Optimal Actuator Design for Nonlinear Systems

by

M. Sajjad Edalatzadeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Dr. Irena Lasiecka Professor, Department of Mathematical Sciences, University of Memphis, 3720 Alumni Ave, Memphis, TN 38152, USA
Internal-External Member:	Dr. Sriram Narasimhan Professor, Departments of Civil and Environmental Engineering, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
Internal Member:	Dr. Brian Ingalls Associate Professor, Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
Internal Member:	Dr. Xinzhi Liu Professor, Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
Supervisor(s):	Dr. Kirsten Morris Professor, Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

1. M. S. Edalatzadeh and K. A. Morris. *Concurrent optimal control and actuator design for semi-linear systems*. The 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS), Hong Kong University of Science and Technology, Hong Kong, pp. 212-215, 2018. I wrote the extended abstract and was the corresponding author. Kirsten Morris gave comments on the arrangement of the paper as well as the presentation of the paper at the conference.

2. M. S. Edalatzadeh and K. A. Morris. *Stability and well-posedness of a nonlinear railway track model*. IEEE Control Systems Letters, 3(1):162-167, 2019. I wrote the paper and was the corresponding author. Kirsten Morris gave comments on the arrangement of the paper as well as the validation of the proofs (see Chapter 5).

3. M. S. Edalatzadeh and K. A. Morris. *Optimal actuator design for semilinear systems*. SIAM Journal on Control and Optimization, 57(4):2992-3020, 2019. I wrote the paper and was the corresponding author. Kirsten Morris gave comments on the arrangement of the paper and the validation of the proofs. She also edited parts of the paper (see Chapter 3 and Section 2.4).

4. M. S. Edalatzadeh and K. A. Morris. *Optimal actuator design for nonlinear partial differential equations*. CPDE-CDPS'19, Oaxaca, Mexico, available on IFAC PapersOnLine 52(2):128–131, 2019. I wrote the extended abstract and was the corresponding author. Kirsten Morris gave comments on the arrangement of the paper as well as the presentation of the paper at the conference (see Chapter 6).

5. M. S. Edalatzadeh and K. A. Morris. *Optimal controller and actuator design for nonlinear parabolic systems*. To be submitted, 2019. I wrote the paper and was the corresponding author. Kirsten Morris gave comments on the arrangement of the paper and the validation of the proofs. She also edited parts of the paper (see Chapter 4 and Section 2.4).

6. M. S. Edalatzadeh, D. Kalise, K. A. Morris, and K. Sturm. *Optimal actuator design for vibration control based on LQR performance and shape calculus*. To be submitted, available on arXiv: 1903.07572, 2019. I wrote the first two sections: the introduction and abstract formulation of the problem. Section 3 and 4 was written partly by me and partly by Dante Kalise. Dante Kalise wrote section 5. Kirsten Morris provided a computer solver for the beam equation, and Dante Kalise wrote the optimization code and conducted the simulations.

Abstract

For systems modeled by partial differential equations (PDE's), the location and shape of the actuators can be regarded as a design variable and included as part of the controller synthesis procedure. Optimal actuator location is a special case of optimal design. Appropriate actuator location and design can improve performance and significantly reduce the cost of the control in a distributed parameter system.

For linear partial differential equations, the existence of an optimal actuator design for a number of cost functions has been established. However, many dynamics are affected by nonlinearities and linearization of the PDE can neglect some important aspects of the model. The existing literature uses the finite-dimensional approximation of nonlinear PDE's to address the optimal actuator design problem. There are new theoretical results on the optimal actuator design of nonlinear PDE's in Banach spaces.

This thesis describes new results on optimal actuator design for abstract nonlinear systems on reflexive Banach spaces. Two classes of PDE's have been studied. In the first class, semi-linear systems, a weak continuity assumption on nonlinearities is imposed to establish optimality results. Two examples are provided for this class including nonlinear railway track model and nonlinear wave equation in two space dimensions. The second class is nonlinear parabolic PDE's. For this class, the weak continuity assumption is relaxed at the cost of imposing assumptions on the linear part of the system. The examples provided for this class are Kuramoto-Sivashinsky equation and nonlinear diffusion in two space dimensions.

Furthermore, a thorough study of optimal actuator location for nonlinear railway track model was conducted. The study begins with investigating the well-posedness and stability of solutions to this model. It is shown that under certain conditions on inputs, solutions to the railway track model are stable. Further on, using optimization algorithms and numerical schemes, an optimal input and actuator location are computed. Several simulations are run for various physical parameters. The simulations show that the optimal actuator location is not at the center of the track, contrary to a common belief. They also show that an optimally-located actuator significantly improves the performance of the control system.

Acknowledgements

I would like to thank my supervisor, Kirsten for her guidance and support. I also would like to thank the members of the examining committee, Professor Brian Ingalls, Professor Sriram Narasimhan, Professor Xinzhi Liu, Professor Irena Lasiecka for their helpful insights and time.

Dedication

To my parents, Akram and Kamal, who taught me by example the value of compassion and hard work. This thesis is a product of their love and encouragement.

Table of Contents

List of Figures

List of Symbols

c	Generic constant.
t	Temporal variable.
ξ	Spatial variable.
$\mathbb{X}$	Reflexive Banach space.
$\mathcal{L}(\mathbb{X})$	The space of bounded linear operators on $\mathbb{X}$.
$\mathbb{X}^*$	Dual of $\mathbb{X}$.
$\mathbb{H}$	Hilbert space.
$\lVert\cdot\rVert$	Norm on $\mathbb{X}$ or $\mathcal{L}(\mathbb{X})$.
$\lVert\cdot\rVert_p$	Norm on $L^p(0,\tau;\mathbb{X})$ or $L^p(0,\tau;\mathbb{U})$.
$\langle\cdot,\cdot\rangle$	Inner product on $\mathbb{H}$.
$\to$	Strong convergence on $\mathbb{X}$.
$\rightharpoonup$	Weak convergence on $\mathbb{X}$.
$\mathbb{X}_1 \hookrightarrow \mathbb{X}_2$	Dense and continuous embedding of $\mathbb{X}_1$ into $\mathbb{X}_2$.
$\mathcal{A}$, $\mathcal{Q}$	Linear operators on Banach spaces.
$\mathcal{A}^*$	Adjoint of $\mathcal{A}$.
$\mathcal{A}(\cdot)$, $\mathcal{F}(\cdot)$	Nonlinear operators on Banach spaces.
$\boldsymbol{x}(t)$	State evolving in a Banach space.
$\boldsymbol{x}^o$, $\boldsymbol{u}^o$	Optimal state and optimal input, respectively.
$\tilde{\boldsymbol{x}}$	State of a linearized system.
$C^m(I;\mathbb{X})$	The space of m times continuously differentiable $\mathbb{X}$-valued

	function over the interval I.
$C(0,\tau;\mathbb{X})$	Short notation for $C([0,\tau];\mathbb{X})$.
$L^p_{loc}(I;\mathbb{U})$	The space of strongly measurable functions $\boldsymbol{u}: I \to \mathbb{U}$, $t \to \boldsymbol{u}(t)$, for which $\|\boldsymbol{u}(t)\|_{\mathbb{U}}$ is in $L^p_{loc}(I,\mathbb{R})$.
$L^p(0,\tau;\mathbb{U})$	Short notation for $L^p([0,\tau];\mathbb{U})$.
$PC(\mathbb{R}^+;\mathbb{U})$	The space of all bounded, piecewise continuous $\mathbb{U}$-valued functions over $\mathbb{R}^+$.
$C^s(I;\mathbb{X})$	The space of Hölder continous $\mathbb{X}$-valued functions with the exponent s.
$c^s(I;\mathbb{X})$	The space of little Hölder continous $\mathbb{X}$-valued functions with the exponent s.
$W^{m,p}(I;\mathbb{X})$	The space of all strongly measurable functions $\boldsymbol{x}: I \to \mathbb{X}$ for which $\|\boldsymbol{x}(t)\|_{\mathbb{X}}$ is in $W^{m,p}(I,\mathbb{R})$.

Chapter 1

Introduction

Actuator location and design are important design variables in controller synthesis for distributed parameter systems. Finding the best actuator location to control a distributed parameter system can significantly reduce the cost of the control and improve its effectiveness [43, 97, 98, e.g.]. For example, a static control with a properly located actuator yields better performance on a structure than dynamic control with actuation at a different location [43]. There is a considerable interest in this problem in engineering [48, 117], in particular, in vibration suppression of flexible structures, see Figure 1.1.

Optimal actuator design involves finding the best control input. An optimal control input that can be expressed as a feedback law is more desirable. For linear partial differential equations (PDE's) and for quadratic cost functions, feedback laws can be determined by solving the so-called Riccati equations. The existence of a solution to the Riccati equations and its properties are discussed in [78, 77]. If the objective is to reduce the response of the linear system to disturbances, H_2 and H_∞ cost functions are considered, [122, 118].

The optimal actuator design problem has been discussed by many researchers in various contexts. The existence of an optimal actuator location for a number of cost functions has been established. In [95], it is proven that an linear-quadratic optimal actuator location exists if the control operator continuously depends on actuator locations. Further conditions on operators and cost functions are stated that guarantee the convergence of optimal locations calculated using numerical schemes to an optimal location for the original PDE, see [95]. Similar results have been obtained for H_2 and H_∞ controller design objectives in [96, 68]. Optimal actuator shape design of linear parabolic PDE's is considered in [66]. There are other objectives such as maximizing controllability or stability margin [31, 58, 106, e.g.].

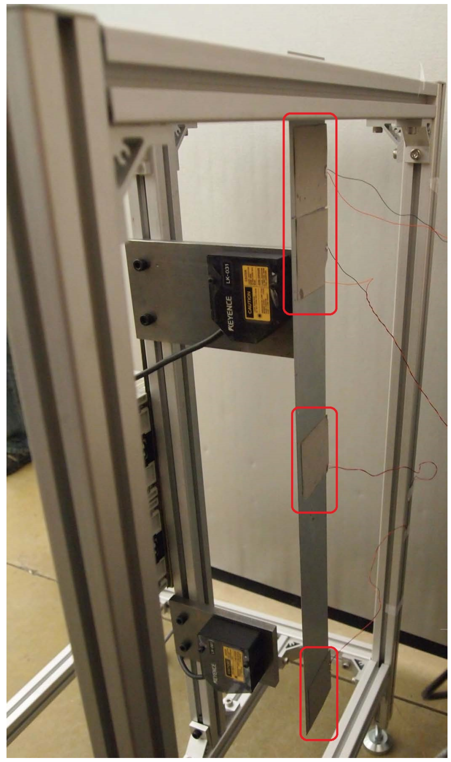

Figure 1.1: Actuators are used to control the vibrations of a flexible structure, [28, Figure 12][1](actuators are indicated in the photo).

Many papers have discussed optimal control for nonlinear distributed parameter systems; but few have looked into actuator design problem of such systems. Using a finite-dimensional approximation of the original PDE model, optimal actuator location in nonlinear PDE's has been addressed for some applications including the Kuramoto-Sivashinsky equation, semilinear transport-reaction equation as well as composite piezo-elastic plates [7, 83, 6, 94, 108]. These papers do not discuss the existence of an optimizer and optimality in infinite-dimensions and consider only a finite-dimensional approximation of the model. To our knowledge, there are no theoretical results on optimal actuator design of nonlinear distributed parameter systems. This thesis develops results on the theory of concurrent optimal actuator/controller design for nonlinear PDE's. This extends earlier results for linear PDE's, as well as insight into optimality equations. It also provides a theoretical framework for numerical studies.

[1]DOI: 10.1088/0964-1726/22/3/035001 ©IOP Publishing. Reproduced with permission. All rights reserved.

One motivator for research into optimal actuator design for nonlinear systems is the significant effect that nonlinearities have on dynamics of some systems. These systems cannot be accurately modelled by linear differential equations. Control of systems modelled by nonlinear PDE's has been studied for a number of applications, including wastewater treatment systems [85], steel cooling plants [116], oil extraction through a reservoir [80], solidification models in metallic alloys [13], thermistors [62], Schlögl model [14, 17], FitzHugh–Nagumo system [17], micro-beam model [36], static elastoplasticity [29], type-II superconductivity [120], Fokker-Planck equation [47], Schrödinger equation with bilinear control [23], Cahn-Hilliard-Navier-Stokes system [60], wine fermentation process [88], time-dependent Kohn-Sham model [112], elastic crane-trolley-load system [70], and railway track model [39]. In these papers, a cost function is being minimized subject to a nonlinear PDE model.

A review of PDE-constrained optimization theory can be found in the books [61, 79, 115]. The theory on optimal control of nonlinear PDE's has focused on PDE models with specific structures. For example, in [15, 107], first order differential equations with elliptic operators was considered. State-constrained optimal control of specific nonlinear PDE models has also been studied. In [12], the authors investigated the structure of Lagrange multipliers for state constrained optimal control problem of linear elliptic PDE's. Second order optimality conditions have also been discussed for specific nonlinear PDE models. In [16], second order optimality conditions are derived for nonlinear first-order elliptic and parabolic PDE models in space dimensions equal or less than three. Optimal control of differential equations in abstract spaces has been discussed only in a few papers. In [89], a nonlinear parabolic system in abstract Banach spaces has been studied. In [16], second order optimality conditions have been derived for a general optimization problem with equality, inequality, and point-wise constraints. This thesis discusses the concurrent optimal control and actuator design for differential equations on abstract reflexive Banach spaces.

Another motivator is the application of actuators on flexible structures, see Figure 1.1. Various models for flexible structures have been suggested including linear and nonlinear Euler-Bernoulli and Timoshenko beam models [69, 74, 36, e.g.]. In nonlinear flexible structures, the nonlinearity typically is on deformations, not on the rate of deformations. The space in which deformations evolve is compactly embedded in that of rate of deformations. As a result, the nonlinear terms are weakly continuous in the underlying state space. This property will be exploited in this thesis to study the optimal actuator design for flexible structures. One application of the results in this study is to the development of an optimal control strategy for the vibration suppression of railway tracks [5, 27, 39]. The railway track model predicts the dynamic behavior of railway tracks and its underneath foundation, see Figure 1.2. The behavior of the foundation introduces nonlinearity into the model. This application is primarily motivated by the need for an optimal control strategy in the vibration suppression of railway tracks [27].

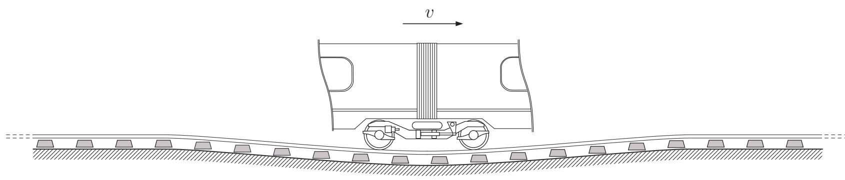

Figure 1.2: Deflection of railway track [72, Figure 7][2].

The contribution of this thesis is threefold. Chapter 3, Chapter 4, and Chapter 5 present each contribution. The content of each chapter is submitted and published in peer-reviewed journals [37, 38, 39]. Chapter 3 describes new results on optimal actuator design for semilinear systems on abstract spaces. The existing literature lacks theoretical results on optimal actuator design of nonlinear systems. In this chapter, a general assumption on the linear part of the system is made that will include both hyperbolic and parabolic systems. The specific examples given in this chapter have not been discussed elsewhere. Chapter 4 presents new results on concurrent optimal control and actuator design of parabolic systems on abstract spaces. Optimal control of abstract parabolic systems has already been studied in the literature; however, this chapter relax some of the existing assumptions, and also extends these results to optimal actuator design. The specific examples given in this chapter have already been studied using finite-dimensional approximations, but not in an infinite-dimensional setting. This chapter derives optimality conditions in infinite-dimensional spaces. In addition, new results on the existence of a solution to abstract nonlinear systems were obtained. These results are presented in Section 2.4 to improve the flow of the contents; however, they are in fact part of the results in Chapters 3 and 4. Chapter 5 studies the optimal control and actuator design of the railway track model. The studies on vibration control of railway tracks are limited to passive vibration strategies not active ones. The literature also lacks study on stability of the nonlinear railway track model in the presence of an input. This chapter begins with results on well-posedness and stability of a nonlinear controlled railway track model. Further on, using optimization algorithms and numerical schemes, the optimal input and actuator location are computed.

The thesis is organized as follows: Chapter 2 is the mathematical background of the thesis. Chapter 3 investigates the optimal actuator design for semilinear systems. Two examples are provided in this section including the nonlinear railway track model and nonlinear wave equation in two space dimensions. Chapter 4 focuses on nonlinear parabolic systems; the examples in this chapter are Kuramoto-Sivashinsky equation and nonlinear diffusion in two space dimensions. These examples do not satisfy the assumptions of the previous chapter. Chapter 5 studies the optimal control and actuator design for the nonlinear railway track model in more detail. Numerical simulations are presented in this chapter. The thesis concludes with Chapter 6 where a summary of results as well as future research directions is mentioned.

Chapter 2

Background

This thesis focuses on abstract differential equations on reflexive Banach spaces over the field of complex numbers. These spaces are denoted by blackboard letters such as $\mathbb{X}$ and $\mathbb{Y}$. This chapter provides a review of these spaces as well as the control systems defined on such spaces. The materials of Section 2.1 and Section 2.3.1 are in [26]; Section 2.2 is in [104], and Section 2.3.2 is in [95]. In addition, new results on the existence of a solution to abstract nonlinear systems were obtained. These results are presented in Section 2.4 to improve the flow of the contents; however, they are in fact part of the results in Chapters 3 and 4.

2.1 Operator Theory

Operators on Banach spaces can be categorized as linear and nonlinear operators.

Definition 2.1.1. *A* linear operator *$\mathcal{A}$ from $\mathbb{X}$ to $\mathbb{Y}$ is a map $\mathcal{A} : D(\mathcal{A}) \subset \mathbb{X} \to \mathbb{Y}$, such that $D(\mathcal{A})$ is a subspace of $\mathbb{X}$, and for all $\boldsymbol{x}_1,\ \boldsymbol{x}_2 \in D(\mathcal{A})$ and scalars α, it holds that*

$$\mathcal{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \mathcal{A}\boldsymbol{x}_1 + \mathcal{A}\boldsymbol{x}_2, \tag{2.1}$$

$$\mathcal{A}(\alpha\boldsymbol{x}) = \alpha\mathcal{A}\boldsymbol{x}. \tag{2.2}$$

A bounded linear operator on a Banach space is defined over the whole space, whereas an unbounded operator is only defined on a subspace of the underlying Banach space.

Definition 2.1.2. *Let $\mathcal{A}$ be a linear operator from $D(\mathcal{A})$ to $\mathbb{Y}$, $\mathcal{A}$ is a* bounded linear operator *if $D(\mathcal{A}) = \mathbb{X}$, and there exists a real number c such that for all $\boldsymbol{x} \in \mathbb{X}$*

$$\|\mathcal{A}\boldsymbol{x}\|_{\mathbb{Y}} \leq c \|\boldsymbol{x}\|_{\mathbb{X}} . \tag{2.3}$$

A bounded linear functional, f, is a bounded linear operator from $\mathbb{X}$ to $\mathbb{C}$.

Definition 2.1.3. *The Banach space $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ is the space of all bounded linear operators from $\mathbb{X}$ to $\mathbb{Y}$ equipped with the norm*

$$\|\mathcal{A}\|_{\mathcal{L}(\mathbb{X},\mathbb{Y})} = \sup_{\|\boldsymbol{x}\|=1} \|\mathcal{A}\boldsymbol{x}\|_{\mathbb{Y}} . \tag{2.4}$$

A consequence of the previous definition is that

$$\|\mathcal{A}\boldsymbol{x}\|_{\mathbb{Y}} \leq \|\mathcal{A}\|_{\mathcal{L}(\mathbb{X},\mathbb{Y})} \|\boldsymbol{x}\|_{\mathbb{X}} . \tag{2.5}$$

Definition 2.1.4. *The Banach spaces $\mathbb{X}$ and $\mathbb{Y}$ are* isometrically isomorphic *if there exists a linear bounded map $\mathcal{R} : \mathbb{X} \to \mathbb{Y}$ with properties*

1. *If $\mathcal{R}\boldsymbol{x} = 0$ then $\boldsymbol{x} = 0$,*
2. *Range$(\mathcal{R}) = \mathbb{Y}$,*
3. *$\|\mathcal{R}\boldsymbol{x}\|_{\mathbb{Y}} = \|\boldsymbol{x}\|_{\mathbb{X}}$ for all $\boldsymbol{x} \in \mathbb{X}$.*

The space $\mathcal{L}(\mathbb{X}, \mathbb{C})$ is referred to as the (algebraic) dual of $\mathbb{X}$, and is denoted by $\mathbb{X}^*$. The dual of $\mathbb{X}^*$ is denoted by $\mathbb{X}^{**}$. Each element $\boldsymbol{x} \in \mathbb{X}$ gives rise to a bounded linear functional $f_{\boldsymbol{x}}$ in $\mathbb{X}^{**}$ by

$$f_{\boldsymbol{x}}(g) = g(\boldsymbol{x}), \quad \forall g \in \mathbb{X}^*. \tag{2.6}$$

It can be shown that the mapping $\boldsymbol{x} \mapsto f_{\boldsymbol{x}}$ is an isometric isomorphism of $\mathbb{X}$ into $\mathbb{X}^{**}$. This mapping is called the *natural embedding* of $\mathbb{X}$ in $\mathbb{X}^{**}$.

Definition 2.1.5. *A Banach space $\mathbb{X}$ is reflexive if its second dual, $\mathbb{X}^{**}$, is isometrically isomorphic to $\mathbb{X}$ under the natural embedding.*

From now on, the Banach spaces $\mathbb{X}$ and $\mathbb{Y}$ are reflexive.

Continuity and boundedness are equivalent concepts for linear operators.

Theorem 2.1.6. *[109, Theorem 5.6.4 and Lemma 5.6.5] If $\mathcal{A} : D(\mathcal{A}) \subset \mathbb{X} \to \mathbb{Y}$ is a linear operator, then*

a. $\mathcal{A}$ is continuous if and only if $\mathcal{A}$ is bounded.

b. If $\mathcal{A}$ is continuous at a single point, it is continuous on $\mathbb{X}$.

Contraction mappings are a particular class of linear or nonlinear operators on Banach spaces.

Definition 2.1.7. *The nonlinear operator $\mathcal{G}(\cdot)$ is a contraction mapping on $\mathbb{X}$, if there are $m \in \mathbb{N}$ and $\alpha < 1$ such that $\mathcal{G}(\cdot)$ satisfies*

$$\|\mathcal{G}^m(\boldsymbol{x}_2) - \mathcal{G}^m(\boldsymbol{x}_2)\| \leq \alpha \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$$

for all $\boldsymbol{x}_1, \ \boldsymbol{x}_2 \in \mathbb{X}$.

Theorem 2.1.8 (Contration Mapping Theorem)**.** *[100, Theorem 3.15.2 and Corollary 3.15.3] If $\mathcal{G}(\cdot)$ is a contraction mapping on $\mathbb{X}$, then there exists a unique $\boldsymbol{x}^* \in \mathbb{X}$ such that $\mathcal{G}(\boldsymbol{x}^*) = \boldsymbol{x}^*$; $\boldsymbol{x}^*$ is the fixed point of $\mathcal{G}(\cdot)$.*

A compact set is a closed bounded set containing only sequences that have at least one convergent subsequence. A set is relatively compact if its closure is compact.

Definition 2.1.9. *An operator $\mathcal{A} \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ is said to be a compact operator if $\mathcal{A}$ maps bounded sets of $\mathbb{X}$ onto relatively compact sets of $\mathbb{Y}$.*

An equivalent definition is that $\mathcal{A}$ is linear and for any bounded sequence $\{\boldsymbol{x}_n\}$ in $\mathbb{X}$, $\{\mathcal{A}\boldsymbol{x}_n\}$ has a convergent subsequence in $\mathbb{Y}$.

2.2 Semigroup of Operators

Semigroup theory provides an abstract setting for studying infinite-dimensional dynamical systems.

Definition 2.2.1 (strongly continuous semi-group)**.** *[26, Definition 2.1.2] A strongly continuous semigroup (also C_0-semigroup or simply semigroup) is an operator-valued function $\mathcal{T}(t)$ from $\mathbb{R}_0^+$ (the non-negative real line) to $\mathcal{L}(\mathbb{X})$ that satisfies the following properties:*

$$\begin{aligned}
&\mathcal{T}(t+s) = \mathcal{T}(t)\mathcal{T}(s), \ \text{for } t, \ s \geq 0, \\
&\mathcal{T}(0) = \mathcal{I}, \\
&\|\mathcal{T}(t)\boldsymbol{x}_0 - \boldsymbol{x}_0\| \to 0 \ \text{as } t \to 0^+, \quad \forall \boldsymbol{x}_0 \in \mathbb{X}.
\end{aligned}$$

Theorem 2.2.2. *[104, Theorem 2.2 and Corollary 2.3] Let $\mathcal{T}(t)$ be a C_0-semigroup.*

1. *There exist constants $\omega \geq 0$ and $M \geq 1$ such that*
$$\|\mathcal{T}(t)\| \leq Me^{\omega t}$$
for $0 \leq t < \infty$.

2. *For every $\boldsymbol{x} \in \mathbb{X}$, $t \mapsto \mathcal{T}(t)\boldsymbol{x}$ is a continuous function from $\mathbb{R}_0^+$ into $\mathbb{X}$.*

In the last theorem, if $M = 1$ and $\omega \leq 0$, the semigroup is said to be a *semigroup of contraction*. If $\omega < 0$, the semigroup is said to be an *exponentially stable semigroup*.

The infinitesimal generator of a C_0-semigroup $\mathcal{T}(t)$ is the linear operator $\mathcal{A}$ defined by

$$D(\mathcal{A}) = \left\{ \boldsymbol{x} \in \mathbb{X} : \lim_{t \downarrow 0} \frac{\mathcal{T}(t)\boldsymbol{x} - \boldsymbol{x}}{t} \text{ exists} \right\}$$

and

$$\mathcal{A}\boldsymbol{x} = \lim_{t \downarrow 0} \frac{\mathcal{T}(t)\boldsymbol{x} - \boldsymbol{x}}{t}, \quad \forall \boldsymbol{x} \in D(\mathcal{A}). \tag{2.7}$$

A C_0-semigroup is not necessarily uniformly bounded; this is, the limit

$$\lim_{t \downarrow 0} \|\mathcal{T}(t) - \mathcal{I}\| = 0.$$

may not hold for some C_0-semigroups.

Theorem 2.2.3. *[104, Theorem 1.2] A linear operator $\mathcal{A}$ is the infinitesimal generator of a uniformly continuous semigroup if and only if $\mathcal{A}$ is a bounded linear operator.*

Definition 2.2.4. *A linear operator $\mathcal{A}$ is dissipative if and only if $\|(\lambda\mathcal{I} - \mathcal{A})\boldsymbol{x}\| \geq \lambda \|\boldsymbol{x}\|$ for all $\boldsymbol{x} \in D(\mathcal{A})$ and $\lambda > 0$.*

Theorem 2.2.5 (Lumer-Phillips)**.** *Let $\mathcal{A}$ be a linear operator with dense domain $D(\mathcal{A})$ in $\mathbb{X}$. If $\mathcal{A}$ is dissipative and there is a $\lambda_0 > 0$ such that the range of $\lambda_0\mathcal{I} - \mathcal{A}$ is $\mathbb{X}$, then $\mathcal{A}$ is the infinitesimal generator of a C_0-semigroup of contraction on $\mathbb{X}$.*

The next lemma on adjoint of a semigroup will be used in future chapters.

Theorem 2.2.6. *[104, Lemma 10.5 and Corrolary 10.6] Let $\mathbb{X}$ be a reflexive Banach space and let $\mathcal{T}(t)$ be a C_0-semigroup on $\mathbb{X}$ with infinitesimal generator $\mathcal{A}$. The adjoint semigroup $\mathcal{T}(t)^*$ of $\mathcal{T}(t)$ is a C_0-semigroup on $\mathbb{X}^*$ whose infinitesimal generator is $\mathcal{A}^*$ the adjoint of $\mathcal{A}$.*

Analytic semigroups are an important class of semigroups and will be used later in this thesis.

Definition 2.2.7. *Let $\Delta = \{z \in \mathbb{C} : \varphi_1 < arg(z) < \varphi_2,\ \varphi_1 < 0 < \varphi_2\}$ be a sector in the complex plane, and for $z \in \Delta$, let $\mathcal{T}(z)$ be a bounded linear operator. The family $\mathcal{T}(z)$, $z \in \Delta$ is an analytic semigroup in Δ if*

1. *$z \to \mathcal{T}(z)$ is analytic in Δ.*
2. *$\mathcal{T}(0) = \mathcal{I}$ and $\lim\limits_{\substack{z \to 0 \\ z \in \Delta}} \mathcal{T}(z)\boldsymbol{x} = \boldsymbol{x}$ for every $\boldsymbol{x} \in \mathbb{X}$.*
3. *$\mathcal{T}(z_1 + z_2) = \mathcal{T}(z_1)\mathcal{T}(z_2)$ for $z_1,\ z_2 \in \Delta$*

A semigroup $\mathcal{T}(t)$ is said to be analytic if it is analytic in some sector Δ containing the non-negative real axis.

2.3 Linear Control Systems

A linear control system with a state trajectory $\boldsymbol{x}(t)$ and input $\boldsymbol{u}(t)$, state operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ and input operator $\mathcal{B} : \mathbb{U} \to \mathbb{X}$ is described by

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{B}\boldsymbol{u}(t), \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{X}. \end{cases} \tag{2.8}$$

Different notions of solution can be defined for a linear system.

Definition 2.3.1 (classical solution)**.** *[26, Definition 3.1.1][104, Definition 4.2.1] A function $\boldsymbol{x} : [0, \tau) \to \mathbb{X}$ is a classical solution to (2.8) on $[0, \tau)$ if $\boldsymbol{x}$ is continuous on $[0, \tau)$, continuously differentiable on $(0, \tau)$, $\boldsymbol{x}(t) \in D(\mathcal{A})$ for $0 < t < \tau$, and (2.8) is satisfied.*

Certain conditions must hold in order to ensure that a linear system has a classical solution.

Theorem 2.3.2. *[104, Theorem 2.4 and Corollary 2.5] Let $\mathcal{A}$ be the infinitesimal generator of a C_0-semigroup $\mathcal{T}(t)$, and let $\mathcal{B} \in \mathcal{L}(\mathbb{U}, \mathbb{X})$. If $\boldsymbol{u}(t)$ is continuously differentiable on $[0, \tau]$ then (2.8) has a classical solution on $[0, \tau]$ for every $\boldsymbol{x}_0 \in D(\mathcal{A})$.*

Definition 2.3.3 (mild solution)**.** *[104, Definition 6.1.1][26, Definition 3.1.4] Let $\mathcal{A}$ be the infinitesimal generator of a strongly continuous semigroup $\mathcal{T}(t)$, $\boldsymbol{x}_0 \in \mathbb{X}$, $\mathcal{B} \in \mathcal{L}(\mathbb{U}, \mathbb{X})$, and $\boldsymbol{u} \in L^1(0, \tau; \mathbb{U})$. The state trajectory $\boldsymbol{x}(t)$ satisfying*

$$\boldsymbol{x}(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)ds, \tag{2.9}$$

defines a mild solution *to (2.8).*

2.3.1 Linear-Quadratic Optimal Control

Consider the linear system (2.8) on a Hilbert space $\mathbb{H}$. Let $\mathcal{A}$ with domain $D(\mathcal{A})$ generates a strongly continuous semigroup $\mathcal{T}(t)$ on $\mathbb{H}$, and let $\mathcal{B} \in \mathcal{L}(\mathbb{U}, \mathbb{H})$. Let $\mathcal{R}$ be a linear, self-adjoint, coercive, bounded operator on $\mathbb{U}$, and $\mathcal{C} \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ be a linear, bounded operator. Linear quadratic optimal control on finite-time horizon aims to minimize the cost function

$$J(\boldsymbol{x}, \boldsymbol{u}) := \int_0^\tau \langle \mathcal{C}\boldsymbol{x}(t), \mathcal{C}\boldsymbol{x}(t)\rangle_{\mathbb{Y}} + \langle \mathcal{R}\boldsymbol{u}(t), \boldsymbol{u}(t)\rangle_{\mathbb{U}} \, dt. \tag{2.10}$$

subject to (2.8) over all inputs in $L^2(0, \tau; \mathbb{U})$. That is,

$$\begin{cases} \min\limits_{\boldsymbol{u} \in L^2(0,\tau;\mathbb{U})} \; J(\boldsymbol{x}, \boldsymbol{u}) \\ \text{s.t.} \;\; \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{B}\boldsymbol{u}(t), \; \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{OP}$$

It is known that there is a unique input $\boldsymbol{u}^o \in L^2(0, \tau; \mathbb{U})$ that solves (OP), see [26, Theorem 6.1.4].

Lemma 2.3.4. *[26, Lemma 6.1.5] Let $\boldsymbol{u}^o(t)$ be the minimizing input function of problem* (OP) *and $\boldsymbol{x}^o(t)$ its corresponding state trajectory. For $t \in [0, \tau]$, the following hold*

$$\boldsymbol{u}^o(t) = -\mathcal{R}^{-1}\mathcal{B}^* \int_t^\tau \mathcal{T}^*(s-t)\mathcal{C}^*\mathcal{C}\boldsymbol{x}^o(s)ds. \tag{2.11}$$

Lemma 2.3.5. *[26, Lemmas 6.1.7 and 6.1.9] Let $\boldsymbol{x}^o(t)$ be the optimal state trajectory in Lemma 2.3.4. For a given $t \in [0, \tau]$, we define the following operator on $\mathbb{X}$:*

$$\Pi(t)\boldsymbol{x}_0 := \int_t^\tau \mathcal{T}^*(s-t)\mathcal{C}^*\mathcal{C}\boldsymbol{x}^o(s)ds. \tag{2.12}$$

This operator has the following properties *a.* $\Pi(t) \in \mathcal{L}(\mathbb{X})$ *for all* $t \in [0, \tau]$

b. The following relationships hold between the optimal state and the optimal input trajectory

$$\boldsymbol{u}^o(t) = -\mathcal{R}^{-1}\mathcal{B}^*\Pi(t)\boldsymbol{x}^o(t). \tag{2.13}$$

c. The following relationship holds between the minimum cost and $\Pi(\tau)$

$$J(\boldsymbol{x}^o, \boldsymbol{u}^o) = \langle \boldsymbol{x}_0, \Pi(\tau)\boldsymbol{x}_0 \rangle \tag{2.14}$$

d. $\Pi(\tau)$ *is a self-adjoint, non-negative operator*

e. If $0 \leq t_1 \leq t_2 \leq \tau$, *then* $\Pi(t_2) \leq \Pi(t_1)$

f. $\|\Pi(t_2)\|_{\mathcal{L}(\mathbb{X})} \leq \|\Pi(t_1)\|_{\mathcal{L}(\mathbb{X})}$

g. $\Pi(\cdot)$ *is strongly continuous from the right in* $[0, \tau]$, *i.e.,*

$$\lim_{h \downarrow 0} \Pi(t+h)\boldsymbol{x}_0 = \Pi(t)\boldsymbol{x}_0$$

for all $\boldsymbol{x}_0 \in \mathbb{X}$ *and* $t \in [0, \tau]$.

h. For every $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in D(\mathcal{A})$ *and* $t \in (0, \tau)$ *the function* $\langle \boldsymbol{x}_1, \Pi(t)\boldsymbol{x}_2 \rangle$ *is differentiable and satisfies the differential Riccati equation*

$$\begin{aligned} \frac{d}{dt}\langle \boldsymbol{x}_2, \Pi(t)\boldsymbol{x}_1 \rangle &= -\langle \boldsymbol{x}_2, \Pi(t)\mathcal{A}\boldsymbol{x}_1 \rangle - \langle \mathcal{A}\boldsymbol{x}_2, \Pi(t)\boldsymbol{x}_1 \rangle \\ &\quad - \langle \mathcal{C}\boldsymbol{x}_2, \mathcal{C}\boldsymbol{x}_1 \rangle_{\mathbb{Y}} + \left\langle \Pi(t)\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\Pi(t)\boldsymbol{x}_2, \boldsymbol{x}_1 \right\rangle, \\ \Pi(\tau) &= 0. \end{aligned}$$

Furthermore, it is the unique solution of this differential Riccati equation in the class of strongly continuous, self-adjoint operators in $\mathcal{L}(\mathbb{X})$ *such that* $\langle \boldsymbol{x}_1, \Pi(t)\boldsymbol{x}_2 \rangle$ *is differentiable for* $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in D(\mathcal{A})$ *and* $t \in (0, \tau)$.

