) Consider a second VIP where the function Q *and the set* C *are given by:*

$$Q(y) = \begin{bmatrix} 3y_1 - 1 \\ y_1 + 3y_2 - 2 \\ 3y_1 + 3y_2 \\ 4y_3 + 4y_4 \\ 4y_3 + 4y_4 - 1 \\ y_4 - 3 \end{bmatrix}, \quad C = \left\{ y \in \mathbb{R}^n \;\middle|\; \begin{array}{l} y_1 + y_2 = 1, \\ y_3 + y_4 \geq 0, \\ l \leq y \leq h \end{array} \right\},$$

where $p = (0.2, 0, 0, 0.5)^T$ *and* $q = (8, 8, 8, 8)^T$*. This variational inequality problem is nonlinear, and solving the system leads to a unique solution:*

$$y* = (1.5, 0, 0, 0.5)^T \quad \textit{(see [107])}.$$

This example shows how a nonlinear VIP can also have a unique solution under certain conditions, even with additional constraints like the equality $y_1 + y_2 = 1$ *and the inequality* $y_3 + y_4 \geq 0$.

2.5.2 Related problems

Minty Variational Inequality Problems

The Minty variational inequality problem was introduced and studied by Minty in 1967. In his work, Minty explored the relationship between VIPs and MVIPs under the operator's assumptions of continuity and monotonicity Q.

For $C \subseteq H$ *and let* $Q : H \to H$ *be a nonlinear operator. The Minty Variational Inequality Problem (MVIP) is the problem of finding* $y^* \in C$ *such that:*

$$\langle Qy, y - y^* \rangle \geq 0 \quad \forall y \in C. \tag{2.5.2}$$

We denote the solution set of the MVIP by $MVIP(C, Q)$.

Remark 2.5.3. *For any* $\lambda > 0$, y^* *is a solution of VI*(Q, C)(1.1.1) *if and only if* $y^* = P_C(y^* - \lambda Q y^*)$.

Theorem 2.5.4. *[71] Let* H *be a real Hilbert space and* C *a nonempty, closed, and convex subset of* H*. Then, the problems* $MVIP(C, Q)$ *and* $VIP(C, Q)$ *are equivalent if and only if the operator* $Q : H \to H$ *is monotone and continuous.*

2.5.3 Minimization

Since variational forms are often linked to minimization problems, we will first study their connection to minimization. To understand the connection between minimization problems and variational inequalities, we study the following theorem.

Theorem 2.5.5. *Let $C \subset \mathbb{R}^d$ be a nonempty, closed, and convex set. Suppose $Q \in C^1(C)$ and let $y^* \in C$ be the solution to the minimization problem*

$$Q(y^*) = \min_{w \in C} Q(w). \tag{2.5.3}$$

Then $y^* \in C$ is also a solution to the variational inequality: Find $y^* \in C$, such that

$$(\nabla Q(y^*), w - y^*) \geq 0 \quad \text{for all } w \in C. \tag{2.5.4}$$

To prove the reverse direction it requires us to use the concept of convexity which the proof is provided in this book [86]. The following theorem shows that z is the minimiser by solving (2.5.4).

Theorem 2.5.6. *[32] Let $Q \in C^1(L)$ be convex and let $z \in L$ satisfy the inequality (2.5.4). Then $z \in L$ solves the minimization problem (2.5.3).*

Proof. Let $z \in L$ be a solution to the inequality (2.5.4). Since Q is convex, we know by definition that for any $a \in L$ and $k \in [0, 1]$, the following holds:

$$Q(ka + (1-k)z) \leq kQ(a) + (1-k)Q(z).$$

we get that,

$$Q(z + k(a - z)) \leq Q(z) + k(Q(a) - Q(z)),$$

which implies

$$Q(a) \geq Q(z) + \frac{Q(z + k(a-z)) - Q(z)}{k}, \quad \text{for all } k \in (0, 1].$$

Taking the limit as $k \to 0$, we obtain

$$Q(a) \geq Q(z) + (\nabla Q(z), a - z), \tag{2.5.5}$$

Since z satisfies (2.5.4), it follows from (2.5.5) that

$$Q(a) \geq Q(z) \quad \text{for all } a \in L,$$

which shows that z is indeed the minimiser. □

2.6 First order Methods

The first-order method dynamical system forms the fundamental methods used to study problems in many fields like optimization, equilibrium problems, and variational inequality. The method simply incorporates only the velocity term to enhance convergence without the acceleration which you will see later the importance of the acceleration term. We will study the first-order method from the continuous and discrete case for some methods to solve variational inequality and optimization problems. Most first-order methods have the same convergence rate in both discrete and continuous cases, as demonstrated below in the section .

2.6.1 Optimization

In this section, we will study the first-order methods in optimization. These methods, including gradient descent and its accelerated variants like Nesterov's method, are particularly effective for large-scale optimization in solving convex and non-convex problems.

Continuous Gradient Flow

We will use the continuous part to study the behavior of the trajectories. We will start with the continuous gradient flow method, which forms the foundation for studying optimization problems. The following differential equation governs the system:

$$\dot{y}(t) + \lambda \nabla Q(y(t)) = 0, \tag{2.6.1}$$

where $\nabla Q(y(t))$ *is the gradient of the function* Q*. The parameter* $\lambda > 0$ *is the stepsize. This framework is beneficial for studying the asymptotic behavior of optimization algorithms in a continuous-time setting. As the limit of* $t \to \infty$ *the trajectory* $y(t)$ *converges to a critical point of* $Q(y)(t)$*. Specifically, the state* $y(t)$ *converges to* $\bar{y} \in argmin_k Q$*, where* k *is the domain of* Q*. Jorge Cortes [37] have normalised the gradient method to get the convergence of the solution to the optimal solution in the finite time and further Fei Chen and Wei Ren [35] consider it for the case of the convex optimization where they have the assumption that the objective function is strong convexity and additional constructing a discontinuous dynamics with equality constraints they have achieved infinite time convergence to the optimal solution. More work has been done see [82]. Additionally, Kunal Garg and Dimitra Panagou [44] studied the modification of the method to study minimisation - maximum problems and also convex optimization for both constrained and unconstrained. For the application of the method see [36]. The key takeaway is that appropriately selecting* λ*, this approach ensures stability and convergence of the method to the solution set .*

Remark 2.6.1. *For the case of explicit discretization, we have the following method*

$$y_{n+1} = y_n - \lambda \nabla Q(y_n). \tag{2.6.2}$$

The method achieves weak convergence in the discrete case, similar to the continuous case.

Projected Gradient Method

The method is used mostly in constraint optimization problems. The method is the combination of the continuous gradient method with a projection step to ensure that the updated state stays $x(t)$ within the feasible region C. The following differential equation gives the continuous-time formulation of the method,

$$\dot{y}(t) + y(t) = P_C\left(y(t) - \lambda \nabla Q(y(t))\right), \tag{2.6.3}$$

where $y(t)$ is the state at time t, $\lambda > 0$ is the step size, $\nabla Q(y(t))$ is the gradient of the objective function at chi(t), and $P_C(\cdot)$ is the projection operator that maps a point onto the set C. The existence of a solution on the finite time interval is guaranteed since the projection operator is continuous. For the infinite case, we require that $\nabla Q(y(t))$ is Lipschitz continuous, which will ensure that the trajectories not only exist but are unique. Furthermore, when $\nabla Q(y(t))$ is locally Lipschitz, the trajectories converge weakly to the constrained minimizer. If we let the projection operator be the unit matrix, we get (2.6.1), which is considered above. Boris T Polyak [85] has studied the discrete projection method and obtained convergence to the optimal solution. The convergence of (2.6.3) part has been studied by Antipin [99]. Jerome Bolte [22] studied the extension of the method where he included the Tykhonov term that is used to force the trajectories $y(t)$ to reach the equilibrium. The method is defined as follows:

$$\dot{y}(t) + y(t) = P_C\big[y(t) - \lambda \nabla Q(y(t)) - \mu(t) y(t)\big], \tag{2.6.4}$$
$$y(0) = y_0 \in C,$$

where $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ satisfies

$$\mu(t) \to 0 \quad as \quad t \to \infty,$$

and

$$\int_0^\infty \mu(t)\, dt = \infty.$$

This approach shows that the trajectories can be forced to converge strongly toward the minimizer. The convergence speed of the method depends on the choice of the step size λ, and in practice, an adaptive choice of λ can help improve the efficiency of the method.

Proximal Gradient Method

In the literature, they involve non-smooth terms that extend the gradient descent method by the proximal operator. The continuous proximal method is defined as follows:

$$\dot{y}(t) + y(t) = prox_{\lambda g}\left(y(t) - \lambda \nabla Q(y(t))\right), \tag{2.6.5}$$

where $y(t)$ is the state at time t, $\lambda > 0$ is the step size, $\nabla Q(y(t))$ is the gradient of the objective function at $y(t)$, and $prox_{\lambda g}$ is the proximal operator that maps a point onto the *set C. The proximal operator $prox_{\lambda g}$ is defined as:*

$$prox_{\lambda g}(v) = \arg\min_{y} \left(g(y) + \frac{1}{2\lambda}\|y - v\|^2 \right), \tag{2.6.6}$$

which can be interpreted as a regularization step that modifies the update based on the non-smooth function $g(y)$. The method is a forward-backward method, where the first step applies the gradient descent method to update the $y(t) - \lambda\nabla Q(y(t))$, and the second part is to apply the proximal operator to the non-smooth term $g(y)$. When $g(y) = 0$, then the proximal operator is the identity and the method becomes (2.6.1). The study of the proximal gradient method was first introduced by Moghaddam and Jovanovic [53] where they have proven that the trajectories achieve exponential convergence. The idea was further improved by Gokhale, Alexander and Bullo [47] where they introduce a new condition that guarantees exponential convergence of the objective function to the optimal solution. They also proved that the proximal Polyak-Lojasiewicz (PL) condition guarantees exponential convergence of the objective function to the optimal solution. Abbas and Attouch [1] studied the case of the monotone inclusion problem, where they considered both the continuous and discrete parts. The key takeaway is that to achieve convergence the method requires both $Q(y)$ and $g(y)$ to be convex and $Q(y)$ to be smooth then the convergence to the minimizer of the objective function is guaranteed. For the application of the method, see [47, 53, 1].

2.6.2 Variational inequality

Projected Gradient Method

The method is used mostly to solve variational inequality problems. The method includes the projection step to ensure that the updated state stays $y(t)$ within the solution set Ω.The following differential equation gives the continuous-time formulation of the method,

$$\dot{y}(t) + y(t) = P_C\left(y(t) - \lambda Q(y(t))\right), \tag{2.6.7}$$

where $y(t)$ is the state at time t, $\lambda > 0$ is the step size, $Q(y(t))$ is the monotone operator, and $P_C(\cdot)$ is the projection operator that maps a point onto the set C. Youshen Xia, Henry Leung [104] studied (2.6.7) when Q is monotone, he have demonstrated the global convergence and global asymptotic stability of the trajectories. Cavazzuti and Pappalardo [31] studied the case when the underlying operator is μ strongly monotone, they show that the dynamical system converges exponentially. Phan Tu Vuong [50] considers the case when Q is strong pseudo-monotonicity and Lipschitz continuity assumptions, he proves that the dynamical system has a unique equilibrium solution. Pham Duy Khanh and Phan Tu Vuong [65] have studied the explicit discretisation of the method, which is Korpelevich's extragradient method for the case when the underlying operator is the strong pseudomonotone, and they have obtained the linear convergence. The results obtained for projection show that there is a connection between the algorithmic and dynamical system approaches, which yield similar results.

The projected method serves as a fundamental approach for solving variational inequalities. As stated above, we will introduce Tseng's method to improve convergence behavior by combining projection steps with additional correction terms.

The continuous Tseng method, which includes the correction term, is given below:

Forward-Backward-Forward dynamical system

The forward-backward-forward dynamical system is described as:

$$\begin{cases} w(t) = P_C\big(y(t) - \lambda Q(y(t))\big), \\ \dot{y}(t) + y(t) - w(t) - \lambda\big[Q(y(t)) - Q(w(t))\big] = 0, \\ y(0) = y_0, \end{cases} \tag{2.6.8}$$

where $\lambda > 0$ and $x_0 \in H$. Here Q is the underlying operator that can be monotone or pseudomonotone. Radu I Botand Erno Robert Csetnek [25] studied the system when the operator is pseudomonotone and with the additional assumption that the function is weak to weak sequentially. Using those assumptions, they were able to prove the weak convergence of the trajectories to the solution set of (1.1.1)*. In the case of explicit time discretization of* (2.6.8) *they have also shown linear convergence. For the case when we replace the projection operator with the resolvent operator, we get the more generalized problem of the variational inequality problem which is the monotone inclusion problem where Radu Ioan Bot and Erno Robert Csetnek [26] have proven not only the existence and uniqueness of the trajectories but the weak convergence to the solution set of Monotone inclusion problem.*

In our paper, we are interested in extending the first-order methods to the second-order methods. We will study the methods of second-order dynamical systems to solve variational inequalities problems and optimization problems. The methods provided below show that including the acceleration term indeed improves the rate of convergence. We will also show the rate of convergence. We will start with the method of Polyak, followed by the subsequence methods of the second-order dynamical system.

2.7 Second order methods

In this section, we will study methods that include an acceleration term in their dynamical system from a continuous perspective. We will compare their convergence rate with the corresponding first-order method, highlighting the impact of the acceleration term methods. The methods in this section have constant damping, asymptotic damping, and vanishing damping. The damping force influences the convergence speed of dynamical systems and ensures that the velocity is controlled over time, leading the trajectory to converge to a steady-state solution, ideally close to the optimal point.

2.7.1 Optimization

The heavy ball with friction method of Polyak

We will use the continuous part to study the behavior of the trajectories. Since the method includes the acceleration term, it is used to accelerate convergence. We will begin with the extension of the continuous gradient flow method, where Polyak [84] studied the second-order version given below.

$$\ddot{y}(t) + \alpha \dot{y}(t) = -\nabla Q(y(t)) \tag{2.7.1}$$

where, $\alpha > 0$ is a fixed parameter. Alvarez [6] has proven that the trajectories converge to the minimizer of Q. For the μ strongly convex case, we have the following :

$$\ddot{y}(t) + 2\sqrt{\mu}\dot{y}(t) = -\nabla Q(y(t)), \tag{2.7.2}$$

where μ is a positive constant. Attouch and Csetnek [11] in their study have established minimal assumptions on the damping potential Q to ensure the convergence of trajectories and further improve the convergence rate from linear rate to exponential rate given below by,

$$Q(y(t)) - \inf_H Q = O\left(\frac{1}{t}\right) \quad \textit{as } t \to +\infty. \tag{2.7.3}$$

to

$$Q(y(t)) - \inf_H Q = O\left(e^{-\sqrt{\mu}t}\right) \quad \textit{as } t \to +\infty. \tag{2.7.4}$$

The value of $Q(y(t))$ does not generally decrease along the trajectories; rather, it is the energy $s(t) = \frac{1}{2}|\dot{y}(t)|^2 + Q(y(t))$ that decreases over time. Other authors studied [69].

Projection operator viscous damping

The second-order method for the equation (2.6.3) is given below :

$$\ddot{y}(t) + \alpha \dot{y}(t) + y(t) - P_C(y(t) - \epsilon \nabla Q(y(t))) = 0 \tag{2.7.5}$$

where $\epsilon > 0$ and Q is the underlying operators used to study the system. The convergence properties of trajectories toward a global minimizer of Q over the set C have been thoroughly studied and are well known in the literature. Antipin [99] has proven the asymptotic and exponential stability of the generated trajectories by dynamical system (2.7.5).

For the general case, when we replace the gradient operator by Q, which is a non-expansive operator, we get,

$$\ddot{y}(t) + \alpha \dot{y}(t) = (I - Q)y(t), \tag{2.7.6}$$

where I is the identity operator. Attouch and Alvarez [8] studied the case when $\alpha^2 > 2$, then the trajectories generated by (2.7.6) converge weakly to the element that is in the fixed point set of Q. The presence of the α damping factor plays an important role in convergence. α should not be larger since it will discard the impact of the acceleration term, and $\alpha\dot{y}(t)$ will dominate and lead to the first-order method. If we choose alpha to be small, the system will keep on oscillating and lead to slow convergence. We require α to depend on time such that it approaches zero but not too quickly. Below, we study a method that addresses this issue by allowing α to be $\alpha(t)$.

Asymptotic vanishing damping

The study of dynamical systems with time-dependent damping coefficients is

$$\ddot{y}(t) + \alpha(t)\dot{y}(t) = -\nabla Q(y(t)), \quad t \geq t_0, \tag{2.7.7}$$

where $\alpha(t)$ *represents a time-dependent damping coefficient and* $Q(y(t))$ *is a potential . When* $a(t) = \frac{1}{t+1}$*, the resulting trajectories correspond to those of an averaged gradient system. Alexandre and Engler [29] used the optimization property given by* $\int_{t_0}^{+\infty} \alpha(t)\, dt = +\infty$*. In their paper, they have proven that this ensures that the objective function converges toward its minimum. The property below assists in ensuring that the trajectories converge towards the minimum value of the function.*

$$\int_0^{\infty} e^{-\int_0^t \alpha(s)\, ds} dt < \infty. \tag{2.7.8}$$

Alexandre and Engler [29] studied the stronger condition sufficient for convergence. Su, Boyd, and Candés [94] considered $\alpha = \frac{\beta}{t}$*, which will lead to the Nesterov accelerated method given below.*

Nesterov accelerated method

The Nesterov accelerated gradient method can be studied using continuous-time dynamics. The modified system is expressed by the following equation:

$$\ddot{y}(t) + \frac{\alpha}{t}\dot{y}(t) + \nabla Q(y(t)) = 0, \quad t > 0, \tag{2.7.9}$$

where $\alpha > 0$ *is a parameter that influences the rate of damping. The term* $\frac{\alpha}{t}\dot{y}(t)$ *represents a time-varying damping coefficient. The parameter* α *assists in determining the system's convergence properties. Su-Boyd-Candès [49] studied it for the case when* $\alpha = 3$*, the system achieves an optimal convergence rate given below by,*

$$O\left(\frac{1}{t^2}\right) \quad \text{as } t \to +\infty. \tag{2.7.10}$$

Attouch-Chbani-Peypouquet-Redont [12] have studied the case when $\alpha > 3$*, the convergence rate have improved to ,*

$$o\left(\frac{1}{t^2}\right), \quad \text{as } t \to +\infty. \tag{2.7.11}$$

They have achieved the weak convergence of the trajectories generated by (2.7.9). Attouch, Hedy and Chbani [13] studied the case when $\alpha \leq 3$*, the system converges at the rate ,*

$$O\left(\frac{1}{t^{\frac{2\alpha}{3}}}\right) \quad \text{as } t \to +\infty. \tag{2.7.12}$$

Finally, the combination Polyak method (2.7.1) and Nesterov accelerated method (2.7.9) gives the following equation,

$$\ddot{y}(t) + (\frac{\alpha}{t} + \mu)\dot{y}(t) + \nabla Q(y(t)) = 0, \tag{2.7.13}$$

where μ *is a viscous damping term. The combination of the methods improves the rate of convergence.*

2.7.2 Variational inequality

Phan Tu Vuong [100] extended the projection method proposed by Antipin [99] from solving optimization problems to variational inequality by replacing the gradient operator with a strong pseudomonotone operator. The equation is given by

$$\begin{cases} \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta(t)\left[y(t) - P_C\left(y(t) - \lambda Q y(t)\right)\right] = 0, \\ \quad y(0) = y_0, \quad \dot{y}(0) = y_0. \end{cases} \tag{2.7.14}$$

However, the dynamical system given by (2.7.14) does not produce trajectories that converge to $C \subseteq H$ if Q is monotone. The result obtained for the second-order dynamical system shows better convergence than the first-order dynamical system. The overall reason is that the first-order methods rely solely on gradient information, which may result in slow convergence, especially in ill-conditioned problems, and the second-order method relies on constant damping, asymptotic damping, and more to improve convergence. Rapid speed alone is not sufficient when solving variational inequality or optimization, we are required to control our trajectories, which the next section will deal with. Before we move to the methods using the Hessian damping term to manage the trajectories, here is an example in mechanics that uses the dynamical system to model the behavior of objects.

We provide an example of Newton's method applied to the spring problem.

Example 2.7.1. *Newton's second law states that the total force acting on an object is equal to the product of its mass and acceleration. We have that $x(t)$, $\dot{x}(t)$, $\ddot{x}(t)$ represent displacement, velocity and acceleration respectively and m is mass of the object, $f(t)$ is the external force, $c^2x(t)$ corresponds to the restoring force due to the compression of the spring, and $2k\dot{x}(t)$ denotes the damping force exerted by the absorber*

$$m\ddot{x}(t) = f(t) - c^2x(t) - 2k\dot{x}(t) \tag{2.7.15}$$

Rearranging the equation, we obtain:

$$m\ddot{x}(t) + 2k\dot{x}(t) + c^2x(t) = f(t) \tag{2.7.16}$$

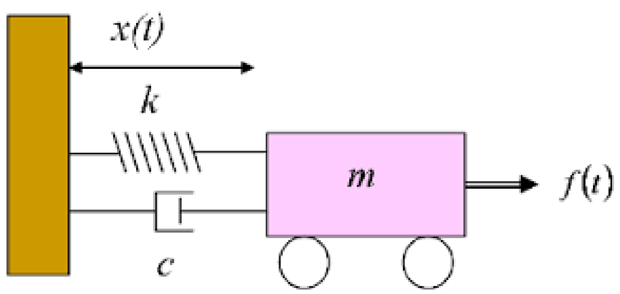

Figure 1 - Spring system. In this system, we can study the force of the spring applied to the wall and the object of mass m. Furthermore, we can also find the kinetic energy, which is equal to the mechanical energy, since the potential energy is equal to zero.

We are going to extend our study to include the Hessian damping factor. The convergence of the trajectories generated by (2.7.14) can be improved by including the Hessian damping factor alongside the Tseng forward-backward forward, which includes the correction forward step term. In algorithms, damping helps regulate the momentum that propels the iterates forward. With too little damping, the iterates may overshoot, oscillate, or even diverge. With too much damping, the system will have slow convergence. We will modify our system to incorporate Hessian damping to address the oscillations. The Hessian method is more effective than the standard damping method. Below, we study the well-known methods that use the Hessian damping factor to neutralise oscillation.

2.8 Hessian damping

In this section, we will study methods that include a Hessian damping term in their dynamical system from a continuous perspective. We will compare their convergence rate with the corresponding second-order method without the Hessian damping term, highlighting the impact of the Hessian damping term. The Hessian damping + standard damping influences the convergence speed of dynamical systems and ensures that the velocity is controlled over time with less oscillation, leading the trajectory to converge to a steady-state solution, ideally close to the optimal point.

2.8.1 Optimization

Heavy Ball with Friction (HBF)

Alvarez [6] studied the following equation.

$$\ddot{y}(t) + X\dot{y}(t) + \nabla Q(y(t)) = 0, \tag{2.8.1}$$

where $X : H \to H$ is a positive definite matrix called the anisotropic. This formulation models physical systems with inertia and friction and has been extensively studied in optimization. Replacing the anisotropic by a Hessian matrix, we study the second-order dynamical inertial Newtonian system, which is a well-posed formulation of the continuous Newton method.

Dynamical Inertial Newtonian (DIN)

The DIN method introduces Hessian-based corrections to improve stability and convergence rates:

$$\ddot{y}(t) + \alpha\dot{y}(t) + \beta\nabla^2 Q(y(t))\dot{y}(t) + \nabla Q(y(t)) = 0, \tag{2.8.2}$$

where, $\beta > 0$ governs the contribution of the Hessian $\nabla^2 Q(y(t)$ and $\beta\nabla^2 Q(y(t))\dot{y}(t)$ serves to neutralise oscillations, providing a smoother and faster convergence, similar to Nesterov's accelerated gradient method. The method analysed by Alvarez, Attouch, Bolte, and Redonte [7] showed that when Q is convex, then each trajectory of (2.8.2) weakly converges to a minimiser of Q. Furthermore, the trajectories converge to the critical value of Q if Q is real analytic. For the application of the method, see [96].

Remark 2.8.1. *The above system can be written in its equivalent first-order method. The system below represents the equivalence of (2.8.2).*

$$\begin{cases} \dot{y}(t) + \beta\nabla Q(y(t)) - \left(\frac{1}{\beta} - \frac{\alpha}{t}\right) y(t) + \frac{1}{\beta} z(t) = 0, \\ \dot{z}(t) - \left(\frac{1}{\beta} - \frac{\alpha}{t} + \frac{\alpha\beta}{t^2}\right) y(t) + \frac{1}{\beta} z(t) = 0, \end{cases} \tag{2.8.3}$$

where β is the relaxation factor, α is the viscous damping and Q is our operator . The system has the same convergence as (2.8.2).

Dynamical Inertial Accelerated Variational Dynamics (DIN-AVD)

The DIN-AVD method combines inertial acceleration with Hessian-based corrections:

$$\text{(DIN-AVD)}_{\alpha,\beta} \quad \ddot{y}(t) + \frac{\alpha}{t}\dot{y}(t) + \beta\nabla^2 Q(y(t))\dot{y}(t) + \nabla Q(y(t)) = 0, \tag{2.8.4}$$

where $\alpha > 0$ introduces time-dependent damping. This approach was studied by Attouch [14] to solve a monotone inclusion problem, and he achieved weak convergence.

Dynamical system including three factors

The following dynamical system for the inertial gradient system includes three factors, which are the Hessian-driven damping, the asymptotic vanishing term, and the time-varying term. The system is given by the following differential equation,

$$\ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta(t)\nabla^2 Q(y(t))\dot{y}(t) + b(t)\nabla Q(y(t)) = 0, \tag{2.8.5}$$

where $\alpha(t)$ is the time-dependent asymptotic damping coefficient, $\beta(t)$ is the scaling coefficient for the Hessian-driven damping, and $b(t)$ is the coefficient associated with the gradient term driving convergence.

Hedy Attouch [10] demonstrates that the proposed dynamics achieve exponential convergence rates for function values under certain assumptions, even without requiring the strong convexity of $Q(y)$.

Our proposed method, which includes the Hessian damping factor, is presented below and will be the focus of this paper. The method is define as follows:

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta\frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Qy(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0 \end{cases} \tag{2.8.6}$$

where P_C is the projection operator, and let $\alpha(t)$ be the asymptomatic vanishing damping which is a Lebesgue measurable function, and Q can be a monotone or strongly monotone operator and Lipschitz continuous.

Chapter 3

Existence and Uniqueness of Strong Global Solution

3.1 Introduction

In this chapter, we begin by assuming that the underlying operator is only Lipschitz continuous. Using the Cauchy-Lipschitz-Picard theorem, we establish the existence and uniqueness of the trajectories generated by the dynamical system (3.1.1). Furthermore, we demonstrate that the acceleration term of the dynamical system is absolutely continuous, an essential property that ensures the integrability and smoothness of the system's evolution over time. In addition to these foundational results, we derive an explicit estimate for the third-order derivative of the dynamical system (3.1.1). This analysis assists in understanding the stability and convergence properties of the system. Additionally, we close the chapter by providing two examples. We are going to solve equation (1.1.1) using the proposed system, which is:

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)), \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Qy(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0 \end{cases} \tag{3.1.1}$$

where P_C is the projection operator, and let $\alpha(t)$ be the asymptomatic vanishing damping, which is a Lebesgue measurable function, and Q be Lipschitz continuous.

Let $R : H \to H$ be defined as,

$$Ry(t) = w(t) + \lambda \left[Q(y(t)) - Q(w(t)) \right]. \tag{3.1.2}$$

Then we have (3.1.1) can be rewritten as follows.

$$\ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) - Ry(t) = 0. \tag{3.1.3}$$

3.2 Preliminaries

We are going to present some lemmas and definitions required to establish this chapter's main results. We denote the strong and weak convergence of a sequence $\{y_n\}$ to a point $y \in H$ by $y_n \to y$ and $y_n \rightharpoonup y$, respectively.

Definition 3.2.1. *Let (Q, d) be a metric space with a distance function d. A mapping $Q : N \to N$ is termed a* contraction mapping *if:*

$$d(Q(\tau), Q(\mu)) \leq \beta d(\tau, \mu) \quad \text{for all } \tau, \mu \in N, \text{ where } \beta \in [0, 1).$$

When $\beta = 1$, the mapping is called non-expansive.

Theorem 3.2.2. *[18] Let N be a complete metric space, and let $Q : N \to N$ be a contraction mapping. Then there is exactly one point $p \in N$ such that $Q(p) = p$.*

Lemma 3.2.3. *[21]*
Let $y \in C^1[a, b]$ and if $R"[a, b]$ denotes the set of Riemann integrable functions on $[a, b]$, $\beta \in R"[a, b]$, and

$$y'(t) \leq \beta(t) \cdot y(t), \quad \forall t \in [a, b],$$

then

$$y(t) \leq y(a) \cdot e^{\int_a^t \beta(s)\, ds}, \quad \forall t \in [a, b].$$

The natural generalization of the Gronwall inequality.

Lemma 3.2.4. *[41] Let $y \in C^1[c, d]$, $y, \delta \in R"[c, d]$. If*

$$y'(t) \leq y(t) + \delta(t)y(t), \quad \forall t \in [c, d],$$

then

$$y(t) \leq \left(y(c) + \int_c^t e^{-\int_c^s \delta(\tau)\, d\tau} y(s)\, ds \right) \cdot e^{\int_c^t \delta(s)\, ds}, \quad \forall t \in [c, d].$$

Definition 3.2.5. *[25] Let $y : [t_0, +\infty) \to \mathbb{R}^n$ be a strong global solution of (3.2.5), if the properties below are satisfied,*

(i) $y, \dot{y} : [t_0, +\infty) \to \mathbb{R}^n$ *are locally absolutely continuous, which means that absolutely continuous on each interval $[t_0, T]$ where $t_0 < T < +\infty$;*

(ii)

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) = R(y(t)) \end{cases}$$

for almost every $t \geq t_0$;

(iii) $y(t_0) = u_0$ *and* $\dot{y}(t_0) = v_0$.

Definition 3.2.6. *A function $y : [t_0, +\infty) \to \mathbb{R}^n$ is said to be* absolutely continuous *on an interval $[t_0, T]$ if there exists an integrable function $x : [t_0, T] \to \mathbb{R}^n$ satisfying*

$$y(t) = y(t_0) + \int_{t_0}^t x(s)\, ds \quad \text{for all } t \in [t_0, T]. \tag{3.2.1}$$

Theorem 3.2.7 (Rademacher's Theorem). *Let $Q : C \to \mathbb{R}^n$ where C is the subset of $\mathbb{R}^n$ be a Lipschitz continuous function. Then, Q is differentiable almost everywhere in $\mathbb{R}^n$.*

Lemma 3.2.8 (Young's Inequality). *Let x and y be two positive real numbers such that*

$$\frac{1}{x} + \frac{1}{y} = 1. \tag{3.2.2}$$

Then, for any positive real numbers a and b,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq ab. \tag{3.2.3}$$

Equality holds if and only if $a^x = b^y$.

Proof. Using the (2.1.3) and let the weight be given by $m_1 = \frac{1}{x}$ and $m_2 = \frac{1}{y}$. we have the following equation.

$$\frac{m_1 a^x + m_2 b^y}{m_1 + m_2} \geq \left((a^x)^{m_1} (b^y)^{m_2}\right)^{\frac{1}{m_1+m_2}}. \tag{3.2.4}$$

Having that,

$$\frac{1}{x} + \frac{1}{y} = 1 \implies m_1 + m_2 = 1, \tag{3.2.5}$$

substituting this into (3.2.4) it gives the following equation ,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq (a^x)^{\frac{1}{x}} (b^y)^{\frac{1}{y}}. \tag{3.2.6}$$

Given that this equality holds,

$$(a^x)^{\frac{1}{x}} = a, \quad (b^y)^{\frac{1}{y}} = b, \tag{3.2.7}$$

Therefore,

$$\frac{a^x}{x} + \frac{b^y}{y} \geq ab. \tag{3.2.8}$$

This is true if and only if $a^x = b^y$, therefore the Young's inequality. □

Lemma 3.2.9. *[103] Let $(\tau_1, \tau_2, \ldots, \tau_n)$ and $(\mu_1, \mu_2, \ldots, \mu_n)$ be two sequences of real numbers. Then,*

$$\left(\sum_{i=1}^{n} \tau_i^2\right)\left(\sum_{i=1}^{n} \mu_i^2\right) \geq \left(\sum_{i=1}^{n} \tau_i \mu_i\right)^2, \tag{3.2.9}$$

with equality if and only if the sequences $(\tau_1, \tau_2, \ldots, \tau_n)$ and $(\mu_1, \mu_2, \ldots, \mu_n)$ are proportional, i.e., there exists a constant λ such that

$$\tau_k = \lambda \mu_k \quad \text{for each } k \in \{1, 2, \ldots, n\}.$$

Definition 3.2.10. $L^1([0, b], H \times H)$

$$L^1([0, b], H \times H) = \left\{ Q : [0, b] \to H \times H \;\middle|\; \int_0^b \|Q(t)\|_{H \times H}\, dt < \infty \right\}$$

Theorem 3.2.11 (Cauchy Picard)**.** *[41]*
Let $C \subset H$ and let $Q : H \to H$ be a Lipschitz map, i.e., there is a constant L such that.

$$\|Q(u) - Q(v)\| \le L\|u - v\|, \quad \forall u, v \in H. \tag{3.2.10}$$

Then, for any given $u_0 \in H$, there exists a unique solution $u \in C^1([0, +\infty), H)$ of the problem:

$$\begin{cases} \frac{du(t)}{dt} = Q(u(t)), & \forall t \in (0, +\infty), \\ u(0) = u_0. \end{cases}$$

A function $u(t)$ is a solution of the differential equation if and only if it satisfies the integral equation:

$$u(t) = u_0 + \int_0^t Q(s, u(s))\, ds. \tag{3.2.11}$$

This integral equation defines a mapping Φ as:

$$\Phi(u)(t) = u_0 + \int_0^t Q(s, u(s))\, ds. \tag{3.2.12}$$

Theorem 3.2.12. *[41] Let $Q : [0, +\infty) \times H \to H$ be a map satisfying*

$$\|Q(t, u) - Q(t, v)\| \le L(t) \cdot p(\|u - v\|), \quad \forall t \in [0, +\infty),\ u, v \in H, \tag{3.2.13}$$

where $L : [0, +\infty) \to [0, +\infty)$ and $p : (0, +\infty) \to [0, +\infty)$ are continuous functions. Assume there exist constants $0 \le a < +\infty$ such that

$$L(t) \le a, \quad p(s) \le s,$$

and $p(s)$ is an increasing function. Then, for any given $u_0 \in H$, there exists a unique solution $u \in C^1([0, +\infty); H)$ for the problem

$$\frac{du(t)}{dt} = Q(t, u(t)).$$

Definition 3.2.13. *The space $L^1_{loc}(\mathbb{R}^n)$ consists of all locally integrable functions on $\mathbb{R}^n$. A function $Q : \mathbb{R}^n \to \mathbb{R}$ belongs to $L^1_{loc}(\mathbb{R}^n)$ if Q is measurable and integrable over every compact subset of $\mathbb{R}^n$. Formally, this means*

$$Q \in L^1_{loc}(\mathbb{R}^n) \iff \int_K |Q(x)|\, dx < \infty \quad \textit{for every compact set } K \subset \mathbb{R}^n.$$

Lemma 3.2.14 (Banach Fixed-Point Theorem)**.** *Let R be a nonempty complete metric space, and let $Q : R \to R$ be a strict contraction. That is, there exists a constant $0 < c < 1$ such that*

$$d(Q(\tau), Q(\mu)) \le c\, d(\tau, \mu), \quad \forall \tau, \mu \in R. \tag{3.2.14}$$

Then, there exists a unique point $\eta \in R$ such that $Q(\eta) = \eta$.

3.3 Existence and uniqueness of the trajectory

Proposition 3.3.1. *Let $Q : C \to H$ Lipschitz continuous with constant $L > 0$. Then,*

(i) $Ry(t)$ is Lipschitz continuous with the following constant $(1+\lambda L)^2 + \lambda L > 0$.

(ii) $(I-R)y(t)$ is Lipschitz continuous with the following constant $(1+\lambda L)^2 + (1+\lambda L) > 0$.

(iii) $Ey(t)$ is Lipschitz continuous with the following constant $(2 + \lambda L_E) > 0$

Proof. Using the Cauchy-Schwarz inequality and the Lipschitz continuity of Q, we obtain for all $y_1, y_2 \in H$ and $\lambda > 0$,

$$\begin{aligned} \|Ry_1(t) - Ry_2(t)\| &= \|w_1(t) + \lambda[Q(y_1(t)) - Q(w_1(t))] - (w_2(t) + \lambda[Q(y_2(t)) - Q(w_2(t))])\| \\ &\leq \|w_1(t) - w_2(t)\| + \lambda\|Q(y_1(t)) - Q(y_2(t))\| + \lambda\|Q(w_2(t)) - Q(y_1(t))\| \\ &\leq \|w_1(t) - w_2(t)\| + \lambda L\|y_1(t) - y_2(t)\| + \lambda L\|w_2(t) - w_1(t)\| \\ &= (1+\lambda L)\|w_1(t) - w_2(t)\| + \lambda L\|y_1(t) - y_2(t)\|. \end{aligned} \tag{3.3.1}$$

Since $w_1(t) = P_C(y_1(t) - \lambda Q(y_1(t)))$ and $w_2(t) = P_C(y_2(t) - \lambda Q(y_2(t)))$,

$$\begin{aligned} \|w_1(t) - w_2(t)\| &= \|P_C(y_1(t) - \lambda Q(y_1(t))) - P_C(y_2(t) - \lambda Q(y_2(t)))\| \\ &\leq \|(y_1(t) - \lambda Q(y_1(t))) - (y_2(t) - \lambda Q(y_2(t)))\| \\ &\leq \|y_1(t) - y_2(t)\| + \lambda\|Q(y_1(t)) - Q(y_2(t))\| \\ &\leq (1+\lambda L)\|y_1(t) - y_2(t)\|. \end{aligned} \tag{3.3.2}$$

Putting (3.3.2) into (3.3.1) , we obtain the following equation ,

$$\begin{aligned} \|Ry_1(t) - Ry_2(t)\| &\leq (1+\lambda L)^2\|y_1(t) - y_2(t)\| + \lambda L\|y_1(t) - y_2(t)\| \\ &= \left((1+\lambda L)^2 + \lambda L\right)\|y_1(t) - y_2(t)\|. \end{aligned} \tag{3.3.3}$$

This shows that Ry(t) is Lipschitz continuous with constant $(1+\lambda L)^2 + \lambda L$.
(ii) Using (i) we have that for $(I-R)y(t) = y(t) - w(t) + \lambda(Qy(t) - Qw(t))$,

$$\|(I-R)y_1(t) - (I-R)y_2(t)\| \leq \left(1 + (1+\lambda L)^2 + \lambda L\right)\|y_1(t) - y_2(t)\|. \tag{3.3.4}$$

This shows that $(I-R)$y(t) is Lipschitz continuous with constant $(1+\lambda L)^2 + (1+\lambda L)$.

(iii) For all $y \in H$, if given that $E : H \to H$ by

$$Ey(t) = y(t) - P_C(y(t) - \lambda Qy(t)), \tag{3.3.5}$$

where P_C is the projection operator and λ is a scalar.

$$
\begin{aligned}
\|Ey(t)-E\tilde{y}(t)\| &= \|y(t)-P_C(y(t)-\lambda Qy(t))-\tilde{y}(t)+P_C(\tilde{y}(t)-\lambda Q\tilde{y}(t))\| \\
&\leq \|y(t)-\tilde{y}(t)\|+\|P_C(y(t)-\lambda Qy(t))-P_C(\tilde{y}(t)-\lambda Q\tilde{y}(t))\| \\
&\leq \|y(t)-\tilde{y}(t)\|+\|y(t)-\lambda Qy(t)-\tilde{y}(t)+\lambda Q\tilde{y}(t)\| \\
&\leq 2\|y(t)-\tilde{y}(t)\|+\lambda\|Qy(t)-Q\tilde{y}(t)\| \\
&\leq (2+\lambda L)\|y(t)-\tilde{y}(t)\|.
\end{aligned} \tag{3.3.6}
$$

□

Using Rademacher's theorem see (3.2.7), we conclude that the operator E is differentiable almost everywhere in H. Furthermore, if E is continuously differentiable, this implies that $\frac{d}{dt}(Ey(t)) = By(t)$ is continuous. Let $By(t)$ be continuously differentiable such that $By(t)$ is Lipschitz continuous with constant L_B.
The following Lemma converts the second-order dynamical system into an equivalent first-order system.

Lemma 3.3.2. *Given a second-order dynamical system of the form (4.1.1), it can be reformulated as a first-order dynamical system within the product space $H \times H$. Let*

$$Y : \mathbb{R}_{\geq 0} \to H \times H, \quad Y(t) = (y(t), \dot{y}(t))$$

and

$$G : \mathbb{R}_{\geq 0} \times H \times H \to H \times H,$$

then the product space equations are given by,

$$\begin{cases} \dot{Y}(t) = G(t, Y(t)), \\ Y(0) = (u_0, v_0), \end{cases} \tag{3.3.7}$$

where, $G(t,u,v) = (v, -\alpha(t)v-(I-R)u-\beta\frac{d}{dt}(u(t)-\tau(t)))$ and $\tau(t) = P_C(u(t)-\lambda Q(u(t))$.