Since the time interval considered in (OP) is finite, the problem is also referred to as finite-time horizon optimal control problem. The cost function of an infinite-time horizon problem is

$$J(\boldsymbol{x}, \boldsymbol{u}) := \int_0^\infty \langle \mathcal{C}\boldsymbol{x}(t), \mathcal{C}\boldsymbol{x}(t) \rangle_{\mathbb{Y}} + \langle \mathcal{R}\boldsymbol{u}(t), \boldsymbol{u}(t) \rangle_{\mathbb{U}} \, dt. \tag{2.15}$$

The stability of the system is important in the study of the infinite-time optimal control problem.

Definition 2.3.6. *The system* (2.8) *with cost* (2.10) *is* optimizable *if for every* $\boldsymbol{x}_0$ *there exists* $\boldsymbol{u} \in L^2(0,\infty;\mathbb{U})$ *such that the cost is finite.*

Definition 2.3.7. *The pair* $(\mathcal{A},\mathcal{B})$ *is stabilizable if there exists* $\mathcal{P} \in \mathcal{L}(\mathbb{X},\mathbb{U})$ *such that* $\mathcal{A}-\mathcal{B}\mathcal{P}$ *generates an exponentially stable semigroup.*

It is straightforward to show that the assumption of optimizability in the last theorem is equivalent to stabilizability.

Definition 2.3.8. *Let* $\mathcal{C} \in \mathcal{L}(\mathbb{X},\mathbb{Y})$. *The pair* $(\mathcal{A},\mathcal{C})$ *is detectable if there exists* $\mathcal{P} \in \mathcal{L}(\mathbb{Y},\mathbb{X})$ *such that* $\mathcal{A}-\mathcal{P}\mathcal{C}$ *generates an exponentially stable semigroup.*

Theorem 2.3.9. *[26, Thms. 6.2.4 and 6.2.7] If system* (2.8) *with cost* (2.10) *is optimizable and detectable, then the cost function has a minimum for every* $\boldsymbol{x}_0 \in \mathbb{X}$. *Let* $\boldsymbol{u}^o(t)$ *be the optimal input and* $\boldsymbol{x}^o(t)$ *be the corresponding optimal state trajectory. There exists a self-adjoint non-negative operator* $\Pi \in \mathcal{L}(\mathbb{H})$ *such that*

$$J(\boldsymbol{u}^o,\boldsymbol{x}^o) = \langle \boldsymbol{x}_0, \Pi \boldsymbol{x}_0\rangle . \tag{2.16}$$

The operator Π *is the minimal non-negative solution to the equation*

$$\langle \mathcal{A}\boldsymbol{x}_1, \Pi\boldsymbol{x}_2\rangle + \langle \Pi\boldsymbol{x}_1, \mathcal{A}\boldsymbol{x}_2\rangle + \langle \mathcal{C}\boldsymbol{x}_1, \mathcal{C}\boldsymbol{x}_2\rangle - \left\langle \mathcal{B}^*\Pi\boldsymbol{x}_1, \mathcal{R}^{-1}\mathcal{B}^*\Pi\boldsymbol{x}_1\right\rangle, \quad \forall \boldsymbol{x}_1,\ \boldsymbol{x}_2 \in D(\mathcal{A}). \tag{2.17}$$

Moreover, defining $\mathcal{K} = \mathcal{R}^{-1}\mathcal{B}^*\Pi$, *the corresponding optimal control is* $\boldsymbol{u}(t) = -\mathcal{K}\boldsymbol{x}(t)$ *and* $\mathcal{A}-\mathcal{B}\mathcal{K}$ *generates an exponentially stable semigroup.*

2.3.2 Linear-Quadratic Optimal Actuator Design

Let $\boldsymbol{r}$ be an actuator design parameter taking values in some topological space. Consider a linear system with a control operator $\mathcal{B}(\boldsymbol{r})$ that depends on the actuator location $\boldsymbol{r}$:

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.18}$$

There are different ways of defining a cost function to be minimized over actuator choices [95]. Linear-quadratic optimal actuator location minimizes the cost (2.15) with respect to the worst case initial condition. That is, choose $\boldsymbol{r}$ to minimize

$$\mu(\boldsymbol{r}) = \max_{\substack{\boldsymbol{x}_0\in\mathbb{X} \\ \|\boldsymbol{x}_0\|=1}} \min_{\boldsymbol{u}\in L^2(0,\infty;\mathbb{U})} J(\boldsymbol{x},\boldsymbol{u}), \tag{2.19}$$

where $\boldsymbol{x}(t)$ solves (2.18). Applying Theorem 2.3.9 and noting that the Riccati operator will also depend on the actuator choice yield

$$\max_{\substack{\boldsymbol{x}_0 \in \mathbb{X} \\ \|\boldsymbol{x}_0\|=1}} \min_{\boldsymbol{u} \in L^2(0,\infty;\mathbb{U})} J(\boldsymbol{x}, \boldsymbol{u}) = \|\Pi(\boldsymbol{r})\| . \tag{2.20}$$

If the actuator design parameter $\boldsymbol{r}$ belongs to a compact subset Ω in $\mathbb{R}^n$, then the optimal performance is

$$\inf_{\boldsymbol{r} \in \Omega} \|\Pi(\boldsymbol{r})\| . \tag{2.21}$$

Theorem 2.3.10. *[95, Theorem 2.6] Let $\mathcal{B}(\boldsymbol{r}) \in \mathcal{L}(\mathbb{U}, \mathbb{X})$ be a family of compact input operators such that for any $\boldsymbol{r}_0 \in \Omega$*

$$\lim_{\boldsymbol{r} \to \boldsymbol{r}_0} \|\mathcal{B}(\boldsymbol{r}) - \mathcal{B}(\boldsymbol{r}_0)\| = 0. \tag{2.22}$$

If $(\mathcal{A}, \mathcal{B}(\boldsymbol{r}))$ is stabilizable for all $\boldsymbol{r} \in \Omega$ and $(\mathcal{A}, \mathcal{C})$ is detectable, then the Riccati operator $\Pi(\boldsymbol{r})$ is continuous function of $\boldsymbol{r}$ in the operator norm as follows:

$$\lim_{\boldsymbol{r} \to \boldsymbol{r}_0} \|\Pi(\boldsymbol{r}) - \Pi(\boldsymbol{r}_0)\| = 0 \tag{2.23}$$

and there exists an optimal actuator location $\boldsymbol{r}^o$ such that

$$\|\Pi(\boldsymbol{r}^o)\| = \inf_{\boldsymbol{r} \in \Omega} \|\Pi(\boldsymbol{r})\| . \tag{2.24}$$

A different cost is considered for random initial conditions. If the initial condition is random, with zero mean and variance V, the cost to be minimized over the actuator choice is

$$\mu(\boldsymbol{r}) = \operatorname{trace}(V^{1/2}\Pi(\boldsymbol{r})V^{1/2}). \tag{2.25}$$

Since the Riccati operator is self-adjoint and non-negative,

$$\operatorname{trace}(V^{1/2}\Pi(\boldsymbol{r})V^{1/2}) = \left\|V^{1/2}\Pi(\boldsymbol{r})V^{1/2}\right\|_1 ,$$

where $\|\cdot\|_1$ indicates the nuclear norm. Assuming that the variance is unity, $V = 1$, the cost to be minimized becomes

$$\|\Pi(\boldsymbol{r})\|_1 . \tag{2.26}$$

Theorem 2.3.11. *[95, Theorem 2.10] Let $\mathcal{B}(\boldsymbol{r}) \in \mathcal{L}(\mathbb{U}, \mathbb{X})$, $\boldsymbol{r} \in \Omega$, be a family of input operators such that for any $\boldsymbol{r}_0 \in \Omega$*

$$\lim_{\boldsymbol{r} \to \boldsymbol{r}_0} \|\mathcal{B}(\boldsymbol{r}) - \mathcal{B}(\boldsymbol{r}_0)\| = 0. \tag{2.27}$$

Assume that $(\mathcal{A}, \mathcal{B}(\boldsymbol{r}))$ are all stabilizable and that $(\mathcal{A}, \mathcal{C})$ is detectable. If $\mathbb{U}$ and $\mathbb{Y}$ are finite-dimensional, then the corresponding Riccati operator $\Pi(\boldsymbol{r})$ is continuous function of $\boldsymbol{r}$ in the nuclear norm:

$$\lim_{\boldsymbol{r}\to\boldsymbol{r}_0} \|\Pi(\boldsymbol{r}) - \Pi(\boldsymbol{r}_0)\|_1 = 0, \tag{2.28}$$

and there exists an optimal actuator location $\boldsymbol{r}^o$ such that

$$\|\Pi(\boldsymbol{r}^o)\|_1 = \inf_{\boldsymbol{r}\in\Omega} \|\Pi(\boldsymbol{r})\|_1 . \tag{2.29}$$

The optimal actuator design also involves designing the shape of the actuators. In this case, the actuator parameter design belongs to a topological space. Papers [99, 66] discuss actuator shape design and present various examples.

2.4 Nonlinear Control Systems

An important class of nonlinear control systems consists of a nonlinear operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$, control operator $\mathcal{B} : \mathbb{U} \to \mathbb{X}$, state $\boldsymbol{x}(t)$, $t > 0$, input $\boldsymbol{u}(t)$, and initial condition $\boldsymbol{x}_0 \in \mathbb{X}$. It is described by

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}(\boldsymbol{x}(t)) + \mathcal{B}\boldsymbol{u}(t), \quad t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.30}$$

In this section, it is assumed that the operator $\mathcal{B}$ is bounded. In what follows, different types of nonlinear control systems are discussed.

2.4.1 Semilinear Systems

A semilinear system is often defined as a system with $\mathcal{A}(\boldsymbol{x}) = \mathcal{A}\boldsymbol{x} + \mathcal{F}(\boldsymbol{x})$ where the operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is linear and generates a C_0-semigroup, and the operator $\mathcal{F}(\boldsymbol{x}) : \mathbb{X} \to \mathbb{X}$ is nonlinear and continuous [104]. System (2.30) becomes

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}\boldsymbol{u}(t), \quad t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.31}$$

A common notion of solution for such systems is the *mild solution.*

Definition 2.4.1 (mild solution)**.** *[104, Definition 6.1.1] A function $\boldsymbol{x} : [0, \tau] \to \mathbb{X}$ is said to be a mild solution of (2.31) if $\boldsymbol{x}$ is in $C(0, \tau; \mathbb{X})$ and satisfies*

$$\boldsymbol{x}(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)ds. \tag{2.32}$$

The following theorem modifies [104, Theorem 6.1.4] which only allows continuous inputs.

Theorem 2.4.2. *[37, Thm 3.1] Let $\mathcal{A}$ be the infinitesimal generator of a C_0-semigroup, $\mathcal{F}(\cdot)$ be locally Lipschitz continuous on $\mathbb{X}$, and $\mathcal{B} : \mathbb{U} \to \mathbb{X}$ be a bounded linear operator, $\boldsymbol{x}_0 \in \mathbb{X}$, and $\boldsymbol{u} \in L^p_{loc}(0, \infty; \mathbb{U})$. For any $R > 0$, $\|\boldsymbol{u}\|_p \leq R$, there is $\tau > 0$ such that IVP (2.31) admits a mild solution over $[0, \tau]$.*

Proof. The idea of the proof is similar to [104, Theorem 6.1.4], with a slight modification that here $\boldsymbol{u}(t)$ is in $L^p(0, \tau; \mathbb{U})$. For any $\boldsymbol{x}_0 \in \mathbb{X}$ choose constants $\delta_0 > 0$ and $\tau > 0$ such that for all $t \in [0, \tau]$

$$\|\mathcal{T}(t)\boldsymbol{x}_0 - \boldsymbol{x}_0\| \leq \delta_0.$$

Let $\mathbb{S}$ be the closed bounded subset of $C(0, \tau; \mathbb{X})$ defined as

$$\mathbb{S} = \{\boldsymbol{x} \in C(0, \tau; \mathbb{X}) |\ \boldsymbol{x}(0) = \boldsymbol{x}_0,\ \|\boldsymbol{x}(t) - \boldsymbol{x}_0\| \leq 2\delta_0,\ \forall t \in [0, \tau]\}. \tag{2.33}$$

Define the operator $\mathcal{G}$ on $\mathbb{S}$ to be

$$\mathcal{G}(\boldsymbol{x})(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))\, ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)\, ds. \tag{2.34}$$

It will be shown that for sufficiently small τ, $\mathcal{G}$ maps $\mathbb{S}$ into $\mathbb{S}$ and is a contraction on $\mathbb{S}$.

Use the triangle inequality and write

$$\begin{aligned} \|\mathcal{G}(\boldsymbol{x})(t) - \boldsymbol{x}_0\| \leq \|\mathcal{T}(t)\boldsymbol{x}_0 - \boldsymbol{x}_0\| + \left\| \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))\, ds \right\| \\ + \left\| \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)\, ds \right\|. \end{aligned} \tag{2.35}$$

There exist a number $M_\mathcal{T} > 0$ that $\|\mathcal{T}(t)\| \leq M_\mathcal{T}$ for all $t \in [0, \tau]$. Also, there is $L_{\mathcal{F}\delta} > 0$ so that $\|\mathcal{F}(\boldsymbol{x}(s))\| \leq L_{\mathcal{F}\delta}\|\boldsymbol{x}(s)\|$ on a ball of radius $\delta = \|\boldsymbol{x}_0\| + 2\delta_0$ centered at the origin. This gives a bound for the second term on the right hand side of the inequality (2.35)

$$\left\| \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))\, ds \right\| \leq M_\mathcal{T} L_{\mathcal{F}\delta}\delta\tau. \tag{2.36}$$

An upper bound for the third right hand side term is

$$\left\| \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)\, ds \right\| \le M_{\mathcal{T}} \left\|\mathcal{B}\right\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \left\|\boldsymbol{u}\right\|_p \tau^{(p-1)/p}. \tag{2.37}$$

Applying these bounds to inequality (2.35), it follows for all $\boldsymbol{u} \in B_R$ that

$$\|\mathcal{G}(\boldsymbol{x})(t) - \boldsymbol{x}_0\| \le \delta_0 + M_{\mathcal{T}} L_{\mathcal{F}\delta}\delta\tau + M_{\mathcal{T}} \left\|\mathcal{B}\right\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} R\tau^{(p-1)/p}. \tag{2.38}$$

Choose τ small enough that the right hand side in (2.38) is less than $2\delta_0$. For such τ, $\mathcal{G} : \mathbb{S} \to \mathbb{S}$.

Because of the local Lipschitz continuity of $\mathcal{F}(\cdot)$

$$\begin{aligned} \|\mathcal{G}(\boldsymbol{x}_2) - \mathcal{G}(\boldsymbol{x}_1)\|_{C(0,\tau;\mathbb{X})} &\le \sup_{t\in[0,\tau]} \left\| \int_0^t \mathcal{T}(t-s) \left(\mathcal{F}(\boldsymbol{x}_2(s)) - \mathcal{F}(\boldsymbol{x}_1(s))\right) ds \right\| \\ &\le M_{\mathcal{T}} L_{\mathcal{F}\delta}\tau \left\|\boldsymbol{x}_2 - \boldsymbol{x}_1\right\|_{C(0,\tau;\mathbb{X})}. \end{aligned} \tag{2.39}$$

Choosing τ so $M_{\mathcal{T}} L_{\mathcal{F}\delta}\tau < 1$ yields that $\mathcal{G}$ is a contraction on $\mathbb{S}$. Thus, the operator $\mathcal{G}$ has a unique fixed point in $\mathbb{S}$ that satisfies

$$\boldsymbol{x}(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))\, ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}\boldsymbol{u}(s)\, ds\,. \tag{2.40}$$

Therefore, $\boldsymbol{x}(t)$ is the unique local mild solution of (3.1). □

2.4.2 Nonlinear Parabolic Systems

In nonlinear parabolic systems, the operator $\mathcal{F}(x)$ is defined on the space $D(\mathcal{F})$, where $D(\mathcal{F})$ is densely embedded in $\mathbb{X}$. Stronger assumption on the linear part of the system will be imposed. It is assumed that $\mathcal{A}$ is the generator of an analytic semigroup $e^{\mathcal{A}t}$ on $\mathbb{X}$.

For every $p \in [0, \infty]$ and $\alpha \in (0, 1)$, the interpolation space $D_{\mathcal{A}}(\alpha, p)$ is defined as [84, Section 2.2.1]

$$\begin{cases} D_{\mathcal{A}}(\alpha,p) = \left\{\boldsymbol{x} \in \mathbb{X} : t \mapsto v(t) := \left\|t^{1-\alpha-1/p}\mathcal{A}e^{t\mathcal{A}}\boldsymbol{x}\right\| \in L^p(0,1)\right\}, \\ \|\boldsymbol{x}\|_{D_{\mathcal{A}}(\alpha,p)} = \|\boldsymbol{x}\| + \|v\|_{L^p(0,1)}\,. \end{cases} \tag{2.41}$$

Theorem 2.4.3. *[35, Theorem 4.1] Let $\mathcal{A}$ generate an analytic semigroup on $\mathbb{X}$. Then, for every $\boldsymbol{u} \in L^p(0,\tau;\mathbb{X})$ and $\boldsymbol{x}_0 \in D_{\mathcal{A}}(1/p,p)$, the linear system*

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \boldsymbol{u}(t), & a.e.\ on\ (0,\tau), \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.42}$$

admits a unique solution $\boldsymbol{x}(t)$ in $L^p(0,\tau;D(\mathcal{A})) \cap W^{1,p}(0,\tau;\mathbb{X})$.

Definition 2.4.4 (strict solution)**.** *[11, Definition 3.1.i] If the mild solution to (2.30) is in $W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau;D(\mathcal{A}))$, and for almost every $t \in (0,\tau]$: $\boldsymbol{x}(t)$ is in $D(\mathcal{A})$ and satisfies (2.31), it is said to be a strict solution of (2.31).*

Lemma 2.4.5. *[24, Proposition 2.2 and Corollary 2.3] Let $\tau_0 > \tau$ and $p \in (1,\infty)$ be given. If $\mathcal{A}$ generate an analytic semigroup on $\mathbb{X}$, then there exists a constant c_{τ_0} independent of τ such that for all $\tau \in (0,\tau_0]$ and $\boldsymbol{v} \in W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau;D(\mathcal{A}))$,*

$$\begin{aligned} &\|\dot{\boldsymbol{v}}\|_{L^2(0,\tau;\mathbb{X})} + \|\mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})} \\ &\qquad \le M_{\tau_0}\left(\|\dot{\boldsymbol{v}} + \mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})} + \|\boldsymbol{v}(0)\|_{D_{\mathcal{A}}(1/p,p)}\right). \end{aligned}$$

Furthermore, if $\boldsymbol{v}(0) = 0$,

$$\|\boldsymbol{v}\|_{C(0,\tau;D_{\mathcal{A}}(1/p,p))} \le M_{\tau_0}\left(\|\dot{\boldsymbol{v}}\|_{L^2(0,\tau;\mathbb{X})} + \|\mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})}\right).$$

Let $\mathbb{V}$ be a reflexive subspace of $\mathbb{X}$. For any positive numbers R_1 and R_2, define the sets

$$B_{L^p(0,\tau;\mathbb{U})}(R_1) = \left\{\boldsymbol{u} \in L^p(0,\tau;\mathbb{U}) : \|\boldsymbol{u}\|_p \le R_1\right\}, \tag{2.43}$$

$$B_{\mathbb{V}}(R_2) = \left\{\boldsymbol{x}_0 \in \mathbb{V} : \|\boldsymbol{x}_0\|_{\mathbb{V}} \le R_2\right\}. \tag{2.44}$$

The following theorem modifies [24, Theorem 2.1] which only considers a fixed initial condition.

Theorem 2.4.6. *[38, Theorem 2] Let $\mathcal{A}$ generate an analytic semigroup on $\mathbb{X}$, and $\mathcal{F}(\cdot) : \mathbb{V} \to \mathbb{X}$ be locally Lipschitz continuous, and $D_{\mathcal{A}}(1/p,p) \hookrightarrow \mathbb{V}$. For every pair $R_1 > 0$, $R_2 > 0$, there is $\tau > 0$ and $\delta > 0$ such that (2.31) admits a unique strict solution $\boldsymbol{x} \in \mathbb{W}(0,\tau)$, $\|\boldsymbol{x}\|_{\mathbb{W}(0,\tau)} \le \delta$ for all $(\boldsymbol{u},\boldsymbol{r},\boldsymbol{x}_0) \in B_{L^p(0,\tau;\mathbb{U})}(R_1) \times K \times B_{\mathbb{V}}(R_2)$.*

Proof. The proof of this theorem follows the same line as that of [24, Theorem 2.1] with some modifications. Let $\boldsymbol{w}$ solve the linear equation

$$\begin{cases} \dot{\boldsymbol{w}}(t) = \mathcal{A}\boldsymbol{w}(t) + \mathcal{F}(\boldsymbol{x}_0) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad t \in (0,\tau], \\ \boldsymbol{w}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.45}$$

Define for an arbitrary number $\rho > 0$ the set

$$\Sigma_{\rho,\tau} = \left\{ \boldsymbol{v} \in \mathbb{W}(0,\tau) : \boldsymbol{v}(0) = \boldsymbol{x}_0,\ \|\boldsymbol{v} - \boldsymbol{w}\|_{\mathbb{W}(0,\tau)} \le \rho \right\}. \tag{2.46}$$

Because $\boldsymbol{w}(\cdot) \in \mathbb{W}(0,\tau)$, $\boldsymbol{w}(\cdot) \in C(0,\tau;\mathbb{V})$. Define $\phi(\tau;R_1,R_2) = \|\boldsymbol{w} - \boldsymbol{x}_0\|_{C(0,\tau;\mathbb{V})}$ where here $\boldsymbol{x}_0$ indicates the constant function in $C(0,\tau;\mathbb{V})$ that equals $\boldsymbol{x}_0$. Note that

$$\lim_{\tau\to 0} \phi(\tau;R_1,R_2) = 0. \tag{2.47}$$

According to Lemma 4.1.3, there is a constant M independent of τ such that

$$\|\boldsymbol{v} - \boldsymbol{x}_0\|_{C(0,\tau;\mathbb{V})} \le M\rho + \phi(\tau;R_1,R_2), \quad \forall \boldsymbol{v} \in \Sigma_{\rho,\tau}. \tag{2.48}$$

Consider the mapping $\gamma : \mathbb{W}(0,\tau) \to \mathbb{W}(0,\tau)$, $\boldsymbol{x}(\cdot) \mapsto \boldsymbol{v}(\cdot)$ defined by

$$\begin{cases} \dot{\boldsymbol{v}}(t) = \mathcal{A}\boldsymbol{v}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t),\ t \in (0,\tau], \\ \boldsymbol{v}(0) = \boldsymbol{x}_0. \end{cases} \tag{2.49}$$

It will now be shown that for some numbers ρ and τ the mapping γ defines a contraction on $\Sigma_{\rho,\tau}$ and hence has a unique fixed point.

Consider the linear equation

$$\begin{cases} \dot{\boldsymbol{v}}(t) - \dot{\boldsymbol{w}}(t) = \mathcal{A}(\boldsymbol{v}(t) - \boldsymbol{w}(t)) + \mathcal{F}(\boldsymbol{x}(t)),\ t \in (0,\tau], \\ (\boldsymbol{v} - \boldsymbol{w})(0) = 0, \end{cases}$$

Use Lemma 4.1.3 together with Lipschitz continuity of $\mathcal{F}$, let $L_{\mathcal{F}}$ be the Lipschitz constant of $\mathcal{F}$ over the ball $B(\boldsymbol{x}_0, M\rho + \phi(\tau;R_1,R_2))$. It follows that

$$\|\boldsymbol{v} - \boldsymbol{w}\|_{\mathbb{W}(0,\tau)} \le M^2 L_{\mathcal{F}} \tau^{\frac{1}{p}} (M\rho + \phi(\tau;R_1,R_2)). \tag{2.50}$$

Furthermore, for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Sigma_{\rho,\tau}$, define $\boldsymbol{v}_1 = \gamma(\boldsymbol{x}_1)$ and $\boldsymbol{v}_2 = \gamma(\boldsymbol{x}_2)$, then Lemma 4.1.3 yields

$$\|\boldsymbol{v}_2 - \boldsymbol{v}_1\|_{\mathbb{W}(0,\tau)} \leq M^2 L_{\mathcal{F}} \tau^{\frac{1}{p}} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_{\mathbb{W}(0,\tau)} . \tag{2.51}$$

Choose ρ and τ so that

$$M^2 L_{\mathcal{F}} \tau^{\frac{1}{p}} < 1,$$
$$M^2 L_{\mathcal{F}} \tau^{\frac{1}{p}} (M\rho + \phi(\tau; R_1, R_2)) \leq \rho.$$

The Contraction Mapping Theorem ensures that the mapping γ has a unique fixed point in $\Sigma_{\rho,\tau}$. This fixed point is the unique solution $\boldsymbol{x}$ to (2.31). Also, from the definition (2.46), every $\boldsymbol{x}$ in $\Sigma_{\rho,\tau}$ satisfies

$$\|\boldsymbol{x}\|_{\mathbb{W}(0,\tau)} \leq \|\boldsymbol{w}\|_{\mathbb{W}(0,\tau)} + \rho. \tag{2.52}$$

Let $L_{\mathcal{F}}$ be the Lipschitz constant of $\mathcal{F}$ over the ball $B(0, \|\boldsymbol{x}_0\|)$. Proposition 2.2 in [24] yields

$$\begin{aligned}
\|\boldsymbol{w}\|_{\mathbb{W}(0,\tau)} &\leq M(\|\boldsymbol{x}_0\|_{\mathbb{V}} + \|\mathcal{F}(\boldsymbol{x}_0) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t)\|_p) \\
&\leq M(\|\boldsymbol{x}_0\|_{\mathbb{V}} + \tau^{\frac{1}{p}} L_{\mathcal{F}} \|\boldsymbol{x}_0\|_{\mathbb{V}} + \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})} \|\boldsymbol{u}(t)\|_p) \\
&\leq \underbrace{M(R_2 + \tau^{\frac{1}{p}} L_{\mathcal{F}} R_2 + R_1 \max_{\boldsymbol{r}\in K} \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})})}_{\delta} .
\end{aligned}$$

Defining

$$\delta = M(R_2 + \tau^{\frac{1}{p}} L_{\mathcal{F}} R_2 + R_1 \max_{\boldsymbol{r}\in K} \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{X},\mathbb{U})}),$$

yields the required upper-bound on $\|\boldsymbol{x}\|_{\mathbb{W}(0,\tau)}$. □

Chapter 3

Optimal Actuator Design for Semilinear Systems

Semilinear partial differential equations model a wide spectrum of physical systems with distributed parameters. Theory for concurrent optimal control and actuator design of a class of controlled semilinear PDE's is described in this chapter. The research described extends previous work on optimal control of PDE's in that the system is in a general form and the linear part of the equation is not necessarily an elliptic differential operator. A general class of PDE's with weakly continuous nonlinear part is considered in this chapter. It is shown that under certain conditions on the nonlinearity and the cost function, an optimal control input together with an optimal actuator choice exists. Optimality equations explicitly characterizing the optimal control and actuator are obtained. The results are applied to optimal actuator and controller design in a nonlinear railway track model as well as semilinear wave models.

The chapter is organized as follows. After a short paragraph on notation, the problem definition as well as the main results are stated in Section 3.1. Section 3.2 briefly discusses the existence of a solution to the semilinear equation as well as an estimate on the solution. The existence of an optimizer is established in Section 3.3. First-order necessary conditions for the optimizer are provided in Section 3.4. In Section 3.5, the results of the previous sections are applied to the railway track model and semilinear wave models, respectively. Concluding remarks are made in Section 3.6.

3.1 Main Results

Consider a semilinear system with state $\boldsymbol{x}(t)$ on a separable reflexive Banach space $\mathbb{X}$:

$$\dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{X}, \tag{3.1}$$

The function $\boldsymbol{u}(t)$ is the input to the system, and takes values in a reflexive Banach space $\mathbb{U}$. Control operator $\mathcal{B}(\boldsymbol{r})$ depends on a parameter $\boldsymbol{r}$ that takes values in an open set K in a topological space $\mathbb{K}$. The parameter $\boldsymbol{r}$ typically has interpretation as possible actuator designs. The operators $\mathcal{A}$, $\mathcal{F}(\cdot)$, and $\mathcal{B}(\cdot)$ satisfy the following assumptions.

Assumption A.

1. *The state operator $\mathcal{A}$ with domain $D(\mathcal{A})$ generates a strongly continuous semigroup $\mathcal{T}(t)$ on $\mathbb{X}$.*
2. *Let $\mathcal{F}(0) = 0$; the nonlinear operator $\mathcal{F}(\cdot)$ is locally Lipschitz continuous on $\mathbb{X}$; that is, for every positive number δ, there exists $L_{\mathcal{F}\delta} > 0$ such that*

$$\|\mathcal{F}(\boldsymbol{x}_2) - \mathcal{F}(\boldsymbol{x}_1)\| \le L_{\mathcal{F}\delta} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|,$$

   for all $\|\boldsymbol{x}_2\| \le \delta$ and $\|\boldsymbol{x}_1\| \le \delta$.

3. *For each $\boldsymbol{r} \in K$, the input operator $\mathcal{B}(\boldsymbol{r})$ is a linear bounded operator that maps the input space $\mathbb{U}$ into the state space $\mathbb{X}$. This family of operators is uniformly bounded over K, i.e., there exist a positive number $M_{\mathcal{B}}$ such that $\|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \le M_{\mathcal{B}}$ for all $\boldsymbol{r} \in K$.*

In some cases, due to lack of regularity of the input $\boldsymbol{u}$, a classical solution to (3.1) is not assured.

Definition 3.1.1. *If $\boldsymbol{x} \in C(0,\tau;\mathbb{X})$ satisfies*

$$\boldsymbol{x}(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r})\boldsymbol{u}(s)ds, \tag{3.2}$$

for every $\boldsymbol{x}_0 \in \mathbb{X}$, it is said to be a mild solution *to (3.1).*

In Theorem 2.4.2, the existence of a unique mild solution to the initial value problem (IVP) (3.1) is proven for $\boldsymbol{u}(t)$ in the ball

$$B_R = \{\boldsymbol{u} \in L^p(0,\tau;\mathbb{U}) : \|\boldsymbol{u}\|_p \le R\},$$

where $1 < p < \infty$.

Theorem 2.4.2: *Under assumption A, for each* $\boldsymbol{x}_0 \in \mathbb{X}$ *and positive number* R*, there exists* $\tau > 0$ *such that (3.1) admits a unique local mild solution* $\boldsymbol{x} \in C(0, \tau; \mathbb{X})$ *for all* $\boldsymbol{u} \in B_R$ *and all* $\boldsymbol{r} \in K$*.*

For functionals $\phi(\boldsymbol{x})$ on $\mathbb{X}$ and $\psi(\boldsymbol{u})$ on $\mathbb{U}$, consider the cost function

$$J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) = \int_0^\tau \phi(\boldsymbol{x}(t)) + \psi(\boldsymbol{u}(t))\, dt.$$

The optimization problem is to minimize $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ over an admissible control input set U_{ad} and an admissible actuator design set K_{ad}, subject to (3.1) with a fixed initial condition $\boldsymbol{x}_0 \in \mathbb{X}$. That is,

$$\begin{cases} \min & J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) \\ \text{s.t.} & \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad \forall t \in (0, \tau], \\ & \boldsymbol{x}(0) = \boldsymbol{x}_0, \\ & \boldsymbol{u} \in U_{ad}, \\ & \boldsymbol{r} \in K_{ad}. \end{cases} \tag{P}$$

To guarantee the existence of an optimizer, further assumptions are needed on the operators $\mathcal{F}(\cdot)$, $\mathcal{B}(\cdot)$, the sets U_{ad} and K_{ad}, and on the cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$.

Assumption B.

1. *Let* $\boldsymbol{x}_n(t)$ *be a bounded sequence in* $C(0, \tau; \mathbb{X})$ *such that* $\boldsymbol{x}_n(t) \rightharpoonup \boldsymbol{x}(t)$ *in* $L^p(0, \tau; \mathbb{X})$*. Then,* $\mathcal{F}(\boldsymbol{x}_n(t)) \rightharpoonup \mathcal{F}(\boldsymbol{x}(t))$ *in* $L^p(0, \tau; \mathbb{X})$*.*

2. *The set* K_{ad} *is a compact subset of* K*, and the set* U_{ad} *is a closed and convex subset of* $B_R \backslash \partial B_R$*. Also, letting* $\boldsymbol{r}_2 \to \boldsymbol{r}_1$ *with respect to the topology on* $\mathbb{K}$*, the family of input operators* $\mathcal{B}(\cdot) : K \to \mathcal{L}(\mathbb{U}, \mathbb{X})$ *are continuous with respect to* $\boldsymbol{r}$ *in the operator norm topology:*
$$\lim_{\boldsymbol{r}_2 \to \boldsymbol{r}_1} \|\mathcal{B}(\boldsymbol{r}_2) - \mathcal{B}(\boldsymbol{r}_1)\|_{\mathcal{L}(\mathbb{U}, \mathbb{X})} = 0.$$

3. *The functionals* $\phi(\cdot)$ *and* $\psi(\cdot)$ *are weakly lower semi-continuous non-negative functionals on* $\mathbb{X}$ *and* $\mathbb{U}$*, respectively.*

It is shown in Section 3.3 that under these assumptions, an optimal control and actuator design exist.

Theorem 3.3.1: *For initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, let τ be such that the mild solution exists for all $\boldsymbol{u} \in B_R$ and all $\boldsymbol{r} \in K$. Under assumptions A and B, there exists a control input $\boldsymbol{u}^o \in U_{ad}$ together with an actuator design $\boldsymbol{r}^o \in K_{ad}$, that solves the optimization problem P.*

To characterize an optimizer to the optimization problem, further assumptions on differentiability of the nonlinear operators $\mathcal{F}(\cdot)$ and $\mathcal{B}(\cdot)$, and the cost function are needed.

Assumption C.