Proof. Let,

$$u(t) = y(t), \quad v(t) = \dot{y}(t),$$
$$\dot{u}(t) = \dot{y}(t) = v(t),$$
$$\dot{v(t)} = \ddot{y}(t) = -\alpha(t)\dot{y}(t) - (I-R)y(t) - \beta\frac{d}{dt}(u(t)-\tau(t)).$$

We have that ,

$$Y(t) = (y(t), \dot{y}(t)) = (u, v).$$

We have the following equation given by,

$$\dot{Y}(t) = (\dot{y}(t), \ddot{y}(t)) = (\dot{u}, \dot{v}) = (v, -\alpha(t)v(t) - (I-R)u(t) - \beta v(t) + \beta\dot{\tau}(t))$$

where $\dot{\tau}(t) = P_C(v(t) - \dot{Q}(u(t)v(t))$. Then

$$\dot{Y}(t) = G(t, (u,v)) = G(t, Y(t))$$

and $Y(0) = (u_0, v_0)$ where we have that the function G is,

$$G(t, u, v) = (v, -\alpha(t)v(t) - (I - R)u(t) - \beta v(t) + \beta\dot{\tau}(t)).$$

$\square$

Lemma 3.3.3. *For all starting points $u_0, v_0 \in H$, we endow $H \times H$ with the scalar product $X = \{a(t) \in L^1([0, +\infty); H \times H) \,|\|a(t)\| < +\infty\}$.*

$$\langle (a, y), (a', y') \rangle_{H\times H} = \langle a, a' \rangle + \langle y, y' \rangle$$

and the corresponding norm

$$\|(a, y)\|_{H\times H} = \sqrt{\|a\|^2 + \|y\|^2}.$$

Then,

(i) We show that the Lipschitz constant of $G(t, \cdot, \cdot)$ is locally integrable.

(ii) We show that $\forall u, v \in H, \forall b > 0$ then $\quad G(\cdot, u, v) \in L^1([0, b], H \times H)$.

(iii) We show that, for every $Y \in X$, $\Phi(Y(t))$ also belongs to X

Proof. **(i)** Let $y(t)$ be a function satisfying the relations

$$(I - R)y(t) = Ay(t) \quad \text{and} \quad By(t) = \frac{d}{dt}(Ey(t)),$$

where A and B are Lipschitz continuous operators. Then, for any $u, u', v, v' \in H$, the following result holds. Then we get the following:

$$\begin{aligned}
\|G(t, u, v) - G(t, u', v')\|_{H\times H} &\leq \sqrt{\begin{array}{l}(\|\alpha(t)(v' - v)\| + \|Au' - Au\| + \beta\|Bu - Bu'\|)^2 \\ + \|v - v'\|^2\end{array}} \\
&\leq \sqrt{\begin{array}{l}\big(\|\alpha(t)\|\|v' - v\|)^2 + \|A(u' - u)\| + \beta\|B(u - u')\|\big)^2 \\ + \|v - v'\|^2\end{array}} \\
&\leq \sqrt{\begin{array}{l}\big(\|\alpha(t)\|\|v - v'\|)^2 + L_A\|u - u'\| + \beta L_B\|u - u'\|\big)^2 \\ + \|v - v'\|^2\end{array}} \\
&= \sqrt{\begin{array}{l}\big(\|\alpha(t)\|\|v - v'\|)^2 + (L_A + \beta L_B)\|u - u'\|\big)^2 \\ + \|v - v'\|^2\end{array}} \\
&\leq (1 + \|\alpha(t)\|)\|v - v'\| + (L_A + \beta L_B)\|u - u'\|, \quad \forall t \geq 0.
\end{aligned} \tag{3.3.8}$$

Hence, the Lipschitz constant for G is

$$L(t)_G = 1 + \sqrt{2}\alpha(t) + 2(L_A + \beta L_B). \tag{3.3.9}$$

Considering that $\alpha \in L^1_{\text{loc}}([0, +\infty))$, it follows that the Lipschitz constant of $G(t, \cdot, \cdot)$ is integrable over any bounded interval, ensuring that the function behaves in a controlled manner within such intervals.

(ii) We now demonstrate that the function G is in the space $L^1([0, b], H \times H)$ that is

$$G(\cdot, u, v) \in L^1([0, b], H \times H).$$

$\forall u, v \in H, \forall b > 0.$
By applying the norm property outlined earlier, we can derive the following result,

$$\begin{aligned}
\int_0^b \|G(t, u, v)\|_{H\times H}\, dt &= \int_0^b \sqrt{\|v\|^2 + \|\alpha(t)v + A(u) + B(u)\|^2}\, dt \\
&\leq \int_0^b \sqrt{(1 + \|2\alpha^2(t))\|v\|^2 + (2\|Au + Bu\|)^2}\, dt \\
&\leq \int_0^b \sqrt{(1 + \|2\alpha^2(t))\|v\|^2 + (4\|Au\|^2 + 4\|Bu\|)^2}\, dt \\
&< \infty.
\end{aligned} \tag{3.3.10}$$

Given the condition $\|u(t)\| < \infty$, which ensures that $u(t)$ remains bounded over the interval $[0, b]$ and the assumption that α is locally integrable, it follows that the $\|G(t, u, v)\|_{H\times H} < \infty$.

(iii) Let $\Phi(t)$ be given by

$$\Phi(Y)(t) = Y_0 + \int_0^t G(s, y(s))\, ds \tag{3.3.11}$$

When we apply the norm to both sides of equation (3.3.11), we obtain the following equation:

$$\|\Phi(Y)(t)\| = \|Y_0\| + \int_0^t \|G(t, u, v)\|_{H\times H}\, dt \tag{3.3.12}$$

Since in (3.3.10) we have that ,

$$\int_0^b \|G(t, u, v)\|_{H\times H}\, dt < \infty \tag{3.3.13}$$

and $\|Y_0\|$ is a constant.

Therefore,

$$\|\Phi(Y(t))\| < \infty, \tag{3.3.14}$$

we have that $\|\Phi(Y(t))\|$ belongs to the space X.

□

We are now ready to present the key result of this section. This outcome builds on the previously established theoretical framework and assumptions, integrating essential lemmas and intermediate findings.

3.4 Main results

Theorem 3.4.1. *For all starting points $u_0, v_0 \in H$ with norm $\|\cdot\|$). Let $\Phi : [0, +\infty) \times H \times H \to H \times H$ be a map satisfying*

$$\|\Phi((t, u, u') - \Phi(t, v, v')\|_{H\times H} \leq \ L(t)_\Phi \|(u, u') - (v, v')\|_{H\times H}, \tag{3.4.1}$$

where $L(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a continuous function with

$$L(t)_\Phi < k,$$

where k is an integer. Then, for any given $u_0, v_0 \in H$, there exists a unique strong global solution of the dynamical system

$$\begin{cases} \frac{dY(t)}{dt} = G(t, u(t)), & \forall t \in (0, +\infty), \\ Y(0) = (u_0, v_0). \end{cases} \tag{3.4.2}$$

Proof. Using (3.3.11)

$$\begin{aligned} \|\Phi(t, u, v) - \Phi(t, u', v')\|_{H\times H} &\leq \sqrt{\begin{matrix}(\|\alpha(t)(v' - v)\| + \|Au' - Au\| + \beta\|Bu - Bu'\|)^2 \\ + \|v - v'\|^2\end{matrix}} \\ &\leq \sqrt{\begin{matrix}\left(\|\alpha(t)\|\|v' - v\|)^2 + \|A(u' - u)\| + \beta\|B(u - u')\|\right)^2 \\ + \|v - v'\|^2\end{matrix}} \\ &\leq \sqrt{\begin{matrix}\left(\|\alpha(t)\|\|v - v'\|)^2 + L_A\|u - u'\| + \beta L_B\|u - u'\|\right)^2 \\ + \|v - v'\|^2\end{matrix}} \\ &= \sqrt{\begin{matrix}\left(\|\alpha(t)\|\|v - v'\|)^2 + (L_A + \beta L_B)\|u - u'\|\right)^2 \\ + \|v - v'\|^2\end{matrix}} \\ &\leq (1 + \|\alpha(t)\|)\|v - v'\| + (L_A + \beta L_B)\|u - u'\|, \quad \forall t \geq 0. \end{aligned} \tag{3.4.3}$$

Hence we have $L(t)_\Phi = (1 + \sqrt{2}\alpha(t) + 2L_A + 2\beta L_B)$. Since $L(t)_\Phi$ is locally integral, we get that there exists k such that $k > L(t)_\Phi$, the function $L(t)_\Phi$ is bounded.

$$\|\Phi(t, u, u') - \Phi(t, v, v')\|_{H\times H} \leq \frac{L_\Phi(t)}{k}\|(u, u') - (v, v')\|_X, \quad \forall u, v \in H. \tag{3.4.4}$$

Then we have that Φ is a contraction on the product space $H \times H$, the Contraction mapping theorem guarantees the existence of a fixed point $u^* \in H$ such that,

$$\Phi(Y^*) = Y^*. \tag{3.4.5}$$

That means that Y^* satisfies the equation (3.3.7) . Therefore, there exists at least one solution to the differential equation on $[0, +\infty)$.

We will use Gronwall's inequality (3.2.3) for the uniqueness. Let Y_1, Y_2 be two solutions of (3.3.7), then

$$\|Y_1(t) - Y_2(t)\| \leq \int_0^t \|G(s, Y_1(s)) - G(s, Y_2(s))\| \, ds. \tag{3.4.6}$$

Using the Lipschitz condition on G, we get

$$\|Y_1(t) - Y_2(t)\| \leq \int_0^t L(s)_G(\|Y_1(s) - Y_2(s)\|) \, ds. \tag{3.4.7}$$

By Gronwall's inequality, we have

$$\|Y_1(t) - Y_2(t)\| \leq 0.$$

Hence, for any $t \in [0, +\infty)$, we have $Y_1(t) = Y_2(t)$.

The Cauchy-Lipschitz-Picard Theorem ensures that a strong global solution to the first-order dynamical system (3.3.7) exists and is unique. □

This shows that we have achieved the existence and uniqueness of the strong global solution for the generated trajectories in a Hilbert space under the assumption that Q is just only Lipschitz continuous. Moreover we show that $\ddot{y}(t)$ is locally absolutely continuous on $[t_0, +\infty)$. The following proposition addresses the issue.

Proposition 3.4.2. *For the starting points $u_0, v_0 \in H$, let y be the unique strong global solution of the dynamical system. Then $\ddot{y}(t)$ is locally absolutely continuous on $[t_0, +\infty)$.*

Proof. Given that $Y(t) := (y(t), \dot{y}(t))$ is absolutely continuous on $[t_0, T]$ and $T > 0$ is fixed, we endow the product space $H \times \; H$ with the 1-norm $(\| \cdot \|)$. For arbitrary $s, t \in [t_0, T]$, we have

$$\begin{aligned}
\|\dot{Y}(s) - \dot{Y}(t)\| &= \|G(s, y(s)) - G(t, y(t))\| \\
&= \|(\dot{y}(s) - \dot{y}(t), \alpha(t)\dot{y}(t) - \alpha(t)\dot{y}(s) - Ay(t) + Ay(s) + \beta By(s) - \beta By(t)\| \\
&\leq \|\dot{y}(s) - \dot{y}(t)\| + \|\alpha(t)\dot{y}(s) - \alpha(t)\dot{y}(t)\| + \|Ay(s) - Ay(t)\| + \|\beta By(s) - \beta By(t)\| \\
&\leq \|\dot{y}(s) - \dot{y}(t)\| + \alpha(t)\|\dot{y}(s) - \dot{y}(t)\| + L_A\|y(s) - y(t)\| + \beta L_B\|y(s) - y(t)\| \\
&\leq L_1\|\dot{y}(s) - \dot{y}(t)\| + L_A\|y(s) - y(t)\| + \beta L_B\|y(s) - y(t)\|,
\end{aligned}$$

The Lipschitz continuous is given by

$$L_Y \leq L_1 + L_A + \beta L_B$$

where $L_1 = 1 + \alpha(t)$.

For any finite collection of the $I_k = (a_k, b_k) \subset [t_0, T]$ and let $\epsilon > 0$ on the interval $[t_0, T]$ then there exists $\eta > 0$ such that

$$I_k \cap I_j = \emptyset \quad \text{and} \quad \sum_k |b_k - a_k| < \eta.$$

Since $\dot{y}(t)$, $\alpha(t)$, and $y(\cdot)$ are absolutely continuous we have that ,

$$\sum_k |\frac{\alpha}{b_k} - \frac{\alpha}{a_k}| < \frac{\epsilon}{3L_1}, \quad \text{and} \quad \sum_k \|y(b_k) - y(a_k)\| < \frac{\epsilon}{3L_A}.$$

$$\sum_k \|\dot{y}(b_k) - \dot{y}(a_k)\| < \frac{\epsilon}{3L_B}$$

This implies that we have,

$$\sum_k \|\dot{Y}(a_k) - \dot{Y}(b_k)\| < \frac{\epsilon}{3L_1} + \frac{\epsilon}{3L_B} + \frac{\epsilon}{3L_A} < \epsilon.$$

Therefore, $\dot{Y}(y(t)) = (\dot{y}(t), \ddot{y}(t))$ is absolutely continuous on $[t_0, T]$, which implies that $\ddot{y}(t)$ is also absolutely continuous on $[t_0, T]$. □

The third-order derivative of the dynamical system (3.1.1) *in terms of its first- and second-order derivatives is given by the following results.*

Lemma 3.4.3. *For the starting points $u_0, v_0 \in H$, let $y(t)$ be the unique strong global solution of* (5.1.1)*. Then, for almost every $t \in [t_0, +\infty)$,*

$$\| \dddot{y}(t)\| \leq \alpha(t))\|\ddot{y}(t)\| + (L_A + \beta L_B)\|\dot{y}(t)\|. \tag{3.4.8}$$

Proof. Let $t \in [t_0, +\infty)$ be such that $\dot{Y}(t) = G(t, Y(t))$. For almost every $t > 0$, we have:

$$\begin{aligned}
\|\dot{Y}(t+1) - \dot{Y}(t)\| &= \|G(s, y(t+1)) - G(t, y(t))\| \\
= \|(\dot{y}(t+1) - \dot{y}(t), \alpha(t)\dot{y}(t) &- \alpha(t)\dot{y}(t+1) - Ay(t) + Ay(s) + \beta By(s) - \beta By(t)\| \\
\leq \|\dot{y}(t+1) - \dot{y}(t)\| + \|\alpha(t)\dot{y}(s) &- \alpha(t)\dot{y}(t)\| + \|Ay(s) - Ay(t)\| + \|\beta By(t+1) - \beta By(t)\| \\
\leq \|\dot{y}(t+1) - \dot{y}(t)\| + \alpha(t)\|\dot{y}(t+1) &- \dot{y}(t)\| + L_A\|y(t+1) - y(t)\| + \beta L_B\|y(t+1) - y(t)\|.
\end{aligned}$$

This implies that,

$$\left\|\frac{\dot{Y}(t+1) - \dot{Y}(t)}{h}\right\| \leq \left\|1 + \alpha(t)\right\|\left\|\frac{\dot{y}(t+1) - \dot{y}(t)}{h}\right\| + L_A\left\|\frac{y(t+1) - y(t)}{h}\right\| \tag{3.4.9}$$

$$+ \beta L_B\left\|\frac{y(t+1) - y(t)}{h}\right\|. \tag{3.4.10}$$

Taking the limit of both sides, we obtain

$$\begin{aligned}
\lim_{h\to 0}\left\|\frac{\dot{Y}(t+1) - \dot{Y}(t)}{h}\right\| \leq \lim_{h\to 0}\Bigg(\left\|1 + \alpha(t)\right\|&\left\|\frac{\dot{y}(t+1) - \dot{y}(t)}{h}\right\| \\
+ (L_A + \beta L_B)&\left\|\frac{y(t+1) - y(t)}{h}\right\|\Bigg).
\end{aligned} \tag{3.4.11}$$

Hence,

$$\|\ddot{Y}(t)\| \leq (1+\alpha(t))\|\ddot{y}(t)\| + (L_A + \beta L_B)\|\dot{y}(t)). \tag{3.4.12}$$

We have that,

$$\ddot{Y}(t) = (\ddot{y}(t),\ \dddot{y}(t)),$$

and using the norm-1 we get that,

$$\|\ddot{Y}(t)\| = \|\dddot{y}(t)\| + \|\ddot{y}(t)\|, \tag{3.4.13}$$

Putting (3.4.12) into (3.4.13) we get,

$$\|\dddot{y}(t)\| + \|\ddot{y}(t)\| \leq (1+\alpha(t))\|\ddot{y}(t)\| + (L_A + \beta L_B)\|\dot{y}(t)\|. \tag{3.4.14}$$

Simplifying further we get,

$$\|\dddot{y}(t)\| \leq \alpha(t))\|\ddot{y}(t)\| + (L_A + \beta L_B)\|\dot{y}(t)\|. \tag{3.4.15}$$

□

Example 3.4.4. *To solve the Ordinary differential equation for the existence of the strong global solution.*

$$\dot{y} = Q(t, y) = 2ty \tag{3.4.16}$$

with initial condition

$$y(0) = 1$$

using Picard iteration, we proceed with the following steps,

Initial Guess

$$y_0(t) = 1$$

First Iteration

$$y_1(t) = 1 + \int_0^t 2sy_0(s)\, ds = 1 + \int_0^t 2s \cdot 1\, ds = 1 + t^2$$

Second Iteration

$$y_2(t) = 1 + \int_0^t 2sy_1(s)\, ds = 1 + \int_0^t 2s(1+s^2)\, ds$$

$$y_2(t) = 1 + \int_0^t (2s + 2s^3)\, ds = 1 + t^2 + \frac{t^4}{2}$$

Third Iteration

$$y_3(t) = 1 + \int_0^t 2sy_2(s)\, ds = 1 + \int_0^t 2s\left(1 + s^2 + \frac{s^4}{2}\right)\, ds$$

$$y_3(t) = 1 + \int_0^t (2s + 2s^3 + s^5)\, ds = 1 + t^2 + \frac{t^4}{2} + \frac{t^6}{6}$$

General Formula

$$y_n(t) = 1 + t^2 + \frac{t^4}{2} + \frac{t^6}{6} + \cdots = e^{t^2}$$

Using the fixed point theorem shows that the sequence of the Picard iteration converges to the solution of the initial value problem and that implies that the solution exists. The Gronwall inequality guarantees the uniqueness of the solution.

Example 3.4.5. *Consider the differential equation*

$$\dot{y}(t) = 1 + y(t)^2, \tag{3.4.17}$$

with the initial condition $y(t_0) = y_0 = 0$ *at* $t_0 = 0$. *The exact solution is given by* $y(t) = \tan(t)$. *Using Picard iteration, we start with* $\varphi_0(t) = 0$ *and define the iterative sequence:*

$$\varphi_{k+1}(t) = \int_0^t \left(1 + (\varphi_k(s))^2\right)\, ds, \tag{3.4.18}$$

which converges as $\varphi_n(t) \to y(t)$.

The first few iterations are as follows:

$$\varphi_1(t) = \int_0^t \left(1 + 0^2\right)\, ds = t,$$

$$\varphi_2(t) = \int_0^t \left(1 + s^2\right)\, ds = t + \frac{t^3}{3},$$

$$\varphi_3(t) = \int_0^t \left(1 + \left(s + \frac{s^3}{3}\right)^2\right)\, ds = t + \frac{t^3}{3} + \frac{2t^5}{15} + \frac{t^7}{63}.$$

Thus, the functions $\varphi_n(t)$ *compute the terms of the Taylor series expansion of the known solution* $y(t) = \tan(t)$.

Therefore, the Picard iteration converges to a local solution only for $|t| < \frac{\pi}{2}$ *and is not valid for all* $t \in \mathbb{R}$ *since* $\tan(t)$ *has poles at* $t = \pm\frac{\pi}{2}$, *where it is not Lipschitz continuous. Using the Gronwall inequality, we get the uniqueness of the solution.*

Chapter 4

Weak Convergence

4.1 Introduction

In this chapter, we analyze a dynamical system involving the projection operator. We aim to establish the Lipschitz continuity of a certain mapping derived from the system. We will perform the asymptotic analysis of the trajectories generated by (4.1.1) and outline the key assumptions required for convergence analysis. We further constructed the Lyapunov function to analyse the trajectories. We will use the monotonicity and Lipschitz continuity of the operator Q. This helps us to establish weak convergence of the system. We are going to solve equation (1.1.1) using the proposed system, which is:

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Q(y(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0. \end{cases} \tag{4.1.1}$$

where P_C is the projection operator, and let $\alpha(t)$ be the asymptomatic vanishing damping which is the Lebesgue measurable function, and Q is monotone operator and Lipschitz continuous.
Let $R : H \to H$ be defined as,

$$Ry(t) = w(t) + \lambda \left[Q(y(t)) - Q(w(t))\right], \tag{4.1.2}$$

Then, the system can be rewritten as follows.

$$\ddot{y}(t) + a(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) - Ry(t) = 0. \tag{4.1.3}$$

4.2 Preliminaries

We are going to present some lemmas and definitions required to establish this chapter's main results. We denote the strong and weak convergence of a sequence $\{y_n\}$ to a point $y \in H$ by $y_n \to y$ and $y_n \rightharpoonup y$, respectively.

The Minty variational inequality problem was introduced and studied by Minty in 1967. In his work, Minty explored the relationship between VIPs and MVIPs under the operator's assumptions of continuity and monotonicity Q.

Lemma 4.2.1. *[64] Given that $Q : H \to H$ is a Lipschitz continuous and monotone operator. Then we have that y^* is a solution of (1.1.1) if and only if y^* is a solution of the following problem: find $y^* \in C$ such that*

$$\langle Qy, y - y^* \rangle \geq 0, \quad \forall y \in C. \tag{4.2.1}$$

Lemma 4.2.2. *[26] Let $h : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be locally absolutely continuous and bounded below, and assume that there exists $\varphi \in L^1([0, +\infty))$ such that for almost every $s \in [0, +\infty)$,*

$$\frac{d}{ds} h(s) \leq \varphi(s). \tag{4.2.2}$$

Then, $\lim_{s \to +\infty} h(s) \in \mathbb{R}$ exists.

Definition 4.2.3. *Let $L^2([0, +\infty); H)$ denote the space of square-integrable functions from the interval $[0, +\infty)$ to a Hilbert space H. This means that a function $Q : [0, +\infty) \to H$ belongs to $L^2([0, +\infty); H)$ if*

$$\int_0^{+\infty} \|Q(t)\|_H^2 \, dt < \infty.$$

Here, $\| \cdot \|_H$ denotes the norm in the Hilbert space H.

Lemma 4.2.4. *[26] Suppose that $Q : [0, +\infty) \to [0, +\infty)$ is locally absolutely continuous and $Q \in L^p([0, +\infty))$, $1 \leq p < \infty$, and that there exists $G : [0, +\infty) \to \mathbb{R}$, $G \in L^r([0, +\infty))$, $1 \leq r \leq \infty$, such that for almost every $t \in [0, +\infty)$,*

$$\frac{d}{dt} Q(t) \leq G(t) \tag{4.2.3}$$

then $\lim_{t \to +\infty} Q(t) = 0$.

Lemma 4.2.5 (Opial Lemma)**.** *Let $C \subseteq H$ be a nonempty set, and let $y : [0, +\infty) \to H$ be a given map. Suppose that*

(i) for each $y^* \in C$, $\lim_{t \to +\infty} \|y(t) - y^*\|$ exists;

(ii) every weak sequential cluster point of the map y lies in C.

Then there exists $y^* \in C$ such that $y(t)$ converges weakly to y^* as $t \to +\infty$.

Lemma 4.2.6. *Suppose that $Q : [0, +\infty) \to \mathbb{R}$ is locally absolutely continuous and bounded from below, and that $G \in L^1([0, +\infty))$ such that for almost every $t \in [0, +\infty)$,*

$$\frac{d}{dt} Q(t) \leq G(t). \tag{4.2.4}$$

Then, there exists $\lim_{t \to +\infty} Q(t) \in \mathbb{R}$.

Proposition 4.2.7. *Let $Q : C \to H$ be monotone and Lipschitz continuous with constant $L > 0$. Then for all $y^* \in C$ and all $y(t) \in H$, it holds that*

$$\langle (I - R)y(t), y(t) - y^* \rangle \geq \frac{1 - \lambda L}{(1 + \lambda L)^2} \|(I - R)y(t)\|^2. \tag{4.2.5}$$

Proof. Let $y^* \in \Omega$, and by the characterization of the projection P_C, we get,

$$\langle w(t) - y(t) + \lambda Q(w(t)), y^* - w(t) \rangle \geq 0. \tag{4.2.6}$$

Since $w(t) \in C$, we obtain from Lemma(4.2.1) that

$$\lambda \langle Q(w(t)), w(t) - y^* \rangle \geq 0. \tag{4.2.7}$$

Adding (4.2.6) and (4.2.7) gives,

$$\langle y(t) - \lambda Q(y(t)) - w(t) + \lambda Q(w(t)), w(t) - y^* \rangle \geq 0. \tag{4.2.8}$$

This implies that,

$$\langle y(t) - w(t) - \lambda(Qy(t) - Qw(t), y(t) - y^* \rangle \geq \langle y(t) - Rx(t), y(t) - w(t) \rangle.$$

Additionally, using the L Lipschitz continuity of Q, we obtain

$$\begin{aligned} (1 - \lambda L)\|y(t) - w(t)\| &\leq \|(I - R)y(t)\| \\ &= \|y(t) - w(t) - \lambda(Qy(t) - Qw(t))\| \\ &\leq (1 + \lambda L)\|y(t) - w(t)\| \end{aligned} \tag{4.2.9}$$

This implies that

$$\begin{aligned} \langle (I - R)y(t), y(t) - y^* \rangle &\geq \langle y(t) - w(t) - \lambda(Q(y(t)) - Q(w(t)), y(t) - w(t) \rangle \\ &= \|y(t) - w(t)\|^2 - \lambda \langle Qy(t) - Qw(t), y(t) - w(t) \rangle \\ &\geq \|y(t) - w(t)\|^2 - \lambda \|Qy(t) - Qw(t)\| \|y(t) - w(t)\| \\ &\geq (1 - \lambda L)\|y(t) - w(t)\|^2 \\ &\geq \frac{1 - \lambda L}{(1 + \lambda L)^2} \|(I - R)y(t)\|^2. \end{aligned} \tag{4.2.10}$$

Assumption 4.2.1. *Let our asymptotic vanishing damping $\alpha : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be locally absolutely continuous. If there exists $\nu > 0$, $\beta > 0$ and $K > 0$ such that for almost every $t \in \mathbb{R}_0^+$ it holds*

$$\dot{\alpha}(t) \leq 0,$$

and

$$(\alpha(t) + \beta)^2 \geq \frac{(1 + \nu)(1 + \lambda L)^2}{(1 - \lambda L)}, \tag{4.2.11}$$

and

$$\int_0^T \left(\beta \frac{d}{dt} \langle y(t) - y^*, w(t) \rangle + 2\kappa\beta \|\dot{w}(t)\| \|\ddot{y} + c(t)\dot{y}(t)\| \right) dt < K \tag{4.2.12}$$

where $c(t) = \alpha(t) + \beta$. Since β is a constant, we have that $c(t)$ is locally absolutely continuous.

4.3 Asymptotic and convergence analysis

This section aims to establish the weak convergence of the trajectories generated by (5.1.1) to elements of Ω. Our analysis will focus on the asymptotic behavior of these trajectories.

Proposition 4.3.1. *Let $Q : C \to H$ be monotone, Lipschitz continuous, and Frechet differentiable. Additionally, we assume that $\|\dot{y}(t)\|^2$ and $\|\ddot{y}(t)\|^2$ belong to L^2 space and trajectories are bounded . Let $Q(0) = 0$ and $c(t) = (\alpha(t) + \beta)$ where $\alpha(t)$ is the asymptotic damping term and β is a positive constant. Then we have the following estimates given below by,*

(i) $\displaystyle\int_0^\infty \|\dot{y}(t)\| \, \|\dot{w}(t)\| \, dt < \infty,$

(ii) $\displaystyle\int_0^\infty \|\ddot{y}(t)\| \, \|\dot{w}(t)\| \, dt < \infty,$

(iii) $\displaystyle\int_0^\infty \|\dot{y}(t)\| \, \|w(t)\| \, dt < \infty.$

Proof. From the definition of $w(t)$, we have

$$w(t) = P_C\big(y(t) - \lambda Q(y(t))\big). \tag{4.3.1}$$

Taking the derivative of $w(t)$ with respect to t using the chain rule we obtain,

$$\begin{aligned} \|\dot{w}(t)\| &= \|P_C\big(\dot{y}(t) - \lambda \dot{V}(y(t))\big)\| \\ &\le \|\dot{y}(t) - \lambda \dot{V}(y(t))\| \\ &\le \|\dot{y}(t)\| + \lambda \|\dot{V}(y(t))\|, \end{aligned} \tag{4.3.2}$$

where $\dot{V}(y(t)) = \dot{Q}(y(t))\dot{y}(t)$. Since Q is Fréchet differentiable and Lipschitz continuous, we get the following equation:

$$\|\dot{V}(y(t))\| \le L\|\dot{y}(t)\|. \tag{4.3.3}$$

Substituting (4.3.3) into (4.3.2) we have,

$$\|\dot{w}(t)\| \le (1 + \lambda L)\|\dot{y}(t)\|. \tag{4.3.4}$$

Multiplying through by $\|\dot{x}(t)\|$, we obtain

$$\|\dot{y}(t)\|\|\dot{w}(t)\| \le (1 + \lambda L)\|\dot{y}(t)\|^2. \tag{4.3.5}$$

Integrating both sides, we obtain

$$\int_0^\infty \|\dot{y}(t)\|\|\dot{w}(t)\| \, dt \le \int_0^\infty (1 + \lambda L)\|\dot{y}(t)\|^2 dt < \infty. \tag{4.3.6}$$

(ii) Multiplying (4.3.4) by $\|\ddot{x}(t)\|$, we obtain

$$\|\ddot{y}(t)\|\|\dot{w}(t)\| \leq (1+\lambda L)\|\ddot{y}(t)\|\dot{y}(t)\|. \tag{4.3.7}$$

Integrating both sides, we obtain

$$\begin{aligned}\int_0^\infty \|\ddot{y}(t)\|\|\dot{w}(t)\|\, dt &\leq (1+\lambda L)\int_0^\infty \|\ddot{y}(t)\|\|\dot{y}(t)\|\, dt \\ &\leq (1+\lambda L)\left(\int_0^\infty \|\ddot{y}(t)\|^2\, dt\right)^{\frac{1}{2}}\left(\int_0^\infty \|\dot{y}(t)\|^2\, dt\right)^{\frac{1}{2}} \\ &< \infty.\end{aligned} \tag{4.3.8}$$

(iii) From the definition of $y(t)$, we have

$$w(t) = \|P_C(y(t)) - \lambda Q(y(t))\| \tag{4.3.9}$$

From the property of P_C, we obtain

$$\|w(t)\| \leq \|y(t)\| + \lambda\|Q(y(t))\|. \tag{4.3.10}$$

Using the Lipschitz continuity of Q and the fact that $Q(0) = 0$,we obtain

$$\|Q(y(t))\| \leq L\|y(t)\|. \tag{4.3.11}$$

This shows that the operator Q is bounded. from (4.3.11) and (4.3.10), we obtain

$$\|w(t)\| \leq \|y(t)\| + \lambda L\|y(t)\|. \tag{4.3.12}$$

Multiplying (4.3.12) by $\|\dot{x}(t)\|$, we obtain

$$\|w(t)\|\|\dot{y}(t)\| \leq \|\dot{y}(t)\|\|y(t)\| + \lambda L\|y(t)\|\|\dot{y}(t)\|. \tag{4.3.13}$$

Taking the integral from 0 to T and using the fact that $\|y(t)\|^2$ is bounded, we obtain

$$\int_0^\infty \|w(t)\|\|\dot{y}(t)\|dt \leq \int_0^\infty (1+\lambda L)\|y(t)\|\|\dot{y}(t)\|dt < \infty, \tag{4.3.14}$$

Therefore, we obtain

$$\int_0^\infty \|w(t)\|\|\dot{y}(t)\|dt < \infty. \tag{4.3.15}$$

□

This completes the proof for the asymptotic and convergence analysis of the generated trajectories by the second-order dynamical system (5.1.1). We will use these results to establish our main results.

4.4 Main results

We are now ready to present the key result of this section. This outcome builds on the previously established theoretical framework and assumptions, integrating essential lemmas and intermediate findings.

Theorem 4.4.1. *Let $Q : C \to H$ be a monotone and Lipschitz continuous function with constant $L > 0$, and let $\alpha : [0, +\infty) \to (0, +\infty)$ be a function that satisfies the conditions of Assumption (4.2.1) . Let $y : [0, +\infty) \to H$ be the unique solution to the second-order dynamical system given by (4.1.1) and $\kappa = \frac{1-\lambda L}{(1+\lambda L)^2} > 0$. The following results hold:*

(i) The trajectory $y(t)$ is bounded, and $\dot{y}(t), \ddot{y}(t), (I - R)y(t) \in L^2([0, +\infty); H)$.

(ii) $\lim_{t\to+\infty} \dot{y}(t) = \lim_{t\to+\infty} \ddot{y}(t) = \lim_{t\to+\infty}(I - R)y(t) = 0$.

(iii) As $t \to +\infty$, the trajectory $y(t)$ weakly converges to an element of Ω.

Proof. Let $y^* \in C$ be arbitrary. Let $s(t) = \frac{1}{2}\|y(t) - y^*\|^2$ represent the Lyapunov function for all $t \in [0, +\infty)$. Let κ and c(t) be defined as in Assumption (4.2.1). Taking the first and second derivatives of the Lyapunov function and using the chain rule, we obtain

$$\dot{s}(t) = \langle y(t) - y^*, \dot{y}(t)\rangle, \text{and} \tag{4.4.1}$$

$$\ddot{s}(t) = \|\dot{y}(t)\|^2 + \langle y(t) - y^*, \ddot{y}(t)\rangle. \tag{4.4.2}$$

Now we are in the position to write our trajectories $y(t)$ in terms of the Lyapunov function. Adding $c(t)\dot{s}(t)$ on both sides of the equation (4.4.2) and using (4.2.10), we obtain

$$\begin{aligned}\ddot{s}(t) + c(t)\dot{s}(t) &= \|\dot{y}(t)\|^2 + \langle y(t) - y^*, \ddot{y}(t) + (c(t))\dot{y}(t)\rangle \\ &= \|\dot{y}(t)\|^2 - \langle y(t) - y^*, (I - R)y(t) - \beta\dot{w}(t)\rangle \\ &= \|\dot{y}(t)\|^2 - \langle y(t) - y^*, (I - R)y(t)\rangle + \beta\langle\dot{w}(t), y(t) - y^*\rangle \\ &\le \|\dot{y}(t)\|^2 - \kappa\|(I - R)y(t)\|^2 + \langle\beta\dot{w}(t), y(t) - y^*\rangle.\end{aligned} \tag{4.4.3}$$

Note that

$$\beta \langle\dot{w}(t), y(t) - y^*\rangle = \beta\frac{d}{dt} \langle y(t) - y^*, w(t)\rangle - \beta \langle\dot{y}(t), w(t)\rangle . \tag{4.4.4}$$

Substituting equation (4.4.4) into (4.4.3) and considering equation (4.1.3), we obtain

$$\begin{aligned}&\ddot{s}(t) + c(t)\dot{s}(t) + \beta \langle\dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t) + c(t)\dot{y}(t) - \beta\dot{w}(t)\|^2 \\ &\quad\le \|\dot{y}(t)\|^2 + \beta\frac{d}{dt} \langle y(t) - y^*, w(t)\rangle .\end{aligned} \tag{4.4.5}$$

Simplifying the fourth term in (4.4.5), we obtain

$$\begin{aligned}&\ddot{s}(t) + c(t)\dot{s}(t) + \beta\langle\dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t) + c(t)\dot{y}(t)\|^2 - 2\kappa\beta\|\dot{w}(t)\|(\|\ddot{y} + c(t)\dot{y}(t)\| \\ &\quad+\kappa\beta^2\|\dot{w}(t)\|^2 \le \|\dot{y}(t)\|^2 + \beta\frac{d}{dt} \langle y(t) - y^*, w(t)\rangle .\end{aligned} \tag{4.4.6}$$

Simplifying further, we obtain

$$\begin{aligned}&\ddot{s}(t) + c(t)\dot{s}(t) + \beta\langle \dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 + \kappa c(t)^2\|\dot{y}(t)\|^2 \\ &\quad - 2\kappa\beta\|\dot{w}(t)\|(\|\ddot{y} + c(t)\dot{y}(t)\| + \kappa\beta^2\|\dot{w}(t)\|^2 \\ &\quad \leq \|\dot{y}(t)\|^2 + \beta\frac{d}{dt}\left\langle y(t) - y^*, w(t)\right\rangle .\end{aligned} \tag{4.4.7}$$

Rearranging, we obtain,

$$\begin{aligned}&\ddot{s}(t) + c(t)\dot{s}(t) + \beta\langle \dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 \\ &\quad - 2\kappa\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\| + \kappa\beta^2\|\dot{w}(t)\|^2 + (\kappa c(t)^2 - 1)\|\dot{y}(t)\|^2 \\ &\quad \leq \beta\frac{d}{dt}\left\langle y(t) - y^*, w(t)\right\rangle .\end{aligned} \tag{4.4.8}$$

Recall that

$$c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 = \frac{d}{dt}\left(c(t)\|\dot{y}(t)\|^2\right) - \dot{c}(t)\|\dot{y}(t)\|^2 \geq \frac{d}{dt}\left(c(t)\|\dot{y}(t)\|^2\right). \tag{4.4.9}$$

$$\begin{aligned}c(t)\dot{s}(t) &= \frac{d}{dt}\left(c(t)s(t)\right) - \dot{c}(t)s(t) \\ &\geq \frac{d}{dt}\left(c(t)s(t)\right).\end{aligned} \tag{4.4.10}$$

Combining these identities and (4.4.8), we obtain

$$\begin{aligned}&\ddot{s}(t) + \frac{d}{dt}c(t)s(t) + \beta\langle \dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa\frac{d}{dt}c(t)\|\dot{y}(t)\|^2 \\ &\quad - 2\kappa\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\| + \kappa\beta^2\|\dot{w}(t)\|^2 + (\kappa c(t)^2 - 1)\|\dot{y}(t)\|^2 \\ &\quad \leq \beta\frac{d}{dt}\left\langle y(t) - y^*, w(t)\right\rangle\end{aligned} \tag{4.4.11}$$

Simplifying further, we obtain

$$\begin{aligned}&\ddot{s}(t) + \frac{d}{dt}c(t)s(t) + \beta\langle \dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa\frac{d}{dt}c(t)\|\dot{y}(t)\|^2 \\ &\quad - 2\kappa\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\| + \kappa\beta^2\|\dot{w}(t)\|^2 + (\kappa c(t)^2 - 1)\|\dot{y}(t)\|^2 \\ &\quad \leq \beta\frac{d}{dt}\left\langle y(t) - y^*, w(t)\right\rangle\end{aligned} \tag{4.4.12}$$

Using Assumption 4.2.1 , we obtain

$$(\kappa c(t)^2 - 1)\|\dot{y}(t)\|^2 = \kappa v\|\dot{y}(t)\|^2. \tag{4.4.13}$$

Substituting (4.4.13) into (4.4.12) , we obtain

$$\begin{aligned}&\ddot{s}(t) + \frac{d}{dt}c(t)s(t) + \beta\langle \dot{y}(t), w(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa\frac{d}{dt}c(t)\|\dot{y}(t)\|^2 \\ &\quad - 2\kappa\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\| + \kappa\beta^2\|\dot{w}(t)\|^2 + \kappa v\|\dot{y}(t)\|^2 \\ &\quad \leq \beta\frac{d}{dt}\left\langle y(t) - y^*, w(t)\right\rangle\end{aligned} \tag{4.4.14}$$

Integrating both sides from 0 to T, we obtain

$$\int_0^T \Big(\ddot{s}(t) + \frac{d}{dt}\big(c(t)s(t)\big) + \beta\langle \dot{y}(t), y(t)\rangle + \kappa\|\ddot{y}(t)\|^2 + \kappa\frac{d}{dt}\big(c(t)\|\dot{y}(t)\|^2\big) + \kappa\beta^2\|\dot{y}(t)\|^2 + \kappa v\|\dot{y}(t)\|^2\Big)\, dt \leq \int_0^T \Big(\beta\frac{d}{dt}\langle y(t)-y^*, y(t)\rangle + 2\kappa\beta\|\dot{w}(t)\|\|\ddot{y}+c(t)\dot{y}(t)\|\Big)\, dt. \tag{4.4.15}$$

Since $\beta\langle \dot{y}(t), y(t)\rangle$,$\kappa\|\ddot{y}(t)\|^2$,$\kappa\beta^2\|\dot{y}(t)\|^2$ and $\kappa v\|\dot{y}(t)\|^2$ are positive terms then we have,

$$\dot{s}(t) + c(t)s(t) + \kappa c(t)\|\dot{y}(t)\|^2 \leq K \quad \forall t \in [0, +\infty). \tag{4.4.16}$$

We get that the above function is locally absolutely continuous. This yields the following,

$$\dot{s}(t) + \mu s(t) \leq K \quad \forall t \in [0, +\infty). \tag{4.4.17}$$

Here we have that μ represents the positive lower bound of our damping function $c(t)$. Since our relaxation function is a constant equal to one, our positive upper bound is equal to 1.