1. *The nonlinear operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable on $\mathbb{X}$ ([61, Definition 1.29]). Indicate the Gâteaux derivative of $\mathcal{F}(\cdot)$ at $\boldsymbol{x}$ in the direction $\boldsymbol{p}$ by $\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}$. Furthermore, the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded; that is, bounded sets in $\mathbb{X}$ are mapped to bounded sets in $\mathcal{L}(\mathbb{X})$.*

2. *The space $\mathbb{K}$ is a Hilbert space, K is an bounded subset of $\mathbb{K}$, and K_{ad} is convex. Also, the control operator $\mathcal{B}(\boldsymbol{r})$ is Gâteaux differentiable with respect to $\boldsymbol{r}$ from K to $\mathcal{L}(\mathbb{U}, \mathbb{X})$. Indicate the Gâteaux derivative of $\mathcal{B}(\boldsymbol{r})$ at $\boldsymbol{r}^o$ in the direction $\boldsymbol{r}$ by $\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{r}$. Furthermore, the mapping $\boldsymbol{r}^o \mapsto \mathcal{B}'_{\boldsymbol{r}^o}$ is bounded; that is, bounded sets in K are mapped to bounded sets in $\mathcal{L}(\mathbb{K}, \mathcal{L}(\mathbb{U}, \mathbb{X}))$.*

3. *The spaces $\mathbb{X}$ and $\mathbb{U}$ are Hilbert spaces, and p=2. Also, in the cost function, set*

$$\phi(\boldsymbol{x}) = \langle \mathcal{Q}\boldsymbol{x}, \boldsymbol{x} \rangle, \quad \psi(\boldsymbol{u}) = \langle \mathcal{R}\boldsymbol{u}, \boldsymbol{u} \rangle_{\mathbb{U}}, \tag{3.3}$$

   where the linear operator $\mathcal{Q}$ is a positive semi-definite, self-adjoint bounded operator on $\mathbb{X}$, and the linear operator $\mathcal{R}$ is a coercive, self-adjoint bounded operator on $\mathbb{U}$.

Since $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$ are assumed to be Hilbert spaces, the dual of each of these spaces can be identified with the space itself. The operator $(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^* : \mathbb{X} \to \mathbb{K}$ is defined as

$$\langle (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^*\boldsymbol{p}, \boldsymbol{r} \rangle_{\mathbb{K}} = \langle \boldsymbol{p}, (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{r})\boldsymbol{u} \rangle, \quad \forall (\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{r}) \in \mathbb{U} \times \mathbb{X} \times \mathbb{K}.$$

The following theorem is proven in Section 3.4. In this theorem $\boldsymbol{x} = \mathcal{S}(\boldsymbol{u}; \boldsymbol{r}, \boldsymbol{x}_0)$ denotes the control-to-state map (see Definition 3.4.1).

Theorem 3.4.7: *Suppose assumptions A1 and C hold. For any initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, let the pair $(\boldsymbol{u}^o, \boldsymbol{r}^o) \in U_{ad} \times K_{ad}$ be a local minimizer of the optimization problem P with the optimal trajectory $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0)$ and let $\boldsymbol{p}^o(t)$, the adjoint state, indicate the mild solution of the final value problem*

$$\dot{\boldsymbol{p}}^o(t) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}^o(t)})\boldsymbol{p}^o(t) - \mathcal{Q}\boldsymbol{x}^o(t), \quad \boldsymbol{p}^o(\tau) = 0.$$

Then $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ *satisfies*

$$\left\langle \boldsymbol{u}^o + \mathcal{R}^{-1}\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o, \boldsymbol{u} - \boldsymbol{u}^o \right\rangle_{L^2(0,\tau;\mathbb{U})} \geq 0, \quad \forall \boldsymbol{u} \in U_{ad},$$
$$\left\langle \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^* \boldsymbol{p}^o(t)\, dt, \boldsymbol{r} - \boldsymbol{r}^o \right\rangle_{\mathbb{K}} \geq 0, \quad \forall \boldsymbol{r} \in K_{ad}.$$

If the optimizer $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ is in the interior of $U_{ad} \times K_{ad}$, the optimality conditions become the equality conditions presented in Corollary 3.4.8.

3.2 Estimate on the Solution

In the existing literature, the existence of a unique local solution to (3.1) is guaranteed for continuously differentiable control inputs (see e.g. [104, Theorem 6.1.5]). Requiring that $\boldsymbol{u} \in C^1(0,\tau;\mathbb{U})$ is too restrictive for establishing existence of an optimal control. Theorem 2.4.2 ensures that under assumption A, for each $\boldsymbol{x}_0 \in \mathbb{X}$ and positive number R, there exists $\tau > 0$ such that (3.1) admits a unique local mild solution $\boldsymbol{x} \in C(0,\tau;\mathbb{X})$ for all $\boldsymbol{u} \in B_R$ and all $\boldsymbol{r} \in K$. The following lemma provides an estimate on the solution to be used in the next sections. Gronwall's Inequality is used in the proof.

Lemma 3.2.1 (Gronwall's Inequality)**.** *[121, Theorem 1.4.1] Let c be a number and $g(t)$ be a non-negative continuous function. If $y(t)$ is continuous, real valued, and satisfies*

$$y(t) \leq c + \int_0^t g(s)y(s)ds, \quad t \in [0,\tau], \tag{3.5}$$

then

$$y(t) \leq c\,exp\left(\int_0^t g(s)ds\right), \quad t \in [0,\tau]. \tag{3.6}$$

Lemma 3.2.2. *Under assumption A, for all $\boldsymbol{u} \in B_R$, there exists a positive number c_τ such that the mild solution to (3.1) satisfies*

$$\|\boldsymbol{x}\|_{C(0,\tau;\mathbb{X})} \leq c_\tau \left(\|\boldsymbol{x}_0\| + \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\boldsymbol{u}\|_p \right). \tag{3.7}$$

Proof. Let τ be as in Theorem 2.4.2. Take the norm of both sides of (3.2) and apply assumption A together with the triangle inequality to obtain

$$\begin{aligned} \|\boldsymbol{x}(t)\| &\leq \|\mathcal{T}(t)\boldsymbol{x}_0\| + \left\| \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))\, ds \right\| + \left\| \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r})\boldsymbol{u}(s)\, ds \right\| \\ &\leq M_{\mathcal{T}} \|\boldsymbol{x}_0\| + M_{\mathcal{T}} L_{\mathcal{F}\delta} \int_0^t \|\boldsymbol{x}(s)\|\, ds + M_{\mathcal{T}} \tau^{(p-1)/p} \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\boldsymbol{u}\|_p . \end{aligned} \tag{3.8}$$

Defining the constant

$$c_\tau = M_\mathcal{T} e^{M_\mathcal{T} L_{\mathcal{F}\delta}\tau} \max\left\{1, \tau^{(p-1)/p}\right\},$$

and applying Lemma 3.2.1 to inequality (3.8) yield

$$\|\boldsymbol{x}(t)\| \le c_\tau \left(\|\boldsymbol{x}_0\| + \|\mathcal{B}(\boldsymbol{r})\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\boldsymbol{u}\|_p \right). \tag{3.9}$$

Taking supremum of both side over $[0, \tau]$ results in (3.7). □

3.3 Existence of an Optimizer

In the previous section assumptions A were shown to imply existence of a solution. It is now shown, that if in addition assumptions B are satisfied, the optimization problem P has a solution and there exists an optimal control input $\boldsymbol{u}^o \in U_{ad}$ together with an optimal actuator design $\boldsymbol{r}^o \in K_{ad}$.

Theorem 3.3.1. *For initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, let τ be such that the mild solution exists for all $\boldsymbol{u} \in B_R$ and all $\boldsymbol{r} \in K$. Under assumptions A and B, there exists a control input $\boldsymbol{u}^o \in U_{ad}$ together with an actuator design $\boldsymbol{r}^o \in K_{ad}$, that solves the optimization problem P.*

Proof. The cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ is bounded from below, and thus it has an infimum, say $j(\boldsymbol{x}_0)$. This infimum is finite by assumption. As a result, there is a sequence of inputs $\boldsymbol{u}_n \in U_{ad}$ and actuator design $\boldsymbol{r}_n \in K_{ad}$ such that

$$\lim_{n\to\infty} J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0) = j(\boldsymbol{x}_0). \tag{3.10}$$

The set U_{ad} is a bounded subset of the reflexive space $L^p(0, \tau; \mathbb{U})$, $1 < p < \infty$, and hence it is weakly sequentially compact [119, Theorem 9.4.3]. Since U_{ad} is closed and convex, it is also weakly closed [115, Theorem 2.11.]. These statements mean that there is a subsequence of $\boldsymbol{u}_n$ that converges weakly to some element $\boldsymbol{u}^o \in U_{ad}$. To simplify the notation, we denote the weakly convergent subsequence by $\boldsymbol{u}_n$:

$$\boldsymbol{u}_n(t) \rightharpoonup \boldsymbol{u}^o(t) \quad \text{as} \quad n \to \infty. \tag{3.11}$$

The compactness of K_{ad} implies that there is also a subsequence of $\boldsymbol{r}_n$ that converges to some $\boldsymbol{r}^o$ in K_{ad} with respect to the topology on $\mathbb{K}$. This subsequence is also indicated by $\boldsymbol{r}_n$; that is

$$\boldsymbol{r}_n \to \boldsymbol{r}^o \quad \text{as} \quad n \to \infty. \tag{3.12}$$

Using assumption B2, it follows that

$$\mathcal{B}(\boldsymbol{r}_n)\boldsymbol{u}_n(t) \rightharpoonup \mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}^o(t) \quad \text{in} \quad L^p(0,\tau;\mathbb{X}) \quad \text{as} \quad n \to \infty. \tag{3.13}$$

According to Proposition 1.84 of [9], every continuous linear map is weakly continuous, yielding

$$\int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r}_n)\boldsymbol{u}_n(s)ds \rightharpoonup \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}^o(s)ds \quad \text{in } C(0,\tau;\mathbb{X}). \tag{3.14}$$

Moreover, by Theorem 2.4.2, for every pair $(\boldsymbol{u}_n, \boldsymbol{r}_n)$, there exists a state $\boldsymbol{x}_n(t) \in C(0,\tau;\mathbb{X})$. The sequence $\{\boldsymbol{x}_n(t)\}$ is also bounded in $C(0,\tau;\mathbb{X})$ by Lemma 3.2.2; that is

$$\|\boldsymbol{x}_n\|_{C(0,\tau;\mathbb{X})} \le c_\tau \left(\|\boldsymbol{x}_0\| + M_{\mathcal{B}}R\right). \tag{3.15}$$

The sequence $\boldsymbol{x}_n(t)$ is bounded in $C(0,\tau;\mathbb{X})$ and so in $L^p(0,\tau;\mathbb{X})$ as well. The latter is a reflexive Banach space; this means that a subsequence of $\boldsymbol{x}_n(t)$, denote it by $\boldsymbol{x}_n(t)$ for simplicity, weakly converges to an element of $\boldsymbol{x}^o$ in $L^p(0,\tau;\mathbb{X})$. By assumption B1, it follows that

$$\mathcal{F}(\boldsymbol{x}_n(t)) \rightharpoonup \mathcal{F}(\boldsymbol{x}^o(t)) \quad \text{in } L^p(0,\tau;\mathbb{X}), \tag{3.16}$$

and also by Proposition 1.84 of [9]

$$\int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}_n(s))ds \rightharpoonup \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}^o(s))ds \quad \text{in } C(0,\tau;\mathbb{X}). \tag{3.17}$$

Recall that each $(\boldsymbol{x}_n, \boldsymbol{u}_n, \boldsymbol{r}_n)$ satisfies

$$\boldsymbol{x}_n(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}_n(s))ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r}_n)\boldsymbol{u}_n(s)ds. \tag{3.18}$$

Apply (3.14) and (3.17) to (3.18), it follows that $\boldsymbol{x}^o(t)$ is in $C(0,\tau;\mathbb{X})$. Note that the mild solution is unique; thus, $\boldsymbol{x}^o(t)$ is the mild solution to IVP (3.1) with input $\boldsymbol{u}^o(t)$ and actuator design $\boldsymbol{r}^o$, satisfying

$$\boldsymbol{x}^o(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}^o(s))ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}^o(s)ds. \tag{3.19}$$

It remains to show that $(\boldsymbol{x}^o(t), \boldsymbol{u}^o(t), \boldsymbol{r}^o)$ minimizes $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$. Recall from definition of the sequence $\boldsymbol{u}_n$ and $\boldsymbol{r}_n$ that

$$\begin{aligned} j(\boldsymbol{x}_0) &= \liminf_{n\to\infty} J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0) \\ &= \liminf_{n\to\infty} \int_0^\tau \phi(\boldsymbol{x}_n(t))\, dt + \liminf_{n\to\infty} \int_0^\tau \psi(\boldsymbol{u}_n(t))\, dt. \end{aligned} \tag{3.20}$$

From assumption B3, the cost function is weakly lower semi-continuous in $\boldsymbol{x}$ and $\boldsymbol{u}$. This together with Fatou's Lemma implies

$$j(\boldsymbol{x}_0) \geq \int_0^\tau \phi(\boldsymbol{x}^o(t))\, dt + \int_0^\tau \psi(\boldsymbol{u}^o(t))\, dt = J(\boldsymbol{u}^o, \boldsymbol{r}^o; \boldsymbol{x}_0). \tag{3.21}$$

Since $j(\boldsymbol{x}_0)$ was defined to be the infimum,

$$j(\boldsymbol{x}_0) = J(\boldsymbol{u}^o, \boldsymbol{r}^o; \boldsymbol{x}_0).$$

Therefore, for every initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, there exists an control input $\boldsymbol{u}^o(t)$ together with an actuator design $\boldsymbol{r}^o$, with corresponding mild solution $\boldsymbol{x}^o(t)$ that achieves the minimum value of the cost function. □

For a linear partial differential equation and quadratic cost, the optimal actuator problem may not be convex; see for example [95, Fig. 7]. Uniqueness of the optimal control and actuator is not guaranteed.

3.4 Optimality Conditions

In order to establish the first order optimality condition for an optimizer $(\boldsymbol{u}^o, \boldsymbol{r}^o)$, further regularity on the control-to-state map is needed.

Definition 3.4.1. *For each initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, and actuator design $\boldsymbol{r} \in K$, the control-to-state operator is the operator $\mathcal{S}(\boldsymbol{u}; \boldsymbol{r}, \boldsymbol{x}_0) : B_R \subset (L^p(0,\tau;\mathbb{U})) \to L^p(0,\tau;\mathbb{X})$ that maps every input $\boldsymbol{u} \in B_R$ to the state $\boldsymbol{x} \in L^p(0,\tau;\mathbb{X})$. It is described by*

$$\boldsymbol{x}(t) = \mathcal{T}(t)\boldsymbol{x}_0 + \int_0^t \mathcal{T}(t-s)\mathcal{F}(\boldsymbol{x}(s))ds + \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r})\boldsymbol{u}(s)ds.$$

In next two theorems, it is proven that under certain assumptions, the control-to-state map is Lipschitz continuous in both $\boldsymbol{u}$ and $\boldsymbol{r}$. For the Lipschitz continuity with respect to the actuator design, a stronger assumption on the input operator $\mathcal{B}(\boldsymbol{r})$ than continuity in $\boldsymbol{r}$ is needed.

Proposition 3.4.2. *(a) Under assumption A, for any initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, the control-to-state map $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ is Lipschitz continuous in $\boldsymbol{u}$, i.e., there exists a positive constant $L_{\boldsymbol{u}}$ such that*

$$\|\mathcal{S}(\boldsymbol{u}_2;\boldsymbol{r},\boldsymbol{x}_0) - \mathcal{S}(\boldsymbol{u}_1;\boldsymbol{r},\boldsymbol{x}_0)\|_{C(0,\tau;\mathbb{X})} \leq L_{\boldsymbol{u}} \|\boldsymbol{u}_2 - \boldsymbol{u}_1\|_p , \tag{3.22}$$

for all $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ in B_R, and $\boldsymbol{r} \in K$.

(b) Under extra assumption C2, the control-to-state map $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ is Lipschitz continuous in $\boldsymbol{r}$, i.e., there exists a positive constant $L_{\boldsymbol{r}}$ such that

$$\|\mathcal{S}(\boldsymbol{u};\boldsymbol{r}_2,\boldsymbol{x}_0) - \mathcal{S}(\boldsymbol{u};\boldsymbol{r}_1,\boldsymbol{x}_0)\|_{C(0,\tau;\mathbb{X})} \leq L_{\boldsymbol{r}} \|\boldsymbol{r}_2 - \boldsymbol{r}_1\|_{\mathbb{K}} , \tag{3.23}$$

for all $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in K, and $\boldsymbol{u} \in B_R$.

Proof. For $\boldsymbol{x}_0 \in \mathbb{X}$ and $\boldsymbol{r} \in K$, consider $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(t)$ as the mild solutions to (3.1) corresponding to the inputs $\boldsymbol{u}_1(t)$ and $\boldsymbol{u}_2(t)$, respectively. The inputs are in a ball of radius R contained in $L^p(0,\tau;\mathbb{U})$, $1 < p < \infty$; consequently, by Lemma 3.2.2 and assumption A3, the states $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(t)$ are contained in a ball of radius

$$\delta = c_\tau(\|\boldsymbol{x}_0\| + M_{\mathcal{B}}R). \tag{3.24}$$

From (3.2), it follows that

$$\begin{aligned} \boldsymbol{x}_2(t) - \boldsymbol{x}_1(t) = &\int_0^t \mathcal{T}(t-s)\left(\mathcal{F}(\boldsymbol{x}_2(s)) - \mathcal{F}(\boldsymbol{x}_1(s))\right) ds \\ &+ \int_0^t \mathcal{T}(t-s)\mathcal{B}(\boldsymbol{r})\left(\boldsymbol{u}_2(s) - \boldsymbol{u}_1(s)\right) ds. \end{aligned} \tag{3.25}$$

Recall that $\mathcal{T}(t)$ satisfies $\|\mathcal{T}(t)\| \leq M_{\mathcal{T}}$ for all $t \in [0,\tau]$ and some number $M_{\mathcal{T}} > 0$. Also, remember that the operator $\mathcal{F}(\cdot)$ is locally Lipschitz continuous, and $\mathcal{B}(\boldsymbol{r})$ is uniformly bounded in $\mathbb{X}$ for all $\boldsymbol{r} \in K$. Taking the norm in $\mathbb{X}$ of both sides of this equation yields

$$\begin{aligned} \|\boldsymbol{x}_2(t) - \boldsymbol{x}_1(t)\| \leq &M_{\mathcal{T}} L_{\mathcal{F}\delta} \int_0^t \|\boldsymbol{x}_2(s) - \boldsymbol{x}_1(s)\| \, ds \\ &+ M_{\mathcal{T}} M_{\mathcal{B}} \tau^{(p-1)/p} \|\boldsymbol{u}_2 - \boldsymbol{u}_1\|_p . \end{aligned} \tag{3.26}$$

Define the constant $L_{\boldsymbol{u}}$ as

$$L_{\boldsymbol{u}} = e^{M_{\mathcal{T}} L_{\mathcal{F}\delta} \tau} M_{\mathcal{T}} M_{\mathcal{B}} \tau^{(p-1)/p}. \tag{3.27}$$

By Gronwall's Lemma [121, Theorem 1.4.1], it follows that

$$\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_{C(0,\tau;\mathbb{X})} \leq L_{\boldsymbol{u}} \|\boldsymbol{u}_2 - \boldsymbol{u}_1\|_p . \tag{3.28}$$

This is in fact the inequality (3.22).

Similarly, for a fixed initial condition $\boldsymbol{x}_0 \in \mathbb{X}$ and control input $\boldsymbol{u} \in B_R$, consider $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(t)$ as the mild solutions to (3.1) corresponding to the actuator designs $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, respectively. Use local Lipschitz continuity of $\mathcal{F}(\cdot)$ and growth condition on semigroup $\mathcal{T}(t)$ and obtain

$$\begin{aligned} \|\boldsymbol{x}_2(t) - \boldsymbol{x}_1(t)\| \leq & M_{\mathcal{T}} L_{\mathcal{F}\delta} \int_0^t \|\boldsymbol{x}_2(s) - \boldsymbol{x}_1(s)\| \, ds \\ & + M_{\mathcal{T}} \tau^{(p-1)/p} \|\boldsymbol{u}\|_p \|\mathcal{B}(\boldsymbol{r}_2) - \mathcal{B}(\boldsymbol{r}_1)\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} . \end{aligned} \tag{3.29}$$

Assumption C2 implies that the control operator $\mathcal{B}(\boldsymbol{r})$ is locally Lipschitz continuous with respect to $\boldsymbol{r}$. That is, letting

$$L_{\mathcal{B}} = \sup\{\|\mathcal{B}'_{\boldsymbol{r}}\|_{\mathcal{L}(\mathbb{K},\mathcal{L}(\mathbb{U},\mathbb{X}))} : \boldsymbol{r} \in K\},$$

operator $\mathcal{B}(\boldsymbol{r})$ for all $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in K satisfies

$$\|\mathcal{B}(\boldsymbol{r}_2) - \mathcal{B}(\boldsymbol{r}_1)\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \leq L_{\mathcal{B}} \|\boldsymbol{r}_2 - \boldsymbol{r}_1\|_{\mathbb{K}} . \tag{3.30}$$

Accordingly, the inequality (3.29) can be re-written as

$$\begin{aligned} \|\boldsymbol{x}_2(t) - \boldsymbol{x}_1(t)\| \leq & M_{\mathcal{T}} L_{\mathcal{F}\delta} \int_0^t \|\boldsymbol{x}_2(s) - \boldsymbol{x}_1(s)\| \, ds \\ & + M_{\mathcal{T}} \tau^{(p-1)/p} R L_{\mathcal{B}} \|\boldsymbol{r}_2 - \boldsymbol{r}_1\|_{\mathbb{K}} . \end{aligned} \tag{3.31}$$

Denote the constant $L_{\boldsymbol{r}}$ by

$$L_{\boldsymbol{r}} = e^{M_{\mathcal{T}} L_{\mathcal{F}\delta} \tau} M_{\mathcal{T}} \tau^{(p-1)/p} R L_{\mathcal{B}}. \tag{3.32}$$

Use Gronwall's Lemma [121, Theorem 1.4.1] to derive

$$\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_{C(0,\tau;\mathbb{X})} \leq L_{\boldsymbol{r}} \|\boldsymbol{r}_2 - \boldsymbol{r}_1\|_{\mathbb{K}} . \tag{3.33}$$

This is in fact the inequality (3.23). □

Gâteaux differentiability of the control-to-state map as well as its derivatives need to be formulated in order to characterize an optimizer.

For any $\boldsymbol{x}^o \in C(0, \tau; \mathbb{X})$ define the time-varying operator operator $\mathcal{F}'_{\boldsymbol{x}^o(t)}$. At any $t > 0$, this operator is linear on $\mathbb{X}$. Consider the time-varying IVP

$$\dot{\tilde{\boldsymbol{x}}}(t) = (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\tilde{\boldsymbol{x}}(t) + \mathcal{B}(\boldsymbol{r})\tilde{\boldsymbol{u}}(t), \quad \tilde{\boldsymbol{x}}(0) = 0. \tag{3.34}$$

The mild solution is described by a two-parameter family of operators, say $\mathcal{U}(t, s)$, known as an evolution operator.

The following lemma relies on the existence results: Theorem 5.5.6 and Theorem 5.5.10 in [46].

Lemma 3.4.3. *(a) The mild solution of IVP problem (3.34) is described by*

$$\tilde{\boldsymbol{x}}(t) = \int_0^t \mathcal{U}(t, s)\mathcal{B}(\boldsymbol{r})\tilde{\boldsymbol{u}}(s)\, ds, \tag{3.35}$$

in which $\mathcal{U}(t, s)$ is a strongly continuous evolution operator on $\mathbb{X}$ for $0 \le s \le t \le \tau$.

(b) Let $\boldsymbol{f} \in L^1(0, \tau; \mathbb{X})$, and consider the following final value problem (FVP) backward in time

$$\dot{\tilde{\boldsymbol{p}}}(s) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}^o(s)})\tilde{\boldsymbol{p}}(s) - \boldsymbol{f}(s), \quad \tilde{\boldsymbol{p}}(\tau) = 0, \tag{3.36}$$

then, the mild solution of this evolution equation satisfies

$$\tilde{\boldsymbol{p}}(s) = \int_s^\tau \mathcal{U}^*(t, s)\boldsymbol{f}(t)\, dt, \tag{3.37}$$

where $\mathcal{U}^*(t, s)$ is the adjoint of $\mathcal{U}(t, s)$ on $\mathbb{X}$ for every $0 \le s \le t \le \tau$.

Proof. The time-invariant part of the state operator in (3.34), $\mathcal{A}$, is the generator of an strongly continuous semigroup. According to [46, Theorem 5.5.6], in order for a strongly continuous evolution operator $\mathcal{U}(t, s)$ to exist so that (3.35) is the mild solution to the (3.34), it is sufficient that for every $\tilde{\boldsymbol{x}} \in \mathbb{X}$ the mapping $t \mapsto \mathcal{F}'_{\boldsymbol{x}^o(t)}\tilde{\boldsymbol{x}}$ is strongly measurable and that a function $\alpha(t) \in L^1(0, \tau)$ exists such that

$$\left\|\mathcal{F}'_{\boldsymbol{x}^o(t)}\right\| \le \alpha(t), \quad t \in [0, \tau]. \tag{3.38}$$

By assumption C1, since the state $\boldsymbol{x}^o(t)$ is uniformly bounded, the operator norm of $\mathcal{F}'_{\boldsymbol{x}^o(t)}$ admits an upper bound for all $t \in [0, \tau]$. Consequently, a strongly continuous evolution operator $\mathcal{U}(t, s)$ exists so that (3.35) is the mild solution to (3.34).

Since the state space $\mathbb{X}$ is a separable reflexive Banach space, Theorem 5.5.10 of [46] implies that the mild solution of (3.36) is described by an evolution operator. Moreover, for every $0 \le s \le t \le \tau$, this evolution operator is the adjoint on $\mathbb{X}$ of the evolution operator $\mathcal{U}(t,s)$. □

Proposition 3.4.4. *Under assumptions A and C1, for every initial condition $\boldsymbol{x}_0 \in \mathbb{X}$ and actuator design $\boldsymbol{r} \in K$, the control-to-state map $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ is Gâteaux differentiable with respect to $\boldsymbol{u}$ in U_{ad}. The Gâteaux derivative of $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ at $\boldsymbol{u}^o$ in the direction $\tilde{\boldsymbol{u}}$ is*

$$\mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = \tilde{\boldsymbol{x}}, \quad \forall \tilde{\boldsymbol{u}} \in L^p(0,\tau;\mathbb{U}), \tag{3.39}$$

where, defining $\boldsymbol{x}^o(t) = \mathcal{S}(\boldsymbol{u}^o;\boldsymbol{r},\boldsymbol{x}_0)$, $\tilde{\boldsymbol{x}}$ is the mild solution to the IVP

$$\dot{\tilde{\boldsymbol{x}}}(t) = (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\tilde{\boldsymbol{x}}(t) + \mathcal{B}(\boldsymbol{r})\tilde{\boldsymbol{u}}(t), \quad \tilde{\boldsymbol{x}}(0) = 0. \tag{3.40}$$

The mild solution to this equation is given by the evolution operator $\mathcal{U}(t,s)$ in Lemma 3.4.3(a).

Proof. For sufficiently small ϵ, there is a mild solution to IVP (3.1) with input $\boldsymbol{u}^o + \epsilon\tilde{\boldsymbol{u}}$. Denote by $\boldsymbol{x} = \mathcal{S}(\boldsymbol{u}^o + \epsilon\tilde{\boldsymbol{u}};\boldsymbol{r},\boldsymbol{x}_0)$ the mild solution to the IVP

$$\dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})(\boldsymbol{u}^o(t) + \epsilon\tilde{\boldsymbol{u}}(t)), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{3.41}$$

The state $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}^o;\boldsymbol{r},\boldsymbol{x}_0)$ is by definition the mild solution of the IVP

$$\dot{\boldsymbol{x}}^o(t) = \mathcal{A}\boldsymbol{x}^o(t) + \mathcal{F}(\boldsymbol{x}^o(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}^o(t), \quad \boldsymbol{x}^o(0) = \boldsymbol{x}_0. \tag{3.42}$$

Define $\boldsymbol{x}_e = (\boldsymbol{x} - \boldsymbol{x}^o)/\epsilon - \tilde{\boldsymbol{x}}$, subtract the equations (3.42) and (3.40) from (3.41) to obtain

$$\begin{aligned}\dot{\boldsymbol{x}}_e(t) =& (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{x}_e(t) \\ &+ \frac{1}{\epsilon}\left(\mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}^o(t)) - \mathcal{F}'_{\boldsymbol{x}^o(t)}(\boldsymbol{x}(t) - \boldsymbol{x}^o(t))\right), \quad \boldsymbol{x}_e(0) = 0.\end{aligned} \tag{3.43}$$

Define $\boldsymbol{e}_{\mathcal{F}}(t)$ as

$$\boldsymbol{e}_{\mathcal{F}}(t) \coloneqq \frac{1}{\epsilon}\left(\mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}^o(t)) - \mathcal{F}'_{\boldsymbol{x}^o(t)}(\boldsymbol{x}(t) - \boldsymbol{x}^o(t))\right) \tag{3.44}$$

Assumption C1 ensures that for each $t \in [0,\tau]$, $\boldsymbol{e}_{\mathcal{F}}(t) \to 0$ as $\epsilon \to 0$. It will be shown that $\boldsymbol{e}_{\mathcal{F}}(t)$ is uniformly bounded. By Lemma 3.2.2, the norm of the states $\boldsymbol{x}(t)$ and $\boldsymbol{x}^o(t)$ is uniformly bounded over $[0,\tau]$ by some number δ,

$$\delta \le c_\tau \left(\|\boldsymbol{x}_0\| + M_{\mathcal{B}}R\right). \tag{3.45}$$

Use the local Lipschitz continuity of $\mathcal{F}(\cdot)$ (assumption A2) and Proposition 3.4.2(a) to obtain

$$\begin{aligned}\frac{1}{\epsilon}\left\|\mathcal{F}(\boldsymbol{x}(t))-\mathcal{F}(\boldsymbol{x}^o(t))\right\| &\leq \frac{1}{\epsilon}L_{\mathcal{F}\delta}\left\|\boldsymbol{x}(t)-\boldsymbol{x}^o(t)\right\| \\ &\leq L_{\mathcal{F}\delta}L_{\boldsymbol{u}}\left\|\tilde{\boldsymbol{u}}\right\|_p. \end{aligned} \tag{3.46}$$

Letting $M_{\mathcal{F}'}=\sup\{\|\mathcal{F}'_{\boldsymbol{x}^o(t)}\|:\ t\in[0,\tau]\}$, assumption C1 together with Proposition 3.4.2(a) also yields

$$\frac{1}{\epsilon}\left\|\mathcal{F}'_{\boldsymbol{x}^o(t)}(\boldsymbol{x}(t)-\boldsymbol{x}^o(t))\right\| \leq M_{\mathcal{F}'}L_{\boldsymbol{u}}\left\|\tilde{\boldsymbol{u}}\right\|_p. \tag{3.47}$$

Combining (3.46) and (3.47) leads to

$$\|\boldsymbol{e}_{\mathcal{F}}(t)\| \leq \left(L_{\mathcal{F}\delta}+M_{\mathcal{F}'}\right)L_{\boldsymbol{u}}\left\|\tilde{\boldsymbol{u}}\right\|_p, \quad \forall t\in[0,\tau]. \tag{3.48}$$

Now substitute (3.44) into (3.43). The state $\boldsymbol{x}_e$ is the mild solution to the IVP

$$\dot{\boldsymbol{x}}_e(t)=(\mathcal{A}+\mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{x}_e(t)+\boldsymbol{e}_{\mathcal{F}}(t), \quad \boldsymbol{x}_e(0)=0. \tag{3.49}$$

Recall that the mild solution of this evolution equation is described by an evolution operator $\mathcal{U}(t,s)$ by Lemma 3.4.3(a). Let $M_{\mathcal{U}}$ be an upper bound for the operator norm of $\mathcal{U}(t,s)$ over $0\leq t\leq s\leq\tau$, the mild solution to (3.49) satisfies the estimate

$$\begin{aligned}\|\boldsymbol{x}_e\|_{L^p(0,\tau;\mathbb{X})} &\leq \tau^{1/p}\|\boldsymbol{x}_e\|_{C(0,\tau;\mathbb{X})} \\ &\leq \tau^{1/p}M_{\mathcal{U}}\int_0^\tau \|\boldsymbol{e}_{\mathcal{F}}(t)\|\,dt. \end{aligned} \tag{3.50}$$

Since $\lim_{\epsilon\to0}\|\boldsymbol{e}_{\mathcal{F}}(t)\|=0$ for each $t\in[0,\tau]$ and $\|\boldsymbol{e}_{\mathcal{F}}(t)\|$ is uniformly bounded over $[0,\tau]$ for all ϵ, the bounded convergence theorem implies that the integral in (3.50) converges to zero. Thus,

$$\lim_{\epsilon\to0}\left\|\frac{1}{\epsilon}\left(\mathcal{S}(\boldsymbol{u}^o+\epsilon\tilde{\boldsymbol{u}};\boldsymbol{r},\boldsymbol{x}_0)-\mathcal{S}(\boldsymbol{u}^o;\boldsymbol{r},\boldsymbol{x}_0)\right)-\mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}\right\|_{L^p(0,\tau;\mathbb{X})}=\lim_{\epsilon\to0}\|\boldsymbol{x}_e\|_{L^p(0,\tau;\mathbb{X})}=0.$$

This proves that $\mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}$ is the Gâteaux derivative of $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ at $\boldsymbol{u}^o$ in the direction $\tilde{\boldsymbol{u}}$. □

Proposition 3.4.5. *Under assumptions A, C1, and C2, for every initial condition $\boldsymbol{x}_0\in\mathbb{X}$ and control input $\boldsymbol{u}\in B_R$, the control-to-state map $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ is Gâteaux differentiable in $\boldsymbol{r}$ in K_{ad}. The Gâteaux derivative of $\mathcal{S}(\boldsymbol{u};\boldsymbol{r},\boldsymbol{x}_0)$ at $\boldsymbol{r}^o$ in the direction $\tilde{\boldsymbol{r}}$ is*

$$\mathcal{S}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}=\tilde{\boldsymbol{y}}, \quad \forall\tilde{\boldsymbol{r}}\in\mathbb{K}, \tag{3.51}$$

where, defining $\boldsymbol{x}^o(t)=\mathcal{S}(\boldsymbol{u};\boldsymbol{r}^o,\boldsymbol{x}_0)$, $\tilde{\boldsymbol{y}}$ is the mild solution to the IVP

$$\dot{\tilde{\boldsymbol{y}}}(t)=(\mathcal{A}+\mathcal{F}'_{\boldsymbol{x}^o(t)})\tilde{\boldsymbol{y}}(t)+(\mathcal{B}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}})\,\boldsymbol{u}(t), \quad \tilde{\boldsymbol{y}}(0)=0. \tag{3.52}$$

Proof. Let number ϵ be small enough so that $\boldsymbol{r}^o + \epsilon\tilde{\boldsymbol{r}} \in K$. Denote by $\boldsymbol{x} = \mathcal{S}(\boldsymbol{u}; \boldsymbol{r}^o + \epsilon\tilde{\boldsymbol{r}}, \boldsymbol{x}_0)$ the mild solution to the IVP

$$\dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r}^o + \epsilon\tilde{\boldsymbol{r}})\boldsymbol{u}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{3.53}$$

The state $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}; \boldsymbol{r}^o, \boldsymbol{x}_0)$ is the mild solution of the IVP

$$\dot{\boldsymbol{x}}^o(t) = \mathcal{A}\boldsymbol{x}^o(t) + \mathcal{F}(\boldsymbol{x}^o(t)) + \mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}(t), \quad \boldsymbol{x}^o(0) = \boldsymbol{x}_0. \tag{3.54}$$