Multiplying both sides of (4.4.17) by $e^{\mu t}$, we obtain

$$e^{\mu t}\left[\dot{s}(t) + \mu(t)s(t)\right] \leq Ke^{\mu t}. \tag{4.4.18}$$

Now, integrating both sides from 0 to T, we obtain

$$\int_0^T \frac{d}{dt}\left[e^{\mu t}s(t)\right]\, dt \leq \int_0^T Ke^{\mu t}\, dt. \tag{4.4.19}$$

Which implies that

$$e^{\mu T}s(T) - s(0) \leq \frac{K}{\mu}\left(e^{\mu T} - 1\right) \tag{4.4.20}$$

This further implies that

$$s(T) \leq s(0)e^{-\mu T} + \frac{K}{\mu}\left(1 - e^{-\mu T}\right) \tag{4.4.21}$$

Then, $s(t)$ is bounded. Since $s(t) = \frac{1}{2}\|y(t) - y^*\|^2$ we have that the trajectory $y(t)$ is also bounded.

Using equation(4.4.16), it follows that:

$$\dot{s}(t) + \kappa\mu\|\dot{y}(t)\|^2 \leq K. \tag{4.4.22}$$

Therefore,

$$\langle y(t) - y^*, \dot{y}(t)\rangle + \kappa\mu\|\dot{y}(t)\|^2 \leq K. \tag{4.4.23}$$

Hence, we obtain that $\dot{s}(t)$ is bounded.

$$\langle y(t) - y^*, \dot{y}(t)\rangle \leq K. \tag{4.4.24}$$

Since $y(t)$ is bounded, it follows that $\dot{y}$ is also bounded.

Integrating the (4.4.14), we obtain that there exists a real number $K \in \mathbb{R}$ such that for every $t \in [0, +\infty)$:

$$\dot{s}(t) + c(t)s(t) + \kappa c(t)\|\dot{y}(t)\|^2 + \kappa v \int_0^t \|\dot{y}(s)\|^2\, ds + \kappa \int_0^t \|\ddot{y}(s)\|^2\, ds \leq K. \tag{4.4.25}$$

Thus, it follows that $\dot{y}, \ddot{y} \in L^2([0, +\infty); H)$. Moreover, from (4.1.3) and Assumption 4.2.1, we conclude that $(I - R)y(t) \in L^2([0, +\infty); H)$.

(ii) Calling on the Cauchy-Schwarz and Young inequality, we obtain

$$\frac{d}{dt}\left(\frac{1}{2}\|\dot{y}(t)\|^2\right) = \langle \dot{y}(t), \ddot{y}(t)\rangle \leq \frac{1}{2}\|\dot{y}(t)\|^2 + \frac{1}{2}\|\ddot{y}(t)\|^2. \tag{4.4.26}$$

Using Lemma 4.2.2 and (i) we get that $\lim_{t\to+\infty} \dot{y}(t) = 0$.

using again Cauchy-Schwarz and Young inequality, we obtain

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{2}\|(I - R)(y(t))\|^2\right) &= \langle (I - R)(y(t)), \frac{d}{dt}((I - R)(y(t)))\rangle \\ &\leq \frac{1}{2}\|(I - R)(y(t))\|^2 + \frac{((1 + \lambda L)^2 + (1 + \lambda L))^2}{2}\|\dot{y}(t)\|^2.\end{aligned} \tag{4.4.27}$$

Using Lemma 4.2.2 we get that , $\lim_{t\to+\infty}(I - R)(y(t)) = 0$, Using (4.1.3) we get that ,

$$\lim_{t\to\infty}\left(\ddot{y} + \alpha(t)\dot{y}(t) + \beta\frac{d}{dt}(y(t) - w(t))\right) = \lim_{t\to\infty}(R - I)y(t) = 0. \tag{4.4.28}$$

From Assumption 4.2.1, we get that $\lim_{t\to+\infty} \ddot{y}(t) = 0$.

(iii) From (4.4.16) , we obtain

$$\int_0^T c(t)s(t)\, dt \leq KT - s(T) + s(0) - \kappa \int_0^T c(t)\|\dot{y}(t)\|^2\, dt. \tag{4.4.29}$$

Since $s(t)$ is bounded and $\dot{y}(t)$ is in $L^2([0, \infty); H)$, as $T \to \infty$, the integral

$$\int_0^T c(t)s(t)\, dt$$

converges to some finite limit. By the properties of absolutely continuous functions, this implies that

$$\lim_{t\to\infty} c(t)s(t)$$

exists. Using Assumption 4.2.1,(i) and (ii) , the limit $\lim_{t\to+\infty} c(t)s(t)$ exists and is a finite real number.

Since $\lim_{t\to+\infty} c(t) \in (0,\infty)$ and $\lim_{t\to+\infty} c(t)s(t)$ exists, we get that $\lim_{t\to+\infty} s(t)$ also exists. This implies that $\lim_{t\to+\infty} \|y(t) - y^*\|$ exist and is a real number by Opial Lemma.

Let $\hat{y}$ be a weak sequential cluster point of $y(t)$. This means that there exists a subsequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to \infty$ as $n \to \infty$, and the sequence $(y(t_n))_{n\in\mathbb{N}}$ converges weakly to $\hat{y}$. Consequently, the sequence $(w(t_n))_{n\geq 0}$ also converges weakly to the same limit $\hat{y}$.

According to (5.1.1), we obtain

$$w(t_n) = P_C(y(t_n) - \lambda Q y(t_n)).$$

Since Q is monotone, we obtain

$$\begin{aligned}
0 &\leq \langle w(t_n) - y(t_n) + \lambda Q y(t_n), y - w(t_n)\rangle \quad ,\forall y \in C \\
&= \langle w(t_n) - y(t_n), y - w(t_n)\rangle + \lambda\langle Q y(t_n), y - w(t_n)\rangle \quad , \\
&= \langle w(t_n) - y(t_n), y - w(t_n)\rangle + \lambda\langle Q y(t_n), y(t_n) - w(t_n)\rangle + \lambda\langle Q y(t_n), y - y(t_n)\rangle, \\
&\leq \langle w(t_n) - y(t_n), y - w(t_n)\rangle + \lambda\langle Q y(t_n), y(t_n) - w(t_n)\rangle + \lambda\langle Q y, y - y(t_n)\rangle.
\end{aligned}$$

Now, passing to the limit as $t_n \to \infty$, we obtain

$$\langle Q\hat{y}, y - \hat{y}\rangle \geq 0 \quad \forall y \in C.$$

This shows that $\hat{y}$ satisfies the VIP, i.e.,

$$\hat{y} \in \Omega.$$

Therefore, it follows from the Opial lemma that $y(t)$ converges weakly to an element of Ω. This shows that we have achieved weak convergence for the generated trajectories in a Hilbert space under the assumption that Q is monotone and Lipschitz continuous.

$\square$

Chapter 5

Exponential Convergence

5.1 Introduction

In this chapter, we prove the exponential convergence of the system's trajectories to the unique solution of the variational inequality. We will show the error bounds of the trajectories to the solution. The proposed method is efficient and it is reliable to use in practice. We will provide numerical examples that show the effectiveness of our proposed method. We are going to solve equation (1.1.1) using the proposed system, which is:

$$\begin{cases} w(t) = P_C(y(t) - \lambda Q(y(t)) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Q(y(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0, \end{cases} \tag{5.1.1}$$

where P_C is the projection operator, and let $\alpha(t)$ be the asymptomatic vanishing damping which is the Lebesgue measurable function, and Q is strongly monotone operator and Lipschitz continuous.

5.2 Preliminaries

We are going to present some lemmas and definitions required to establish this chapter's main results. We denote the strong and weak convergence of a sequence $\{y_n\}$ to a point $y \in H$ by $y_n \to y$ and $y_n \rightharpoonup y$, respectively.

Definition 5.2.1. *[83] A point $\tilde{y}^*$ is an equilibrium point for the system (5.1.1) if $h(\tilde{y}^*) = 0$. Moreover, the equilibrium point $\tilde{y}^*$ is characterized as follows:*

(a) The equilibrium point of (5.1.1), $\tilde{y}^*$, is stable if, for any $\epsilon > 0$, there exists $\delta' > 0$ such that for every initial condition $y_0 \in V(\tilde{y}^*, \delta')$, the solution $\tilde{y}(t)$ of the dynamical system with $\tilde{y}(0) = y_0$ exists and stays within $V(\tilde{y}^*, \epsilon)$ for all $t > 0$, where $V(\tilde{y}^*, r)$ denotes the open ball with center $\tilde{y}^*$ and radius r;

(b) $\tilde{y}^*$ *is asymptotically stable if there exists* $\delta' > 0$ *such that for every solution* $\tilde{y}(t)$ *with* $\tilde{y}(0) \in V(\tilde{y}^*, \delta')$, *we have*

$$\lim_{t \to +\infty} \tilde{y}(t) = \tilde{y}^*;$$

(c) $\tilde{y}^*$ *is exponentially stable if there exist* $\delta' > 0$, $\nu > 0$, *and* $\zeta > 0$ *such that for every solution* $\tilde{y}(t)$ *with* $\tilde{y}(0) \in B(\tilde{y}^*, \delta')$, *we have*

$$\|\tilde{y}(t) - \tilde{y}^*\| \leq \nu \|\tilde{y}(0) - \tilde{y}^*\| e^{-\zeta t} \quad \forall t \geq 0;$$

(d) $\tilde{y}^*$ *is globally exponentially stable if the condition in (c) holds for all solutions* $\tilde{y}(t)$ *of the system.*

We build upon the work of Phan Tu Vuong [100], where the author considered the case where the operator Q is strongly pseudo-monotone. In contrast, our study focuses on the scenario where the underlying operator Q is strongly monotone.

Proposition 5.2.2. *[83] Let* Q *be strongly monotone with modulus* ρ *and Lipschitz continuous with constant* L. *Then* $\rho \leq L$.

Proof. Using the Cauchy–Schwarz inequality we get that,

$$\begin{aligned} \rho\|a_1 - a_2\|^2 &\leq (Q(a_1) - Q(a_2))(a_1 - a_2) \\ &\leq \|Q(a_1) - Q(a_2)\| \|a_1 - a_2\| \\ &\leq L\|a_1 - a_2\|^2, \end{aligned} \tag{5.2.1}$$

for all $\mathrm{a}_1, a_2 \in H$, which implies that $\rho \leq L$.

□

Theorem 5.2.3. *[32] If* $Q : H \to H$ *is a strongly monotone and* L *Lipschitz continuous operator, then the variational inequality* (1.1.1) *has a unique solution* $y^* \in C$.

5.2.1 Convergence Analysis

We present the convergence analysis of trajectories generated by our proposed method (5.1.1). We start with the following proposition which establishes the boundedness of the trajectories $\{x(t)\}$ generated by (5.1.1).

Proposition 5.2.4. *[100] Let* $Q : C \to H$ *be* ρ*-strongly monotone and* L*-Lipschitz continuous. Assume that* y^* *is the unique solution to the VIP* (1.1.1). *For any* $\lambda > 0$ *and* $y(t) \in H$, *and* $w(t)$.

$$w(t) = P_C(y(t) - \lambda Q(y(t))). \tag{5.2.2}$$

Then, the following holds:

$$\langle (I - R)y(t), y(t) - y(t)^* \rangle \geq \left(1 - \frac{\lambda L^2}{\rho} - \rho\lambda\right) \|(I - R)y(t)\|^2. \tag{5.2.3}$$

$$\|y(t) - y^*\| \leq \frac{1 + \lambda\rho + \lambda L}{\lambda\rho}\|y(t) - w(t)\|. \tag{5.2.4}$$

Proof. Given that $w(t) = P_C(y(t) - \lambda Qy(t)$ where $w(t) \in C$ and y^* is the unique solution of (1.1.1) we have ,

$$\langle y(t) - \lambda Qy(t) - w(t), y^* - w(t)\rangle \leq 0 \quad \forall y^* \in C. \tag{5.2.5}$$

Using Lipschitz continuity of Q, we obtain the following.

$$\begin{aligned}\langle y(t) - w(t), w(t) - y^*\rangle &\geq \lambda\langle Q(y(t)), w(t) - y^*\rangle \\ &= \lambda\langle Qy(t) - Qw(t), w(t) - y^*\rangle + \lambda\langle Q(w(t)) \\ &- Q(y^*), w(t) - y^*\rangle + \lambda\langle Q(y^*), w(t) - y^*\rangle.\end{aligned} \tag{5.2.6}$$

Since x^* is the unique solution of (1.1.1) and $y(t) \in C$, we have,

$$\langle Qy^*, w(t) - y^*\rangle \geq 0. \tag{5.2.7}$$

From the ρ- strongly monotonicity of Q, we have

$$\langle Q(w(t)) - Q(y^*), w(t) - y^*\rangle \geq \rho\|w(t) - y^*\|^2. \tag{5.2.8}$$

From (5.2.6), (5.2.7) and (5.2.8), we have

$$\langle y(t) - w(t), w(t) - y^*\rangle \geq \lambda\rho\|w(t) - y^*\|^2 - \lambda L\|y(t) - w(t)\|\|w(t) - y^*\|. \tag{5.2.9}$$

Hence, from (5.2.9), we have

$$\begin{aligned}\langle (I - R)y(t), y(t) - y^*\rangle &= \langle y(t) - w(t) - \lambda(Qy(t) - Q(w(t)), y(t) - w(t) + w(t) - y^*\rangle \\ &= \langle y(t) - w(t), y(t) - w(t)\rangle + \langle y(t) - w(t), w(t) - y^*\rangle \\ &- \lambda\langle Q(y(t) - Q(w(t)), y(t) - w(t)\rangle \\ &- \lambda\langle Qy(t) - Qw(t), w(t) - y^*\rangle \\ &= \|y(t) - w(t)\|^2 + \langle y(t) - w(t), w(t) - y^*\rangle \\ &- \lambda\langle Q(y(t) - Q(w(t)), y(t) - w(t)\rangle \\ &- \lambda\langle Qy(t) - Qw(t), w(t) - y^*\rangle \\ &\geq \|y(t) - w(t)\|^2 + \lambda\rho\|w(t) - y^*\|^2 \\ &- \lambda L\|y(t) - w(t)\|\|w(t) - y^*\| - \lambda\rho\|y(t) - w(t)\|^2 \\ &- \lambda L\|y(t) - w(t)\|\|w(t) - y^*\| \\ &= (1 - \lambda\rho)\|y(t) - w(t)\|^2 - 2\lambda L\|y(t) - w(t)\|\|w(t) - y^*\| \\ &+ \lambda\rho\|w(t) - y^*\|^2.\end{aligned} \tag{5.2.10}$$

Consider the term

$$-2\lambda L\|y(t) - w(t)\|\|w(t) - y^*\|, \tag{5.2.11}$$

which can be rewritten as

$$-2\lambda L\|y(t)-w(t)\|\|w(t)-y^*\| = -L\frac{2}{\sqrt{2}}\sqrt{\lambda}\|y(t)-w(t)\|\sqrt{\lambda}\sqrt{2}\|w(t)-y^*\|. \tag{5.2.12}$$

Using the Young inequalities we have the following equation,

$$ab \leq \frac{a^2}{2}+\frac{b^2}{2}. \tag{5.2.13}$$

From (5.2.12) let

$$\begin{aligned} a &= \frac{2}{\sqrt{2\rho}}\sqrt{\lambda}L\|y(t)-w(t)\|, \\ b &= \sqrt{2\rho\lambda}\|w(t)-y^*\|. \end{aligned} \tag{5.2.14}$$

Then, using (5.2.14), we obtain

$$-2\lambda L\|y(t)-w(t)\|\|w(t)-y^*\| \geq -\frac{\lambda L^2\|y(t)-w(t)\|^2}{\rho} - \frac{2\rho\lambda\|w(t)-y^*\|^2}{2}. \tag{5.2.15}$$

Substituting (5.2.15) into (5.2.10), we obtain,

$$\langle (I-R)y(t), y(t)-y^*\rangle \geq \left(1-\lambda\rho-\frac{\lambda L^2}{\rho}\right)\|y(t)-w(t)\|^2. \tag{5.2.16}$$

(ii) Using (5.2.9) , we obtain

$$\begin{aligned} \lambda\rho\|w(t)-y^*\|^2 &\leq \|y(t)-w(t)\|\|w(t)-y^*\| + \lambda L\|y(t)-w(t)\|\|w(t)-y^*\| \\ &\leq (1+\lambda L)\|y(t)-w(t)\|\|w(t)-y^*\|. \end{aligned}$$

Simplifying further, we have

$$\|w(t)-y^*\| \leq \frac{1+\lambda L}{\lambda\rho}\|y(t)-w(t)\|. \tag{5.2.17}$$

Using triangular inequality, we have

$$\|y(t)-y^*\| \leq \|y(t)-w(t)\| + \|w(t)-y^*\|. \tag{5.2.18}$$

Putting (5.2.18) into (5.2.17), we obtain

$$\|y(t)-y^*\| \leq \frac{1+\lambda\rho+\lambda L}{\lambda\rho}\|y(t)-w(t)\|. \tag{5.2.19}$$

□

This concludes the proof of the error bounds for the trajectories generated by the second-order dynamical system (5.1.1). The following remark comes from Phan Tu Vuong [100], where we obtain a similar consequence but for a different range of step sizes.

Remark 5.2.5. *[100] The term $\|x(t)-y(t)\|$ in (5.2.4), provides an upper bound on the error in the distance from any vector $y \in H$ to the unique solution y^* of the VIP (1.1.1). Additionally, as a direct consequence of inequality (5.2.3), and by using the Cauchy-Schwarz inequality, we deduce that if $\lambda \in \left(0, \frac{\rho}{\rho^2+L^2}\right)$, then $\|y(t) - w(t)\|$ also serves as a lower bound for the distance between any vector $y \in H$ and the unique solution y^* of (1.1.1).*

We are now ready to present the key result of this section. This outcome builds on the previously established theoretical framework and assumptions, integrating essential lemmas and intermediate findings. The proof of exponential convergence will follow a similar structure provided by Phan Tu Vuong [100] with additional properties. We will start with the following key assumptions.

Assumption 5.2.1. *Let $\alpha(t) : [0, +\infty) \to [0, +\infty)$ be a locally absolutely continuous function that fulfills, for every $t \in [0, +\infty)$, where the trajectories $\|y(t)\|$ are bounded and $\mu, K, n \geq 0$, $\kappa^* = 1 - \lambda\rho - \frac{\lambda L^2}{\rho}$, $k_1^2 = \frac{\lambda^2\rho^2}{(1+\lambda\rho+\lambda L)^2}$ and $c(t) = \alpha(t) + \beta$ then we have the following condition,*

(i) $0 < \mu \leq c(t)$,

(ii) $\dot{c}(t) \leq 0$,

(iii) $\kappa^* c(t)^2 - \kappa^* c(t) - 1 \geq 0$,

(iv) $(1 - c(t) + k_1^2) \geq 0$,

(v) $\int_0^T \left(\beta \frac{d}{dt} \langle y(t) - y^*, w(t)\rangle + 2\kappa\beta \|\dot{w}(t)\| \|\ddot{y}(t) + c(t)\dot{y}(t)\| + \|y(t)\|^2\right) dt \leq K$.

Theorem 5.2.6. *[100] Let C be a nonempty, closed, and convex subset of a real Hilbert space H. Let Q be a mapping that is ρ-strongly monotone and L-Lipschitz continuous on C, and $\alpha(t) : [0, +\infty) \to [0, +\infty)$ be a locally absolutely continuous function that satisfies Assumption (5.2.1). Assume that y^* is the unique solution to the VIP (1.1.1), and let λ satisfy the following condition:*

$$0 < \lambda < \frac{\rho}{\rho^2 + L^2}. \tag{5.2.20}$$

Then, the trajectories $y(t)$ generated by the dynamical system (5.1.1) are bounded and converge exponentially to y^* as $t \to \infty$. That is, there exist positive constants τ and η such that

$$\|y(t) - y^*\| \leq \tau \|y(0) - y^*\| e^{-\eta t}, \quad \forall t \geq 0. \tag{5.2.21}$$

Proof. Let $y^* \in C$ be arbitrary. Let $s(t) = \frac{1}{2}\|y(t) - y^*\|^2$ represent the Lyapunov function for all $t \in [0, +\infty)$. Let κ^* , k_1^2 and c(t) be defined as in Assumption(5.2.1). Taking the first and second derivatives of the Lyapunov function and applying the chain rule, we obtain

$$\dot{s}(t) = \langle y(t) - y^*, \dot{y}(t)\rangle, \text{and} \tag{5.2.22}$$

$$\ddot{s}(t) = \|\dot{y}(t)\|^2 + \langle y(t) - y^*, \ddot{y}(t)\rangle. \tag{5.2.23}$$

Now we are in the position to write our trajectories $y(t)$ in terms of the Lyapunov function. Adding $c(t)\dot{s}(t)$ on both sides of (5.2.23) , we obtain

$$
\begin{aligned}
\ddot{s}(t)+(c(t))\dot{s}(t) &= \|\dot{y}(t)\|^2+\langle y(t)-y^*,\ddot{y}(t)+(c(t)\dot{y}(t)\rangle \\
&= \|\dot{y}(t)\|^2-\langle y(t)-y^*,(I-R)y(t)-\beta\dot{w}(t)\rangle \\
&= \|\dot{y}(t)\|^2-\langle y(t)-y^*,(I-R)y(t)\rangle+\beta\langle\dot{w}(t),y(t)-y^*\rangle \\
&\leq \|\dot{y}(t)\|^2-\kappa^*\|(I-R)y(t)\|^2+\langle\beta\dot{w}(t),y(t)-y^*\rangle.
\end{aligned} \tag{5.2.24}
$$

Where we have that,

$$
\beta\langle\dot{w}(t),y(t)-y^*\rangle=\beta\frac{d}{dt}\langle y(t)-y^*,w(t)\rangle-\beta\langle\dot{y}(t),w(t)\rangle. \tag{5.2.25}
$$

Putting equation (5.2.25) into (5.2.24) and considering equation (4.1.3) we get,

$$
\begin{aligned}
&\ddot{s}(t)+c(t)\dot{s}(t)+\beta\left\langle\dot{y}(t),w(t)\right\rangle+\kappa^*\|\ddot{y}(t)+c(t)\dot{x}(t)-\beta\dot{w}(t)\|^2 \\
&\leq \|\dot{y}(t)\|^2+\beta\frac{d}{dt}\left\langle y(t)-y^*,w(t)\right\rangle
\end{aligned} \tag{5.2.26}
$$

Simplifying the fourth term, we obtain

$$
\begin{aligned}
&\ddot{s}(t)+c(t)\dot{s}(t)+\beta\langle\dot{y}(t),w(t)\rangle+\kappa^*\|\ddot{y}(t)+c(t)\dot{y}(t)\|^2-2\kappa^*\beta\|\dot{w}(t)\|(\|\ddot{y}+c(t)\dot{y}(t)\| \\
&\quad+\kappa^*\beta^2\|\dot{w}(t)\|^2 \\
&\quad\leq \|\dot{y}(t)\|^2+\beta\frac{d}{dt}\left\langle y(t)-y^*,w(t)\right\rangle
\end{aligned} \tag{5.2.27}
$$

equivalently,

$$
\begin{aligned}
&\ddot{s}(t)+c(t)\dot{s}(t)+\beta\langle\dot{y}(t),w(t)\rangle+\kappa^*\|\ddot{y}(t)\|^2+\kappa^*c(t)\frac{d}{dt}\|\dot{y}(t)\|^2+\kappa^*c(t)^2\|\dot{y}(t)\|^2 \\
&\quad-2\kappa^*\beta\|\dot{w}(t)\|(\|\ddot{y}+c(t)\dot{y}(t)\|+\kappa^*\beta^2\|\dot{w}(t)\|^2 \\
&\quad\leq \|\dot{y}(t)\|^2+\beta\frac{d}{dt}\left\langle y(t)-y^*,w(t)\right\rangle.
\end{aligned} \tag{5.2.28}
$$

Rearranging we obtain,

$$
\begin{aligned}
&\ddot{s}(t)+c(t)\dot{s}(t)+\beta\langle\dot{y}(t),w(t)\rangle+\kappa^*\|\ddot{y}(t)\|^2+\kappa^*c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 \\
&\quad-2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t)+c(t)\dot{y}(t)\|+\kappa^*\beta^2\|\dot{w}(t)\|^2+(\kappa^*c(t)^2-1)\|\dot{y}(t)\|^2 \\
&\quad\leq \beta\frac{d}{dt}\left\langle y(t)-y^*,w(t)\right\rangle.
\end{aligned} \tag{5.2.29}
$$

Multiplying both sides of (5.2.29) by $e^t>0$,we obtain

$$
\begin{aligned}
&e^t\ddot{s}(t)+e^tc(t)\dot{s}(t)+e^t\beta\langle\dot{y}(t),w(t)\rangle+e^t\kappa^*\|\ddot{y}(t)\|^2+e^t\kappa^*c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 \\
&\quad-e^t2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t)+c(t)\dot{y}(t)\|+e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 \\
&\quad+e^t(\kappa^*c(t)^2-1)\|\dot{y}(t)\|^2 \\
&\quad\leq e^t\beta\frac{d}{dt}\left\langle y(t)-y^*,w(t)\right\rangle.
\end{aligned} \tag{5.2.30}
$$

Let the following terms be rewritten as follows

$$c(t)\frac{d}{dt}\|\dot{y}(t)\|^2 = \frac{d}{dt}\left(c(t)\|\dot{y}(t)\|^2\right) - \dot{y}(t)\|\dot{y}(t)\|^2 \geq \frac{d}{dt}\left(c(t)\|\dot{y}(t)\|^2\right). \tag{5.2.31}$$

$$e^t\ddot{s}(t) = \frac{d}{dt}\left(e^t\dot{s}(t)\right) - e^t\dot{s}(t) \tag{5.2.32}$$

$$e^t\dot{s}(t) = \frac{d}{dt}\left(e^t s(t)\right) - e^t s(t) \tag{5.2.33}$$

$$e^t\dot{c(t)} = \frac{d}{dt}(e^t c(t) - e^t c(t). \tag{5.2.34}$$

Putting (5.2.31), (5.2.32) and (5.2.33) into (5.2.30), we obtain

$$\begin{aligned}\frac{d}{dt}e^t\dot{s}(t) - \frac{d}{dt}e^t s(t) + e^t s(t) + c(t)\frac{d}{dt}(e^t s(t)) - c(t)e^t s(t) + e^t\beta\langle\dot{y}(t), w(t)\rangle&\\ + e^t\kappa^*\|\ddot{y}(t)\|^2 + e^t\kappa^*\frac{d}{dt}c(t)\|\dot{y}(t)\|^2&\\ - e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\|&\\ + e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 + e^t(\kappa^* c(t)^2 - 1)\|\dot{y}(t)\|^2&\\ \leq e^t\beta\frac{d}{dt}\langle y(t) - y^*, w(t)\rangle&\end{aligned} \tag{5.2.35}$$

Simplifying, we obtain

$$\begin{aligned}\frac{d}{dt}e^t\dot{s}(t) - \frac{d}{dt}e^t s(t) + e^t s(t) + c(t)\frac{d}{dt}(e^t s(t)) - c(t)e^t s(t) + e^t\beta\langle\dot{y}(t), w(t)\rangle&\\ + e^t\kappa^*\|\ddot{y}(t)\|^2 + e^t\kappa^*\frac{d}{dt}c(t)\|\dot{y}(t)\|^2&\\ - e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y} + c(t)\dot{y}(t)\|&\\ + e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 + e^t(\kappa^* c(t)^2 - 1)\|\dot{y}(t)\|^2&\\ \leq e^t\beta\frac{d}{dt}\langle y(t) - y^*, w(t)\rangle.&\end{aligned} \tag{5.2.36}$$

We have,

$$e^t\frac{d}{dt}\left(c(t)\|\dot{y}(t)\|^2\right) = \frac{d}{dt}\left(e^t c(t)\|\dot{y}(t)\|^2\right) - e^t c(t)\|\dot{y}(t)\|^2. \tag{5.2.37}$$

Putting (5.2.37) into (5.2.36), we obtain

$$\begin{aligned}&\frac{d}{dt}e^t\dot{s}(t) + (c(t)-1)\frac{d}{dt}e^t s(t) + (1 - c(t))e^t s(t) + e^t\beta\langle\dot{y}(t), w(t)\rangle + e^t\kappa^*\|\ddot{y}(t)\|^2\\ &\quad + \kappa^*\frac{d}{dt}e^t c(t)\|\dot{y}(t)\|^2 - \kappa^* e^t c(t)\|\dot{y}(t)\|^2 + e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 + e^t(\kappa^* c(t)^2 - 1)\|\dot{y}(t)\|^2\\ &\leq e^t\beta\frac{d}{dt}\langle y(t) - y^*, w(t)\rangle + e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t) + c(t)\dot{y}(t)\|.\end{aligned} \tag{5.2.38}$$

Simplifying, we obtain

$$\begin{aligned}
&\frac{d}{dt}e^t\dot{s}(t) + (c(t)-1)\frac{d}{dt}e^t s(t) + (1-c(t))e^t s(t) + e^t\beta\langle \dot{y}(t), w(t)\rangle + e^t\kappa^*\|\ddot{y}(t)\|^2 \\
&\quad + \kappa^*\frac{d}{dt}e^t c(t)\|\dot{y}(t)\|^2 + e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 + e^t(\kappa^* c(t)^2 - \kappa^* c(t) - 1)\|\dot{y}(t)\|^2 \\
&\quad + e^t\|y(t)-w(t)\|^2 \\
&\leq e^t\beta\frac{d}{dt}\langle y(t)-w^*, w(t)\rangle + e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t)+c(t)\dot{y}(t)\| + e^t\|y(t)\|^2.
\end{aligned} \tag{5.2.39}$$

When we use (5.2.4) then, we obtain

$$(c(t)-1)\frac{d}{dt}e^t s(t) = \frac{d}{dt}((c(t)-1)e^t s(t) - \dot{c}(t)e^t s(t).$$

$$k_1^2\|y(t)-y^*\|^2 \leq \|y(t)-w(t)\|^2.$$

$$k_1^2 s(t) \leq \|y(t)-w(t)\|^2. \tag{5.2.40}$$

Applying Assumption (5.2.1) and putting inequality (5.2.40) into (5.2.39) then, we obtain

$$\begin{aligned}
&\frac{d}{dt}e^t\dot{s}(t) + \frac{d}{dt}(c(t)-1)e^t s(t) + (1-c(t)+k_1^2)e^t s(t) + e^t\beta\langle \dot{y}(t), w(t)\rangle + e^t\kappa^*\|\ddot{y}(t)\|^2 \\
&\quad + \kappa^*\frac{d}{dt}e^t c(t)\|\dot{y}(t)\|^2 + e^t(\kappa^* c(t)^2 - \kappa^* c(t) - 1)\|\dot{y}(t)\|^2 + e^t\kappa^*\beta^2\|\dot{w}(t)\|^2 \\
&\leq e^t\beta\frac{d}{dt}\langle y(t)-y^*, w(t)\rangle + e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t)+c(t)\dot{y}(t)\| + e^t\|y(t)\|^2.
\end{aligned} \tag{5.2.41}$$

We have that this terms $(1-c(t)+k_1^2)e^t s(t)$, $e^t\beta\langle \dot{y}(t), w(t)\rangle$, $e^t\kappa^*\|\ddot{y}(t)\|^2$, $e^t(\kappa^* c(t)^2 - \kappa^* c(t) - 1)\|\dot{y}(t)\|^2$ and $e^t\kappa^*\beta^2\|\dot{w}(t)\|^2$) are positive terms therefore integrating both sides of inequality (5.2.42), then we get the following inequality,

$$\begin{aligned}
&\int_0^T (\frac{d}{dt}e^t\dot{s}(t) + \frac{d}{dt}(c(t)-1)e^t s(t) + \kappa^*\frac{d}{dt}e^t c(t)\|\dot{y}(t)\|^2 \\
&\leq \int_0^T (e^t\beta\frac{d}{dt}\langle y(t)-y^*, w(t)\rangle + e^t 2\kappa^*\beta\|\dot{w}(t)\|\|\ddot{y}(t)+c(t)\dot{y}(t)\| + e^t\|y(t)\|^2).
\end{aligned}$$

From Assumption (5.2.1) ,we have

$$e^t\dot{s}(t) + c(t)e^t s(t) + \kappa^* e^t c(t)\|\dot{y}(t)\|^2 \leq K. \tag{5.2.42}$$

Since $c(t)\|\dot{y}(t)\| \geq 0$ and the above function is locally absolutely continuous, we obtain

$$e^t\dot{s}(t) + e^t(c(t)s(t)) \leq K. \tag{5.2.43}$$

Multiplying (5.2.43) by $\mathrm{e}^{-t}, we have \dot{s}(t) + (\mu s(t)) \leq Ke^{-t}.(5.2.44)$ *Here, μ represents the positive lower bound of our damping function $c(t)$. Since our relaxation function is a*

constant equal to one, our positive upper bound is equal to 1.
Integrating this differential inequality and multiplying by $e^{\mu t}$*, we get the following,*

$$e^{\mu t}\dot{s}(t) + e^{\mu t}(\mu s(t)) \leq Ke^{(\mu-1)t}. \tag{5.2.45}$$

This implies,

$$\frac{d}{dt}\left(e^{(\mu)t}s(t)\right) \leq Ke^{(\mu-1)t}. \tag{5.2.46}$$

Integrating both sides, we have the solutions for different cases of μ *given below,*

$$s(t) \leq \begin{cases} e^{-\mu t}s(0) + \frac{K}{1-\mu}\left(e^{-t} - e^{-\mu t}\right), & \text{if } 0 < \mu < 1, \\ e^{-\mu t}s(0) + \frac{K}{\mu-1}\left(e^{(\mu-1)t} - 1\right), & \text{if } \mu > 1, \\ e^{-\mu t}s(0) + Kte^{-\mu t}, & \text{if } \mu = 1. \end{cases} \tag{5.2.47}$$

This implies that $y(t)$ *converges exponentially to* y^**. This shows that we have achieved exponential convergence for the generated trajectories in a Hilbert space under the assumption that* Q *is strongly monotone and Lipschitz continuous. To demonstrate the effectiveness of our proposed method, we present the following example, accompanied by relevant graphs.*

5.3 Numerical solution

In this section, we provide several examples to demonstrate the effectiveness of the theoretical results discussed earlier.

Example 5.3.1. *Let* $C = \{y \in \mathbb{R}^n \mid \|y\| \leq 5\}$*. Let our underlying operator* $Q : \Omega \to \mathbb{R}^n$ *be defined as,*

$$Qy(t) = Ay(t). \tag{5.3.1}$$

Let $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$*, where* A *is a positive definite matrix. Then, the operator* Q *is defined as:*

$$Q(y) = \begin{pmatrix} 2y_1 + 1 \\ 2y_2 + 1 \end{pmatrix}. \tag{5.3.2}$$

We are going to prove the strong monotonicity of our operator Q*. for any* $x, y \in C$*:*

$$(Q(x) - Q(y))^T(x - y) = \begin{pmatrix} 2(x_1 - y_1) \\ 2(x_2 - y_2) \end{pmatrix}^T \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix}. \tag{5.3.3}$$

Simplifying further by applying the dot product we get the following equation,

$$(Q(x) - Q(y))^T(x - y) = 2((x_1 - y_1)^2 + (x_2 - y_2)^2) \geq 2\|x - y\|^2. \tag{5.3.4}$$

This implies that Q *is strongly monotone with* $\rho = 2$*.*

Now we show that Q *is Lipschitz Continuity for any* $x, y \in \Omega$*:*

$$\|Q(x) - Q(y)\| = \|A(x - y)\| \leq \|A\|\|x - y\| = 2\|x - y\|, \tag{5.3.5}$$

where $\|A\| = 2$*. This shows that* Q *is Lipschitz continuous with the constant* $L = 2$*.*

All computations were conducted on a personal computer with an Intel(R) Core(TM) i5-2600 CPU running at 2.30 GHz and 8.00 GB of RAM. The simulations consistently demonstrate that the trajectory of the dynamical systems (5.1.1) *converges globally and exponentially to the unique solution y^* of the problem. For instance, Figures 1, 2,3, and 4 present results for example 1 across various dimensional spaces with different initial points and parameters.*

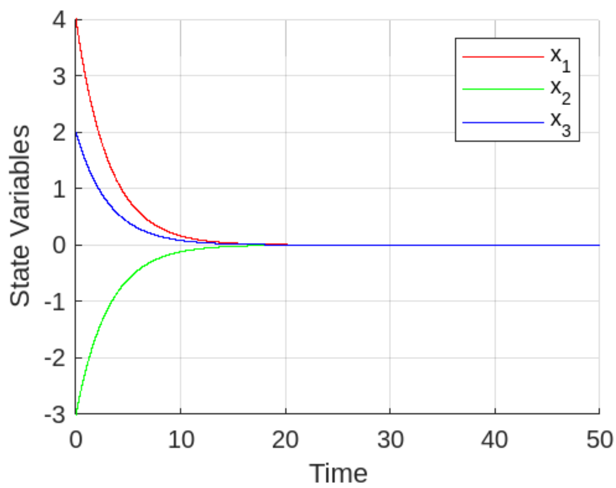

Figure 1 illustrates the trajectories produced by the dynamical system described in equation (5.1.1)*, with parameters set to $n = 3$, $\alpha(t) = (1 - \frac{\beta^2}{t})$, and $\lambda = 1$. The system starts with the initial condition $(4, 2, -3)^T$. This system's convergence is significantly faster than the case without the Hessian damping term.*

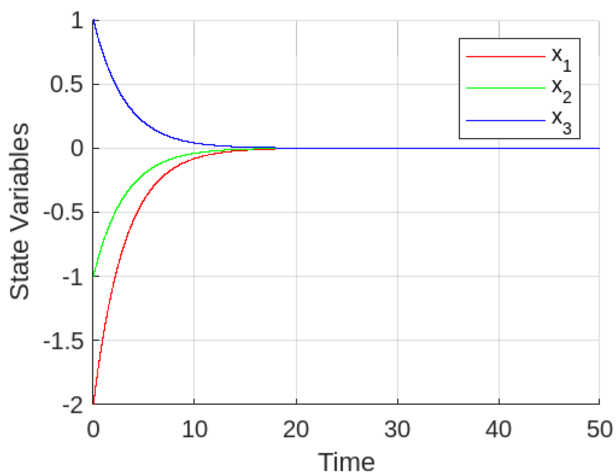

Figure 2 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 3$, $\alpha(t) = (1 - \frac{\beta^2}{t})$, and $\lambda = 1$. The system starts with the initial condition $(-2, -1, 1)^T$. This system's convergence is significantly faster than the case without the Hessian damping term.

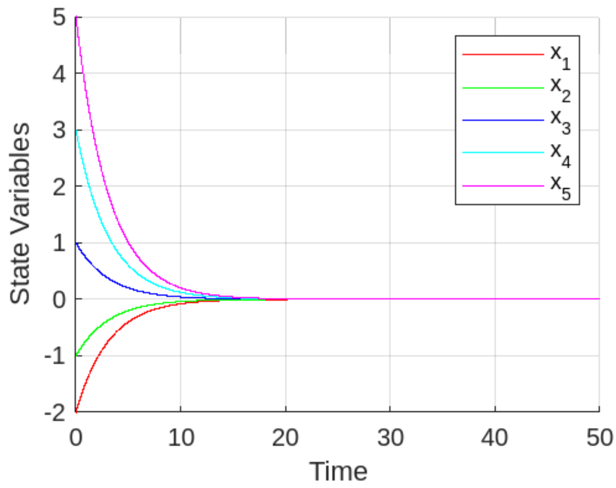

Figure 3 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 5$, $\alpha(t) = (1 - \frac{\beta^2}{t})$, and $\lambda = 1$. The system starts with the initial condition $(-2, -1, 1, 3, 5)^T$. This system's convergence is significantly faster than the case without the Hessian damping term.

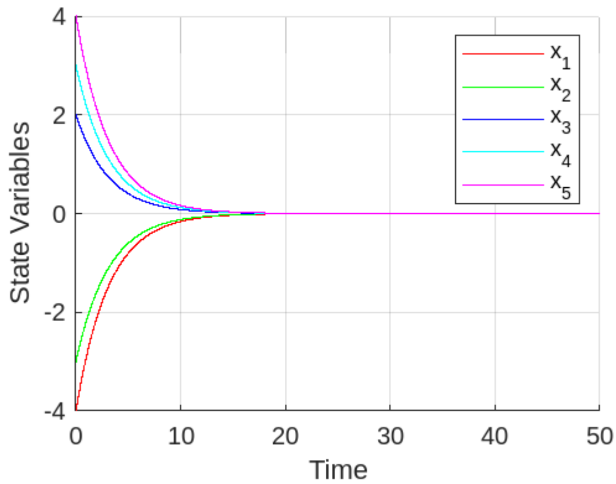

Figure 4 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 5$, $\alpha(t) = (1 - \frac{\beta^2}{t})$, and $\lambda = 1$. The system starts with the initial condition $(-4, -3, 2, 3, 4)^T$. This system's convergence is significantly faster than the case without the Hessian damping term.