Define $\boldsymbol{x}_e = (\boldsymbol{x} - \boldsymbol{x}^o)/\epsilon - \tilde{\boldsymbol{y}}$, subtract the equations (3.54) and (3.52) from (3.53), obtain

$$\begin{aligned}\dot{\boldsymbol{x}}_e(t) =& (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{x}_e(t) + \frac{1}{\epsilon}\left(\mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}^o(t)) - \mathcal{F}'_{\boldsymbol{x}^o(t)}(\boldsymbol{x}(t) - \boldsymbol{x}^o(t))\right) \\ &+ \left(\frac{1}{\epsilon}\left(\mathcal{B}(\boldsymbol{r}^o + \epsilon\tilde{\boldsymbol{r}}) - \mathcal{B}(\boldsymbol{r}^o)\right) - \mathcal{B}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}\right)\boldsymbol{u}(t), \quad \boldsymbol{x}_e(0) = 0. \end{aligned} \tag{3.55}$$

Define $\boldsymbol{e}_{\mathcal{F}}(t)$ and $\boldsymbol{e}_{\mathcal{B}}$ as

$$\boldsymbol{e}_{\mathcal{F}}(t) \coloneqq \frac{1}{\epsilon}\left(\mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}^o(t)) - \mathcal{F}'_{\boldsymbol{x}^o(t)}(\boldsymbol{x}(t) - \boldsymbol{x}^o(t))\right), \tag{3.56a}$$

$$\boldsymbol{e}_{\mathcal{B}} \coloneqq \frac{1}{\epsilon}\left(\mathcal{B}(\boldsymbol{r}^o + \epsilon\tilde{\boldsymbol{r}}) - \mathcal{B}(\boldsymbol{r}^o)\right) - \mathcal{B}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}. \tag{3.56b}$$

Assumption C1 and C2 ensure that as $\epsilon \to 0$

$$\|\boldsymbol{e}_{\mathcal{F}}(t)\| \to 0, \quad \forall t \in [0, \tau], \tag{3.57a}$$

$$\|\boldsymbol{e}_{\mathcal{B}}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \to 0. \tag{3.57b}$$

Also, similar to inequality (3.48), using Proposition 3.4.2(b), and letting $\delta = c_\tau(\|\boldsymbol{x}_0\| + M_{\mathcal{B}}R)$ and $M_{\mathcal{F}'} = \sup\{\|\mathcal{F}'_{\boldsymbol{x}^o(t)}\| : \ t \in [0, \tau]\}$, the following upper bounded can be obtained

$$\|\boldsymbol{e}_{\mathcal{F}}(t)\| \le (L_{\mathcal{F}\delta} + M_{\mathcal{F}'})\, L_{\boldsymbol{r}}\, \|\tilde{\boldsymbol{r}}\|_{\mathbb{K}}, \quad \forall t \in [0, \tau]. \tag{3.58}$$

Rewrite (3.55) as follows

$$\dot{\boldsymbol{x}}_e(t) = (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{x}_e(t) + \boldsymbol{e}_{\mathcal{F}}(t) + \boldsymbol{e}_{\mathcal{B}}\boldsymbol{u}(t), \quad \boldsymbol{x}_e(0) = 0. \tag{3.59}$$

According to Lemma 3.4.3(a), the mild solution of this evolution equation is described by an evolution operator $\mathcal{U}(t,s)$ as follows

$$\dot{\boldsymbol{x}}_e(t) = \int_0^t \mathcal{U}(t,s)\boldsymbol{e}_{\mathcal{F}}(s)ds + \int_0^t \mathcal{U}(t,s)\boldsymbol{e}_{\mathcal{B}}\boldsymbol{u}(s)ds. \tag{3.60}$$

Let $M_{\mathcal{U}}$ be an upper bound for the operator norm of $\mathcal{U}(t,s)$ over $0 \le t \le s \le \tau$,

$$\begin{aligned} \|\boldsymbol{x}_e\|_{L^p(0,\tau;\mathbb{X})} &\le \tau^{1/p} \|\boldsymbol{x}_e\|_{C(0,\tau;\mathbb{X})} \\ &\le \tau^{1/p} M_{\mathcal{U}} \int_0^\tau \|\boldsymbol{e}_{\mathcal{F}}(t)\| + \tau^{1/p} M_{\mathcal{U}} \|\boldsymbol{e}_{\mathcal{B}}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\boldsymbol{u}\|_1 . \end{aligned} \tag{3.61}$$

As a result of (3.57a) and (3.58), the first integral in (3.61) tends to zero by the bounded convergence theorem. The second term of (3.61) also converges to zero using (3.57b). It follows that

$$\lim_{\epsilon \to 0} \left\| \frac{1}{\epsilon} \left(\mathcal{S}(\boldsymbol{u}; \boldsymbol{r}^o + \epsilon \tilde{\boldsymbol{r}}, \boldsymbol{x}_0) - \mathcal{S}(\boldsymbol{u}; \boldsymbol{r}^o, \boldsymbol{x}_0) \right) - \mathcal{S}'_{\boldsymbol{r}^o} \tilde{\boldsymbol{r}} \right\|_{L^p(0,\tau;\mathbb{X})} = \lim_{\epsilon \to 0} \|\boldsymbol{x}_e\|_{L^p(0,\tau;\mathbb{X})} = 0.$$

This shows that $\mathcal{S}'_{\boldsymbol{r}^o} \tilde{\boldsymbol{r}}$ is the Gâteaux derivative of $\mathcal{S}(\boldsymbol{u}; \boldsymbol{r}, \boldsymbol{x}_0)$ at $\boldsymbol{r}^o$ in the direction $\tilde{\boldsymbol{r}}$. □

In order to place the problem in a Hilbert space, assumptions C2 and C3 are used, assuming that the spaces are Hilbert spaces and defining a cost function. It will also be assumed that $p = 2$, considering inputs in $L^2(0,\tau;\mathbb{U})$. It is shown in the following lemma that this cost function is consistent with previous assumptions on the cost function (assumption B3).

Lemma 3.4.6. *The cost function in assumption C3 satisfies assumption B3; that is, it is weakly lower semi-continuous in $\boldsymbol{x}$ and $\boldsymbol{u}$.*

Proof. The cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ in assumption C3 is continuous and convex function in both $\boldsymbol{x}$ and $\boldsymbol{u}$. That is, letting $\lambda \in (0,1)$,

$$\begin{gathered} \int_0^\tau \langle \mathcal{Q}\boldsymbol{x}_n(t), \boldsymbol{x}_n(t) \rangle \, dt \to \int_0^\tau \langle \mathcal{Q}\boldsymbol{x}(t), \boldsymbol{x}(t) \rangle \, dt \quad \text{as} \quad \boldsymbol{x}_n \to \boldsymbol{x} \quad \text{in} \quad L^2(0,\tau;\mathbb{X}), \\ \langle \lambda \mathcal{Q}\boldsymbol{x}_1 + (1-\lambda)\mathcal{Q}\boldsymbol{x}_2, \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \rangle \le \lambda \langle \mathcal{Q}\boldsymbol{x}_1, \boldsymbol{x}_1 \rangle + (1-\lambda) \langle \mathcal{Q}\boldsymbol{x}_2, \boldsymbol{x}_2 \rangle , \end{gathered}$$

and a similar argument for $\boldsymbol{u}$. According to Theorem 13.2.2 in [119] and the corollary thereafter, if a functional defined on a Banach space is continuous and convex; then, it is also weakly lower semi-continuous. Therefore, the cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ is weakly lower semi-continuous in both $\boldsymbol{x}$ and $\boldsymbol{u}$. □

The next theorem derives the first order necessary conditions for an optimizer of the optimization problem P.

Theorem 3.4.7. *Suppose assumptions A1 and C hold. For any initial condition* $\boldsymbol{x}_0 \in \mathbb{X}$*, let the pair* $(\boldsymbol{u}^o, \boldsymbol{r}^o) \in U_{ad} \times K_{ad}$ *be a local minimizer of the optimization problem P with the optimal trajectory* $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0)$*. Let* $\boldsymbol{p}^o(t)$*, the adjoint state, indicate the mild solution of the final value problem*

$$\dot{\boldsymbol{p}}^o(t) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}^o(t)})\boldsymbol{p}^o(t) - \mathcal{Q}\boldsymbol{x}^o(t), \quad \boldsymbol{p}^o(\tau) = 0.$$

Then $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ *satisfies*

$$\left\langle \boldsymbol{u}^o + \mathcal{R}^{-1}\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o, \boldsymbol{u} - \boldsymbol{u}^o \right\rangle_{L^2(0,\tau;\mathbb{U})} \geq 0, \quad \forall \boldsymbol{u} \in U_{ad}, \tag{3.62a}$$

$$\left\langle \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^* \boldsymbol{p}^o(t)\, dt, \boldsymbol{r} - \boldsymbol{r}^o \right\rangle_{\mathbb{K}} \geq 0, \quad \forall \boldsymbol{r} \in K_{ad}. \tag{3.62b}$$

Proof. To derive the optimality conditions (3.62), the Gâteaux derivative of the cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ with respect to $\boldsymbol{u} \in U_{ad}$ and $\boldsymbol{r} \in K_{ad}$ is calculated. Using assumption C3, the cost function is sum of two inner products in the Hilbert spaces $L^2(0, \tau; \mathbb{X})$ and $L^2(0, \tau; \mathbb{U})$; that is

$$J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) = \langle \mathcal{Q}\boldsymbol{x}, \boldsymbol{x} \rangle_{L^2(0,\tau;\mathbb{X})} + \langle \mathcal{R}\boldsymbol{u}, \boldsymbol{u} \rangle_{L^2(0,\tau;\mathbb{U})}. \tag{3.63}$$

Thus, the Gâteaux derivative of J at $\boldsymbol{u}^o$ in the direction $\boldsymbol{h}_{\boldsymbol{u}}$ is

$$\begin{aligned} J'_{\boldsymbol{u}^o}\boldsymbol{h}_{\boldsymbol{u}} &= 2\,\langle \mathcal{Q}\mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0), \mathcal{S}'_{\boldsymbol{u}^o}\boldsymbol{h}_{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{X})} + 2\,\langle \mathcal{R}\boldsymbol{u}^o, \boldsymbol{h}_{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{U})} \\ &= 2\,\langle \mathcal{S}'^*_{\boldsymbol{u}^o}\mathcal{Q}\mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0) + \mathcal{R}\boldsymbol{u}^o, \boldsymbol{h}_{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{U})}. \end{aligned} \tag{3.64}$$

To calculate the adjoint operator $\mathcal{S}'^*_{\boldsymbol{u}^o}$, let $\tilde{\boldsymbol{u}}(t) \in L^2(0, \tau; \mathbb{U})$, $\tilde{\boldsymbol{x}}(t) \in L^2(0, \tau; \mathbb{X})$ be arbitrary. Using Lemma 3.4.3,

$$\begin{aligned} \langle \tilde{\boldsymbol{x}}, \mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{X})} &= \int_0^\tau \left\langle \tilde{\boldsymbol{x}}(t), \int_0^t \mathcal{U}(t,s)\mathcal{B}(\boldsymbol{r}^o)\tilde{\boldsymbol{u}}(s)\, ds \right\rangle dt \\ &= \int_0^\tau \int_s^\tau \langle \tilde{\boldsymbol{x}}(t), \mathcal{U}(t,s)\mathcal{B}(\boldsymbol{r}^o)\tilde{\boldsymbol{u}}(s) \rangle \, dtds \\ &= \int_0^\tau \left\langle \mathcal{B}^*(\boldsymbol{r}^o) \int_s^\tau \mathcal{U}^*(t,s)\tilde{\boldsymbol{x}}(t)dt, \tilde{\boldsymbol{u}}(s) \right\rangle_{\mathbb{U}} ds. \end{aligned} \tag{3.65}$$

Thus,

$$(\mathcal{S}'^*_{\boldsymbol{u}^o}\tilde{\boldsymbol{x}})(s) = \mathcal{B}^*(\boldsymbol{r}^o) \int_s^\tau \mathcal{U}^*(t,s)\tilde{\boldsymbol{x}}(t)dt.$$

Define $\tilde{\boldsymbol{p}}(s) = \int_s^\tau \mathcal{U}^*(t,s)\tilde{\boldsymbol{x}}(t)dt$. By Lemma 3.4.3(b), $\tilde{\boldsymbol{p}}(s)$ is the mild solution of the following FVP solved backward in time

$$\dot{\tilde{\boldsymbol{p}}}(s) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}^o(s)})\tilde{\boldsymbol{p}}(s) - \tilde{\boldsymbol{x}}(s), \quad \tilde{\boldsymbol{p}}(\tau) = 0. \tag{3.66}$$

It follows that

$$(\mathcal{S}'^*_{\boldsymbol{u}^o}\tilde{\boldsymbol{x}})(s) = \mathcal{B}^*(\boldsymbol{r}^o)\tilde{\boldsymbol{p}}(s). \tag{3.67}$$

Use (3.67) in (3.64) to obtain

$$J'_{\boldsymbol{u}^o}\boldsymbol{h}_{\boldsymbol{u}} = 2\left\langle \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o(s) + \mathcal{R}\boldsymbol{u}^o, \boldsymbol{h}_{\boldsymbol{u}}\right\rangle_{L^2(0,\tau;\mathbb{U})}. \tag{3.68}$$

Applying [61, Theorem 1.46] yields the optimality condition

$$\left\langle \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o + \mathcal{R}\boldsymbol{u}^o, \boldsymbol{u} - \boldsymbol{u}^o\right\rangle_{L^2(0,\tau;\mathbb{U})} \geq 0, \quad \forall \boldsymbol{u} \in U_{ad}, \tag{3.69}$$

where $\boldsymbol{p}^o(s)$ solves

$$\dot{\boldsymbol{p}}^o(s) = -(\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(s)})^*\boldsymbol{p}^o(s) - \mathcal{Q}\boldsymbol{x}^o(s), \quad \boldsymbol{p}^o(\tau) = 0. \tag{3.70}$$

Since $\mathcal{R}$ is positive-definite, and hence, invertible, inequality (3.62a) follows.

Taking the directional derivative of $J(\boldsymbol{u}^o, \cdot\,; \boldsymbol{x}_0)$ at $\boldsymbol{r}^o$ in the direction $\boldsymbol{h}_{\boldsymbol{r}}$ yields

$$\begin{aligned} J'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}} &= 2\left\langle \mathcal{Q}\mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0), \mathcal{S}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}}\right\rangle_{L^2(0,\tau;\mathbb{X})} \\ &= 2\left\langle \mathcal{S}'^*_{\boldsymbol{r}^o}\mathcal{Q}\mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0), \boldsymbol{h}_{\boldsymbol{r}}\right\rangle_{\mathbb{K}}. \end{aligned} \tag{3.71}$$

To calculate the adjoint operator $\mathcal{S}'^*_{\boldsymbol{r}^o}$, use Lemma 3.4.3(b), and proceed as follows

$$\begin{aligned} \left\langle \mathcal{Q}\mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0), \mathcal{S}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}}\right\rangle_{L^2(0,\tau;\mathbb{X})} &= \int_0^\tau \left\langle \mathcal{Q}\boldsymbol{x}^o(t), \int_0^t \mathcal{U}(t,s)(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}})\boldsymbol{u}^o(s)\,ds \right\rangle dt \\ &= \int_0^\tau \left\langle \int_s^\tau \mathcal{U}^*(t,s)\mathcal{Q}\boldsymbol{x}^o(t)\,dt, (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}})\boldsymbol{u}^o(s) \right\rangle ds \\ &= \int_0^\tau \left\langle \boldsymbol{p}^o(s), (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}})\boldsymbol{u}^o(s)\right\rangle ds. \end{aligned} \tag{3.72}$$

For each $\boldsymbol{u} \in \mathbb{U}$, $(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}})u$ is an element of $\mathbb{X}$. This defines a bounded linear map from $\boldsymbol{h}_{\boldsymbol{r}} \in \mathbb{K}$ to $\mathbb{X}$. There exists a bounded linear operator $(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^*\colon \mathbb{X} \to \mathbb{K}$ satisfying

$$\left\langle (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^*\boldsymbol{p}, \boldsymbol{h}_{\boldsymbol{r}}\right\rangle_{\mathbb{K}} = \left\langle \boldsymbol{p}, (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{h}_{\boldsymbol{r}})u\right\rangle. \tag{3.73}$$

Incorporate this into (3.72) to obtain

$$\langle \mathcal{Q}\mathcal{S}(\boldsymbol{u}^o;\boldsymbol{r}^o,\boldsymbol{x}_0), \mathcal{S}'_{\boldsymbol{r}^o}\boldsymbol{h_r}\rangle_{L^2(0,\tau;\mathbb{X})} = \int_0^\tau \langle (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(s))^*\boldsymbol{p}^o(s), \boldsymbol{h_r}\rangle_{\mathbb{K}}\, ds$$
$$= \left\langle \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(s))^*\boldsymbol{p}^o(s)\, ds, \boldsymbol{h_r}\right\rangle_{\mathbb{K}}. \tag{3.74}$$

This gives an explicit form of the adjoint operator $\mathcal{S}'^*_{\boldsymbol{r}^o}$. Again, by Theorem 1.46 of [61], the inner product (3.18) must be non-negative for any direction $\boldsymbol{r}-\boldsymbol{r}^o$ in K_{ad} yielding (3.62b). □

Corollary 3.4.8. *If the minimizer $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ is in the interior of $U_{ad}\times K_{ad}$, then the following set of equations characterizes $(\boldsymbol{x}^o, \boldsymbol{p}^o, \boldsymbol{u}^o, \boldsymbol{r}^o)$:*

$$\begin{cases} \dot{\boldsymbol{x}}^o(t) = \mathcal{A}\boldsymbol{x}^o(t) + \mathcal{F}(\boldsymbol{x}^o(t)) + \mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}^o(t), & \boldsymbol{x}^o(0) = \boldsymbol{x}_0, \\ \dot{\boldsymbol{p}}^o(t) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}^o(t)})\boldsymbol{p}^o(t) - \mathcal{Q}\boldsymbol{x}^o(t), & \boldsymbol{p}^o(\tau) = 0, \\ \boldsymbol{u}^o(t) = -\mathcal{R}^{-1}\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o(t), \\ \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)\, dt = 0. \end{cases} \tag{3.75}$$

Proof. In the interior of $U_{ad}\times K_{ad}$, all directions $\boldsymbol{u}-\boldsymbol{u}^o \in L^2(0,\tau;\mathbb{U})$ and $\boldsymbol{r}-\boldsymbol{r}^o\in\mathbb{K}$ are permitted. Thus, the inner products in (3.62) are non-negative only if

$$\boldsymbol{u}^o(t) + \mathcal{R}^{-1}\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o(t) = 0,$$
$$\int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)\, dt = 0.$$

□

If the control space $\mathbb{U}$ and actuator design space $\mathbb{K}$ are separable Hilbert spaces, the optimizing control and actuator can be characterized further. Let $\boldsymbol{e}_j^{\mathbb{K}}$, $\boldsymbol{e}_i^{\mathbb{U}}$, and $\boldsymbol{e}_k^{\mathbb{X}}$ be orthonormal bases for $\mathbb{K}$, $\mathbb{U}$, and $\mathbb{X}$, respectively. Then there exists $\boldsymbol{b}_i(\boldsymbol{r})\in\mathbb{X}$, $\boldsymbol{r}\in\mathbb{K}$ so that for any $\boldsymbol{u}\in\mathbb{U}$,

$$\mathcal{B}(\boldsymbol{r})\boldsymbol{u} = \sum_{i=1}^{\infty} \left\langle \boldsymbol{u}, \boldsymbol{e}_i^{\mathbb{U}}\right\rangle_{\mathbb{U}} \boldsymbol{b}_i(\boldsymbol{r}). \tag{3.76}$$

Since the operator $\mathcal{B}(\cdot)\boldsymbol{u}:\mathbb{K}\to\mathbb{X}$ is Gâteaux differentiable with respect to $\boldsymbol{r}$, each $\boldsymbol{b}_i(\cdot)$ is a Gâteaux differentiable map from $\mathbb{K}$ to $\mathbb{X}$. Denote the Gâteaux derivative of $\boldsymbol{b}_i(\boldsymbol{r})$ at $\boldsymbol{r}^o$ by $\boldsymbol{b}'_{i,\boldsymbol{r}^o}:\mathbb{K}\to\mathbb{X}$, then

$$(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{r})\boldsymbol{u} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \left\langle \boldsymbol{u}, \boldsymbol{e}_i^{\mathbb{U}}\right\rangle_{\mathbb{U}} \left\langle \boldsymbol{r}, \boldsymbol{e}_j^{\mathbb{K}}\right\rangle_{\mathbb{K}} \boldsymbol{b}'_{i,\boldsymbol{r}^o}\boldsymbol{e}_j^{\mathbb{K}}. \tag{3.77}$$

Corollary 3.4.9. *Assume further that the input space $\mathbb{U}$ and actuator design space $\mathbb{K}$ are separable. Let $\boldsymbol{e}_i^{\mathbb{U}}$, $\boldsymbol{e}_j^{\mathbb{K}}$ and $\boldsymbol{e}_k^{\mathbb{X}}$ be orthonormal bases for $\mathbb{K}$, $\mathbb{U}$, and $\mathbb{X}$, respectively. Define $\boldsymbol{u}_j^o(t)$ and $\boldsymbol{p}_k(t)$ as*

$$\boldsymbol{u}_j^o(t) \coloneqq \left\langle \boldsymbol{u}^o(t), \boldsymbol{e}_j^{\mathbb{U}} \right\rangle_{\mathbb{U}}, \tag{3.78a}$$

$$\boldsymbol{p}_k^o(t) \coloneqq \left\langle \boldsymbol{p}^o(t), \boldsymbol{e}_k^{\mathbb{X}} \right\rangle. \tag{3.78b}$$

The optimality conditions (3.62) in the interior of $U_{ad} \times K_{ad}$ can be written as

$$\boldsymbol{u}_j^o(t) + \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \left\langle \boldsymbol{b}_i(\boldsymbol{r}^o), \boldsymbol{e}_k^{\mathbb{X}} \right\rangle \left\langle \mathcal{R}^{-1} \boldsymbol{e}_i^{\mathbb{U}}, \boldsymbol{e}_j^{\mathbb{U}} \right\rangle_{\mathbb{U}} \boldsymbol{p}_k^o(t) = 0, \quad \textit{for each } j, \tag{3.79a}$$

$$\sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \left\langle \boldsymbol{b}'_{i,\boldsymbol{r}^o} \boldsymbol{e}_j^{\mathbb{K}}, \boldsymbol{e}_k^{\mathbb{X}} \right\rangle \int_0^{\tau} \boldsymbol{u}_i^o(s) \boldsymbol{p}_k^o(s)\, ds = 0, \quad \textit{for each } j. \tag{3.79b}$$

Proof. For every $\boldsymbol{p} \in \mathbb{X}$, the element $\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p} \in \mathbb{U}$ can by obtained by using (3.76), and doing the calculation

$$\begin{aligned} \langle \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}, \boldsymbol{u} \rangle_{\mathbb{U}} &= \langle \boldsymbol{p}, \mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u} \rangle \\ &= \sum_{i=1}^{\infty} \left\langle \boldsymbol{u}, \boldsymbol{e}_i^{\mathbb{U}} \right\rangle_{\mathbb{U}} \langle \boldsymbol{p}, \boldsymbol{b}_i(\boldsymbol{r}^o) \rangle \\ &= \left\langle \sum_{i=1}^{\infty} \langle \boldsymbol{b}_i(\boldsymbol{r}^o), \boldsymbol{p} \rangle\, \boldsymbol{e}_i^{\mathbb{U}}, \boldsymbol{u} \right\rangle_{\mathbb{U}}. \end{aligned} \tag{3.80}$$

This yields

$$\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p} = \sum_{i=1}^{\infty} \langle \boldsymbol{b}_i(\boldsymbol{r}^o), \boldsymbol{p} \rangle\, \boldsymbol{e}_i^{\mathbb{U}}. \tag{3.81}$$

Similarly, using (3.77), for every $\boldsymbol{u} \in \mathbb{U}$, the operator $(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^*$ maps $\boldsymbol{p} \in \mathbb{X}$ to $\mathbb{K}$ as follows

$$(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o)^*\boldsymbol{p} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \left\langle \boldsymbol{u}^o, \boldsymbol{e}_i^{\mathbb{U}} \right\rangle_{\mathbb{U}} \left\langle \boldsymbol{b}'_{i,\boldsymbol{r}^o} \boldsymbol{e}_j^{\mathbb{K}}, \boldsymbol{p} \right\rangle \boldsymbol{e}_j^{\mathbb{K}}. \tag{3.82}$$

Substituting these elements into the optimality conditions (3.62) and using (3.78) leads to (3.79). □

3.5 Examples

3.5.1 Railway Track Model

Railway tracks are rested on ballast which are known for exhibiting nonlinear viscoelastic behavior [5]. If a track beam is made of a Kelvin-Voigt material, then the railway track model will be a semilinear partial differential equation on $\xi \in [0, \ell]$ as follows:

$$\begin{cases} \rho a \dfrac{\partial^2 w}{\partial t^2} + \dfrac{\partial}{\partial \xi^2}(EI \dfrac{\partial^2 w}{\partial \xi^2} + C_d \dfrac{\partial^3 w}{\partial \xi^2 \partial t}) + \mu \dfrac{\partial w}{\partial t} + kw + \alpha w^3 = b(\xi; r)u(t), & \xi \in (0, \ell), \\ w(\xi, 0) = w_0(\xi), \quad \dfrac{\partial w}{\partial t}(\xi, 0) = v_0(\xi), & \xi \in (0, \ell), \\ w(0, t) = w(\ell, t) = 0, & t \geq 0, \\ EI \dfrac{\partial^2 w}{\partial \xi^2}(0, t) + C_d \dfrac{\partial^3 w}{\partial \xi^2 \partial t}(0, t) = EI \dfrac{\partial^2 w}{\partial \xi^2}(\ell, t) + C_d \dfrac{\partial^3 w}{\partial \xi^2 \partial t}(\ell, t) = 0, & t \geq 0. \end{cases}$$

where the positive constants E, I, ρ, a, and ℓ are the modulus of elasticity, second moment of inertia, density of the beam, cross-sectional area, and length of the beam, respectively. The linear and nonlinear parts of the foundation elasticity correspond to the coefficients k and α, respectively. The constant $\mu \geq 0$ is the viscous damping coefficient of the foundation, and $C_d \geq 0$ is the coefficient of Kelvin-Voigt damping in the beam. The track deflection is controlled by an external force $u(t)$; $u(t)$ will be assumed to be a scalar input in order to simplify the exposition. The shape influence function $b(\xi; r)$ is a continuous function over $[0, \ell]$ parametrized by the parameter r that describes its dependence on actuator location. For example, as shown in Figure 3.1, the control force is typically localized at some point r and $b(\xi; r)$ models the distribution of the force $u(t)$ along the beam. The function $b(\xi; r)$ is assumed continuously differentiable with respect to r over $\mathbb{R}$ (assumptions B2 and C2); a suitable function for the case of actuator location is illustrated in Figure 3.1.

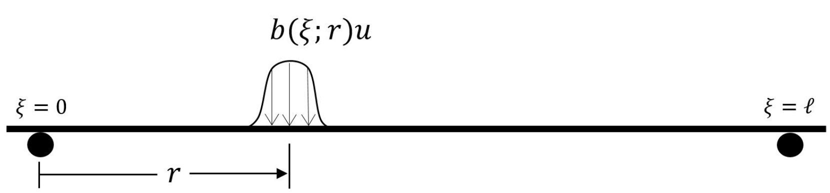

Figure 3.1: Schematic of an actuator on the railway track beam.

Define the closed self-adjoint positive operator $\mathcal{A}_0$ on $L^2(0,\ell)$ as:

$$\mathcal{A}_0 w := w_{\xi\xi\xi\xi},$$
$$D(\mathcal{A}_0) := \left\{ w \in H^4(0,\ell) |\, w(0) = w(\ell) = 0,\ w_{\xi\xi}(0) = w_{\xi\xi}(\ell) = 0 \right\}, \tag{3.83}$$

where the subscript $\cdot_\xi$ denote the derivative with respect to the spatial variable. As a result, the state operator associated with the Kelvin-Voigt beam is

$$\mathcal{A}_{KV}(w,v) := \left(v, -\frac{1}{\rho a}\mathcal{A}_0(EIw + C_d v) \right), \tag{3.84}$$

which is defined on the state space $\mathbb{X} = H^2(0,\ell) \cap H_0^1(0,\ell) \times L^2(0,\ell)$ equipped with the norm

$$\|(w,v)\|^2 = \int_0^\ell EIw_{\xi\xi}^2 + kw^2 + \rho a v^2 \, d\xi. \tag{3.85}$$

Accordingly, the domain of the state operator is

$$D(\mathcal{A}_{KV}) := \left\{ (w,v) \in \mathbb{X} |\, v \in H^2(0,\ell) \cap H_0^1(0,\ell),\ EIw + C_d v \in D(\mathcal{A}_0) \right\}. \tag{3.86}$$

The underlying state space $\mathbb{X}$ is separable since the spaces $H^2(0,\ell) \cap H_0^1(0,\ell)$ and $L^2(0,\ell)$ are separable. Furthermore, define the linear operators $\mathcal{K}$, $\mathcal{B}(r)$, and the nonlinear operator $\mathcal{F}(\cdot)$ as

$$\mathcal{K}(w,v) := (0, -\frac{1}{\rho a}(\mu v + kw)), \tag{3.87}$$

$$\mathcal{B}(r)u := (0, \frac{1}{\rho a} b(\xi; r)u), \tag{3.88}$$

$$\mathcal{F}(w,v) := (0, -\frac{\alpha}{\rho a} w^3). \tag{3.89}$$

The operator $\mathcal{K}$ is a bounded linear operator on $\mathbb{X}$. For each r, operator $\mathcal{B}(r)$ is also a bounded operator that maps an input $u \in \mathbb{R}$ to the state space $\mathbb{X}$. Since the space $H^2(0,\ell)$ is contained in the space of continuous functions over $[0,\ell]$, the the nonlinear term w^3 is in $L^2(0,\ell)$. Thus, the nonlinear operator $\mathcal{F}(\cdot)$ is well-defined on $\mathbb{X}$. Lastly, define the operator $\mathcal{A} = \mathcal{A}_{KV} + \mathcal{K}$, with the same domain as $\mathcal{A}_{KV}$. With these definition and by setting $\boldsymbol{x} = (w,v)$, the state space representation of the railway model is

$$\dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(r)u(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \in D(\mathcal{A}). \tag{3.90}$$

It is straightforward to show that the operator $\mathcal{A}_0$ is closed, densely-defined, self-adjoint, and positive; it also has a compact resolvent. As a result, the operator $\mathcal{A}_{KV}$ will be a special case of the operator $\mathscr{A}_B$ in [21] with $\alpha = 1$. According to Theorem 1.1 in [21], such operators are generator of an analytic semigroup (also see [8, Section 3] for a different approach). Furthermore, the operator $\mathcal{A}_{KV} + \mathcal{K}$ is a bounded perturbation of the operator $\mathcal{A}_{KV}$. By Corollary 3.2.2 in [104], $\mathcal{A}_{KV} + \mathcal{K}$ also generates an analytic semigroup.

The railway track model in [5] neglects the Kelvin-Voigt damping in the beam (i.e. $C_d = 0$), and only includes Kelvin-Voigt damping in the ballast. In this case, the semigroup generated by $\mathcal{A}$ is not analytic. The results of this chapter hold true for both models.

The following result is due to Simon [110, Theorem 3], and will be used to check assumption B1.

Theorem 3.5.1. *[110, Theorem 3] Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces, and $\mathbb{Y} \hookrightarrow \mathbb{X}$ with compact embedding. Assume $X \subset L^p(0,\tau;\mathbb{X})$ where $1 \le p \le \infty$, and*

$$X \text{ is bounded in } L^1_{loc}(0,\tau;\mathbb{Y}), \tag{3.91}$$

$$\int_0^{\tau-h} \|\boldsymbol{x}(t+h) - \boldsymbol{x}(t)\|^p_{\mathbb{X}}\, dt \to 0 \text{ as } h \to 0 \text{ uniformly for } \boldsymbol{x} \in X. \tag{3.92}$$

Then, X is relatively compact in $L^p(0,\tau;\mathbb{X})$ (and in $C(0,\tau;\mathbb{X})$ if $p = \infty$).

Lemma 3.5.2. *The operator $\mathcal{F}(\cdot)$*

1. *is continuously Fréchet differentiable on $\mathbb{X}$; the Fréchet derivative of this operator at $\boldsymbol{x} = (w,v)$ maps every $\boldsymbol{p} = (f,g)$ to $\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p} = (0, -3\alpha w^2 f/\rho a)$,*
2. *the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded, and*
3. *$\mathcal{F}(\cdot)$ satisfies assumption B1.*

Proof. If $\mathcal{F}'_{\boldsymbol{x}}$ is the Fréchet derivative of $\mathcal{F}(\cdot)$ at $\boldsymbol{x}$, this operator must satisfy

$$\lim_{\boldsymbol{p}\to 0} \frac{\|\mathcal{F}(\boldsymbol{x}+\boldsymbol{p}) - \mathcal{F}(\boldsymbol{x}) - \mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}\|}{\|\boldsymbol{p}\|} = 0. \tag{3.93}$$

Recall the definition of the operator $\mathcal{F}$ and that of norm on the space $\mathbb{X}$, above limit simplifies to

$$\lim_{\|f\|_{H^2}\to 0} \frac{\|f^3 + 3f^2 w\|_{L^2}}{\|f\|_{H^2}} = 0. \tag{3.94}$$

Notice that functions f and w are in $H^2(0,\ell)$, and thus, continuous on $[0,\ell]$. Use triangle inequality, and Hölder's inequality to obtain

$$\begin{aligned}\|f^3+3f^2w\|_{L^2} &\le \|f^3\|_{L^2}+\|3f^2w\|_{L^2}\\ &\le \|f\|_{L^6}^3+3\|f\|_{L^8}^2\|w\|_{L^4}.\end{aligned} \tag{3.95}$$

Apply the Sobolev embedding result $H^2(0,\ell)\hookrightarrow L^p(0,\ell)$ and let c_p be the embedding constant

$$\|f^3+3f^2w\|_{L^2}\le c_6^3\|f\|_{H^2}^3+3c_8^2c_4\|f\|_{H^2}^2\|w\|_{H^2}, \tag{3.96}$$

As a result, the expression in (3.94) is bounded above according to

$$\frac{\|f^3+3f^2w\|_{L^2}}{\|f\|_{H^2}}\le c_6^3\|f\|_{H^2}^2+3c_8^2c_4\|w\|_{H^2}\|f\|_{H^2}. \tag{3.97}$$

This shows that the limit in (3.94) holds, and the operator $\mathcal{F}(\cdot)$ is indeed Fréchet differentiable.