5.3.1 Example

The simulation involves solving two dynamical systems one governed by a classical second-order differential equation with a time-dependent damping term $\alpha(t)$, and another enhanced by an additional inertial correction parameterized by β. Both models simulate the behavior of a projected dynamical system with a constraint-enforcing projection operator P_C, and a nonlinear feedback through a logistic function $Q(x)$.

Example 5.3.2. *Let $C = \{y \in \mathbb{R}^n \mid \|y\| \leq 5\}$. Let our underlying operator $Q : \Omega \to \mathbb{R}^n$ be defined as,*

$$Qy(t) = y(t) - sin(x(t)). \tag{5.3.6}$$

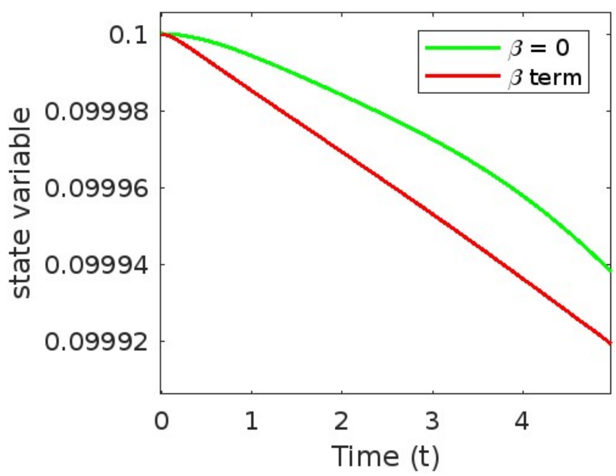

Figure 5 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.5\sin(t))$, $\beta = 10$, and $\lambda = 0.1$. The system starts with the initial condition $(0.1)^T$. The convergence of this system is significantly faster compared to the case without the Hessian damping term.

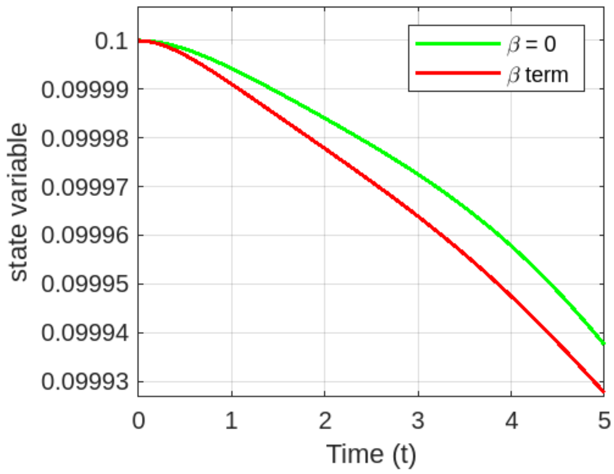

Figure 6 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.5\sin(t))$, $\beta = 1$, and $\lambda = 0.1$. The system starts with the initial condition $(0.1)^T$. The convergence of this system is significantly faster compared to the case without the Hessian damping term.

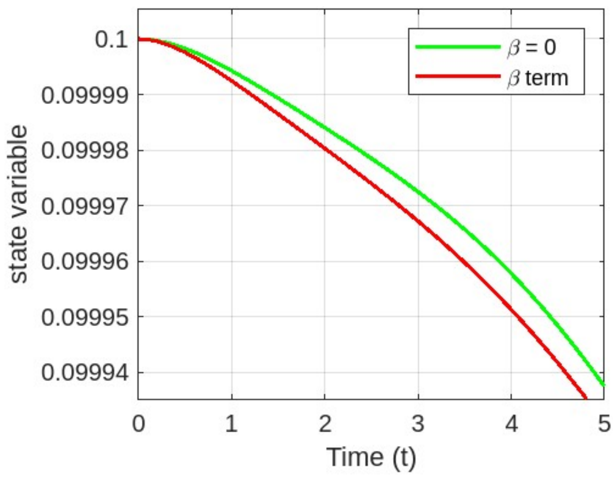

Figure 7 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.5\sin(t))$, $\beta = \frac{1}{2}$, and $\lambda = 0.1$. The system starts with the initial condition $(0.1)^T$. The convergence of this system is significantly faster compared to the case without the Hessian damping term.

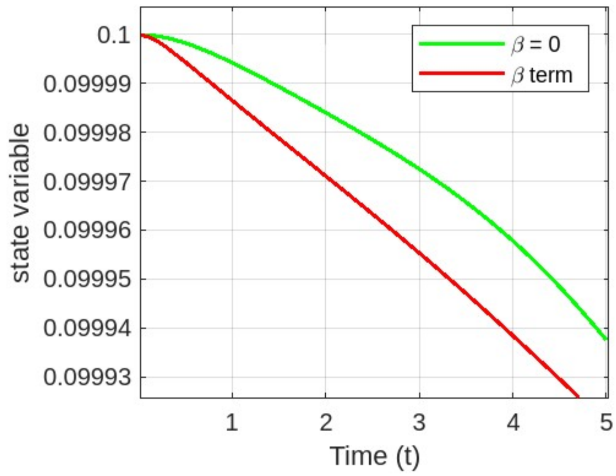

Figure 8 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.5\sin(t))$, $\beta = 5$, and $\lambda = 0.1$. The system starts with the initial condition $(0.1)^T$. The convergence of this system is significantly faster compared to the case without the Hessian damping term.

5.3.2 Example

Example 5.3.3. *Let $C = \{y \in \mathbb{R}^n \mid \|y\| \leq 5\}$. Let our underlying operator $Q : \Omega \to \mathbb{R}^n$ be defined as,*

$$Qy(t) = \frac{1}{1 + e^{-x(t)}}. \tag{5.3.7}$$

The function $Q(x)$, known as the sigmoid or logistic function, appears in both dynamical systems as part of the feedback mechanism. In the context of this simulation, the sigmoid function contributes to shaping the system's nonlinear response by compressing input values into a bounded range $[0, 1]$.

In our simulation, the use of the sigmoid function serves a dual role:

(a) It models a saturating response to changes in the state variable, akin to neuron activation.

(b) It offers a machine-learning-inspired feedback control in the optimization dynamics, representing a soft-threshold or probabilistic transition in the adjustment of the state.

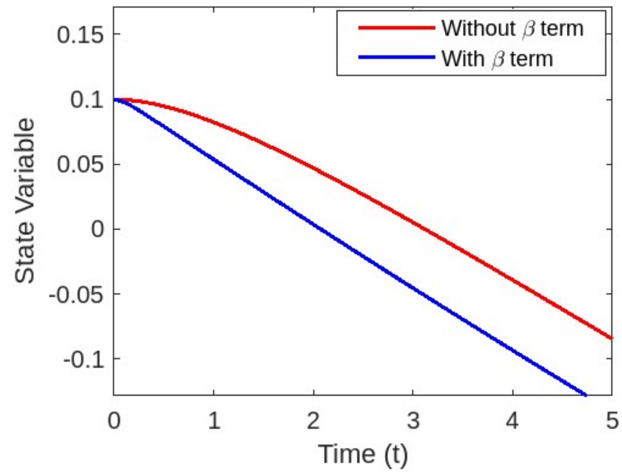

Figure 10 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.3e^{-0.5t})$, $\beta = 1$, *and* $\lambda = 0.1$. *The system starts with the initial condition* $(0.1)^T$. *The convergence of this system is significantly faster compared to the case without the Hessian damping term.*

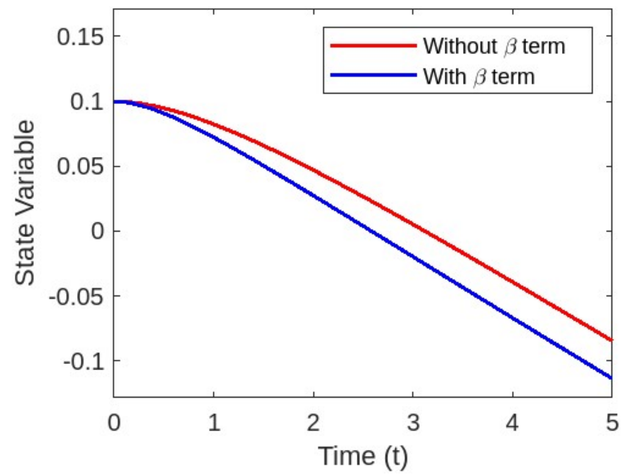

Figure 11 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.3e^{-0.5t})$, $\beta = 1$, *and* $\lambda = 0.1$. *The system starts with the initial condition* $(0.1)^T$. *The convergence of this system is significantly faster compared to the case without the Hessian damping term.*

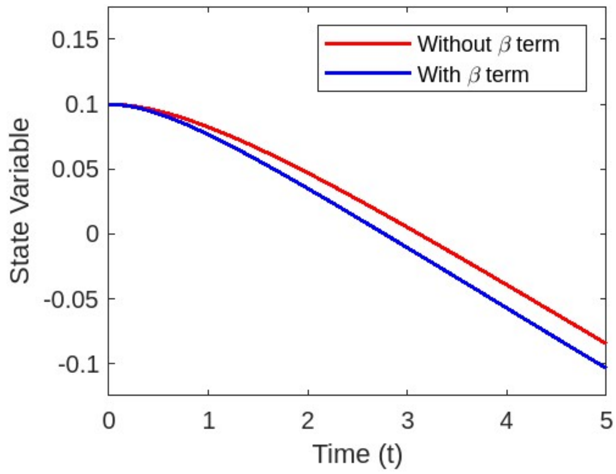

Figure 12 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.3e^{-0.5t})$, $\beta = 1$, *and* $\lambda = 0.1$. *The system starts with the initial condition* $(0.1)^T$. *The convergence of this system is significantly faster compared to the case without the Hessian damping term.*

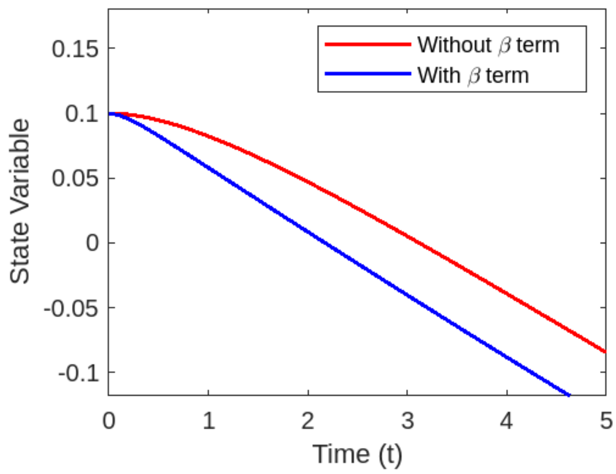

Figure 13 illustrates the trajectories produced by the dynamical system described in equation (5.1.1), with parameters set to $n = 1$, $\alpha(t) = (x(t) + 0.3e^{-0.5t})$, $\beta = 1$, *and* $\lambda = 0.1$. *The system starts with the initial condition* $(0.1)^T$. *The convergence of this system is significantly faster compared to the case without the Hessian damping term.*

Chapter 6

Conclusion, Contribution to Knowledge and Future Research

6.1 Conclusion

This dissertation studied one of the first second-order dynamical forward backward forward systems that combine asymptotic vanishing damping and Hessian damping to solve the fast convergence for variational inequality problems. We studied the asymptotic analysis of the trajectories generated by the dynamical system. First, we provided a brief background on the key ideas and reviewed important previous work in this area. Then, we built a foundation for our study by explaining key theorems and concepts. In chapters 3, 4, and 5, we presented our main results, which extend the results in the literature. In chapter 3, we showed the existence and uniqueness of strong global convergence in Hilbert space for (3.1.1)*. In Chapter 4, we achieved weak convergence for the case where the underlying operator was Lipschitz continuous and monotone. Our results extended earlier findings of Bot, Radu I, and Sedlmayer [27]. In chapter 5, we showed that when the operator is strongly monotone and Lipschitz continuous, the system converges exponentially. The trajectories generated by this system* (5.1.1) *have minimal oscillation. We provided numerical examples and graphs highlighting how well this method performs. The rate of convergence has improved after adding the Hessian damping term. Overall, this dissertation provides clear work for understanding and using second-order dynamical systems to solve variational inequality problems in Hilbert space. Our results improve on previous research.*

6.2 Contribution to knowledge

Our results expand on and improve some recent studies in the literature, especially those that motivated our work, by making the following contributions to the field of variational inequalities and optimizations.

1. *The author in [25] studied the first-order system of* (4.1.1) *in continuous and discrete*

cases to solve a variational inequality problem. We have extended the first-order method to the second-order method to solve a variational inequality problem. The results we obtain in Chapters 4 and 5 converge much faster than the results the author provided .

2. Phan Tu Vuong [100] utilised the Antipin equation proposed by Antipin [99] and replaced the gradient operator with a strongly pseudo-monotone operator. He obtained exponential convergence for (1.2.1)*, and the explicit discretisation of* (1.2.1) *has been shown to converge linearly. However, the dynamical system given by Phan Tu Vuong [100] does not produce trajectories that converge to the solution of the variational inequality problem if Q is monotone. We modified the system by adding the Tseng forward backward forward term, and we were able to achieve weak convergence and exponential convergence in chapters 4 and 5 for the case where the operator Q is monotone and strongly monotone.*

3. Radu I. Bot and Michael Sedlmayer [27] extend the work by developing a second-order dynamical system, specifically designed to approximate the solution of monotone inclusion problems and provide an asymptotic analysis of its trajectories. They have achieved weak convergence, but for the case when the operator is a pseudo-monotone. In this dissertation, their results have been improved by incorporating the Hessian damping factor to solve the variational inequality problem. In our results, we not only achieved weak convergence but also extended the results to have exponential convergence for the generated trajectories to solve the variational inequality problem (1.1.1)*.*

The following paper is submitted for potential publication.
***Tumelo Ranoto** and Chinedu Izuchukwu, Second-order differential equation with Hessian-driven damping for variational inequality problem. Submitted to Computational Optimization and Applications.*

6.3 Future research

First, in the future, we will study the discrete version of the second-order dynamical system (4.1.1)*. This approach will allow us to check the results obtained in the discrete and continuous parts. It is shown that the explicit discretisation by [25] leads to the Tseng [98] forward-backward-forward algorithm having the relaxation parameters, and they have proven linear convergence for the case when the underlying operator Q is pseudo-monotone. The Nesterov discretisation of our dynamical system* (5.1.1) *is:*

$$(\forall k \geq 1) \begin{cases} z_k = y_k + \tau_k(y_k - y_{k-1}) \\ w_k = P_C(I - \lambda_k Q) z_k \\ y_{k+1} = (1-\rho_k) z_k + \beta(w_k - w_{k-1}) + \rho_k(w_k - \lambda_k(Q y_k - Q z_k)), \end{cases} \tag{6.3.1}$$

where $y_0, y_1 \in H$ are starting points, $(\lambda_k)_{k\geq 1}$, $(\rho_k)_{k\geq 1}$ and $(\tau_k)_{k\geq 1}$ are sequences of positive numbers . For the case when $\tau_k = 0$. We get the following explicit equation.

The explicit case is given by the following equation,

$$(\forall k \geq 1) \begin{cases} z_k = y_k \\ w_k = P_C(I - \lambda_k Q) z_k \\ y_{k+1} = (1-\rho_k) z_k + \beta(w_k - w_{k-1}) + \rho_k(w_k - \lambda_k(Qy_k - Qz_k)), \end{cases} \tag{6.3.2}$$

where $y_0, y_1 \in H$ are starting points, $(\lambda_k)_{k\geq 1}$ and $(\rho_k)_{k\geq 1}$ are sequences of positive numbers .The explicit discretisation leads to the relaxed forward-backward forward with Hessian-driven term.

The implicit case is given below by the following equation,

$$(\forall k \geq 1) \begin{cases} z_k = y_{k+1} \\ w_k = P_C(I - \lambda_k Q) z_k \\ y_{k+1} = (1-\rho_k) z_k + \beta(w_k - w_{k-1}) + \rho_k(w_k - \lambda_k(Qy_k - Qz_k)), \end{cases} \tag{6.3.3}$$

where $y_0, y_1 \in H$ are starting points, $(\lambda_k)_{k\geq 1}$ and $(\rho_k)_{k\geq 1}$ are sequences of positive numbers. We will study the behavior of the sequence generated in these different cases, hoping to achieve convergence similar to the result obtained in the continuous case.

Secondly, the results obtained in this paper may be extended to solve the variational inequalities of the second kind. That is extending the results obtained in [39, 17, 93]. The problem is formulated as follows finding $y, v \in H$ such that

$$\langle Q(y), v - y \rangle + G(v) - G(y) \geq 0, \quad \forall v \in H, \tag{6.3.4}$$

where $G : H \mapsto \mathbb{R}$ is the given functional and $Q : H \mapsto H$ can be monotone , strong monotone and more. More interesting, we can extend them to solve the strong monotone for the strong mixed variational inequality problem given by considering the problem of finding $y \in H$ such that

$$\langle Q(y), v - y \rangle + Q(v) - Q(y) \geq \mu \|v - y\|^2, \quad \forall v \in H, \tag{6.3.5}$$

where $\mu > 0$. The problem is called a strongly mixed variational inequality. It has been shown that the minimum of the sum of a differentiable convex function and a non-differentiable strongly convex function can be characterized by strongly mixed variational inequalities.

We would like to generalize our result to study the monotone inclusion problem. We will replace the projection operator P_C with the resolvent operator $J_{\tau Q}$. First, we will define the resolvent operator .

Definition 6.3.1. *Let $Q : H \to 2^H$ be a set-valued operator, and let $\tau \in \mathbb{R}^{++}$ (i.e., τ is a positive real number). The resolvent of Q is defined as*

$$J_Q y(t) = (y(t) + \tau Q y(t))^{-1} \tag{6.3.6}$$

and the Yosida approximation of Q with parameter τ is given by the following equation ,

$$Qy(t) = \frac{1}{\tau}\left(y(t) - J_{\tau Q y(t)}\right). \tag{6.3.7}$$

The equation is given below. See similar work done by Attouch and László [15].

$$\text{Find } y^* \in H \text{ such that } 0 \in Qy^* + Gy^*, \tag{6.3.8}$$

where y^* is the solution to the problem, our system will be extended to the following system:

$$\begin{cases} w(t) = J_{\lambda G}(I - \lambda Q)y(t) \\ \ddot{y}(t) + \alpha(t)\dot{y}(t) + \beta \frac{d}{dt}(y(t) - w(t)) + y(t) = w(t) + \lambda[Q(y(t)) - Q(w(t))] \\ y(0) = y_0 \quad \dot{y}(0) = \dot{y}_0, \end{cases} \tag{6.3.9}$$

where $J_{\lambda G}$ is the resolvent operator, and $\alpha(t) : [0, +\infty) \to [0, +\infty)$ is a Lebesgue measurable function.
See similar work done by Attouch and László [15]. When we have that $G = 0$, the problem will lead to a monotone inclusion problem given by the following equation:

$$\text{find } y \in H \text{ such that } 0 \in Q(y), \tag{6.3.10}$$

where $Q(y)$ is a maximally monotone operator in H. For the case where we have that $Qy = 0$ the problem was studied by [30, 4]. The problem is formulate to find $y \in H$ such that,

$$G(y) = 0. \tag{6.3.11}$$

where $G(y)$ is the cocoercive operator .

Finally, if ∇Q is the gradient of Q and it is Lipschitz continuous, then we will study the following convex optimization problem,

$$\min_{x \in H} \left(Q(y) + G(y)\right), \tag{6.3.12}$$

where the function $Q : H \to \mathbb{R} \cup \{+\infty\}$ is a function that is, lower semicontinuous, proper, and convex, and we have that $G : H \to \mathbb{R}$ is a convex function with the additional property that is Fréchet derivative. This problem can be reformulated to monotone inclusion, that is, finding $y \in H$ such that the sum of the subdifferential of Q and the gradient of G at y.

$$0 \in \left(\partial Q + \nabla G\right)(y). \tag{6.3.13}$$

where we have that ∂Q represents the subdifferential of Q .

In summary, in the future, we will consider problems in optimization, variational inequalities of the first and second kind, and monotone inclusion problems to see the applicability of our second-order dynamical system in different fields.

Bibliography

[1] *Boushra Abbas and Hedy Attouch. Dynamical systems and forward–backward algorithms associated with the sum of a convex subdifferential and a monotone cocoercive operator.* Optimization*, 64(10):2223–2252, 2015.*

[2] *Seydamet S Ablaev, Alexander A Titov, Fedor S Stonyakin, Mohammad S Alkousa, and Alexander Gasnikov. Some adaptive first-order methods for variational inequalities with relatively strongly monotone operators and generalized smoothness. In* International Conference on Optimization and Applications*, pages 135–150. Springer, 2022.*

[3] *Artem Agafonov, Petr Ostroukhov, Roman Mozhaev, Konstantin Yakovlev, Eduard Gorbunov, Martin Takáč, Alexander Gasnikov, and Dmitry Kamzolov. Exploring jacobian inexactness in second-order methods for variational inequalities: Lower bounds, optimal algorithms and quasi-newton approximations.* arXiv preprint arXiv:2405.15990*, 2024.*

[4] *Rais Ahmad, Mohd Dilshad, Mu-Ming Wong, and Jen-Chin Yao. H $(\cdot,\cdot)$-cocoercive operator and an application for solving generalized variational inclusions. In* Abstract and Applied Analysis*, volume 2011, page 261534. Wiley Online Library, 2011.*

[5] *Iusem Alfredo. An iterative algorithm for the variational inequality problem.* Computational and Applied Mathematics*, 13:103–114, 1994.*

[6] *Felipe Alvarez. On the minimizing property of a second order dissipative system in hilbert spaces.* SIAM Journal on Control and Optimization*, 38(4):1102–1119, 2000.*

[7] *Felipe Alvarez, Hedy Attouch, Jérôme Bolte, and Patrick Redont. A second-order gradient-like dissipative dynamical system with hessian-driven damping.: Application to optimization and mechanics.* Journal de mathématiques pures et appliquées*, 81(8):747–779, 2002.*

[8] *Anatoly Antipin. Minimization of convex functions on convex sets by means of differential equations.* Differential equations*, 30(9):1365–1375, 1994.*

[9] *Stuart Antman. The influence of elasticity on analysis: modern developments.* Bulletin of the American Mathematical Society*, 9(3):267–291, 1983.*

[10] *Hedy Attouch, Aïcha Balhag, Zaki Chbani, and Hassan Riahi. Fast convex optimization via inertial dynamics combining viscous and hessian-driven damping with time rescaling.* arXiv preprint arXiv:2009.07620*, 2020.*

[11] *Hedy Attouch, Radu Ioan Boţ, and Ernö Robert Csetnek. Fast optimization via inertial dynamics with closed-loop damping.* Journal of the European Mathematical Society*, 25(5):1985–2056, 2022.*

[12] *Hedy Attouch, Zaki Chbani, Juan Peypouquet, and Patrick Redont. Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity.* Mathematical Programming*, 168:123–175, 2018.*

[13] *Hedy Attouch, Zaki Chbani, and Hassan Riahi. Rate of convergence of the nesterov accelerated gradient method in the subcritical case α 3.* ESAIM: Control, Optimisation and Calculus of Variations*, 25:2, 2019.*

[14] *Hedy Attouch and Szilárd Csaba László. Newton-like inertial dynamics and proximal algorithms governed by maximally monotone operators.* SIAM Journal on Optimization*, 30(4):3252–3283, 2020.*

[15] *Hedy Attouch and Szilárd Csaba László. Continuous newton-like inertial dynamics for monotone inclusions.* Set-valued and variational analysis*, 29:555–581, 2021.*

[16] *Hedy Attouch, Juan Peypouquet, and Patrick Redont. Fast convex optimization via inertial dynamics with hessian driven damping.* Journal of Differential Equations*, 261(10):5734–5783, 2016.*

[17] *IB Badriev, OA Zadvornov, and LN Ismagilov. On iterative regularization methods for variational inequalities of the second kind with pseudomonotone operators.* Computational Methods in Applied Mathematics*, 3(2):223–234, 2003.*

[18] *Heinz H Bauschke, Patrick L Combettes, Heinz H Bauschke, and Patrick L Combettes.* Correction to: convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2017.*

[19] *Amir Beck.* First-order methods in optimization*. SIAM, 2017.*

[20] *Aleksandr Beznosikov, Sergey Samsonov, Marina Sheshukova, Alexander Gasnikov, Alexey Naumov, and Eric Moulines. First order methods with markovian noise: from acceleration to variational inequalities.* Advances in Neural Information Processing Systems*, 36, 2024.*

[21] *Bandar Bin-Mohsin, Muhammad Aslam Noor, Khalida Inayat Noor, and Rafia Latif. Resolvent dynamical systems and mixed variational inequalities.* J. Nonl. Sci. Appl*, 10:2925–2933, 2017.*

[22] *Jérôme Bolte. Continuous gradient projection method in hilbert spaces.* Journal of Optimization Theory and Applications*, 119:235–259, 2003.*

[23] *Radu Ioan Boţ and Ernö Robert Csetnek. Second-order dynamical systems associated to variational inequalities.* Applicable Analysis*, 96(5):799–809, 2017.*

[24] *Radu Ioan Boţ and Ernö Robert Csetnek. A second-order dynamical system with hessian-driven damping and penalty term associated to variational inequalities.* Optimization*, 68(7):1265–1277, 2019.*

[25] *Radu Ioan Boţ, Ernö Robert Csetnek, and Phan Tu Vuong. The forward–backward–forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in hilbert spaces.* European Journal of Operational Research*, 287(1):49–60, 2020.*

[26] *Radu Ioan Bot and Erno Robert Csetnek. Second order forward-backward dynamical systems for monotone inclusion problems.* SIAM Journal on Control and Optimization*, 54(3):1423–1443, 2016.*

[27] *Radu Ioan Bot, Michael Sedlmayer, and Phan Tu Vuong. A relaxed inertial forward-backward-forward algorithm for solving monotone inclusions with application to gans.* Journal of Machine Learning Research*, 24(8):1–37, 2023.*

[28] *RI Boţ, Ernö Robert Csetnek, and Phan Tu Vuong. The forward–backward–forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in hilbert spaces.* European Journal of Operational Research*, 287(1):49–60, 2020.*

[29] *Alexandre Cabot, Hans Engler, and Sébastien Gadat. On the long time behavior of second order differential equations with asymptotically small dissipation.* Transactions of the American Mathematical Society*, 361(11):5983–6017, 2009.*

[30] *Le-cai Cai, Heng-you Lan, and Yunzhi Zou. Perturbed algorithms for solving nonlinear relaxed cocoercive operator equations with general a-monotone operators in banach spaces.* Communications in Nonlinear Science and Numerical Simulation*, 16(10):3923–3932, 2011.*

[31] *Ennio Cavazzuti, Massimo Pappalardo, Mauro Passacantando, et al. Nash equilibria, variational inequalities, and dynamical systems.* Journal of optimization theory and applications*, 114(3):491–506, 2002.*

[32] *Andrzej Cegielski.* Iterative methods for fixed point problems in Hilbert spaces*, volume 2057. Springer, 2012.*

[33] *Yair Censor, Aviv Gibali, and Simeon Reich. The subgradient extragradient method for solving variational inequalities in hilbert space.* Journal of Optimization Theory and Applications*, 148(2):318–335, 2011.*

[34] *Yair Censor, Aviv Gibali, and Simeon Reich. A von neumann alternating method for finding common solutions to variational inequalities.* Nonlinear Analysis: Theory, Methods & Applications*, 75(12):4596–4603, 2012.*

[35] *Fei Chen and Wei Ren. Convex optimization via finite-time projected gradient flows. In* 2018 IEEE Conference on Decision and Control (CDC)*, pages 4072–4077. IEEE, 2018.*

[36] *Ulrich Clarenz, Stefan Henn, Martin Rumpf, and Kristian Witsch. Relations between optimization and gradient flow methods with applications to image registration. In* Proceedings of the 18th GAMM Seminar Leipzig on Multigrid and Related Methods for Optimisation Problems*, pages 11–30, 2002.*

[37] *Jorge Cortés. Finite-time convergent gradient flows with applications to network consensus.* Automatica*, 42(11):1993–2000, 2006.*

[38] *Stella Dafermos. Traffic equilibrium and variational inequalities.* Transportation science*, 14(1):42–54, 1980.*

[39] *Juan Carlos De los Reyes. Optimal control of a class of variational inequalities of the second kind.* SIAM Journal on Control and Optimization*, 49(4):1629–1658, 2011.*

[40] *Paul Dupuis and Anna Nagurney. Dynamical systems and variational inequalities.* Annals of Operations Research*, 44:7–42, 1993.*

[41] *Zonghong Feng, Fengying Li, Ying Lv, and Shiqing Zhang. A note on cauchy-lipschitz-picard theorem.* Journal of Inequalities and Applications*, 2016:1–6, 2016.*

[42] *Gaetano Fichera. The signorini elastostatics problem with ambiguous boundary conditions. In* Proc. Int. Conf. Application of the theory of functions in continuum mechanics*, volume 1, pages 91–140, 1963.*

[43] *Masao Fukushima. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems.* Mathematical programming*, 53:99–110, 1992.*

[44] *Kunal Garg and Dimitra Panagou. Fixed-time stable gradient flows: Applications to continuous-time optimization.* IEEE Transactions on Automatic Control*, 66(5):2002–2015, 2020.*

[45] *Gauthier Gidel, Hugo Berard, Gaëtan Vignoud, Pascal Vincent, and Simon Lacoste-Julien. A variational inequality perspective on generative adversarial networks.* arXiv preprint arXiv:1802.10551*, 2018.*

[46] *Gauthier Gidel, Hugo Berard, Gaëtan Vignoud, Pascal Vincent, and Simon Lacoste-Julien. A variational inequality perspective on generative adversarial networks.* arXiv preprint arXiv:1802.10551*, 2018.*

[47] *Anand Gokhale, Alexander Davydov, and Francesco Bullo. Proximal gradient dynamics: Monotonicity, exponential convergence, and applications.* IEEE Control Systems Letters*, 2024.*

[48] *Jin-Lin Guan, Lu-Chuan Ceng, and Bing Hu. Strong convergence theorem for split monotone variational inclusion with constraints of variational inequalities and fixed point problems.* Journal of Inequalities and Applications*, 2018:1–29, 2018.*

[49] *Kanan Gupta, Jonathan W Siegel, and Stephan Wojtowytsch. Nesterov acceleration despite very noisy gradients. In* The Thirty-eighth Annual Conference on Neural Information Processing Systems*.*

[50] *Nguyen Thi Thu Ha, Jean-Jacques Strodiot, and Phan Tu Vuong. On the global exponential stability of a projected dynamical system for strongly pseudomonotone variational inequalities.* Optimization Letters*, 12:1625–1638, 2018.*

[51] *Pham Viet Hai and Trinh Ngoc Hai. Second-order dynamical systems with a smoothing effect for solving paramonotone variational inequalities.* arXiv preprint arXiv:2411.14651*, 2024.*

[52] *Philip Hartman and Guido Stampacchia. On some non-linear elliptic differential-functional equations. 1966.*

[53] *Sepideh Hassan-Moghaddam and Mihailo R Jovanović. Proximal gradient flow and douglas–rachford splitting dynamics: Global exponential stability via integral quadratic constraints.* Automatica*, 123:109311, 2021.*

[54] *Bingsheng He and Hai Yang. Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities.* Operations research letters*, 23(3-5):151–161, 1998.*

[55] *BS He, Hai Yang, Qiang Meng, and DR Han. Modified goldstein–levitin–polyak projection method for asymmetric strongly monotone variational inequalities.* Journal of Optimization Theory and Applications*, 112:129–143, 2002.*

[56] *René Henrion, Boris Mordukhovich, and Nguyen Mau Nam. Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities.* SIAM Journal on Optimization*, 20(5):2199–2227, 2010.*

[57] *René Henrion, Boris S Mordukhovich, and Nguyen Mau Nam. Second-order analysis of polyhedral systems in finite and infinite dimensions with applications to robust stability of variational inequalities.* SIAM Journal on Optimization*, 20(5):2199–2227, 2010.*

[58] *Ivan Hlavácek, Jaroslav Haslinger, Jindrich Necas, and Jan Lovisek.* Solution of variational inequalities in mechanics*, volume 66. Springer Science & Business Media, 2012.*

[59] *Xiaolin Hu and Jun Wang. Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network.* IEEE Transactions on Neural Networks*, 17(6):1487–1499, 2006.*

[60] *Kevin Huang and Shuzhong Zhang. A unifying framework of accelerated first-order approach to strongly monotone variational inequalities.* arXiv preprint arXiv:2103.15270*, 2021.*

[61] *Kevin Huang and Shuzhong Zhang. New first-order algorithms for stochastic variational inequalities.* SIAM Journal on Optimization*, 32(4):2745–2772, 2022.*

[62] *Radu Ioan Bot and Ernö Robert Csetnek. A second order dynamical system with hessian-driven damping and penalty term associated to variational inequalities.* arXiv e-prints*, pages arXiv–1608, 2016.*

[63] *Alfredo Iusem. An iterative algorithm for the variational inequality problem.* Computational and Applied Mathematics*, 13:103–114, 1994.*

[64] *Chinedu Izuchukwu, Akindele Adebayo Mebawondu, and Oluwatosin Temitope Mewomo. A new method for solving split variational inequality problems without co-coerciveness.* Journal of Fixed Point Theory and Applications*, 22:1–23, 2020.*

[65] *Pham Duy Khanh and Phan Tu Vuong. Modified projection method for strongly pseudomonotone variational inequalities.* Journal of Global Optimization*, 58(2):341–350, 2014.*

[66] *David Kinderlehrer and Guido Stampacchia.* An introduction to variational inequalities and their applications*. SIAM, 2000.*

[67] *Igor Konnov.* Equilibrium models and variational inequalities*. Elsevier, 2007.*

[68] *Qihang Lin, Mingrui Liu, Hassan Rafique, and Tianbao Yang. Solving weakly-convex-weakly-concave saddle-point problems as successive strongly monotone variational inequalities.* arXiv*, 2018, 2018.*

[69] *Yu Malitsky. Projected reflected gradient methods for monotone variational inequalities.* SIAM Journal on Optimization*, 25(1):502–520, 2015.*

[70] *Patrice Marcotte and Jean-Pierre Dussault. A sequential linear programming algorithm for solving monotone variational inequalities.* SIAM Journal on Control and Optimization*, 27(6):1260–1278, 1989.*

[71] *George J Minty. On the generalization of a direct method of the calculus of variations. 1967.*

[72] *Lawan Bulama Mohammed. Iterative methods for solving split common fixed point problems in hilbert spaces. 2016.*

[73] *Renato Monteiro and Benar Svaiter. Iteration-complexity of a newton proximal extragradient method for monotone variational inequalities and inclusion problems.* SIAM Journal on Optimization*, 22(3):914–935, 2012.*

[74] *Boris S Mordukhovich.* Second-order variational analysis in optimization, variational stability, and control: theory, algorithms, applications*. Springer Nature, 2024.*

[75] *Kanikar Muangchoo, Nasser Aedh Alreshidi, and Ioannis K Argyros. Approximation results for variational inequalities involving pseudomonotone bifunction in real hilbert spaces.* Symmetry*, 13(2):182, 2021.*

[76] *Sam B Nadler Jr. Multi-valued contraction mappings. 1969.*

[77] *Muhammad Aslam Noor. General variational inequalities.* Applied mathematics letters*, 1(2):119–122, 1988.*

[78] *Muhammad Aslam Noor, Khalida Inayat Noor, and Michael Rassias. New trends in general variational inequalities.* Acta Applicandae Mathematicae*, 170:981–1064, 2020.*

[79] *Muhammad Aslam Noor, Khalida Inayat Noor, and Michael Th Rassias. New trends in general variational inequalities.* Acta Applicandae Mathematicae*, 170:981–1064, 2020.*

[80] *John Tinsley Oden and Noboru Kikuchi. Theory of variational inequalities with applications to problems of flow through porous media.* International Journal of Engineering Science*, 18(10):1173–1284, 1980.*

[81] *Jolaoso Lateef Olakunle, Taiwo Adeolu, Alakoya Timilehin Opeyemi, and Mewomo Oluwatosin Temitope. A strong convergence theorem for solving pseudo-monotone variational inequalities using projection methods.* Journal of Optimization Theory and Applications*, 185(3):744–766, 2020.*

[82] *Xiaowei Pan, Zhongxin Liu, and Zengqiang Chen. Distributed optimization with finite-time convergence via discontinuous dynamics. In* 2018 37th Chinese Control Conference (CCC)*, pages 6665–6669. IEEE, 2018.*

[83] *Massimo Pappalardo, Mauro Passacantando, et al. Stability for equilibrium problems: from variational inequalities to dynamical systems.* Journal of Optimization Theory and Applications*, 113(3):567–582, 2002.*

[84] *Boris Polyak. Some methods of speeding up the convergence of iteration methods.* Ussr computational mathematics and mathematical physics*, 4(5):1–17, 1964.*

[85] *Boris Polyak. Introduction to optimization. 1987.*

[86] *Roman A Polyak et al.* Introduction to continuous optimization*, volume 172. Springer, 2021.*

[87] *Ryazantseva. On continuous first-order methods and their regularized versions for mixed variational inequalities.* Differential Equations*, 48:1005–1017, 2012.*

[88] *Hassan Saoud. Semistability of first-order evolution variational inequalities.* Electronic Journal of Differential Equations*, 2015(265):1–10, 2015.*

[89] *Gesualdo Scutari, Daniel P Palomar, Francisco Facchinei, and Jong-Shi Pang. Convex optimization, game theory, and variational inequality theory.* IEEE Signal Processing Magazine*, 27(3):35–49, 2010.*

[90] *Michael Sedlmayer, Dang-Khoa Nguyen, and Radu Ioan Bot. A fast optimistic method for monotone variational inequalities. In* International Conference on Machine Learning*, pages 30406–30438. PMLR, 2023.*

[91] *Uday V Shanbhag. Stochastic variational inequality problems: Applications, analysis, and algorithms. In* Theory driven by influential applications*, pages 71–107. Informs, 2013.*

[92] *Yekini Shehu, Qiao-Li Dong, and Dan Jiang. Single projection method for pseudo-monotone variational inequality in hilbert spaces.* Optimization*, 68(1):385–409, 2019.*

[93] *Mikhail V Solodov. Merit functions and error bounds for generalized variational inequalities.* Journal of Mathematical Analysis and Applications*, 287(2):405–414, 2003.*

[94] *Weijie Su, Stephen Boyd, and Emmanuel J Candes. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights.* Journal of Machine Learning Research*, 17(153):1–43, 2016.*

[95] *Juhe Sun, Jein-Shan Chen, and Chun-Hsu Ko. Neural networks for solving second-order cone constrained variational inequality problem.* Computational optimization and applications*, 51:623–648, 2012.*

[96] *Kouichi Taji, Masao Fukushima, and Toshihide Ibaraki. A globally convergent newton method for solving strongly monotone variational inequalities.* Mathematical programming*, 58:369–383, 1993.*

[97] *Duong Viet Thong and Phan Tu Vuong. R-linear convergence analysis of inertial extragradient algorithms for strongly pseudo-monotone variational inequalities.* Journal of Computational and Applied Mathematics*, 406:114003, 2022.*

[98] *Paul Tseng. A modified forward-backward splitting method for maximal monotone mappings.* SIAM Journal on Control and Optimization*, 38(2):431–446, 2000.*

[99] *Phan Tu Vuong. Minimization of convex functions on convex sets by means of differential equations.* Differential equations*, 30(9):1365–1375, 1994.*

[100] *Phan Tu Vuong. A second order dynamical system and its discretization for strongly pseudo-monotone variational inequalities.* SIAM Journal on Control and Optimization*, 59(4):2875–2897, 2021.*

[101] *Li Wang, Xingxu Chen, and Juhe Sun. The second-order differential equation system with the controlled process for variational inequality with constraints.* Complexity*, 2021(1):9936370, 2021.*

[102] *Joa Weber. Introduction to sobolev spaces.* Lecture Notes MM692*, 2, 2018.*

[103] *Hui-Hua Wu and Shanhe Wu. Various proofs of the cauchy-schwarz inequality.* Octogon mathematical magazine*, 17(1):221–229, 2009.*

[104] *Youshen Xia, Henry Leung, and Jun Wang. A projection neural network and its application to constrained optimization problems.* IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 49(4):447–458, 2002.*

[105] *Jun Yang and Hongwei Liu. Strong convergence result for solving monotone variational inequalities in hilbert space.* Numerical Algorithms*, 80:741–752, 2019.*

[106] *Tong Yang, Michael I Jordan, and Tatjana Chavdarova. Solving constrained variational inequalities via a first-order interior point-based method.* arXiv preprint arXiv:2206.10575*, 2022.*

[107] *Maryam Yashtini and Alaeddin Malek. Solving complementarity and variational inequalities problems using neural networks.* Applied Mathematics and Computation*, 190(1):216–230, 2007.*

[108] *Haiyun Zhou, Yu Zhou, and Guanghui Feng. Iterative methods for solving a class of monotone variational inequality problems with applications.* Journal of Inequalities and Applications*, 2015:1–17, 2015.*