Furthermore, select $\boldsymbol{x}_1=(w_1,v_1)$, $\boldsymbol{x}_2=(w_2,v_2)$, and $\boldsymbol{p}=(f,g)$ as generic elements of $\mathbb{X}$. The Fréchet derivative of $\mathcal{F}(\cdot)$ at $\boldsymbol{x}_2-\boldsymbol{x}_1$ is

$$\mathcal{F}'_{\boldsymbol{x}_2-\boldsymbol{x}_1}\boldsymbol{p}=(0,-\frac{3\alpha}{\rho a}(w_2-w_1)^2f).$$

Take the norm of $\mathcal{F}'_{\boldsymbol{x}_2-\boldsymbol{x}_1}\boldsymbol{p}$, and use Hölder's inequality to obtain

$$\begin{aligned}\left\|\mathcal{F}'_{\boldsymbol{x}_2-\boldsymbol{x}_1}\boldsymbol{p}\right\| &= \frac{3\alpha}{\sqrt{\rho a}}\left(\int_0^\ell (w_2(\xi)-w_1(\xi))^4 f^2(\xi)\,d\xi\right)^{\frac{1}{2}}\\ &\le \frac{3\alpha}{\sqrt{\rho a}}\left\|w_2-w_1\right\|_{L^8}^2\left\|f\right\|_{L^4}.\end{aligned} \tag{3.98}$$

Applying the Sobolev embedding result $H^2(0,\ell)\hookrightarrow L^p(0,\ell)$ yields

$$\begin{aligned}\left\|\mathcal{F}'_{\boldsymbol{x}_2-\boldsymbol{x}_1}\boldsymbol{p}\right\| &\le \frac{3\alpha}{\sqrt{\rho a}}c_8^2c_4\left\|w_2-w_1\right\|_{H^2}^2\left\|f\right\|_{H^2}\\ &\le \frac{3\alpha}{\sqrt{\rho a}}c_8^2c_4\left\|\boldsymbol{x}_2-\boldsymbol{x}_1\right\|^2\left\|\boldsymbol{p}\right\|.\end{aligned} \tag{3.99}$$

The last inequality indicates that the operator norm of $\mathcal{F}'_{\boldsymbol{x}}$ continuously depends on $\boldsymbol{x}$.

Inequality (3.99) also yields

$$\|\mathcal{F}'_{\boldsymbol{x}}\|_{\mathcal{L}(\mathbb{X})} \leq \frac{3\alpha}{\sqrt{\rho a}} c_8^2 c_4 \|\boldsymbol{x}\|^2 . \tag{3.100}$$

This shows that the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded.

To show that the nonlinear operator $\mathcal{F}(\cdot)$ satisfies assumption B1, consider a bounded sequence $\boldsymbol{x}_n(t) = (w_n(t), v_n(t))$ in $C(0,\tau;\mathbb{X})$ weakly converging to some element $\boldsymbol{x}(t) = (w(t), v(t))$ in $L^p(0,\tau;\mathbb{X})$. It is shown that the sequence $w_n(t)$ satisfies conditions of Theorem 3.5.1. The sequence $w_n(t)$ is by assumption bounded in $C(0,\tau;H^2(0,\ell)\cap H_0^1(0,\ell))$, and so in $C(0,\tau;L^6(0,\ell))$. This ensures that for all $p \in [1,\infty)$

$$\int_0^{\tau-h} \|w_n(t+h) - w_n(t)\|^p_{L^6(0,\ell)}\, dt \to 0 \text{ as } h \to 0 \text{ uniformly for all } n. \tag{3.101}$$

Also, the space $H^2(0,\ell) \cap H_0^1(0,\ell)$ is compactly embedded in $L^6(0,\ell)$ by the Rellich-Kondrachov Theorem [1, Chapter 6]. According to Theorem 3.5.1, $w_n(t)$ has a strongly convergent subsequence in $L^p(0,\tau;L^6(0,\ell))$. Recall that $w_n(t)$ weakly converges to $w(t)$ in $L^2(0,\tau;H^2(0,\ell)\cap H_0^1(0,\ell))$ as well. A weak limit is unique; thus, $w_n \to w$ in $L^p(0,\tau;L^6(0,\ell))$. This further implies that $w_n^3 \to w^3$ in $L^p(0,\tau;L^2(0,\ell))$. The nonlinear operator $\mathcal{F}(\cdot)$ maps $\boldsymbol{x}_n(t)$ to

$$\mathcal{F}(w_n(t), v_n(t)) = (0, \frac{\alpha}{\rho a} w_n^3(t)). \tag{3.102}$$

Thus, the sequence $\mathcal{F}(w_n(t), v_n(t))$ strongly (and so weakly) converges to $\mathcal{F}(w(t), v(t))$ in $L^2(0,\tau;\mathbb{X})$. □

The previous lemma implies that the nonlinear operator $\mathcal{F}(\cdot)$ of the railway track model satisfies assumption A2. Therefore, Theorem 2.4.2 ensures that for every initial condition $\boldsymbol{x}_0 \in \mathbb{X}$ and positive number R, there exists $\tau > 0$ such that for all $u \in L^2(0,\tau)$, $\|u\|_2 \leq R$, the railway track model has a unique mild solution $\boldsymbol{x} \in C(0,\tau;\mathbb{X})$. If the viscous damping coefficient μ is positive, then the existence of a mild solution for all time intervals, initial conditions and inputs was established in [39].

Theorem 3.5.3. *[39, Theorem 4] If $\mu > 0$, then the railway track model admits a unique mild solution $\boldsymbol{x} \in C(0,\tau;\mathbb{X})$ for all $\tau > 0$, all $\boldsymbol{x}_0 \in \mathbb{X}$, and all $u \in L^2(0,\tau)$.*

The nonlinear operator in the railway track model also satisfies assumptions B1 and C1. As a result, the existence of an optimal pair (u^o, r^o) together with an optimal trajectory $\boldsymbol{x}^o$ follows from Theorem 3.3.1. Also, using Theorem 3.4.7, the optimal pair (u^o, r^o) satisfies the optimality conditions (3.62). In order to characterize the optimizers (3.75), some adjoint operators need to be calculated. Calculation of the operator $\mathcal{A}^*$ is straightforward; it is

$$\mathcal{A}^*(f,g) = \left(-g, \frac{1}{\rho a}\mathcal{A}_0(EIf - C_d g) + \frac{k}{\rho a} f - \frac{\mu}{\rho a} g\right), \tag{3.103}$$

for all (f,g) in the domain

$$D(\mathcal{A}^*) = \left\{(f,g) \in \mathbb{X} | \, g \in H^2(0,\ell) \cap H_0^1(0,\ell), \; EIf - C_d g \in D(\mathcal{A}_0)\right\}. \tag{3.104}$$

Let $\boldsymbol{x}^o(t) = (w^o, v^o)$ be the optimal trajectory evaluated at time t. To calculate the adjoint of the operator $\mathcal{F}'_{\boldsymbol{x}^o(t)}$ for every t on $\mathbb{X}$, take the inner product $\mathcal{F}'_{\boldsymbol{x}^o(t)}(w,v)$ with $(f,g) \in \mathbb{X}$; that is

$$\left\langle \mathcal{F}'_{\boldsymbol{x}^o(t)}(w,v), (f,g) \right\rangle = \int_0^\ell -3\alpha (w^o(\xi))^2 w(\xi) g(\xi) \, d\xi. \tag{3.105}$$

For any $g \in L^2(0,\ell)$, consider the function $h \in H^2(0,\ell) \cap H_0^1(0,\ell)$ satisfying the differential equation

$$\begin{aligned} &EIh_{\xi\xi\xi\xi}(\xi) + kh(\xi) = -3\alpha(w^o(\xi))^2 g(\xi), \\ &h(0) = h(\ell) = 0, \\ &h_{\xi\xi}(0) = h_{\xi\xi}(\ell) = 0. \end{aligned} \tag{3.106}$$

An explicit solution $h(\xi)$ to (3.106) can be calculated using a Green's function:

$$\begin{aligned} h(\xi) &= -3\alpha \int_0^\ell G(\xi,\eta)(w^o(\eta))^2 g(\eta) \, d\eta, \\ G(\xi,\eta) &= \frac{1}{6\ell} \begin{cases} (2\ell^2\eta - 3\ell\eta^2 + \eta^3)\xi + (\eta - \ell)\xi^3, & \xi \leq \eta \\ (\eta^3 - \ell^2\eta)\xi + \eta\xi^3, & \xi > \eta \,. \end{cases} \end{aligned} \tag{3.107}$$

With this calculation, for any $(w,v) \in \mathbb{X}$,

$$\begin{aligned} \langle (w,v),(h,0)\rangle &= \int_0^\ell EIw_{\xi\xi}(\xi) h_{\xi\xi}(\xi) + kw(\xi)h(\xi) \, d\xi \\ &= EI[h_{\xi\xi} w_\xi]_0^\ell - EI[h_{\xi\xi\xi} w]_0^\ell + \int_0^\ell (EIh_{\xi\xi\xi\xi}(\xi) + kh(\xi)) w(\xi) \, d\xi \\ &= \int_0^\ell -3\alpha(w^o(\xi))^2 w(\xi) g(\xi) \, d\xi. \end{aligned} \tag{3.108}$$

Comparing this equation to (3.105); the adjoint of $\mathcal{F}'_{\boldsymbol{x}^o(t)}$ is defined by

$$\mathcal{F}'^*_{\boldsymbol{x}^o(t)}(f,g) = (h,0). \tag{3.109}$$

The adjoint of the operator $\mathcal{B}(r)$ for every $(f,g) \in \mathbb{X}$ is

$$\mathcal{B}^*(r)(f,g) = \int_0^\ell b(\xi;r)g(\xi)\,d\xi. \tag{3.110}$$

Also, denote $b_r(\xi;r)$ to be the derivative of $b(\xi;r)$ with respect to r and let $\boldsymbol{p}^o(t) = (f,g)$ at time t. Use Corollary 3.4.9 to find

$$(\mathcal{B}'_r u)^* \boldsymbol{p}^o(t) = u \int_0^\ell b_r(\xi;r)g(\xi)d\xi, \quad \forall (f,g) \in \mathbb{X}. \tag{3.111}$$

Furthermore, let $q_1 \in C^2([0,\ell])$ and $q_2 \in C([0,\ell])$ be some non-negative functions. Set $\mathcal{Q}(w,v) = (q_1 w, q_2 v)$ and $\mathcal{R} = 1$ in the cost function of assumption C3.

Now applying Theorem 3.4.7 and assuming that (u^o, r^o) is in the interior of $U_{ad} \times K_{ad}$, the following set of equations yields an optimizer for every initial condition $\boldsymbol{x}_0 = (w_0, v_0) \in \mathbb{X}$:

$$\begin{cases} \rho a w^o_{tt} + (EIw^o_{\xi\xi} + C_d w^o_{t\xi\xi})_{\xi\xi} + \mu w^o_t + k w^o + \alpha (w^o)^3 & \\ \qquad = b(\xi;r^o)u^o(t), & \xi \in (0,\ell), \\ w^o(0,t) = w^o(\ell,t) = 0, & t \geq 0, \\ EIw^o_{\xi\xi}(0,t) + C_d w^o_{t\xi\xi}(0,t) = 0, & t \geq 0, \\ EIw^o_{\xi\xi}(\ell,t) + C_d w^o_{t\xi\xi}(\ell,t) = 0, & t \geq 0, \\ w^o(\xi,0) = w_0(\xi),\ w^o_t(\xi,0) = v_0(\xi), & \xi \in (0,\ell), \end{cases} \tag{IVP}$$

$$\begin{cases} \rho a f^o_t - \rho a g^o + 3\alpha \int_0^\ell G(\xi,\eta)(w^o(\eta))^2 g^o(\eta)\,d\eta & \\ \qquad = -\rho a q_1(\xi) w^o, & \xi \in (0,\ell), \\ \rho a g^o_t + (EIf^o_{\xi\xi} - C_d g^o_{\xi\xi})_{\xi\xi} - \mu g^o + k f^o = -\rho a q_2(\xi) w^o_t, & \xi \in (0,\ell), \\ f^o(0,t) = f^o(\ell,t) = 0,\ g^o(0,t) = g^o(\ell,t) = 0, & t \geq 0, \\ EIf^o_{\xi\xi}(0,t) - C_d g^o_{\xi\xi}(0,t) = EIf^o_{\xi\xi}(\ell,t) - C_d g^o_{\xi\xi}(\ell,t) = 0, & t \geq 0, \\ f^o(\xi,\tau) = 0,\ g^o(\xi,\tau) = 0, & \xi \in (0,\ell), \end{cases} \tag{FVP}$$

$$\begin{cases} u^o(t) = -\int_0^\ell b(\xi;r^o) g^o(\xi,t)\,d\xi, \\ \int_0^\tau \int_0^\ell u^o(t) b_r(\xi;r^o) g^o(\xi,t)\,d\xi dt = 0. \end{cases} \tag{OPT}$$

3.5.2 Nonlinear Waves

Nonlinear waves occur in many applications, including fluid mechanics, electromagnetism, elasticity, and also relativistic quantum mechanics. Let the wave evolve on a region Ω that is a bounded, open, connected subset of $\mathbb{R}^2$. It is assumed that Ω has a Lipschitz boundary separated into $\partial\Omega = \overline{\Gamma_0 \cup \Gamma_1}$ where $\Gamma_0 \cap \Gamma_1 = \emptyset$ and $\Gamma_0 \neq \emptyset$. Denote by ν the unit outward normal vector field on $\partial\Omega$. Figure 3.2 illustrates the region and the shape of an actuator. Define $\mathbb{K} = L^2(\Omega)$ and let $r(\xi)$, $\xi \in \Omega$, be the actuator shape design. There are many possible choices of admissible shapes. One is

$$K_{ad} = \{r \in C^1(\overline{\Omega}) : \|r\|_{C^1(\overline{\Omega})} \leq 1\}.$$

A nonlinear wave model with initial conditions $w_0(\xi)$ and $v_0(\xi)$ is

$$\begin{cases} \dfrac{\partial^2 w}{\partial t^2}(\xi, t) = \Delta w(\xi, t) + F(w(\xi, t)) + r(\xi)u(t), & (\xi, t) \in \Omega \times (0, \tau), \\ w(\xi, 0) = w_0(\xi),\ \dfrac{\partial w}{\partial t}(\xi, 0) = v_0(\xi), & \xi \in \Omega, \\ w(\xi, t) = 0, & (\xi, t) \in \Gamma_0 \times [0, \tau), \\ \dfrac{\partial w}{\partial \nu}(\xi, t) = 0, & (\xi, t) \in \Gamma_1 \times [0, \tau). \end{cases}$$

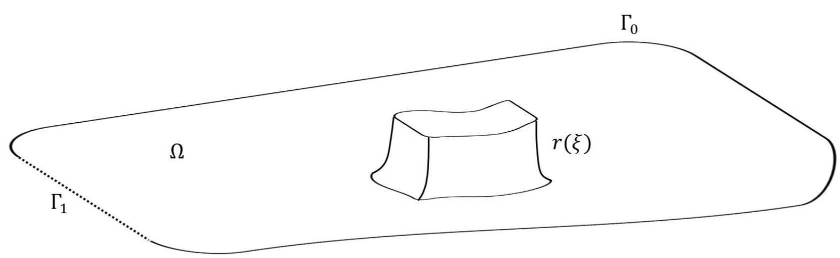

Figure 3.2: Schematic of an actuator on the wave region.

Let $\mathbb{X} = H^1_{\Gamma_0}(\Omega) \times L^2(\Omega)$ and define $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ as

$$\mathcal{A}(w, v) = (v, \Delta w), \tag{3.112}$$

$$D(\mathcal{A}) = \left\{ (w, v) \in \mathbb{X} \,|\, w \in H^2(\Omega) \cap H^1_{\Gamma_0}(\Omega),\ v \in H^1_{\Gamma_0}(\Omega),\ \frac{\partial w}{\partial \nu}\Big|_{\Gamma_1} = 0 \right\}.$$

The operator $\mathcal{A}$ is skew-adjoint and generates a strongly continuous unitary group on $\mathbb{X}$; see for example, [41, Theorem 3.24].

Assumption D.

1. *The function $F(\zeta)$ is twice continuously differentiable over $\mathbb{R}$; denote its derivatives by $F'(\zeta)$ and $F''(\zeta)$.*

2. *There are numbers $a_0 > 0$ and $b > 1/2$ such that $|F''(\zeta)| \leq a_0(1 + |\zeta|^b)$.*

The nonlinear operator $\mathcal{F}(\cdot) : \mathbb{X} \to \mathbb{X}$ is defined as

$$\mathcal{F}(w, v) = (0, F(w)). \tag{3.113}$$

Assumption D is needed to ensure that $\mathcal{F}(\cdot) : \mathbb{X} \to \mathbb{X}$ and satisfies assumption B1 and that the Gâteaux derivative of $\mathcal{F}(\cdot)$ is also an operator on $\mathbb{X}$. Some examples of $F(\cdot)$ satisfying this assumption are $F(w) = \sin(w)$ in the Sine-Gordon equation and $F(w) = |w|^k w,\ k \geq 2$ in the Klein-Gordon equation [109, Section 5.2].

Lemma 3.5.4. *Under assumption D,*

1. *the operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable on $\mathbb{X}$, with the Gâteaux derivative at $\boldsymbol{x} = (w, v)$ in the direction $\boldsymbol{p} = (f, g)$ given by $\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p} = (0, F'(w)f)$,*

2. *the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded, and*

3. *$\mathcal{F}(\cdot)$ satisfies assumption B1.*

Proof. To prove the first part of the lemma, it must be shown that for any variation $f \in H^1_{\Gamma_0}(\Omega)$,

$$\lim_{\epsilon \to 0} \left\| \frac{1}{\epsilon} \left(F(w + \epsilon f) - F(w) \right) - F'(w) f \right\|_{L^2(\Omega)} = 0, \tag{3.114}$$

or

$$\lim_{\epsilon \to 0} \int_\Omega \left| \frac{1}{\epsilon} \left(F(w(\xi) + \epsilon f(\xi)) - F(w(\xi)) \right) - F'(w(\xi)) f(\xi) \right|^2 d\xi = 0. \tag{3.115}$$

Recall that because of the continuous embedding $H^1_{\Gamma_0}(\Omega) \hookrightarrow L^p(\Omega)$, the functions f and w belong to $L^p(\Omega)$ for all $p \in [1, \infty)$. Use of assumption D, applying Taylor's theorem with integral reminder to $F(\cdot)$, and using Jensen's inequality, the integral in (3.115) becomes

$$\begin{aligned}
&\int_\Omega \left(\int_0^1 \epsilon (1 - \eta) F''(w(\xi) + \eta \epsilon f(\xi)) f^2(\xi) d\eta \right)^2 d\xi \\
&\leq \int_\Omega \int_0^1 \epsilon^2 (1 - \eta)^2 F''^2(w(\xi) + \eta \epsilon f(\xi)) f^4(\xi) d\eta d\xi \\
&\leq \int_\Omega \int_0^1 \epsilon^2 (1 - \eta)^2 a_0^2 \left(2 + 2^{2b} |w(\xi)|^{2b} + 2^{2b} \eta^{2b} |\epsilon|^{2b} |f(\xi)|^{2b} \right) f^4(\xi) d\eta d\xi.
\end{aligned} \tag{3.116}$$

Applying Hölder's inequality shows that integral (3.116) is bounded above by a number, and also converges to zero as $\epsilon \to 0$.

Furthermore, the operator $\mathcal{F}'_{\boldsymbol{x}}$ satisfies, for any $\boldsymbol{x} = (w, v)$ and $\boldsymbol{p} = (f, g)$ in $\mathbb{X}$,

$$\|\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}\|^2 = \int_\Omega F'^2(w(\xi))f^2(\xi)d\xi. \tag{3.117}$$

Assumption D2 ensures that there is a number $a_1 > 0$ such that $|F'(\zeta)| \leq a_1(1 + |\zeta|^{b+1})$. Use this together with Hölder's inequality to obtain

$$\begin{aligned}\|\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}\|^2 &\leq \int_\Omega 2a_1^2(1 + w^{2b+2}(\xi))f^2(\xi)d\xi. \\ &\leq 2a_1^2\left(\|f\|^2_{L^2(\Omega)} + \|w\|^{2b+2}_{L^{4b+4}(\Omega)}\|f\|^2_{L^4(\Omega)}\right)\end{aligned} \tag{3.118}$$

Apply the embeddings $H^1_{\Gamma_0}(\Omega) \hookrightarrow L^p(\Omega)$; letting c_p indicates the embedding constants for each $p \in [1, \infty)$,

$$\begin{aligned}\|\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}\|^2 &\leq 2a_1^2\left(c_2^2 + c_{4b+4}^{2b+2}c_4^2\|w\|^{2b}_{H^1_{\Gamma_0}(\Omega)}\right)\|f\|^2_{H^1_{\Gamma_0}(\Omega)} \\ &\leq 2a_1^2\left(c_2^2 + c_{4b+4}^{2b+2}c_4^2\|\boldsymbol{x}\|^{2b}\right)\|\boldsymbol{p}\|^2.\end{aligned} \tag{3.119}$$

Inequality (3.119) implies that

$$\|\mathcal{F}'_{\boldsymbol{x}}\|^2_{\mathcal{L}(\mathbb{X})} \leq 2a_1^2\left(c_2^2 + c_{4b+4}^{2b+2}c_4^2\|\boldsymbol{x}\|^{2b}\right). \tag{3.120}$$

This inequality shows that the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded.

It will now be shown that $\mathcal{F}(\cdot)$ satisfies assumption B1. Consider a bounded sequence $\boldsymbol{x}_n(t) = (w_n(t), v_n(t))$ in $C(0, \tau; \mathbb{X})$ that weakly converges to some element $\boldsymbol{x}(t) = (w(t), v(t))$ in $L^p(0, \tau; \mathbb{X})$. The sequence $w_n(t)$ is bounded in $C(0, \tau; H^1_{\Gamma_0}(\Omega))$ and so it is in $C(0, \tau; L^q(\Omega))$ for all $q \in [1, \infty)$. This together with the bounded convergence theorem ensures that for every $p \in [1, \infty)$

$$\int_0^{\tau-h}\|w_n(t+h) - w_n(t)\|^p_{L^q(\Omega)}\,dt \to 0 \text{ as } h \to 0 \text{ uniformly for all } n. \tag{3.121}$$

The space $H^1_{\Gamma_0}(\Omega)$ is compactly embedded in $L^q(\Omega)$ by Rellich-Kondrachov compact embedding theorem [1, Chapter 6]. By Theorem 3.5.1, this embedding together with (3.121) ensures that $w_n(t)$ has a strongly convergent subsequence in $L^p(0, \tau; L^q(\Omega))$. The sequence

$w_n(t)$ by assumption converges weakly to $w(t)$ in $L^p(0,\tau;H^1_{\Gamma_0}(\Omega))$; a weak limit is unique, so $w_n(t)$ converges strongly to $w(t)$ in $L^p(0,\tau;L^q(\Omega))$. The nonlinear operator $\mathcal{F}(\cdot)$ maps $\boldsymbol{x}_n(t)$ to

$$\mathcal{F}(w_n(t),v_n(t)) = (0, F(w_n(t))). \tag{3.122}$$

Use Taylor's theorem with integral reminder, and let $h(\xi,t;\eta) = w(\xi,t)+\eta(w_n(\xi,t)-w(\xi,t))$, $\eta\in[0,1]$, to obtain

$$|F(w_n(\xi,t)) - F(w(\xi,t))| \le \left(\int_0^1 |F'(h(\xi,t;\eta))|d\eta\right)|w_n(\xi,t)-w(\xi,t)|. \tag{3.123}$$

Let

$$M_1(\xi,t) = a_1\left(1+\int_0^1 |h(\xi,t;\eta)|^{b+1}d\eta\right).$$

Taking integral of both side of (3.123) and using Hölder inequality yield

$$\begin{aligned}\|F(w_n)-F(w)\|^p_{L^p(0,\tau;L^2(\Omega))} &\le \int_0^\tau\left(\int_\Omega M_1^2(\xi,t)|w_n(\xi,t)-w(\xi,t)|^2 d\xi\right)^{\frac{p}{2}}dt\\ &\le \|M_1\|^p_{L^{2p}(0,\tau;L^4(\Omega))}\|w_n-w\|^p_{L^{2p}(0,\tau;L^4(\Omega))}.\end{aligned} \tag{3.124}$$

Note that $\|M_1\|_{L^{2p}(0,\tau;L^4(\Omega))}<\infty$ since $w(\xi,t)$ and $w_n(\xi,t)$ are in $L^p(0,\tau;L^q(\Omega))$ for all $p\in[1,\infty)$ and $q\in[1,\infty)$. From (3.124), it follows that $F(w_n)$ strongly converges to $F(w)$ in $L^p(0,\tau;L^2(\Omega))$. Therefore, the sequence $\mathcal{F}(w_n(t),v_n(t))$ strongly (and so weakly) converges to $\mathcal{F}(w(t),v(t))$ in $L^p(0,\tau;\mathbb{X})$. □

The first part of Lemma 3.5.4 implies that the nonlinear operator $\mathcal{F}(\cdot)$ satisfies assumption A2. Therefore, Theorem 2.4.2 implies that for every initial condition $\boldsymbol{x}_0\in\mathbb{X}$ and positive number R, there exists $\tau>0$ such that for all $u\in L^2(0,\tau)$, $\|u\|_2\le R$, the nonlinear wave equation has a unique mild solution $\boldsymbol{x}\in C(0,\tau;\mathbb{X})$. The existence of a solution over an infinite time interval when damping is present has been discussed in the literature; see for example, [51, 52, 53].

Let $\boldsymbol{x}^o(t) = (w^o,v^o)$ at time $t\in[0,\tau]$. As for the railway track example, in order to obtain an expression for the adjoint of the operator $\mathcal{F}'_{\boldsymbol{x}^o(t)}$, the following boundary-value problem needs to be solved:

$$\begin{cases}\Delta h(\xi) = -F'(w^o(\xi))g(\xi), & \xi\in\Omega,\\ h(\xi)=0, & \xi\in\Gamma_0,\\ \dfrac{\partial h}{\partial\nu}(\xi)=0, & \xi\in\Gamma_1.\end{cases} \tag{3.125}$$

The adjoint with respect to $\mathbb{X}$ of $\mathcal{F}'_{\boldsymbol{x}^o(t)}$ is

$$\mathcal{F}'^{*}_{\boldsymbol{x}^o(t)}(f,g) = (h,0), \tag{3.126}$$

where h solves (3.125). Define $\mathbb{U} = \mathbb{R}$ and the input operator $\mathcal{B}(r) \in \mathcal{L}(\mathbb{U}, \mathbb{X})$ by

$$\mathcal{B}(r)u = (0, r(\xi)u). \tag{3.127}$$

The adjoint of this operator is

$$\mathcal{B}^*(r)(f,g) = \int_\Omega r(\xi)g(\xi)d\xi, \quad \forall (f,g) \in \mathbb{X}. \tag{3.128}$$

Let $\boldsymbol{p}^o(t) = (f,g)$ at time $t \in [0,\tau]$. Use Corollary 3.4.9 to find

$$(\mathcal{B}'_r u)^* \boldsymbol{p}^o(t) = ug. \tag{3.129}$$

Furthermore, let $q_1 \in C^1(\overline{\Omega})$ and $q_2 \in C(\overline{\Omega})$ be some non-negative functions. Set $\mathcal{Q}(w,v) = (q_1 w, q_2 v)$ and $\mathcal{R} = 1$ in the cost function of assumption C3.

If the optimal control u^o and optimal actuator design r^o are in the interior of $U_{ad} \times K_{ad}$, then by Corollary 3.4.9 the following equations are satisfied:

$$\begin{cases} \dfrac{\partial^2 w^o}{\partial t^2}(\xi,t) = \Delta w^o(\xi,t) + F(w^o(\xi,t)) + r^o(\xi)u^o(t), & (\xi,t) \in \Omega \times (0,\tau], \\ w^o(\xi,0) = w_0(\xi), \ \dfrac{\partial w^o}{\partial t}(\xi,0) = v_0(\xi), & \xi \in \Omega, \\ w^o(\xi,t) = 0, & (\xi,t) \in \Gamma_0 \times [0,\tau], \\ \dfrac{\partial w^o}{\partial \nu}(\xi,t) = 0, & (\xi,t) \in \Gamma_1 \times [0,\tau], \end{cases} \tag{IVP}$$

$$\begin{cases} \dfrac{\partial f^o}{\partial t}(\xi,t) = -g^o(\xi,t) - h^o(\xi,t) - q_1(\xi)w^o(\xi,t), & (\xi,t) \in \Omega \times (0,\tau], \\ \dfrac{\partial g^o}{\partial t}(\xi,t) = -\Delta f^o(\xi,t) - q_2(\xi)\dfrac{\partial w^o}{\partial t}(\xi,t), & (\xi,t) \in \Omega \times (0,\tau], \\ f^o(\xi,\tau) = 0, \ g^o(\xi,\tau) = 0, & \xi \in \Omega, \\ f^o(\xi,t) = 0, & (\xi,t) \in \Gamma_0 \times [0,\tau], \\ \dfrac{\partial f^o}{\partial \nu}(\xi,t) = 0, & (\xi,t) \in \Gamma_1 \times [0,\tau], \end{cases} \tag{FVP}$$

$$\begin{cases} u^o(t) = -\displaystyle\int_\Omega r^o(\xi)g^o(\xi,t)d\xi, & t \in [0,\tau], \\ \displaystyle\int_0^\tau u^o(t)g^o(\xi,t)dt = 0, & \xi \in \Omega. \end{cases} \tag{OPT}$$

3.6 Concluding Remarks

Optimal control of semilinear infinite-dimensional systems was considered in this chapter where the optimal controller design involves both the controlled input and the actuator design. It was shown that the existence of an optimal control together with an optimal actuator design is guaranteed under some assumptions. Moreover, first-order necessary optimality conditions were obtained. The theory was illustrated with several applications.

Future work is concerned with developing numerical methods for solution of the optimality equations and also the consideration of a wider class of nonlinearities. Extension of these problems to situations where the input operator is not bounded on the state space is also of interest.

Chapter 4

Optimal Actuator Design for Nonlinear Parabolic Systems

Many physical systems are modeled by nonlinear parabolic differential equations. One example is the Kuramoto-Sivashinsky (KS) equation. In this chapter, the existence of a concurrent optimal controller and actuator design is established for nonlinear parabolic systems. Optimality equations are provided. Also, unlike most work on the optimal control of nonlinear systems, this chapter includes the effect of the worst initial condition on the cost.

The results of this study are also applied to Kuramoto-Sivashinsky (KS) equation. This equation was derived by Kuramoto to model angular phase turbulence in reaction–diffusion systems [73], and by Sivashinsky for modeling plane flame propagation [111]. It also models film layer flow on an inclined plane [25], directional solidification of dilute binary alloys [102], growth and saturation of the potential of dissipative trapped-ion [76], and terrace edge evolution during step-flow growth [10]. From system theoretic perspective, Christofides and Armaou studied the global stabilization of KS equation using distributed output feedback control [22]. Lou and Christofides investigated the optimal actuator/sensor placement for control of KS equation by approximating the model with a finite dimensional system [83]. Gomes et al. also studied the actuator placement problem for KS equation using numerical algorithms [54]. Controllability of KS equation has also been studied [18, 19]. Optimal control of KS equation using maximum principle was studied in [113]. Optimal control of KS equation with point-wise state and mixed control-state constraints was studied in [50]. Liu and Krstic studied boundary control of KS equation in [82]. Al Jamal and Morris studied the relationship between stability and stabilization of linearized and nonlinear KS equation [3].

Previous research on optimal control of PDE's, such as [15, 107], has focused on partial differential equations with certain structures. Optimal control of differential equations in abstract spaces has rarely been discussed [89]. This chapter extends previous results to abstract differential equations without an assumption of stability.

The chapter is organized as follows. Section 4.1 introduces the nonlinear abstract parabolic equations, and discusses the conditions for existence of a solution. Section 4.2 discusses the existence of an optimal input together with an optimal actuator design. Examples are presented in Section 4.3 which includes Kuramoto-Sivashinsky equation and nonlinear heat equation. Concluding remarks are made in Section 4.4.

4.1 Nonlinear Parabolic Systems

Let $\boldsymbol{x}(t)$ and $\boldsymbol{u}(t)$ be the state and input taking values in reflexive Banach spaces $\mathbb{X}$ and $\mathbb{U}$, respectively. Also, let $\boldsymbol{r}$ denote the actuator design parameter that takes value in a compact set K_{ad} of a topological space $\mathbb{K}$. Consider the following initial value problem (IVP):

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{4.1}$$

Definition 4.1.1. *The operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is said to have maximal L^p regularity if for every $\boldsymbol{f} \in L^p(0,\tau;\mathbb{X})$, $1 < p < \infty$, the equation*

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \boldsymbol{f}(t), \quad t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0, \end{cases} \tag{4.2}$$

admits a unique solution in $\mathbb{W}(0,\tau)$ that satisfies (4.2) almost everywhere on $[0,\tau]$.

The linear operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is assumed to have maximal L^p regularity. In particular, if $\mathcal{A}$ is associated with a sesquilinear form that is bounded and coercive with respect to $\mathbb{V} \hookrightarrow \mathbb{X}$, it generates an analytic semigroup on $\mathbb{X}$ [109, Lemma 36.5 and Theorem 36.6].

Let $\mathcal{A}$ be the generator of an analytic semigroup $e^{\mathcal{A}t}$ on $\mathbb{X}$. For every $p \in [1,\infty]$ and $\alpha \in (0,1)$, the interpolation space $D_{\mathcal{A}}(\alpha,p)$ is defined as the set of all $\boldsymbol{x}_0 \in \mathbb{X}$ such that the function

$$t \mapsto v(t) := \left\| t^{1-\alpha-1/p} \mathcal{A} e^{t\mathcal{A}} \boldsymbol{x}_0 \right\| \tag{4.3}$$

belongs to $L^p(0,1)$ [84, Section 2.2.1]. The norm on this space is

$$\|\boldsymbol{x}_0\|_{D_{\mathcal{A}}(\alpha,p)} = \|\boldsymbol{x}_0\| + \|v\|_{L^p(0,1)} .$$

The Banach space $\mathbb{W}(0,\tau)$ is the set of all $\boldsymbol{x}(\cdot) \in W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau;D(\mathcal{A}))$ with norm [11, Section II.2]

$$\|\boldsymbol{x}\|_{\mathbb{W}(0,\tau)} = \|\dot{\boldsymbol{x}}\|_{L^p(0,\tau;\mathbb{X})} + \|\mathcal{A}\boldsymbol{x}\|_{L^p(0,\tau;\mathbb{X})} .$$

The nonlinear operator $\mathcal{F}(\cdot)$ maps a reflexive Banach space $\mathbb{V}$ to $\mathbb{X}$ where $D_{\mathcal{A}}(1/p,p) \hookrightarrow \mathbb{V} \hookrightarrow \mathbb{X}$. The operator $\mathcal{F}(\cdot)$ is locally Lipschitz continuous; that is, for every bounded set D in $\mathbb{V}$, there is a positive number $L_{\mathcal{F}}$ such that

$$\|\mathcal{F}(\boldsymbol{x}_2) - \mathcal{F}(\boldsymbol{x}_1)\|_{\mathbb{X}} \le L_{\mathcal{F}} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_{\mathbb{V}} ,\ \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D. \tag{4.4}$$

When there is no ambiguity, the norm on $\mathbb{X}$ will not be explicitly indicated.

For each $\boldsymbol{r} \in \mathbb{K}$, the input operator $\mathcal{B}(\boldsymbol{r})$ is a linear bounded operator that maps the input space $\mathbb{U}$ into the state space $\mathbb{X}$ and it is continuous with respect to $\boldsymbol{r}$:

$$\lim_{\boldsymbol{r}_n \to \boldsymbol{r}_0} \|\mathcal{B}(\boldsymbol{r}_n) - \mathcal{B}(\boldsymbol{r}_0)\| = 0, \tag{4.5}$$

where the convergence $\boldsymbol{r}_n \to \boldsymbol{r}_0$ is with respect to the topology on $\mathbb{K}$.

For any positive numbers R_1 and R_2, define the sets

$$B_{L^p(0,\tau;\mathbb{U})}(R_1) = \left\{ \boldsymbol{u} \in L^p(0,\tau;\mathbb{U}) : \|\boldsymbol{u}\|_p \le R_1 \right\}, \tag{4.6}$$

$$B_{\mathbb{V}}(R_2) = \{\boldsymbol{x}_0 \in \mathbb{V} : \|\boldsymbol{x}_0\|_{\mathbb{V}} \le R_2\} . \tag{4.7}$$

Definition 4.1.2. *[11, Definition 3.1.i](strict solution) The function $\boldsymbol{x}(\cdot)$ is said to be a* strict solution *of (4.1) if $\boldsymbol{x}(0) = \boldsymbol{x}_0$, $\boldsymbol{x} \in \mathbb{W}(0,\tau)$, and $\boldsymbol{x}(t)$ satisfies (4.1) for almost every $t \in [0,\tau]$.*

Lemma 4.1.3. *[24, Proposition 2.2 and Corollary 2.3] Let $\tau_0 > \tau$ and $p \in (1,\infty)$ be given. If $\mathcal{A}$ has maximal L^p regularity, then there exists a constant c_{τ_0} independent of τ such that for all $\tau \in (0,\tau_0]$ and $\boldsymbol{v} \in W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau;D(\mathcal{A}))$,*

$$\|\dot{\boldsymbol{v}}\|_{L^2(0,\tau;\mathbb{X})} + \|\mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})} \le M_{\tau_0} \left(\|\dot{\boldsymbol{v}} + \mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})} + \|\boldsymbol{v}(0)\|_{D_{\mathcal{A}}(1/p,p)} \right) .$$

Furthermore, if $\boldsymbol{v}(0) = 0$,

$$\|\boldsymbol{v}\|_{C(0,\tau;D_{\mathcal{A}}(1/p,p))} \le M_{\tau_0} \left(\|\dot{\boldsymbol{v}}\|_{L^2(0,\tau;\mathbb{X})} + \|\mathcal{A}\boldsymbol{v}\|_{L^2(0,\tau;\mathbb{X})} \right) .$$

Theorem 2.4.6 ensures that for every pair $R_1 > 0$, $R_2 > 0$, there is $\tau > 0$ and $\delta > 0$ such that the IVP (4.1) admits a unique strict solution $\boldsymbol{x} \in \mathbb{W}(0,\tau)$, $\|\boldsymbol{x}\|_{\mathbb{W}(0,\tau)} \le \delta$ for all $(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0) \in B_{L^p(0,\tau;\mathbb{U})}(R_1) \times K_{ad} \times B_{\mathbb{V}}(R_2)$.

Definition 4.1.4. *Let $\boldsymbol{x}(t)$ be the strict solution to (4.1). The mapping $\mathcal{S}(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0)$: $B_{L^p(0,\tau;\mathbb{U})}(R_1) \times K_{ad} \times B_{\mathbb{V}}(R_2) \to \mathbb{W}(0,\tau)$, $(\boldsymbol{u}(t), \boldsymbol{r}, \boldsymbol{x}_0) \mapsto \boldsymbol{x}(t)$, is called the solution map.*

An embedding $D(\mathcal{A}) \hookrightarrow \mathbb{X}$ where $D(\mathcal{A})$ is compact in $\mathbb{X}$ ensures that the space $W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau, D(\mathcal{A}))$ is compactly embedded in $c^s(0,\tau;\mathbb{V})$, $0 \le s < 1$ [4, Theorem 5.2]. Since $c^s(0,\tau;\mathbb{V}) \hookrightarrow C(0,\tau;\mathbb{V})$, it follows that the space $W^{1,p}(0,\tau;\mathbb{X}) \cap L^p(0,\tau, D(\mathcal{A}))$ is compactly embedded in $C(0,\tau;\mathbb{V})$.

Theorem 4.1.5. *If the embedding $D(\mathcal{A}) \hookrightarrow \mathbb{X}$ is compact then the solution map is weakly continuous in $(\boldsymbol{u}(t),\ \boldsymbol{r},\ \boldsymbol{x}_0)$.*

Proof. The weak continuity of the solution map with respect to $\boldsymbol{u}(t)$ is shown in [89, Lemma 2.12]. Weak continuity with respect to $(\boldsymbol{u}(t), \boldsymbol{r}, \boldsymbol{x}_0)$ follows from a similar proof. Choose any weakly convergent sequences $\{\boldsymbol{u}_n(t)\} \subset L^p(0,\tau;\mathbb{U})$, $\{\boldsymbol{x}_0^n\} \subset \mathbb{V}$, and $\{\boldsymbol{r}_n\} \subset \mathbb{K}$. Since sets $B_{L^p(0,\tau;\mathbb{U})}(R_1)$ and $B_{\mathbb{V}}(R_2)$ are bounded, closed, convex subsets of Banach spaces $L^p(0,\tau;\mathbb{U})$ and $\mathbb{V}$, respectively; these sets are weakly closed [115, Theorem 2.11]. This implies that there are $\boldsymbol{u}^o \in B_{L^p(0,\tau;\mathbb{U})}(R_1)$ and $\boldsymbol{x}_0 \in B_{\mathbb{V}}(R_2)$ such that

$$\boldsymbol{u}_n \rightharpoonup \boldsymbol{u}^o \text{ in } B_{L^p(0,\tau;\mathbb{U})}(R_1), \tag{4.8}$$

$$\boldsymbol{x}_0^n \rightharpoonup \boldsymbol{x}_0 \text{ in } B_{\mathbb{V}}(R_2). \tag{4.9}$$

Since the set K_{ad} is a compact subset of $\mathbb{K}$

$$\boldsymbol{r}_n \to \boldsymbol{r}^o \text{ in } K_{ad}. \tag{4.10}$$

It will be shown that $\mathcal{B}_{\boldsymbol{r}_n}\boldsymbol{u}_n(t)$ converges weakly to $\mathcal{B}_{\boldsymbol{r}^o}\boldsymbol{u}^o(t)$ in $L^p(0,\tau;\mathbb{X})$. For every $\boldsymbol{z} \in L^q(0,\tau;\mathbb{X})$, $1/q = 1 - 1/p$,

$$\begin{aligned} I :=& \langle \boldsymbol{z}, \mathcal{B}_{\boldsymbol{r}_n}\boldsymbol{u}_n - \mathcal{B}_{\boldsymbol{r}^o}\boldsymbol{u}^o \rangle_{L^q(0,\tau;\mathbb{X}^*),L^p(0,\tau;\mathbb{X})} \\ =& \langle \boldsymbol{z}, \mathcal{B}_{\boldsymbol{r}_n}\boldsymbol{u}_n - \mathcal{B}_{\boldsymbol{r}^o}\boldsymbol{u}_n \rangle_{L^q(0,\tau;\mathbb{X}^*),L^p(0,\tau;\mathbb{X})} \\ &+ \langle \boldsymbol{z}, \mathcal{B}_{\boldsymbol{r}^o}\boldsymbol{u}_n - \mathcal{B}_{\boldsymbol{r}^o}\boldsymbol{u}^o \rangle_{L^q(0,\tau;\mathbb{X}^*),L^p(0,\tau;\mathbb{X})}. \end{aligned} \tag{4.11}$$

Taking the adjoint and norm yield

$$\begin{aligned} I \le& \|\mathcal{B}_{\boldsymbol{r}_n} - \mathcal{B}_{\boldsymbol{r}^o}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \int_0^\tau \|\boldsymbol{u}_n(t)\|_{\mathbb{U}} \|\boldsymbol{z}(t)\|\, dt \\ &+ \Big| \int_0^\tau \langle \mathcal{B}_{\boldsymbol{r}^o}^* \boldsymbol{z}(t), \boldsymbol{u}_n(t) - \boldsymbol{u}^o(t) \rangle_{\mathbb{U}^*,\mathbb{U}}\, dt \Big|. \end{aligned}$$

Use Hölder inequality and let $\boldsymbol{v}(t) = \mathcal{B}_{\boldsymbol{r}^o}^* \boldsymbol{z}(t)$, it follows that

$$\begin{aligned} I \leq \|\mathcal{B}_{\boldsymbol{r}_n} - \mathcal{B}_{\boldsymbol{r}^o}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \|\boldsymbol{u}_n\|_{L^p(0,\tau;\mathbb{U})} \|\boldsymbol{z}\|_{L^q(0,\tau;\mathbb{X})} \\ + | \langle \boldsymbol{u}_n - \boldsymbol{u}^o, \boldsymbol{v} \rangle_{L^p(0,\tau;\mathbb{U}),L^q(0,\tau;\mathbb{U}^*)} |. \end{aligned}$$

The convergence of the first term follows from (4.5). The second term converges to zero because $\boldsymbol{u}_n \rightharpoonup \boldsymbol{u}^o$ in $L^p(0,\tau;\mathbb{U})$. Combining these yields

$$\mathcal{B}_{\boldsymbol{r}_n} \boldsymbol{u}_n \rightharpoonup \mathcal{B}_{\boldsymbol{r}^o} \boldsymbol{u}^o \text{ in } L^p(0,\tau;\mathbb{X}). \tag{4.12}$$

Using Theorem 2.4.6, the corresponding solution $\boldsymbol{x}_n(t)$ is a bounded sequence in the reflexive Banach space $L^p(0,\tau;D(\mathcal{A})) \cap W^{1,p}(0,\tau;\mathbb{X})$. Thus, there is a subsequence of $\boldsymbol{x}_n(t)$ such that

$$\boldsymbol{x}_{n_k} \rightharpoonup \boldsymbol{x} \text{ in } \mathbb{W}(0,\tau). \tag{4.13}$$

This in turn implies that the sequence $\boldsymbol{x}_n(t)$ strongly converges to $\boldsymbol{x}(t)$ in $C(0,\tau;\mathbb{V})$. This together with Lipschitz continuity of $\mathcal{F}(\cdot)$ yields

$$\mathcal{F}(\boldsymbol{x}_{n_k}(t)) \to \mathcal{F}(\boldsymbol{x}(t)) \quad \text{in } L^p(0,\tau;\mathbb{X}). \tag{4.14}$$

This strong convergence also yields weak convergence in the same space, that is

$$\mathcal{F}(\boldsymbol{x}_{n_k}(t)) \rightharpoonup \mathcal{F}(\boldsymbol{x}(t)) \quad \text{in } L^p(0,\tau;\mathbb{X}). \tag{4.15}$$

Now apply (4.8), (4.12), (4.13), and (4.15) to the IVP (4.1); take the limit; notice that a solution to the IVP is unique; it follows that $\boldsymbol{x} = \mathcal{S}(\boldsymbol{u},\boldsymbol{r},\boldsymbol{x}_0)$. Deleting elements $\{\boldsymbol{x}_{n_k}(t)\}$ from $\{\boldsymbol{x}_n(t)\}$ and repeating the previous processing, knowing that a weak limit is unique, it follow that $\boldsymbol{x}_n(t) \rightharpoonup \boldsymbol{x}(t)$ in $\mathbb{W}(0,\tau)$. □

4.2 Optimal Actuator Design

Consider a cost function $J(\boldsymbol{x},\boldsymbol{u},\boldsymbol{r}) : \mathbb{W}(0,\tau) \times L^p(0,\tau;\mathbb{U}) \times \mathbb{K} \to \mathbb{R}$ that is bounded below and weakly lower-semicontinuous with respect to $\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{r}$. For a fixed initial condition $\boldsymbol{x}_0 \in B_{\mathbb{V}}(R_2)$, consider the following optimization problem over the admissible input set U_{ad} and actuator design set K_{ad}

$$\begin{cases} \min & J(\boldsymbol{x},\boldsymbol{u},\boldsymbol{r}) \\ \text{s.t.} & \boldsymbol{x} = \mathcal{S}(\boldsymbol{u},\boldsymbol{r},\boldsymbol{x}_0), \\ & (\boldsymbol{u},\boldsymbol{r}) \in U_{ad} \times K_{ad}. \end{cases} \tag{P}$$

The set U_{ad} will be assumed a convex and closed subset of $B_{L^p(0,\tau;\mathbb{U})}(R_1) \backslash \partial B_{L^p(0,\tau;\mathbb{U})}(R_1)$.

Theorem 4.2.1. *For every $\boldsymbol{x}_0 \in B_{\mathbb{V}}(R_2)$, there exists a control input $\boldsymbol{u}^o \in U_{ad}$ together with an actuator design $\boldsymbol{r}^o \in K_{ad}$ that solve the optimization problem P.*

Proof. The proof of this theorem follows from standard analysis; see for example, [61, Theorem 1.45] and [37, Theorem 4.1] for a similar argument. Define

$$j(\boldsymbol{x}_0) := \inf_{(\boldsymbol{u},\boldsymbol{r})\in U_{ad}\times K_{ad}} J(\mathcal{S}(\boldsymbol{u},\boldsymbol{r},\boldsymbol{x}_0),\boldsymbol{u},\boldsymbol{r}). \tag{4.16}$$

and let $(\boldsymbol{u}_n,\boldsymbol{r}_n)$ be the minimizing sequence:

$$\lim_{n\to\infty} J(\mathcal{S}(\boldsymbol{u}_n,\boldsymbol{r}_n,\boldsymbol{x}_0),\boldsymbol{u}_n,\boldsymbol{r}_n) = j(\boldsymbol{x}_0). \tag{4.17}$$

The set U_{ad} is closed and convex in the reflexive Banach space $L^p(0,\tau;\mathbb{U})$, so it is weakly closed. This implies that there is a subsequence of $\boldsymbol{u}_n$, denote it by the same symbol, that converges weakly to some elements $\boldsymbol{u}^o$ in U_{ad}. Because of compactness of K_{ad}, there is also a subsequence of $\boldsymbol{r}_n$, denote it by the same symbol, that strongly converges to $\boldsymbol{r}^o$. Theorem 2.4.6 and Theorem 4.1.5 state that the solution map is bounded and weakly continuous in each variable. Thus, the corresponding state $\boldsymbol{x}_n = \mathcal{S}(\boldsymbol{u}_n,\boldsymbol{r}_n;\boldsymbol{x}_0)$ also weakly converges to $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}^o,\boldsymbol{r}^o;\boldsymbol{x}_0)$ in $\mathbb{W}(0,\tau)$. The cost function is weakly lower semi-continuous with respect to each $\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{r}$, this ensures that $(\boldsymbol{x}^o,\boldsymbol{u}^o,\boldsymbol{r}^o)$ minimizes the cost function. Therefore, $(\boldsymbol{u}^o,\boldsymbol{r}^o)$ is a solution to the optimization problem P. □

Definition 4.2.2. *[61, Definition 1.29] The operator $\mathcal{G}:\mathbb{X}\to\mathbb{Y}$ is said to be Gâteaux differentiable at $\boldsymbol{x}\in\mathbb{X}$ in the direction $\boldsymbol{p}\in\mathbb{X}$, if there is a linear bounded operator $\mathcal{G}'_{\boldsymbol{x}}$ such that for all real ϵ*

$$\lim_{\epsilon\to 0} \|\mathcal{G}(\boldsymbol{x}+\epsilon\boldsymbol{p}) - \mathcal{G}(\boldsymbol{x}) - \epsilon\mathcal{G}'_{\boldsymbol{x}}\boldsymbol{p}\|_{\mathbb{Y}} = 0. \tag{4.18}$$

The optimality conditions are derived next after assuming that the problem has certain properties. Consider the assumptions:

A1. The spaces $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$ are Hilbert spaces, and $p=2$.

A2. Let $a:\mathbb{V}\times\mathbb{V}\to\mathbb{C}$ be a sesquilinear form (see [75, Chapter 4]), where $\mathbb{V}\hookrightarrow\mathbb{X}$, and let there be positive numbers α and β such that

$$|a(\boldsymbol{x}_1,\boldsymbol{x}_2)| \le \alpha\,\|\boldsymbol{x}_1\|_{\mathbb{V}}\,\|\boldsymbol{x}_2\|_{\mathbb{V}}\,, \qquad \forall \boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{V},$$
$$\operatorname{Re}\, a(\boldsymbol{x},\boldsymbol{x}) \ge \beta\,\|\boldsymbol{x}\|_{\mathbb{V}}^2\,, \qquad \forall \boldsymbol{x}\in\mathbb{V}.$$

The operator $\mathcal{A}$ has an extension to $\bar{\mathcal{A}} \in \mathcal{L}(\mathbb{V}, \mathbb{V}^*)$ described by

$$\langle \bar{\mathcal{A}} \boldsymbol{v}, \boldsymbol{w} \rangle_{\mathbb{V}^*, \mathbb{V}} = a(\boldsymbol{v}, \boldsymbol{w}), \quad \forall \boldsymbol{v}, \boldsymbol{w} \in \mathbb{V}, \tag{4.19}$$

where $\mathbb{V}^*$ denotes the dual of $\mathbb{V}$ with respect to pivot space $\mathbb{X}$.

A3. The cost function $J(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{r})$ is continuously Fréchet differentiable with respect to each variable.

A4. The nonlinear operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable. Indicate the Gâteaux derivative of $\mathcal{F}(\cdot)$ at $\boldsymbol{x}$ in the direction $\boldsymbol{p}$ by $\mathcal{F}'_{\boldsymbol{x}}\boldsymbol{p}$. Furthermore, the mapping $\boldsymbol{x} \mapsto \mathcal{F}'_{\boldsymbol{x}}$ is bounded; that is, bounded sets in $\mathbb{V}$ are mapped to bounded sets in $\mathcal{L}(\mathbb{V}, \mathbb{X})$.

A5. The control operator $\mathcal{B}(\boldsymbol{r})$ is Gâteaux differentiable with respect to $\boldsymbol{r}$ from K_{ad} to $\mathcal{L}(\mathbb{U}, \mathbb{X})$. Indicate the Gâteaux derivative of $\mathcal{B}(\boldsymbol{r})$ at $\boldsymbol{r}^o$ in the direction $\boldsymbol{r}$ by $\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{r}$. Furthermore, the mapping $\boldsymbol{r}^o \mapsto \mathcal{B}'_{\boldsymbol{r}^o}$ is bounded; that is, bounded sets in $\mathbb{K}$ are mapped to bounded sets in $\mathcal{L}(\mathbb{K}, \mathcal{L}(\mathbb{U}, \mathbb{X}))$.

Using these assumptions, the Gâteaux derivative of the solution map with respect to a trajectory $\boldsymbol{x}(t) = \mathcal{S}(\boldsymbol{u}(t), \boldsymbol{r}, \boldsymbol{x}_0)$ is calculated. The resulting map is a time-varying linear IVP. Let $\boldsymbol{g} \in L^p(0, \tau; \mathbb{X})$, consider the time-varying system

$$\begin{cases} \dot{\boldsymbol{h}}(t) = (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}(t)})\boldsymbol{h}(t) + \boldsymbol{g}(t), \\ \boldsymbol{h}(0) = 0. \end{cases} \tag{4.20}$$

Lemma 4.2.3. *[32, Corollary 5.2] Let assumptions A1 and A2 hold. Also, let $\mathcal{P}(\cdot) : [0, \tau] \to \mathcal{L}(\mathbb{V}, \mathbb{X})$ be such that $\mathcal{P}(\cdot)\boldsymbol{x}$ is weakly measurable for all $\boldsymbol{x} \in \mathbb{V}$, and there exists an integrable function $h : [0, \tau] \to [0, \infty)$ such that $\|\mathcal{P}(t)\|_{\mathcal{L}(\mathbb{V},\mathbb{X})} \leq h(t)$ for all $t \in [0, \tau]$. Then for every $\boldsymbol{x}_0 \in \mathbb{V}$ and $\boldsymbol{g} \in L^2(0, \tau; \mathbb{X})$, there exists a unique $\boldsymbol{x}$ in $\mathbb{W}(0, \tau)$ such that*

$$\begin{cases} \dot{\boldsymbol{x}}(t) = (\mathcal{A} + \mathcal{P}(t))\boldsymbol{x}(t) + \boldsymbol{g}(t), \\ \boldsymbol{x}(0) = \boldsymbol{x}_0. \end{cases} \tag{4.21}$$

Moreover, there exists a constant $c > 0$ independent of $\boldsymbol{x}_0$ and $\boldsymbol{g}(t)$ such that

$$\|\boldsymbol{x}\|^2_{\mathbb{W}(0,\tau)} \leq c \left(\|\boldsymbol{g}\|^2_{L^2(0,\tau;\mathbb{X})} + \|\boldsymbol{x}_0\|^2_{\mathbb{V}} \right). \tag{4.22}$$

Since $\mathbb{W}(0,\tau)$ is embedded in $C(0,\tau;\mathbb{V})$, the state $\boldsymbol{x}(t)$ is bounded in $\mathbb{V}$ for all $t \in [0,\tau]$. This together with Gâteaux differentiablity of $\mathcal{F}(\cdot)$ ensures that there is a positive number $M_{\mathcal{F}}$ such that

$$\sup_{t\in[0,\tau]} \left\|\mathcal{F}'_{\boldsymbol{x}(t)}\right\|_{\mathcal{L}(\mathbb{V},\mathbb{X})} \leq M_{\mathcal{F}}. \tag{4.23}$$

Thus, replacing the operator $\mathcal{P}(t)$ with $\mathcal{F}'_{\boldsymbol{x}(t)}$ and noting that

$$\|\mathcal{P}(t)\|_{\mathcal{L}(\mathbb{V},\mathbb{X})} \leq M_{\mathcal{F}}, \tag{4.24}$$

shows that the conditions of Lemma 4.2.3 hold. Thus, there is a positive number c independent of $\boldsymbol{g}$ such that

$$\|\boldsymbol{h}\|_{\mathbb{W}(0,\tau)} \leq c\,\|\boldsymbol{g}\|_{L^2(0,\tau;\mathbb{X})}\,. \tag{4.25}$$

Proposition 4.2.4. *Under assumption A, the solution map $\mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r};\boldsymbol{x}_0)$ is Gâteaux differentiable with respect to each $\boldsymbol{u}(t)$ and $\boldsymbol{r}$ in $U_{ad}\times K_{ad}$. Let $\boldsymbol{x}(t) = \mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r},\boldsymbol{x}_0)$.*

a. The Gâteaux derivative of $\mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r};\boldsymbol{x}_0)$ at $\boldsymbol{r}$ in the direction $\tilde{\boldsymbol{r}}$ is the mapping $\mathcal{S}'_{\boldsymbol{r}}$: $\mathbb{K} \to L^2(0,\tau;D(\mathcal{A}))\cap W^{1,2}(0,\tau;\mathbb{X})$, $\tilde{\boldsymbol{r}} \mapsto \boldsymbol{z}(t)$, where $\boldsymbol{z}(t)$ is the strict solution to

$$\begin{cases}\dot{\boldsymbol{z}}(t) = (\mathcal{A}+\mathcal{F}'_{\boldsymbol{x}(t)})\boldsymbol{z}(t) + (\mathcal{B}'_{\boldsymbol{r}}\tilde{\boldsymbol{r}})\boldsymbol{u}(t),\\ \boldsymbol{z}(0) = 0.\end{cases} \tag{4.26}$$

b. The Gâteaux derivative of $\mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r};\boldsymbol{x}_0)$ at $\boldsymbol{u}(t)$ in the direction $\tilde{\boldsymbol{u}}(t)$ is the mapping $\mathcal{S}'_{\boldsymbol{u}}$: $L^2(0,\tau;\mathbb{U}) \to L^2(0,\tau;D(\mathcal{A}))\cap W^{1,2}(0,\tau;\mathbb{X})$, $\tilde{\boldsymbol{u}}(t) \mapsto \boldsymbol{h}(t)$, where $\boldsymbol{h}(t)$ is the strict solution to

$$\begin{cases}\dot{\boldsymbol{h}}(t) = (\mathcal{A}+\mathcal{F}'_{\boldsymbol{x}(t)})\boldsymbol{h}(t) + \mathcal{B}(\boldsymbol{r})\tilde{\boldsymbol{u}}(t),\\ \boldsymbol{h}(0) = 0.\end{cases} \tag{4.27}$$

Proof. a) Let ϵ be sufficiently small such that $\boldsymbol{r}+\epsilon\tilde{\boldsymbol{r}} \in K_{ad}$. Define $\boldsymbol{x}_\epsilon(t) = \mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r}+\epsilon\tilde{\boldsymbol{r}},\boldsymbol{x}_0)$, this state solves

$$\begin{cases}\dot{\boldsymbol{x}}_\epsilon(t) = \mathcal{A}\boldsymbol{x}_\epsilon(t) + \mathcal{F}(\boldsymbol{x}_\epsilon(t)) + \mathcal{B}(\boldsymbol{r}+\epsilon\tilde{\boldsymbol{r}})\boldsymbol{u}(t), \quad t>0,\\ \boldsymbol{x}_\epsilon(0) = \boldsymbol{x}_0.\end{cases} \tag{4.28}$$

Similarly, $\boldsymbol{x}(t) = \mathcal{S}(\boldsymbol{u}(t),\boldsymbol{r},\boldsymbol{x}_0)$ solves (4.43) with $\epsilon = 0$. Define $\boldsymbol{e}_{\mathcal{F}}(t)$ and $\boldsymbol{e}_{\mathcal{B}}$ as

$$\boldsymbol{e}_{\mathcal{F}}(t) \coloneqq \frac{1}{\epsilon}\left(\mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}_\epsilon(t)) - \mathcal{F}'_{\boldsymbol{x}(t)}(\boldsymbol{x}(t)-\boldsymbol{x}_\epsilon(t))\right), \tag{4.29a}$$

$$\boldsymbol{e}_{\mathcal{B}} \coloneqq \frac{1}{\epsilon}\left(\mathcal{B}(\boldsymbol{r}+\epsilon\tilde{\boldsymbol{r}}) - \mathcal{B}(\boldsymbol{r})\right) - \mathcal{B}'_{\boldsymbol{r}}\tilde{\boldsymbol{r}}.$$

The state $\boldsymbol{e}(t) = (\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t))/\epsilon - \boldsymbol{z}(t)$ satisfies

$$\begin{cases} \dot{\boldsymbol{e}}(t) = (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}(t)})\boldsymbol{e}(t) + \boldsymbol{e}_{\mathcal{F}}(t) + \boldsymbol{e}_{\mathcal{B}}\boldsymbol{u}(t), \ t > 0, \\ \boldsymbol{e}(0) = 0. \end{cases} \tag{4.30}$$

Assumption A4 and A5 ensure that as $\epsilon \to 0$

$$\|\boldsymbol{e}_{\mathcal{F}}(t)\| \to 0, \quad \forall t \in [0, \tau], \tag{4.31a}$$

$$\|\boldsymbol{e}_{\mathcal{B}}\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \to 0. \tag{4.31b}$$

It will be shown that $\lim_{\epsilon \to 0} \|\boldsymbol{e}\|_{\mathbb{W}(0,\tau)} = 0$. First, consider $\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)$, which satisfies

$$\begin{cases} \dot{\boldsymbol{x}}(t) - \dot{\boldsymbol{x}}_\epsilon(t) = & \mathcal{A}(\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)) + \mathcal{F}(\boldsymbol{x}(t)) - \mathcal{F}(\boldsymbol{x}_\epsilon(t)) \\ & + (\mathcal{B}(\boldsymbol{r}) - \mathcal{B}(\boldsymbol{r} + \epsilon\tilde{\boldsymbol{r}}))\, \boldsymbol{u}(t), \\ \boldsymbol{x}(0) - \boldsymbol{x}_\epsilon(0) = & 0. \end{cases}$$

Lemma 4.1.3 implies that there is a number c_τ depending only on τ such that for all $t \in [0, \tau]$

$$\|\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)\|_{\mathbb{V}} \tag{4.32}$$

$$\leq c_\tau \left(\|\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_\epsilon\|_{L^2(0,t;\mathbb{X})} + \|\mathcal{A}(\boldsymbol{x} - \boldsymbol{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} \right). \tag{4.33}$$

Also, use [24, Proposition 2.2], there is a number d_τ depending only on τ such that for all $t \in [0, \tau]$

$$\begin{aligned} &\|\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_\epsilon\|_{L^2(0,t;\mathbb{X})} + \|\mathcal{A}(\boldsymbol{x} - \boldsymbol{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} \\ &\qquad \leq d_\tau \|\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_\epsilon - \mathcal{A}(\boldsymbol{x} - \boldsymbol{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})}. \end{aligned} \tag{4.34}$$

Combine (4.33) and (4.34) to obtain

$$\begin{aligned} \|\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)\|_{\mathbb{V}} &\leq c_\tau d_\tau \|\mathcal{F}(\boldsymbol{x}) - \mathcal{F}(\boldsymbol{x}_\epsilon)\|_{L^2(0,t;\mathbb{X})} \\ &\quad + c_\tau d_\tau \|(\mathcal{B}(\boldsymbol{r}) - \mathcal{B}(\boldsymbol{r} + \epsilon\tilde{\boldsymbol{r}}))\, \boldsymbol{u}\|_{L^2(0,t;\mathbb{X})}. \end{aligned} \tag{4.35}$$

Theorem 2.4.6 implies that the states $\boldsymbol{x}(t)$ and $\boldsymbol{x}_\epsilon(t)$ belong to some bounded set in $\mathbb{W}(0, \tau)$ and so in $C(0, \tau, \mathbb{V})$. Let $D \subset \mathbb{V}$ be a bounded set that contains the trajectories $\boldsymbol{x}(t)$ and $\boldsymbol{x}_\epsilon(t)$. Let $L_{\mathcal{F}}$ be the Lipschitz constant of $\mathcal{F}(\cdot)$ on D. Since the set K_{ad} is compact and $\mathcal{B}(\boldsymbol{r})$ satisfies assumption A5, the number $L_{\mathcal{B}}$ defined as

$$L_{\mathcal{B}} = \sup_{\boldsymbol{r} \in K_{ad}} \|\mathcal{B}'_{\boldsymbol{r}}\|_{\mathcal{L}(\mathbb{K},\mathcal{L}(\mathbb{U},\mathbb{X}))}. \tag{4.36}$$

is finite. This together with [119, Theorem 12.1.1 and Corollary 3] yields

$$\|\mathcal{B}(\boldsymbol{r}) - \mathcal{B}(\boldsymbol{r}_\epsilon)\|_{\mathcal{L}(\mathbb{U},\mathbb{X})} \leq L_{\mathcal{B}} \|\boldsymbol{r} - \boldsymbol{r}_\epsilon\|_{\mathbb{K}} \leq L_{\mathcal{B}}\epsilon. \tag{4.37}$$

Use these to obtain the inequality

$$\begin{aligned}\|\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)\|_{\mathbb{V}}^2 \leq 2c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2 \int_0^t \|\boldsymbol{x}(s) - \boldsymbol{x}_\epsilon(s)\|_{\mathbb{V}}^2\, ds \\ + 2c_\tau^2 d_\tau^2 L_{\mathcal{B}}^2 \epsilon^2 \|\boldsymbol{u}\|_{L^2(0,\tau;\mathbb{U})}^2\end{aligned} \tag{4.38}$$

Applying Gronwall's lemma yields

$$\|\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t)\|_{\mathbb{V}} \leq \sqrt{2} e^{c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2} c_\tau d_\tau L_{\mathcal{B}} \epsilon \|\boldsymbol{u}\|_{L^2(0,\tau;\mathbb{U})}. \tag{4.39}$$

Define

$$M_{\mathcal{F}} := \sup_{t\in[0,\tau]} \|\mathcal{F}'_{\boldsymbol{x}(t)}\|_{\mathcal{L}(\mathbb{V},\mathbb{X})}.$$

Assumption A4 ensures that $M_{\mathcal{F}}$ is finite. Take the norm of the right side of (4.29a) in $\mathbb{X}$. It follows that

$$\|\boldsymbol{e}_{\mathcal{F}}(t)\| \leq (L_{\mathcal{F}} + M_{\mathcal{F}} c_e) \sqrt{2} e^{c_\tau^2 d_\tau^2 L_{\mathcal{F}}^2} c_\tau d_\tau L_{\mathcal{B}} \|\boldsymbol{u}\|_{L^2(0,\tau;\mathbb{U})}. \tag{4.40}$$

This and (4.31a) together with the Bounded Convergence Theorem ensure that

$$\lim_{\epsilon\to 0} \int_0^\tau \|\boldsymbol{e}_{\mathcal{F}}(t)\|^2\, dt = 0. \tag{4.41}$$

Statements (4.41) and (4.31b), and Lemma 4.2.3 can be applied to conclude

$$\lim_{\epsilon\to 0} \|\boldsymbol{e}\|_{\mathbb{W}(0,\tau)} = 0. \tag{4.42}$$

This shows that $\mathcal{S}(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0)$ is Gâteaux differentiable at $\boldsymbol{r}$ in the direction $\tilde{\boldsymbol{r}}$ with derivative $\boldsymbol{z}(t) = \mathcal{S}'_{\boldsymbol{r}}\tilde{\boldsymbol{r}}$.

b) This part is proven in [89, Theorem 3.4] assuming that $\partial_t + \mathcal{A}$ is invertible. However, the result is still true without assuming the invertibility of $\partial_t + \mathcal{A}$. Let ϵ be sufficiently small such that $\boldsymbol{u} + \epsilon\tilde{\boldsymbol{u}} \in U_{ad}$. Define $\boldsymbol{x}_\epsilon(t) = \mathcal{S}(\boldsymbol{u}(t) + \epsilon\tilde{\boldsymbol{u}}(t), \boldsymbol{r}, \boldsymbol{x}_0)$, this state solves

$$\begin{cases} \dot{\boldsymbol{x}}_\epsilon(t) = \mathcal{A}\boldsymbol{x}_\epsilon(t) + \mathcal{F}(\boldsymbol{x}_\epsilon(t)) + \mathcal{B}(\boldsymbol{r})(\boldsymbol{u}(t) + \epsilon\tilde{\boldsymbol{u}}(t)), & t > 0, \\ \boldsymbol{x}_\epsilon(0) = \boldsymbol{x}_0. \end{cases} \tag{4.43}$$

Let $\boldsymbol{e}(t) = (\boldsymbol{x}(t) - \boldsymbol{x}_\epsilon(t))/\epsilon - \boldsymbol{h}(t)$. Following the same steps as in part (a) yields

$$\lim_{\epsilon \to 0} \|\boldsymbol{e}\|_{\mathbb{W}(0,\tau)} = 0. \tag{4.44}$$

This means that $\mathcal{S}(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0)$ is Gâteaux differentiable at $\boldsymbol{u}$ in the direction $\tilde{\boldsymbol{u}}$ with derivative $\boldsymbol{h}(t) = (\mathcal{S}'_{\boldsymbol{u}}\tilde{\boldsymbol{u}})(t)$. □

Assumption A1 implies that the dual of each of $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$ will be identified with the space itself. The operator $(\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^* : \mathbb{X} \to \mathbb{K}$ is defined as

$$\langle (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u})^* \boldsymbol{p}, \boldsymbol{r} \rangle_{\mathbb{K}} = \langle \boldsymbol{p}, (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{r})\boldsymbol{u} \rangle, \ \forall (\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{r}) \in \mathbb{U} \times \mathbb{X} \times \mathbb{K}.$$

Theorem 4.2.5. *Suppose assumption A holds, and identify the derivatives $J'_{\boldsymbol{x}}$, $J'_{\boldsymbol{u}}$, and $J'_{\boldsymbol{r}}$ by elements $\boldsymbol{j}_{\boldsymbol{x}} \in \mathbb{W}^*(0,\tau)$, $\boldsymbol{j}_{\boldsymbol{u}} \in L^2(0,\tau;\mathbb{U})$ and $\boldsymbol{j}_{\boldsymbol{r}} \in \mathbb{K}$, respectively. For any initial condition $\boldsymbol{x}_0 \in \mathbb{X}$, let the pair $(\boldsymbol{u}^o, \boldsymbol{r}^o) \in U_{ad} \times K_{ad}$ be a local minimizer of the optimization problem P with the optimal trajectory $\boldsymbol{x}^o = \mathcal{S}(\boldsymbol{u}^o; \boldsymbol{r}^o, \boldsymbol{x}_0)$ and let $\boldsymbol{p}^o(t)$ indicate the strict solution in $\mathbb{W}^*(0,\tau)$ of the final value problem*

$$\dot{\boldsymbol{p}}^o(t) = -(\mathcal{A}^* + {\mathcal{F}'_{\boldsymbol{x}^o(t)}}^*)\boldsymbol{p}^o(t) - \boldsymbol{j}_{\boldsymbol{x}^o}(t), \quad \boldsymbol{p}^o(\tau) = 0. \tag{4.45}$$

Then $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ satisfies

$$\begin{aligned} &\langle \boldsymbol{j}_{\boldsymbol{u}^o} + \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o, \boldsymbol{u} - \boldsymbol{u}^o \rangle_{L^2(0,\tau;\mathbb{U})} \geq 0, && \forall \boldsymbol{u} \in U_{ad}, \\ &\left\langle \boldsymbol{j}_{\boldsymbol{r}^o} + \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^* \boldsymbol{p}^o(t)\, dt, \boldsymbol{r} - \boldsymbol{r}^o \right\rangle_{\mathbb{K}} \geq 0, && \forall \boldsymbol{r} \in K_{ad}. \end{aligned}$$

Proof. Let

$$\mathcal{G}(\boldsymbol{u}, \boldsymbol{r}) = J(\mathcal{S}(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0), \boldsymbol{u}, \boldsymbol{r}).$$

The Gâteaux derivative of $\mathcal{G}(\boldsymbol{u}, \boldsymbol{r})$ with respect to $\boldsymbol{u}$ has been obtained in the proof of [89, Proposition 4.13]. Using the chain rule to take the Gâteaux derivative of $\mathcal{G}(\boldsymbol{u}, \boldsymbol{r})$ at $\boldsymbol{u}^o$ in the direction $\tilde{\boldsymbol{u}}$ yields

$$\mathcal{G}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = J'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} + J'_{\boldsymbol{x}^o}\mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}. \tag{4.47}$$

Identify the functionals $\mathcal{G}'_{\boldsymbol{u}^o} : L^2(0,\tau;\mathbb{U}) \to \mathcal{R}$ and $J'_{\boldsymbol{u}^o} : L^2(0,\tau;\mathbb{U}) \to \mathcal{R}$ with elements of $L^2(0,\tau,\mathbb{U})$. That is

$$\mathcal{G}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = \langle \boldsymbol{g}_{\boldsymbol{u}^o}, \tilde{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{U})}, \tag{4.48}$$

$$J'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = \langle \boldsymbol{j}_{\boldsymbol{u}^o}, \tilde{\boldsymbol{u}} \rangle_{L^2(0,\tau;\mathbb{U})}. \tag{4.49}$$

Also, identifying the functional $J'_{\boldsymbol{x}^o} : L^2(0,\tau;\mathbb{X}) \to \mathcal{R}$ with an element of $\mathbb{W}^*(0,\tau) = L^2(0,\tau; D(\mathcal{A}^*)) \cap W^{1,2}(0,\tau;\mathbb{X})$ yields

$$J'_{\boldsymbol{x}^o}\mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = \langle \boldsymbol{j}_{\boldsymbol{x}^o}, \mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{X})}. \tag{4.50}$$

The adjoint operator $\mathcal{S}'^*_{\boldsymbol{u}^o}$ can be obtained as follows. Use (4.45) in the following inner product and let $\boldsymbol{h}(t) = \mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}$

$$\begin{aligned}&\langle \boldsymbol{j}_{\boldsymbol{x}^o}, \mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{X})}\\ &\quad= \int_0^\tau \left\langle -\dot{\boldsymbol{p}}^o(t) - (\mathcal{A}^* + {\mathcal{F}'_{\boldsymbol{x}^o(t)}}^*)\boldsymbol{p}^o(t), \boldsymbol{h}(t)\right\rangle dt.\end{aligned}$$

Taking the adjoint and integration by parts yield

$$\begin{aligned}&\langle \boldsymbol{j}_{\boldsymbol{x}^o}, \mathcal{S}'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{X})}\\ &\quad= \int_0^\tau \left\langle \boldsymbol{p}^o(t), \dot{\boldsymbol{h}}(t) - (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{h}(t)\right\rangle dt\\ &\quad= \int_0^\tau \langle \boldsymbol{p}^o(t), \mathcal{B}(\boldsymbol{r})\tilde{\boldsymbol{u}}(t)\rangle\, dt\\ &\quad= \int_0^\tau \langle \mathcal{B}^*(\boldsymbol{r})\boldsymbol{p}^o(t), \tilde{\boldsymbol{u}}(t)\rangle\, dt\end{aligned}$$

This implies

$$\mathcal{S}'^*_{\boldsymbol{u}^o}\boldsymbol{j}_{\boldsymbol{x}^o} = \mathcal{B}^*(\boldsymbol{r})\boldsymbol{p}^o(t). \tag{4.51}$$

Combine (4.48), (4.49), (4.50) and use (4.51), equation (4.47) is written using the functionals as

$$\langle \boldsymbol{g}_{\boldsymbol{u}}, \tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{U})} = \langle \boldsymbol{j}_{\boldsymbol{u}^o} + \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o, \tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{U})}. \tag{4.52}$$

Applying [61, Theorem 1.46] and letting $\tilde{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{u}^o$ for all $\boldsymbol{u} \in U_{ad}$ yields

$$\langle \boldsymbol{j}_{\boldsymbol{u}^o} + \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o, \boldsymbol{u} - \boldsymbol{u}^o\rangle_{L^2(0,\tau;\mathbb{U})} \geq 0. \tag{4.53}$$

Using the chain rule to take the Gâteaux derivative of $\mathcal{G}(\boldsymbol{u},\boldsymbol{r})$ at $\boldsymbol{r}^o$ in the direction $\tilde{\boldsymbol{r}}$ yields

$$\mathcal{G}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}} = J'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}} + J'_{\boldsymbol{x}^o}\mathcal{S}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}. \tag{4.54}$$

Similarly, identify the functionals $\mathcal{G}'_{\boldsymbol{r}^o} : \mathbb{K} \to \mathcal{R}$ and $J'_{\boldsymbol{r}^o} : \mathbb{K} \to \mathcal{R}$ with elements of $\boldsymbol{g}_{\boldsymbol{r}^o}$ and $\boldsymbol{j}_{\boldsymbol{r}^o}$ in $\mathbb{K}$, and take the adjoint of $\mathcal{S}'_{\boldsymbol{r}^o}$. It follows that

$$\boldsymbol{g}_{\boldsymbol{r}^o} = \mathcal{S}'^*_{\boldsymbol{r}^o}\boldsymbol{j}_{\boldsymbol{x}^o}(t) + \boldsymbol{j}_{\boldsymbol{r}^o}. \tag{4.55}$$

The adjoint operator $\mathcal{S}'^*_{\boldsymbol{r}^o}$ can also be derived explicitly. Consider the inner product

$$\langle \boldsymbol{j}_{\boldsymbol{x}^o}, \mathcal{S}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}\rangle_{L^2(0,\tau;\mathbb{X})} = \int_0^\tau \langle \boldsymbol{j}_{\boldsymbol{x}^o}(t), \mathcal{S}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}\rangle_{\mathbb{X}}\, dt.$$

Let $\boldsymbol{z}(t) = \mathcal{S}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}}$. Substitute for $\boldsymbol{j}_{\boldsymbol{x}^o}(t)$ from (4.45) into this integral. Perform integration by parts to obtain

$$\begin{aligned}
&\int_0^\tau \left\langle -\dot{\boldsymbol{p}}^o(t) - (\mathcal{A}^* + {\mathcal{F}'_{\boldsymbol{x}^o(t)}}^*)\boldsymbol{p}^o(t), \boldsymbol{z}(t)\right\rangle_{\mathbb{X}} dt \\
&\quad = \int_0^\tau \left\langle \boldsymbol{p}^o(t), \dot{\boldsymbol{z}}(t) - (\mathcal{A} + \mathcal{F}'_{\boldsymbol{x}^o(t)})\boldsymbol{z}(t)\right\rangle_{\mathbb{X}} dt \\
&\quad = \int_0^\tau \langle \boldsymbol{p}^o(t), (\mathcal{B}'_{\boldsymbol{r}^o}\tilde{\boldsymbol{r}})\boldsymbol{u}^o(t)\rangle_{\mathbb{X}}\, dt \\
&\quad = \left\langle \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)dt, \tilde{\boldsymbol{r}} \right\rangle_{\mathbb{K}}. \qquad (4.56)
\end{aligned}$$

Thus,

$$\mathcal{S}'^*_{\boldsymbol{r}^o}\boldsymbol{j}_{\boldsymbol{x}^o}(t) = \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)dt. \qquad (4.57)$$

As a result, the Gâteaux derivative of $\mathcal{G}(\boldsymbol{u}, \boldsymbol{r})$ at $\boldsymbol{r}^o$ in the direction $\tilde{\boldsymbol{r}}$ is identified with

$$\boldsymbol{g}'_{\boldsymbol{r}^o} = \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)dt + \boldsymbol{j}_{\boldsymbol{r}^o}. \qquad (4.58)$$

The optimality conditions now follow by substituting the Gâteaux derivatives $\boldsymbol{g}'_{\boldsymbol{r}^o}$ in [61, Theorem 1.46]. □

Corollary 4.2.6. *If the minimizer* $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ *is in the interior of* $U_{ad} \times K_{ad}$*, then the following set of equations characterizes* $(\boldsymbol{x}^o, \boldsymbol{p}^o, \boldsymbol{u}^o, \boldsymbol{r}^o)$*:*

$$\begin{cases}
\dot{\boldsymbol{x}}^o(t) = \mathcal{A}\boldsymbol{x}^o(t) + \mathcal{F}(\boldsymbol{x}^o(t)) + \mathcal{B}(\boldsymbol{r}^o)\boldsymbol{u}^o(t), & \boldsymbol{x}^o(0) = \boldsymbol{x}_0, \\
\dot{\boldsymbol{p}}^o(t) = -(\mathcal{A}^* + {\mathcal{F}'_{\boldsymbol{x}^o(t)}}^*)\boldsymbol{p}^o(t) - \mathcal{Q}\boldsymbol{x}^o(t), & \boldsymbol{p}^o(\tau) = 0, \\
\boldsymbol{u}^o(t) = -\mathcal{R}^{-1}\mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o(t), \\
\int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)\, dt = 0.
\end{cases}$$

Proof. If the optimizer $(\boldsymbol{u}^o, \boldsymbol{r}^o)$ is in the interior of $U_{ad} \times K_{ad}$, then the optimality conditions of Theorem 4.2.5 hold if and only if

$$\boldsymbol{j}_{\boldsymbol{u}^o} + \mathcal{B}^*(\boldsymbol{r}^o)\boldsymbol{p}^o = 0, \qquad (4.59)$$

$$\boldsymbol{j}_{\boldsymbol{r}^o} + \int_0^\tau (\mathcal{B}'_{\boldsymbol{r}^o}\boldsymbol{u}^o(t))^*\boldsymbol{p}^o(t)\, dt = 0. \qquad (4.60)$$

The derivatives $J'_{\boldsymbol{x}^o}(t) : L^2(0,\tau;\mathbb{X}) \to \mathcal{R}$ and $J'_{\boldsymbol{u}^o}(t) : L^2(0,\tau;\mathbb{U}) \to \mathcal{R}$ are

$$J'_{\boldsymbol{x}^o}\tilde{\boldsymbol{x}} = \langle \mathcal{Q}\boldsymbol{x}^o, \tilde{\boldsymbol{x}}\rangle_{L^2(0,\tau;\mathbb{X})}, \tag{4.61}$$

$$J'_{\boldsymbol{u}^o}\tilde{\boldsymbol{u}} = \langle \mathcal{R}\boldsymbol{u}^o, \tilde{\boldsymbol{u}}\rangle_{L^2(0,\tau;\mathbb{U})}. \tag{4.62}$$

Identify these functionals with elements $\boldsymbol{j}_{\boldsymbol{x}^o} = \mathcal{Q}\boldsymbol{x}^o(t)$ and $\boldsymbol{j}_{\boldsymbol{u}^o} = \mathcal{R}\boldsymbol{u}^o(t)$, and notice that $\boldsymbol{j}_{\boldsymbol{r}^o} = 0$. Substituting the derivatives in (4.59) and (4.60) yields the optimality conditions. □

4.3 Examples

4.3.1 Kuramoto–Sivashinsky Equation

For every actuator location $r \in (0,1)$, let the function $b(\cdot;r)$ be in $C^1[0,1]$. Consider the controlled Kuramoto–Sivashinsky equation with Dirichlet boundary conditions and initial condition $w_0(\xi)$ on $\xi \in [0,1]$ and some number λ

$$\begin{cases} \frac{\partial w}{\partial t} + \frac{\partial^4 w}{\partial \xi^4} + \lambda \frac{\partial^2 w}{\partial \xi^2} + w\frac{\partial w}{\partial \xi} = b(\xi;r)u(t), \quad t > 0, \\ w(0,t) = w(1,t) = 0, \quad t \geq 0, \\ \frac{\partial w}{\partial \xi}(0,t) = \frac{\partial w}{\partial \xi}(1,t) = 0, \quad t \geq 0, \\ w(\xi,0) = w_0(\xi), \quad \xi \in [0,1]. \end{cases} \tag{4.63}$$

Define the state $\boldsymbol{x}(t) \coloneqq w(\cdot,t)$, the state space $\mathbb{X} \coloneqq L^2(0,1)$. Let the state operator $\mathcal{A} : D(\mathcal{A})(\subset \mathbb{X}) \to \mathbb{X}$ be

$$\begin{aligned} \mathcal{A}w &\coloneqq -w_{\xi\xi\xi\xi} - \lambda w_{\xi\xi}, \\ D(\mathcal{A}) &= H^4(0,1) \cap H_0^2(0,1). \end{aligned} \tag{4.64}$$

Also, the control space is $\mathbb{U} \coloneqq \mathcal{R}$. The actuator design space is $\mathbb{K} \coloneqq \mathcal{R}$. Define $\mathbb{V} \coloneqq H_0^1(0,1)$; the nonlinear operator $\mathcal{F}(\cdot) : \mathbb{V} \to \mathbb{X}$ and the input operator $\mathcal{B}(\cdot) : \mathbb{K} \to \mathcal{L}(\mathbb{U},\mathbb{X})$ are defined as

$$\mathcal{F}(w) \coloneqq -ww_\xi, \tag{4.65}$$

$$\mathcal{B}(r)u \coloneqq b(\xi,r)u. \tag{4.66}$$

The state space representation of the model will then be (4.1).

The operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is a self-adjoint operator, is bounded from below, and has compact resolvent. According to Theorem [109, Theorem 32.1], $\mathcal{A}$ generates an analytic semigroup on $\mathbb{X}$. Since the operator $\mathcal{A}$ is analytic on a Hilbert space, Theorem 4.1 in [35] ensures that this operator enjoys maximal parabolic regularity. Also, by Rellich-Kondrachov compact embedding theorem [1, Chapter 6], the space $D(\mathcal{A})$ is compactly embedded in $\mathbb{X}$. The operator $\mathcal{A}$ is also associated with a form described in A2.

Lemma 4.3.1. *The nonlinear operator $\mathcal{F}(\cdot)$ is Gâteaux differentiable from $\mathbb{V}$ to $\mathbb{X}$. The Gâteaux derivative of $\mathcal{F}(\cdot)$ at w in the direction f is $\mathcal{F}'_w f = -wf_\xi - w_\xi f$.*

Proof. The operator $\mathcal{F}'_w$, if exists, needs to satisfy

$$\lim_{\epsilon \to 0} \left\| \frac{1}{\epsilon} \left(\mathcal{F}(w + \epsilon f) - \mathcal{F}(w) \right) - \mathcal{F}'_w f \right\|_{L^2} = 0 \tag{4.67}$$

Substituting in (4.65), inside the limit becomes

$$\begin{aligned} &\left\| \frac{1}{\epsilon} (ww_\xi - (w + \epsilon f)(w_\xi + \epsilon f_\xi)) - wf_\xi - w_\xi f \right\|_{L^2} \\ &\qquad = \|\epsilon f f_\xi\|_{L^2} . \end{aligned} \tag{4.68}$$

Note that $f \in H^1_0(0,1)$. Embedding $H^1_0(0,1) \hookrightarrow C[0,1]$ means that f is a continuous function over $[0,1]$. This implies that ff_ξ is in $L^2(0,1)$, thus

$$\lim_{\epsilon \to 0} \epsilon \, \|ff_\xi\|_{L^2} = 0. \tag{4.69}$$

The lemma now follows from the uniqueness of Gâteaux derivative. □

Note that $D_{\mathcal{A}}(1/2, 2) = H^2_0(0,1) \hookrightarrow \mathbb{V}$ (see [20, Corollary 4.10]). The operator $\mathcal{F}(\cdot) : \mathbb{V} \to \mathbb{X}$ is not however weakly continuous, and does not satisfy assumption B1 of [37].

For all functions f and w in $H^1_0(0,1)$ and g in $H^1(0,1)$, the adjoint of $\mathcal{F}'_w$ satisfies

$$\left\langle f, \mathcal{F}'^{*}_w g \right\rangle_{L^2} = \left\langle \mathcal{F}'_w f, g \right\rangle_{L^2} = \int_0^1 (-wf_\xi - w_\xi f) g d\xi. \tag{4.70}$$

Performing integration by parts yields

$$\int_0^1 (-wf_\xi - w_\xi f) g d\xi = -\int_0^1 w g_\xi f d\xi. \tag{4.71}$$

The operator ${\mathcal{F}'_w}^*$ maps $D({\mathcal{F}'_w}^*) = H^1(0,1)$ to $L^2(0,1)$ as follows

$$ {\mathcal{F}'_w}^* g = -w g_\xi. \tag{4.72} $$

In addition,

$$ \mathcal{B}^*(r)w = \int_0^1 b(\xi, r)w(\xi)d\xi, \qquad \forall w \in \mathbb{V}, \tag{4.73} $$

$$ (\mathcal{B}'_r u)^* f = u \int_0^1 b_r(\xi; r) f(\xi) d\xi, \qquad \forall f \in \mathbb{V}. \tag{4.74} $$

Also, define

$$ K_{ad} := \{r \in [a,b] : 0 < a < b < 1\}. \tag{4.75} $$

Global stability of an uncontrolled KS equation has been studied extensively, see e.g. [3, 82, 44, 2]. Theorem 2.1 of [82] proves that for $\lambda < 4\pi^2$, the uncontrolled KS equation is globally exponentially stable. Proof of this theorem can be modified to ensure that there is solution to the controlled KS equation over $[0, \tau]$ for all initial conditions in $\mathbb{V}$. The following lemma ensures that for some parameters λ there is a solution to the KS equation for all initial conditions and inputs over arbitrary time intervals.

Lemma 4.3.2. *Let $\lambda < 4\pi^2$ and $\sigma(\lambda)$ be the smallest eigenvalue of $-\mathcal{A}$. For all initial conditions $w_0 \in \mathbb{V}$ and inputs $u \in L^2(0,\tau)$, the strict solution to the KS system (4.63) satisfies*

$$ \|w(\tau)\|^2 \le \|w_0\|^2 + \frac{1}{\sigma(\lambda)} \|u\|^2_{L^2(0,\tau)} \max_{\xi \in [0,1]} b^2(\xi; r). $$

Proof. Theorem 2.4.6 ensures that there is a solution $w \in \mathbb{W}(0,\tau)$ over $[0,\tau]$ to the KS system with initial condition $w_0 \in \mathbb{V}$ and input $u \in L^2(0,\tau)$. Consider the Lyapunov function

$$ E(t) := \int_0^1 w^2(\xi, t)\, d\xi. \tag{4.76} $$

Since $w \in W^{1,2}(0,\tau;\mathbb{X})$, the function $E(t)$ is differentiable. Taking the derivative of $E(t)$ and applying [82, Lemma 3.1] yield

$$ \dot{E}(t) \le -2\sigma(\lambda)E(t) + 2\int_0^1 w(\xi,t) b(\xi; r) u(t) d\xi. \tag{4.77} $$

Apply Young's inequality to the integral term, for every $\epsilon > 0$,

$$ \dot{E}(t) \le (-2\sigma(\lambda) + \epsilon)E(t) + \frac{1}{\epsilon}\int_0^1 b^2(\xi; r) u^2(t) d\xi. \tag{4.78} $$

Let $\epsilon = \sigma(\lambda)$. Taking an integral over $[0,\tau]$ yields the desired inequality in the lemma. □

Since the KS system satisfies Assumptions A1-A5, Corollary 4.2.6 can be applied to obtain the optimality conditions. The cost function to be optimized is

$$J(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{r}) = \int_0^\tau \int_0^1 w^2(\xi, t) d\xi dt + \int_0^\tau u^2(t) dt. \tag{4.79}$$

Letting $\boldsymbol{p}(t) = f(\cdot, t)$, the optimizer (u^o, r^o, w^o, f^o) with initial condition $w_0(\xi) \in H_0^1(0, 1)$ satisfies

$$\begin{cases} \dfrac{\partial w^o}{\partial t} + \dfrac{\partial^4 w^o}{\partial \xi^4} + \lambda \dfrac{\partial^2 w^o}{\partial \xi^2} + w^o \dfrac{\partial w^o}{\partial \xi} = b(\xi; r^o) u^o(t), & t > 0 \\ w^o(0, t) = w^o(1, t) = 0, & t > 0 \\ \dfrac{\partial w^o}{\partial \xi}(0, t) = \dfrac{\partial w^o}{\partial \xi}(1, t) = 0, & t > 0 \\ w^o(\xi, 0) = w_0(\xi), \end{cases}$$

$$\begin{cases} \dfrac{\partial f^o}{\partial t} - \dfrac{\partial^4 f^o}{\partial \xi^4} - \lambda \dfrac{\partial^2 f^o}{\partial \xi^2} - w^o \dfrac{\partial f^o}{\partial \xi} = -w^o(\xi, t), & t > 0 \\ f^o(0, t) = f^o(1, t) = 0, & t > 0 \\ \dfrac{\partial f^o}{\partial \xi}(0, t) = \dfrac{\partial f^o}{\partial \xi}(1, t) = 0, & t > 0 \\ f^o(\xi, \tau) = 0, \end{cases}$$

$$\begin{cases} u^o(t) = -\displaystyle\int_0^1 b(\xi; r^o) f^o(\xi, t)\, d\xi, \quad t > 0, \\ \displaystyle\int_0^\tau \int_0^1 u^o(t) b_r(\xi; r^o) f^o(\xi, t)\, d\xi dt = 0. \end{cases}$$

4.3.2 Nonlinear Diffusion

Consider the transfer of heat in a bounded, open, connected set $\Omega \subset \mathbb{R}^2$. It is assumed that Ω has a Lipschitz boundary separated into $\partial\Omega = \overline{\Gamma_0 \cup \Gamma_1}$ where $\Gamma_0 \cap \Gamma_1 = \emptyset$ and $\Gamma_0 \neq \emptyset$. Denote by ν the unit outward normal vector field on $\partial\Omega$. The class of nonlinear heat transfer models is, for actuator shape $r \in C^1(\overline{\Omega})$,

$$\begin{cases} \dfrac{\partial w}{\partial t}(\xi,t) = \\ \qquad \Delta w(\xi,t) + F(w(\xi,t)) + r(\xi)u(t), & (\xi,t) \in \Omega \times (0,\tau], \\ w(\xi,t) = 0, & \xi \in \Gamma_0 \times [0,\tau], \\ \dfrac{\partial w}{\partial \nu}(\xi,t) = 0, & \xi \in \Gamma_1 \times [0,\tau], \\ w(\xi,0) = w_0(\xi), & \xi \in \Omega. \end{cases}$$

Defining $\mathbb{K} = L^2(\Omega)$, a set of admissible actuator shapes is

$$K_{ad} = \{r \in C^1(\overline{\Omega}) : \|r\|_{C^1(\overline{\Omega})} \leq 1\}.$$

The set K_{ad} is compact in $\mathbb{K}$ with respect to the norm topology [1, Chapter 6].

Let $\mathbb{X} := L^2(\Omega)$, $\mathbb{U} := \mathcal{R}$, and the state $\boldsymbol{x}(t) := w(\cdot,t)$. The operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is defined as

$$\mathcal{A}w = \Delta w, \tag{4.80a}$$

$$D(\mathcal{A}) = \left\{ w \in H^2(\Omega) \cap H^1_{\Gamma_0} : \frac{\partial w}{\partial \nu} = 0 \text{ on } \Gamma_1 \right\}. \tag{4.80b}$$

The operator $\mathcal{A}$ self-adjoint, non-negative and has compact resolvent. Thus, it generates an analytic semi-group on the Hilbert space $L^2(\Omega)$ [109, Theorem 32.1], and has maximal L^p regularity.

Define $\mathbb{V} = H^1_{\Gamma_0}(\Omega)$ and assume that the nonlinear operator $\mathcal{F}(\cdot) : \mathbb{V} \to \mathbb{X}$. The proof of the following lemma is the same as that of [37, Lemma 7.1.1].

Lemma 4.3.3. *Let $\mathbb{V} = H^1_{\Gamma_0}(\Omega)$. Assume that*

1. *$F(\zeta)$ is twice continuously differentiable over $\mathbb{R}$; denote its derivatives by $\mathcal{F}'(\zeta)$ and $F''(\zeta)$;*
2. *there are numbers $a_0 > 0$ and $b > 1/2$ such that $|F''(\zeta)| \leq a_0(1 + |\zeta|^b)$.*

Then $\mathcal{F}(\cdot)$ is Gâteaux differentiable from $\mathbb{V}$ to $\mathbb{X}$. The Gâteaux derivative of $\mathcal{F}(\cdot)$ at $w(\xi)$ in the direction $f(\xi)$ is ${\mathcal{F}'}_w f = F'(w)f$.

It is straightforward to show that the operator $\mathcal{F}'_w : \mathbb{V}(\subset \mathbb{X}) \to \mathbb{X}$ is self-adjoint, i.e.,

$$\left\langle {\mathcal{F}'_w}^* g, f \right\rangle = \langle g, \mathcal{F}'_w f \rangle , \quad \forall f, g \in \mathbb{V}. \tag{4.81}$$

Define $\mathbb{U} = \mathbb{R}$ and the input operator $\mathcal{B}(r) \in \mathcal{L}(\mathbb{U}, \mathbb{X})$ maps u to $r(\xi)u$. Also, for all f in $\mathbb{X}$

$$\mathcal{B}^*(r) f = \int_\Omega r(\xi) f(\xi) d\xi, \tag{4.82}$$

$$(\mathcal{B}'_r u)^* f = uf. \tag{4.83}$$

For every initial condition in $\mathbb{V}$, a strict solution over $[0, \tau]$ to the nonlinear heat equation is not guaranteed. The following lemma states a condition under which there is a solution to the diffusion equation for all initial conditions and inputs over arbitrary time intervals.

Lemma 4.3.4. *If the function $F(\zeta)$ satisfies $\zeta F(\zeta) \leq 0$ for all $\zeta \in \mathbb{R}$, then there is $c_\Omega > 0$ such that the strict solution to the nonlinear heat equation satisfies*

$$\|w(\tau)\|^2 \leq \|w_0\|^2 + \frac{4}{c_\Omega} \|u\|^2_{L^2(0,\tau)} \|r\|^2_{\mathbb{K}} .$$

Proof. Theorem 1 in [86] proves that the nonlinear equation in one spatial dimension is input-to-state stable. This lemma extends Theorem 1 of [86] to two-spatial dimension. Using the same idea of proof, consider the Lyapunov function

$$E(t) := \int_\Omega w^2(\xi, t)\, d\xi. \tag{4.84}$$

The function $E(t)$ is differentiable since $w \in W^{1,2}(0, \tau; \mathbb{X})$. Take the derivative of this function, substitute for $\dot{w}(\xi, t)$ from the heat equation, and perform integration by parts as follows

$$\begin{aligned}\dot{E}(t) =& 2 \int_\Omega w(\xi,t) \left(\Delta w(\xi,t) + F(w(\xi,t)) + r(\xi) u(t)\right)\, d\xi \\ =& 2 \int_\Gamma w(\xi,t) \frac{\partial w}{\partial \nu}(\xi,t) d\xi - 2 \int_\Omega \left(\nabla w(\xi,t)\right)^2\, d\xi \\ &+ 2 \int_\Omega w(\xi,t) \left(F(w(\xi,t)) + r(\xi)u(t)\right)\, d\xi. \end{aligned} \tag{4.85}$$

Apply the boundary conditions. Use Poincaré inequality and let c_Ω be its constant. Also, use Young's inequality for all $\epsilon > 0$

$$\dot{E}(t) \le -2\left(c_\Omega - \epsilon\right) E(t) + \frac{2}{\epsilon} u^2(t) \left\| r \right\|_2^2 . \tag{4.86}$$

Set $\epsilon = c_\Omega/2$. Taking the integral over $[0, \tau]$ of (4.86) then yields the desired inequality. □

The nonlinear heat equation satisfies Assumptions A1-A5, and thus, Corollary 4.2.6 can be applied to obtain the optimality conditions. The cost function to be optimized is

$$J(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{r}) = \int_0^\tau \int_\Omega w^2(\xi, t) d\xi dt + \int_0^\tau u^2(t) dt. \tag{4.87}$$

Letting $\boldsymbol{p}(t) = f(\cdot, t)$, The optimizer (u^o, r^o, w^o, f^o) with initial condition $w_0 \in H^1_{\Gamma_0}(\Omega)$ satisfies

$$\begin{cases} \dfrac{\partial w^o}{\partial t}(\xi, t) = \\ \qquad \Delta w^o(\xi, t) + F(w^o(\xi, t)) + r^o(\xi) u^o(t), & (\xi, t) \in \Omega \times (0, \tau], \\ w^o(\xi, t) = 0, & \xi \in \Gamma_0 \times [0, \tau], \\ \dfrac{\partial w^o}{\partial \nu}(\xi, t) = 0, & \xi \in \Gamma_1 \times [0, \tau], \\ w^o(\xi, 0) = w_0(\xi), & \xi \in \Omega. \end{cases}$$

$$\begin{cases} \dfrac{\partial f^o}{\partial t}(\xi, t) = -\Delta f^o(\xi, t) \\ \qquad - F'(w(\xi, t)) f^o(\xi, t) - w^o(\xi, t), & (\xi, t) \in \Omega \times (0, \tau], \\ f^o(\xi, t) = 0, & \xi \in \Gamma_0 \times [0, \tau], \\ \dfrac{\partial f^o}{\partial \nu}(\xi, t) = 0, & \xi \in \Gamma_1 \times [0, \tau], \\ f^o(\xi, \tau) = 0, & \xi \in \Omega. \end{cases}$$

$$\begin{cases} u^o(t) = -\displaystyle\int_\Omega r^o(\xi) f^o(\xi, t)\, d\xi, & t \in [0, \tau], \\ \displaystyle\int_0^\tau u^o(t) f^o(\xi, t) dt = 0, & \xi \in \Omega. \end{cases}$$

4.4 Concluding Remarks

Optimal actuator design for quasi-linear infinite-dimensional systems was considered in this chapter. It was shown that the existence of an optimal control together with an optimal actuator design is guaranteed under some assumptions. Moreover, first-order necessary optimality conditions were obtained. The theory was illustrated with application to Kuramoto-Sivashinsky (KS) equation and nonlinear heat equation.

Future work is concerned with developing numerical methods for solution of the optimality equations. Extension of these problems to situations where the input operator is not bounded on the state space is also of interest.

Chapter 5

Optimal Control and Actuator Location for Railway Tracks

Railway tracks rest on a foundation known for exhibiting nonlinear viscoelastic behavior. Railway track deflections are modeled by a semilinear partial differential equation. This chapter studies the stability of solutions to this equation in presence of an input. It furthers applies the results of previous chapters to compute an optimal control and actuator location. The stability results are obtained with the aid of a suitable Lyapunov function. The existence and exponential stability of classical solutions is established for certain inputs. The Lyapunov function is further used to find an a-priori estimate of the solutions, and also to study the input-to-state stability (ISS) of mild solutions.

The chapter is organized as follows: Section 5.1 begins with a review of the literature on stability of nonlinear PDE's that model the vibration of flexible structures, in particular railway tracks. Further on, the well-posedness and stability of different solutions to a railway track model are discussed in this chapter. Section 5.2 discusses the optimal control and actuator location for the railway track model. Numerical schemes and optimization algorithms are presented in this section. The chapter ends with some concluding remarks in Section 5.3.

5.1 Well-posedness and Stability of Railway Track Model

Stability analysis of nonlinear partial differential equations (PDE's) modeling flexible structures has attracted attention in the past few decades. To name but a few of the publications in this field; in [49], boundary stabilization of a nonlinear beam clamped at one end and supported by a nonlinear bearing at the other end is studied. In [45], authors investigated asymptotic behavior of a semilinear viscoelastic beam model including a memory term. The nonlinearity is assumed to satisfy some growth assumption. In [114], asymptotic stability of Falk model of shape memory alloys is studied using energy method. In [40, 36], boundary stabilization of a nonlinear micro-beam model is studied using linearizion technique together with Lyapunov method. The von Karman model of slender beams has also been investigated in many papers. In a recent study, Liu et al. considered asymptotic stability of von Karman beam with thermo-viscoelastic damping [81]. Nonlinear PDE's with a fourth order spatial derivative are not limited to flexible structures; an example is Cahn-Hilliard equation with inertial term which describes the phase separation of binary fluids; see for example, [55].

The nonlinear railway track model in this chapter was described in [5]. The nonlinearity in this model is caused by the railway support ballast which is known for highly nonlinear viscoelastic behavior [27]. Another nonlinear model for railway tracks has been suggested [34]. Unlike the previous model, this model also includes shear deformations in flexible track beams. This railway track model was used to study the effect of passing vehicles on pavements [33]. There are also studies devoted to the vibration monitoring and control of railway tracks [72, 123]. Track deflections induced by train passage are a cause of ride discomfort, fatigue in railway, and disturbances to nearby buildings [72]. Thus, the vibrations need to be carefully monitored and controlled.

This chapter focuses on well-posedness and stability, with respect to both initial conditions, and inputs of this model. Input-to-state stability (ISS) does not generalize in a straightforward way to infinite dimensions; see [91] for counter examples. Input-to-state stability theory has been extended in recent years to include systems of infinite dimension. In [93], a comparison between ISS theory of finite-dimensional systems and that of infinite-dimensional systems is presented. In [64], ISS of a class of linear infinite dimensional control systems with nonlinear feedbacks is discussed. The nonlinear feedback satisfies a sector condition that does not apply to the nonlinearity in this chapter. In [86], strict Lyapunov functions were used to investigate the stability of nonlinear heat equation; also, such Lyapunov functions were used to establish a robust stability of the equation in presence of a convection term and uncertainties. Integral input-to-state stability is a weaker property than ISS. In [92], integral ISS is discussed. In [63], the relation between iISS and ISS is discussed for linear systems with an unbounded control operator. ISS with respect to boundary inputs and disturbances has also been studied in [67, 103, 105]. In [105], ISS of the reaction-diffusion-advection equation with boundary and in-domain point-wise sensing and actuation is considered.

In this section, a Lyapunov function is used to establish ISS for the nonlinear controlled railway track model. To construct the Lyapunov function, the multiplier method [71] is used. Furthermore, a density argument is used to prove the ISS of the model when the inputs are not differentiable. In such cases, the Lyapunov function is also non-differentiable. A density argument was also used in [90, Lemma 2.2.3], where the control system is assumed to have a transition map that continuously depends on both initial conditions and inputs.

5.1.1 Mathematical Model

In this section, Euler-Bernoulli beam model is used for vibrations of railway tracks. Since tracks rest on ballast which is known to exhibit nonlinear viscoelastic behavior, the Kelvin-Voigt damping will be considered in the model. The semilinear partial differential equation governing the motion of the track $w(\xi,t)$ with initial deflection $w_0(\xi)$ and rate of deflection $v_0(\xi)$ on $\xi \in [0,\ell]$ is [5]

$$\begin{cases} \rho a \frac{\partial^2 w}{\partial t^2} + \frac{\partial}{\partial \xi^2}(EI\frac{\partial^2 w}{\partial \xi^2} + C_d \frac{\partial^3 w}{\partial \xi^2 \partial t}) + \mu \frac{\partial w}{\partial t} + kw + \alpha w^3 = u(\xi,t), \\ w(\xi,0) = w_0(\xi), \quad \frac{\partial w}{\partial t}(\xi,0) = v_0(\xi), \\ w(0,t) = w(\ell,t) = 0, \\ EI\frac{\partial^2 w}{\partial \xi^2}(0,t) + C_d \frac{\partial^3 w}{\partial \xi^2 \partial t}(0,t) = 0, \\ EI\frac{\partial^2 w}{\partial \xi^2}(\ell,t) + C_d \frac{\partial^3 w}{\partial \xi^2 \partial t}(\ell,t) = 0. \end{cases} \tag{5.1}$$

where the positive constants E, I, ρ, a, and ℓ are the modulus of elasticity, second moment of inertia, density of the beam, cross-sectional area, and length of the beam, respectively. The linear and nonlinear parts of the foundation elasticity correspond to the positive coefficients k and α, respectively. The constant $\mu > 0$ is the damping coefficient of the foundation, and $C_d \geq 0$ is the coefficient of Kelvin-Voigt damping in the beam. The external force exerted on the railway track by moving trains, active dampers, or other external force, is denoted by $u(\xi,t)$. The model considered here differs from that in [5] by the inclusion of Kelvin-Voigt damping in the beam if C_d has a non-zero value, although $C_d > 0$ is not assumed in the analysis.

Let $v = \partial w/\partial t$. Define the state space $\mathbb{X} = H^2(0,\ell) \cap H_0^1(0,\ell) \times L^2(0,\ell)$ with norm

$$\|(w,v)\|^2 = \int_0^\ell EIw_{\xi\xi}^2 + kw^2 + \rho a v^2 \, d\xi, \tag{5.2}$$

where the subscript $\cdot_\xi$ denotes the derivative with respect to ξ. Define the closed self-adjoint positive operator

$$\begin{aligned}
&\mathcal{A}_0 w := w_{\xi\xi\xi\xi},\\
&D(\mathcal{A}_0) := \{w \in H^4(0,\ell)|\, w(0) = w(\ell) = 0,\\
&\qquad\qquad w_{\xi\xi}(0) = w_{\xi\xi}(\ell) = 0\},
\end{aligned} \tag{5.3}$$

and also define

$$\mathcal{A}_{KV}(w, v) := \left(v, -\frac{1}{\rho a}\mathcal{A}_0(EIw + C_d v)\right), \tag{5.4}$$

with

$$\begin{aligned}
D(\mathcal{A}_{KV}) := \{&(w, v) \in \mathbb{X}|\, v \in H^2(0,\ell) \cap H_0^1(0,\ell),\\
&EIw + C_d v \in D(\mathcal{A}_0)\}.
\end{aligned} \tag{5.5}$$

Also, let $\boldsymbol{u} \in \mathbb{U} := L^2(0,\ell)$, and define the linear operators $\mathcal{K}$, $\mathcal{A}$, $\mathcal{B}$, and the nonlinear operator $\mathcal{F}(\cdot)$ as

$$\mathcal{K}(w, v) := (0, -\frac{1}{\rho a}(\mu v + kw)), \tag{5.6}$$

$$\mathcal{A} := \mathcal{A}_{KV} + \mathcal{K}, \text{ with } D(\mathcal{A}) = D(\mathcal{A}_{KV}), \tag{5.7}$$

$$\mathcal{B}\boldsymbol{u} := (0, \frac{1}{\rho a}\boldsymbol{u}), \tag{5.8}$$

$$\mathcal{F}(w, v) := (0, -\frac{\alpha}{\rho a}w^3). \tag{5.9}$$

With these definitions and by setting the state $\boldsymbol{x}(t) = (w(\cdot,t), v(\cdot,t))$, initial condition $\boldsymbol{x}_0 = (w_0(\cdot), v_0(\cdot))$, and the input $\boldsymbol{u}(t) = u(\cdot,t)$, the state space representation of the railway track IVP is

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}\boldsymbol{u}(t),\ t > 0,\\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{X}. \end{cases}$$

Notice that the nonlinear term w^3 is in $L^2(0,\ell)$ since $H^2(0,\ell) \subset C([0,\ell])$. Thus, the nonlinear operator $\mathcal{F}(\cdot)$ is well-defined on $\mathbb{X}$. It is also locally Lipschitz continuous; see [37, Lem. 6.1] where it is shown to be continuously Fréchet differentiable on $\mathbb{X}$. Also, the bounded operator $\mathcal{B}$ maps an input $\boldsymbol{u} \in L^2(0,\ell)$ into the state space $\mathbb{X}$. This input space is used in many applications.

It is well known that $\mathcal{A}_{KV}$ generates a strongly continuous contraction semigroup on $\mathbb{X}$; see [21]. The operator $\mathcal{K}$ is a bounded linear operator on $\mathbb{X}$ and so the operator $\mathcal{A}$, with the same domain as $\mathcal{A}_{KV}$, generates a strongly continuous semigroup on $\mathbb{X}$ [104, Cor. 3.2.2]. The assumption that $\mu > 0$ implies that $\mathcal{A}$ generates an exponentially stable semigroup.

Theorem 5.1.1. *[37, Theorem 3.1] Let $\mathcal{A}$ be the infinitesimal generator of a strongly continuous semigroup. If the nonlinear operator $\mathcal{F}(\cdot)$ is locally Lipschitz continuous on $\mathbb{X}$, then for every $\boldsymbol{x}_0 \in \mathbb{X}$ and positive number R, there exist $T > 0$ such that (IVP) admits a unique mild solution $\boldsymbol{x} \in C([0,T];\mathbb{X})$ for all $\boldsymbol{u} \in L^p(0,T;\mathbb{U})$, $\|\boldsymbol{u}\|_{L^p(0,T;\mathbb{U})} \le R$, $1 < p < \infty$.*

Thus, by Theorem 5.1.1, a unique local (in time) mild solution to railway IVP is ensured.

If the input admits further regularity, a mild solution is also a classical solution.

Theorem 5.1.2. *[104, Theorem 6.1.5] Let $\mathcal{A}$ be the infinitesimal generator of a strongly continuous semigroup $\mathcal{T}(t)$ on $\mathbb{X}$. If $\boldsymbol{u} \in C^1([0,T];\mathbb{U})$ and the nonlinear operator $\mathcal{F}(\cdot)$ is continuously Fréchet differentiable on $\mathbb{X}$, then the mild solution of (IVP) with $\boldsymbol{x}_0 \in D(\mathcal{A})$ is a classical solution.*

5.1.2 Stability of Classical Solutions

The comparison function sets $\mathcal{K}$, $\mathcal{K}_\infty$, $\mathscr{L}$, and $\mathscr{K}\mathscr{L}$ are defined as

$$\begin{aligned}
\mathcal{K} :=& \{\gamma : \mathbb{R}^+ \to \mathbb{R}^+ | \, \gamma \text{ is continuous,} \\
& \quad \text{strictly increasing, and } \gamma(0) = 0\}, && (5.10)\\
\mathcal{K}_\infty :=& \{\gamma \in \mathcal{K} | \, \gamma \text{ is unbounded}\}, && (5.11)\\
\mathscr{L} :=& \{\gamma : \mathbb{R}^+ \to \mathbb{R}^+ | \, \gamma \text{ is continuous,} \\
& \quad \text{strictly decreasing, and } \lim_{t\to\infty} \gamma(t) = 0\}, && (5.12)\\
\mathscr{K}\mathscr{L} :=& \{\beta : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+ | \, \beta \text{ is continuous,} \\
& \quad \beta(\cdot, t) \in \mathcal{K}, \beta(r, \cdot) \in \mathscr{L}, \, \forall t \ge 0, \, \forall r > 0\}. && (5.13)
\end{aligned}$$

Let $\boldsymbol{x}$ indicate the state, and $\boldsymbol{u}$ the input. For linear operators $\mathcal{A} : D(\mathcal{A}) \subset \mathbb{X} \to \mathbb{X}$, $\mathcal{B} : \mathbb{U} \to \mathbb{X}$, and possibly nonlinear operator $\mathcal{F}(\cdot) : \mathbb{X} \to \mathbb{X}$, consider the initial value problem

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}\boldsymbol{u}(t), \; t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{X}. \end{cases} \tag{IVP}$$

In the following definitions, it is assumed that a unique mild solution to (IVP) exists for any $\boldsymbol{u} \in PC(\mathbb{R}^+;\mathbb{U})$.

Definition 5.1.3. *[93, Definition 9] (Input-to-state Stability) The (IVP) is input-to-state stable (ISS) if there exist $\beta \in \mathscr{KL}$ and $\gamma \in \mathcal{K}$ such that for all $\boldsymbol{x}_0 \in \mathbb{X}$, $\boldsymbol{u} \in PC(\mathbb{R}^+; \mathbb{U})$, and $t > 0$ the mild solution satisfies*

$$\|\boldsymbol{x}(t)\| \leq \beta(\|\boldsymbol{x}_0\|, t) + \gamma\left(\sup_{t\geq 0} \|\boldsymbol{u}(t)\|_{\mathbb{U}}\right). \tag{5.14}$$

Definition 5.1.4. *[93, Definition 11] (ISS Lyapunov Function) Let $\bar{B}_r$ be a closed ball in $\mathbb{X}$ centered at origin with radius r, and $\bar{B}_{r,pc}$ be a closed ball in $PC(\mathbb{R}^+; \mathbb{U})$ centered at origin with radius r. Also, let the function $\dot{V}_{\boldsymbol{u}}(\boldsymbol{x}_0)$,*

$$\dot{V}_{\boldsymbol{u}}(\boldsymbol{x}_0) \coloneqq \limsup_{t\to 0^+} \frac{1}{t}(V(\boldsymbol{x}(t)) - V(\boldsymbol{x}_0)), \tag{5.15}$$

be the derivative of $V : D(\subset \mathbb{X}) \to \mathbb{R}^+$ along trajectories of (IVP). A continuous function $V : D(\subset \mathbb{X}) \to \mathbb{R}^+$ is an ISS Lyapunov function on D, if there exist ψ_1, $\psi_2 \in \mathcal{K}_\infty$, $\alpha \in \mathcal{K}_\infty$, and $\sigma \in \mathcal{K}$ such that

$$\psi_1(\|\boldsymbol{x}_0\|) \leq V(\boldsymbol{x}_0) \leq \psi_2(\|\boldsymbol{x}_0\|), \tag{5.16}$$

and

$$\dot{V}_{\boldsymbol{u}}(\boldsymbol{x}_0) \leq -\alpha(\|\boldsymbol{x}_0\|) + \sigma\left(\sup_{t\geq 0} \|\boldsymbol{u}(t)\|_{\mathbb{U}}\right), \tag{5.17}$$

for all $\boldsymbol{x}_0 \in \bar{B}_r \subset D$ and $\boldsymbol{u} \in \bar{B}_{r,pc}$.

The existing literature does not predict the existence of a global solution to the railway track PDE model. To investigate the existence and stability of a global solution, a Lyapunov method [30] together with the multiplier method [71] is used. Let c be a constant to be determined. Define, for any $\boldsymbol{x} \in \mathbb{X}$ and number c,

$$V(\boldsymbol{x}) = \int_0^\ell EI(w_{\xi\xi})^2 + kw^2 + \frac{\alpha}{2}w^4 + \rho a v^2 + 2cwv \, d\xi. \tag{5.18}$$

Lemma 5.1.5. *If c in (5.18) satisfies*

$$0 < c < \sqrt{\rho k a}, \tag{5.19}$$

then, there exist positive numbers c_l, c_u, and c_h such that for every $\boldsymbol{x} \in \mathbb{X}$,

$$c_l \|\boldsymbol{x}\|^2 \leq V(\boldsymbol{x}) \leq c_u \|\boldsymbol{x}\|^2 + c_h \|\boldsymbol{x}\|^4. \tag{5.20}$$

Proof. Young's inequality implies that for all $\epsilon_1 > 0$

$$\Big| \int_0^\ell 2cwv \, d\xi \Big| \le c \left(\int_0^\ell \epsilon_1 w^2 + \frac{1}{\epsilon_1} v^2 \, d\xi \right). \tag{5.21}$$

This inequality gives the following lower bound on V:

$$\begin{aligned} V(\boldsymbol{x}) \ge \int_0^\ell & EI(w_{\xi\xi})^2 + (k - c\epsilon_1) w^2 + \frac{\alpha}{2} w^4 \\ & + \left(\rho a - \frac{c}{\epsilon_1} \right) v^2 \, d\xi. \end{aligned} \tag{5.22}$$

Define

$$c_l = \min\{1 - \frac{c\epsilon_1}{k}, 1 - \frac{c}{\rho a \epsilon_1}\}. \tag{5.23}$$

The condition (5.19) on c implies that there exists a number ϵ_1 satisfying $c/\rho a < \epsilon_1 < k/c$, which ensures that $c_l > 0$. The inequality (5.22) can then be re-written as

$$V(\boldsymbol{x}) \ge c_l \left\| \boldsymbol{x} \right\|^2. \tag{5.24}$$

Furthermore, apply the inequality (5.21) to (5.18) and define $c_u = 1 + \max\{c\epsilon_1/k, c/(\epsilon_1 \rho a)\}$ to obtain

$$\begin{aligned} V(\boldsymbol{x}) \le \int_0^\ell & EI w_{\xi\xi}^2 + (k + c\epsilon_1) w^2 + \frac{\alpha}{2} w^4 \\ & + (\rho a + \frac{c}{\epsilon_1}) v^2 \, d\xi \\ \le\ & c_u \left\| \boldsymbol{x} \right\|^2 + \frac{\alpha}{2} \int_0^\ell w^4 \, d\xi. \end{aligned} \tag{5.25}$$

Recall the continuous embedding $H^2(0,\ell) \hookrightarrow L^4(0,\ell)$. Letting c_e be the embedding constant,

$$\begin{aligned} V(\boldsymbol{x}) &\le c_u \left\| \boldsymbol{x} \right\|^2 + \frac{\alpha}{2} \int_0^\ell w^4 \, d\xi \\ &\le c_u \left\| \boldsymbol{x} \right\|^2 + \frac{\alpha c_e}{2} \left(\int_0^\ell w_{\xi\xi}^2 \, d\xi \right)^2 \\ &\le c_u \left\| \boldsymbol{x} \right\|^2 + \frac{\alpha c_e}{2(EI)^2} \left\| \boldsymbol{x} \right\|^4. \end{aligned} \tag{5.26}$$

Set $c_h = \alpha c_e / 2(EI)^2$ in the above inequality to complete the proof. □

The derivative of the Lyapunov function along trajectories of the railway track IVP exists for every $\boldsymbol{x}_0 \in D(\mathcal{A})$ and continuously differentiable input. In the next lemma, if $C_d = 0$, set $1/C_d = \infty$.

Lemma 5.1.6. *Let $[0,T]$ be the interval of existence of the classical solution $\boldsymbol{x}(t)$ to the railway track IVP with $\boldsymbol{x}_0 \in D(\mathcal{A})$ and $\boldsymbol{u} \in C^1([0,T];\mathbb{U})$. Then, the Lyapunov function $V(\boldsymbol{x}(t))$ is differentiable with respect to time. Moreover, let c satisfy*

$$0 < c < \min\left\{\frac{4\rho a EI}{C_d}, \frac{4\rho a k\mu}{\mu^2 + 4\rho a k}, \frac{4\rho a k \mu}{1 + 4\rho a k}\right\}. \tag{5.27}$$

Then, there are positive constants ϵ_3 and ω such that the derivative $\dot{V}(\boldsymbol{x}(t))$ satisfies for all $t \in [0,T]$

$$\dot{V}(\boldsymbol{x}(t)) \leq \epsilon_3 \left\|\boldsymbol{u}(t)\right\|_{\mathbb{U}}^2 - \omega V(\boldsymbol{x}(t)). \tag{5.28}$$

Proof. Since $\boldsymbol{x}_0 \in D(\mathcal{A})$ and $\boldsymbol{u} \in C^1([0,T];\mathbb{U})$, the state $\boldsymbol{x}(t)$ is differentiable on $[0,T]$. The derivative of the Lyapunov function along trajectories of the railway track IVP is

$$\begin{aligned}\dot{V}(\boldsymbol{x}(t)) = 2\int_0^\ell & EI w_{\xi\xi}\dot{w}_{\xi\xi} + kw\dot{w} + \alpha w^3 \dot{w} + \rho a \dot{v} v \\ & + c\dot{w}v + cw\dot{v}\, d\xi. \end{aligned} \tag{5.29}$$

Substituting the time derivatives from the railway track model (5.1) leads to

$$\begin{aligned}\dot{V}(\boldsymbol{x}(t)) = 2\int_0^\ell & EI w_{\xi\xi} v_{\xi\xi} + kwv + \alpha w^3 v \\ & - v((EIw + C_d v)_{\xi\xi\xi\xi} + kw + \mu v + \alpha w^3 + u(\xi,t)) \\ & + cv^2 - \frac{c}{\rho a} w((EIw + C_d v)_{\xi\xi\xi\xi} + kw + \mu v \\ & + \alpha w^3 + u(\xi,t))\, d\xi. \end{aligned} \tag{5.30}$$

Performing repeated integration by parts and using the boundary conditions lead to

$$\begin{aligned}\dot{V}(\boldsymbol{x}(t)) = -2 & \left[\left(v + \frac{c}{\rho a} w\right)(EIw + C_d v)_{\xi\xi\xi} \right. \\ & \left. - \left(v_\xi + \frac{c}{\rho a} w_\xi\right)(EIw + C_d v)_{\xi\xi}\right]_0^\ell \\ & - 2\int_0^\ell \frac{EIc}{\rho a}(w_{\xi\xi})^2 + \frac{kc}{\rho a} w^2 + \frac{\alpha c}{\rho a} w^4 + (\mu - c) v^2 \\ & + C_d v_{\xi\xi}^2\, d\xi - 2\int_0^\ell \frac{\mu c}{\rho a} wv + \frac{C_d c}{\rho a} w_{\xi\xi} v_{\xi\xi}\, d\xi \end{aligned}$$

$$- 2\int_0^\ell u(\xi,t)(v+\frac{c}{\rho a}w)\,d\xi. \tag{5.31}$$

Young's inequalities (such as inequality (5.21)) are used to bound the product terms. Letting ϵ_1, ϵ_2 and ϵ_3 be positive constants,

$$\begin{aligned}\dot{V}(\boldsymbol{x}(t)) \leq \epsilon_3\left\|\boldsymbol{u}(t)\right\|_{\mathbb{U}}^2 - \frac{2}{\rho a}\int_0^\ell &(EI-\frac{C_d\epsilon_2}{2})c(w_{\xi\xi})^2\\ &+(k-\frac{\mu\epsilon_1}{2}-\frac{c}{2\rho a\epsilon_3})c\,w^2+\alpha c\,w^4\\ &+(\rho a\mu-\rho ac-\frac{\mu c}{2\epsilon_1}-\frac{\rho a\epsilon_3}{2})v^2\\ &+(\rho a-\frac{c}{2\epsilon_2})C_d(v_{\xi\xi})^2\,d\xi. \end{aligned} \tag{5.32}$$

Define the constant

$$\begin{aligned}\omega_0=\frac{2c}{\rho a}\min\Bigg\{&1-\frac{C_d\epsilon_2}{2EI},1-\frac{\mu\epsilon_1}{2k}-\frac{c}{2\rho a\epsilon_3 k},\\ &\frac{\mu}{c}-1-\frac{\mu}{2\epsilon_1\rho a}-\frac{\epsilon_3}{2c},\frac{\rho a}{c}-\frac{1}{2\epsilon_2}\Bigg\}.\end{aligned} \tag{5.33}$$

The constant ω_0 needs to be positive. Thus, the constants ϵ_1, ϵ_2, and ϵ_3 are required to satisfy

$$0<\epsilon_1<\frac{2k}{\mu}-\frac{c}{\rho a\epsilon_3\mu}, \tag{5.34}$$

$$\frac{c}{2\rho a}<\epsilon_2<\frac{2EI}{C_d}, \tag{5.35}$$

$$0<\epsilon_3<2\mu-2c-\frac{\mu c}{\epsilon_1\rho a}. \tag{5.36}$$

There is a positive number ϵ_2 satisfying (5.35) if

$$0<c<\frac{4\rho aEI}{C_d}. \tag{5.37}$$

Also, inequalities (5.34) and (5.36) have a solution for ϵ_1 and ϵ_3 if

$$0<c<\min\left\{\frac{4\rho ak\mu}{\mu^2+4\rho ak},\frac{4\rho ak\mu}{1+4\rho ak}\right\}. \tag{5.38}$$

Inequality (5.32) can then be re-written as

$$\dot{V}(\boldsymbol{x}(t)) \leq \epsilon_3 \|\boldsymbol{u}(t)\|_{\mathbb{U}}^2 - \omega_0 \int_0^\ell EI(w_{\xi\xi})^2 + kw^2 + \alpha w^4 + \rho a v^2 + C_d(v_{\xi\xi})^2 \, d\xi. \tag{5.39}$$

Using Young's inequality, an upper bound on the Lyapunov function is

$$V(\boldsymbol{x}(t)) \leq r \int_0^\ell EI(w_{\xi\xi})^2 + kw^2 + \alpha w^4 + \rho a v^2 \, d\xi, \tag{5.40}$$

where

$$r = 1 + \max\{\frac{c\epsilon_1}{k}, \frac{c}{\rho a \epsilon_2}\}. \tag{5.41}$$

Use this upper bound in (5.39)

$$\dot{V}(\boldsymbol{x}(t)) \leq \epsilon_3 \|\boldsymbol{u}(t)\|_{\mathbb{U}}^2 - \frac{\omega_0}{r} V(\boldsymbol{x}(t)). \tag{5.42}$$

Set $\omega = \omega_0/r$ to complete the proof. □

The next theorem uses the Lyapunov function to show that a unique classical solution exists on arbitrary intervals of time for a large class of inputs. It also ensures exponential stability of the solution for some inputs.

Theorem 5.1.7. *Let c_l, c_u, c_h, ω, and ϵ_3 be the same constants as in Lemma 5.1.5 and Lemma 5.1.6. If $\boldsymbol{x}_0 \in D(\mathcal{A})$ and $\boldsymbol{u} \in C^1(\mathbb{R}^+; \mathbb{U})$, then the unique classical solution $\boldsymbol{x}(t)$ to the railway track IVP exists for all $t \geq 0$ and satisfies*

$$\|\boldsymbol{x}(t)\|^2 \leq e^{-\omega t}\left(\frac{c_u}{c_l}\|\boldsymbol{x}_0\|^2 + \frac{c_h}{c_l}\|\boldsymbol{x}_0\|^4\right) + \frac{\epsilon_3}{\omega}(1 - e^{-\omega t}) \max_{s\in[0,t]} \|\boldsymbol{u}(s)\|_{\mathbb{U}}^2. \tag{5.43}$$

Moreover, if there are positive constants u_0 and δ so that $\|\boldsymbol{u}(t)\|_{\mathbb{U}} \leq u_0 e^{-\delta t}$, then $\|\boldsymbol{x}(t)\|$ exponentially decays to zero.

Proof. For every $\bar{T} > 0$, consider the input $\boldsymbol{u}$ over the bounded interval $[0, \bar{T}]$ and define $R := \|\boldsymbol{u}\|_{L^2(0,\bar{T};\mathbb{U})}$. According to Theorem 5.1.1 and 5.1.2, for every $\boldsymbol{x}_0 \in D(\mathcal{A})$ and $\boldsymbol{u} \in C^1([0, \bar{T}]; \mathbb{U})$, with $\|\boldsymbol{u}\|_{L^2(0,\bar{T};\mathbb{U})} \leq R$, there is an interval $[0, T]$, $T = T(\boldsymbol{x}_0, R) \leq \bar{T}$, over which a classical solution to the railway track model (5.1) exists.

Now use Lemma 5.1.6, and apply Grönwall's lemma [121, Theorem 1.4.1] to inequality (5.28) to obtain

$$V(\boldsymbol{x}(t)) \leq e^{-\omega t} V(\boldsymbol{x}_0) + \epsilon_3 \int_0^t e^{-\omega(t-s)} \|\boldsymbol{u}(s)\|_{\mathbb{U}}^2 \, ds, \tag{5.44}$$

for all $t \in [0, T]$. This yields

$$V(\boldsymbol{x}(t)) \leq e^{-\omega t} V(\boldsymbol{x}_0) + \frac{\epsilon_3}{\omega}(1 - e^{-\omega t}) \max_{s \in [0,t]} \|\boldsymbol{u}(s)\|_{\mathbb{U}}^2 . \tag{5.45}$$

From Lemma 5.1.5, it follows that for all $t \in [0, T]$,

$$\begin{aligned} \|\boldsymbol{x}(t)\|^2 \leq e^{-\omega t} \left(\frac{c_u}{c_l} \|\boldsymbol{x}_0\|^2 + \frac{c_h}{c_l} \|\boldsymbol{x}_0\|^4 \right) \\ + \frac{\epsilon_3}{\omega}(1 - e^{-\omega t}) \max_{t \in [0, \bar{T}]} \|\boldsymbol{u}(t)\|_{\mathbb{U}}^2 . \end{aligned} \tag{5.46}$$

This classical solution is of course also a mild solution on $[0, T]$. Note that $\|\boldsymbol{x}(t)\| \leq M$ where M is independent of T. Using [104, Theorem 6.1.4] the solution $\boldsymbol{x}(t)$ can be extended to $[0, \bar{T}]$. Since $\bar{T}$ was arbitrary, the mild solution exists for all $t > 0$. Since $\boldsymbol{x}_0 \in D(\mathcal{A})$ and $\boldsymbol{u} \in C^1(\mathbb{R}^+; \mathbb{U})$, Theorem 5.1.2 then implies that this mild solution is also a classical solution.

Furthermore, if there are positive constants u_0 and δ so that $\|\boldsymbol{u}(t)\|_{\mathbb{U}} \leq u_0 e^{-\delta t}$, inequality (5.44) yields

$$V(\boldsymbol{x}(t)) \leq e^{-\omega t} V(\boldsymbol{x}_0) + \epsilon_3 u_0^2 \begin{cases} \frac{e^{-2\delta t} - e^{-\omega t}}{\omega - 2\delta} & \omega \neq 2\delta \\ e^{-\omega t} & \omega = 2\delta \end{cases} .$$

This shows that the Lyapunov function exponentially decays to zero. Since the Lyapunov function $V(\boldsymbol{x})$ bounds the norm of the state by Lemma 5.1.5, the state will also exponentially decay to zero. □

For inputs with $\sup_{t \geq 0} \|\boldsymbol{u}(t)\|_{\mathbb{U}} < \infty$, Lemma 5.1.5 and Lemma 5.1.6 result in

$$\dot{V}(\boldsymbol{x}(t)) \leq -\omega c_l \|\boldsymbol{x}(t)\|^2 + \epsilon_3 \sup_{t \geq 0} \|\boldsymbol{u}(t)\|_{\mathbb{U}}^2 . \tag{5.47}$$

From Definition 5.1.4, this inequality shows that the Lyapunov function is an ISS Lyapunov function on $D = D(\mathcal{A})$.

5.1.3 Stability of Mild Solutions

If the initial condition is not in $D(\mathcal{A})$ or the input is not continuously differentiable, there may be a unique mild solution to (IVP) even though a classical solution may not exist. In this case, the Lyapunov function from Theorem 5.1.7 may not be differentiable. Thus, exponential decay cannot be shown through manipulating the derivative of Lyapunov function. However, the proof of Theorem 5.1.7 can be modified to yield a result that ensures the existence and stability of global mild solutions for initial conditions in $\mathbb{X}$ and inputs in $L^2_{loc}(0,\infty;\mathbb{U})$.

Theorem 5.1.8. *Let c_l, c_u, c_h, ω, and ϵ_3 be the same constants as in Lemma 5.1.5 and Lemma 5.1.6. If $\boldsymbol{x}_0 \in \mathbb{X}$ and $\boldsymbol{u} \in L^2_{loc}(0,\infty;\mathbb{U})$, then the unique mild solution, $\boldsymbol{x}(t)$, to the railway track IVP exists globally. For every $t > 0$, the mild solution satisfies*

$$\begin{aligned} \|\boldsymbol{x}\|^2_{C(0,t;\mathbb{X})} + \omega \|\boldsymbol{x}\|^2_{L^2(0,t;\mathbb{X})} \leq & \frac{c_u}{c_l}\|\boldsymbol{x}_0\|^2 + \frac{c_h}{c_l}\|\boldsymbol{x}_0\|^4 \\ & + \frac{\epsilon_3}{c_l}\|\boldsymbol{u}\|^2_{L^2(0,t;\mathbb{U})}. \end{aligned} \tag{5.48}$$

Proof. For every $\bar{T} > 0$, consider the input $\boldsymbol{u}$ over the bounded interval $[0,\bar{T}]$ and define $R \coloneqq \|\boldsymbol{u}\|_{L^2(0,\bar{T};\mathbb{U})}$. According to Theorem 5.1.1, for every $\boldsymbol{x}_0 \in \mathbb{X}$ and R, there is an interval $[0,T]$, $T \leq \bar{T}$, over which a unique mild solution $\boldsymbol{x}(t)$ to the railway track IVP exists. Pick a sequence $\boldsymbol{u}_n \in C^1([0,T];\mathbb{U})$, with $\|\boldsymbol{u}_n\|_{L^2(0,T;\mathbb{U})} \leq R$, for all $n \in \mathbb{N}$, that converges to $\boldsymbol{u}$ in $L^2(0,T;\mathbb{U})$. Also, pick a sequence of initial conditions $\boldsymbol{x}_0^n \in D(\mathcal{A})$ that converges to $\boldsymbol{x}_0$ in $\mathbb{X}$. Such sequences always exist since $C^1([0,T],\mathbb{U})$ is dense in $L^2([0,T];\mathbb{U})$, and $D(\mathcal{A})$ is densely embedded in $\mathbb{X}$. Corresponding to each initial condition $\boldsymbol{x}_0^n$ and input $\boldsymbol{u}_n(t)$ is a unique classical solution $\boldsymbol{x}_n(t)$, $t \in [0,T]$, ensured by Theorem 5.1.2. This sequence of solutions also satisfies the equation (2.32) of mild solutions. The mild solution (2.32) is continuous with respect to $\boldsymbol{x}_0 \in \mathbb{X}$ and $\boldsymbol{u} \in L^2(0,T;\mathbb{U})$. See [104, Theorem 6.1.2] for Lipschitz continuity with respect to initial conditions, and [37, Proposition 5.2] for Lipschitz continuity with respect to inputs. It follows that

$$\begin{gathered} \boldsymbol{x}_n \to \boldsymbol{x} \text{ in } C([0,T];\mathbb{X}) \\ \text{as } \boldsymbol{x}_0^n \to \boldsymbol{x}_0 \text{ in } \mathbb{X} \text{ and } \boldsymbol{u}_n \to \boldsymbol{u} \text{ in } L^2(0,T;\mathbb{U}). \end{gathered} \tag{5.49}$$

The Lyapunov function is differentiable for every pair $(\boldsymbol{x}_n(t), \boldsymbol{u}_n(t))$, and from Lemma 5.1.6, its derivative satisfies

$$\dot{V}(\boldsymbol{x}_n(t)) \leq \epsilon_3 \|\boldsymbol{u}_n(t)\|^2_{\mathbb{U}} - \omega V(\boldsymbol{x}_n(t)), \tag{5.50}$$

for all $t \in [0, T]$. Taking the integral yields

$$V(\boldsymbol{x}_n(t)) \leq V(\boldsymbol{x}_0^n) + \epsilon_3 \int_0^t \|\boldsymbol{u}_n(s)\|_{\mathbb{U}}^2 \, ds \\ - \omega \int_0^t V(\boldsymbol{x}_n(s)) \, ds. \tag{5.51}$$

From Lemma 5.1.5, the Lyapunov function satisfies

$$|V(\boldsymbol{x}_2) - V(\boldsymbol{x}_1)| \leq c_u \left| \|\boldsymbol{x}_2\|^2 - \|\boldsymbol{x}_1\|^2 \right| \\ + c_h \left| \|\boldsymbol{x}_2\|^4 - \|\boldsymbol{x}_1\|^4 \right|,$$

for all $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in $\mathbb{X}$. After some manipulation, it follows that

$$|V(\boldsymbol{x}_2) - V(\boldsymbol{x}_1)| \leq \\ \left(c_u + c_h(\|\boldsymbol{x}_2\|^2 + \|\boldsymbol{x}_1\|^2)\right) (\|\boldsymbol{x}_2\| + \|\boldsymbol{x}_1\|) \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|. \tag{5.52}$$

Thus, for every $s \in [0, T]$, the sequence $V(\boldsymbol{x}_n(s))$ converges to $V(\boldsymbol{x}(s))$. The convergence is also uniform; that is, define

$$r = \sup_n \|\boldsymbol{x}_n\|_{C([0,T];\mathbb{X})}, \tag{5.53}$$

find that

$$|V(\boldsymbol{x}_n(s)) - V(\boldsymbol{x}(s))| \leq \max_{s \in [0,T]} |V(\boldsymbol{x}_n(s)) - V(\boldsymbol{x}(s))| \\ \leq \left(2c_u r + 4c_h r^3\right) \|\boldsymbol{x}_n - \boldsymbol{x}\|_{C([0,T];\mathbb{X})}.$$

Thus, by the uniform convergence theorem, the integral in (5.51) converges. These together with (5.49) imply that

$$V(\boldsymbol{x}(t)) + \omega \int_0^t V(\boldsymbol{x}(s)) \, ds \leq V(\boldsymbol{x}_0) + \epsilon_3 \int_0^t \|\boldsymbol{u}(s)\|_{\mathbb{U}}^2 \, ds$$

for all $t \in [0, T]$. Apply Lemma 5.1.5 to this inequality and take the maximum of both side over $[0, t]$ to obtain

$$c_l \|\boldsymbol{x}\|_{C(0,t;\mathbb{X})}^2 + \omega c_l \|\boldsymbol{x}\|_{L^2(0,t;\mathbb{X})}^2 \leq c_u \|\boldsymbol{x}_0\|^2 + c_h \|\boldsymbol{x}_0\|^4 \\ + \epsilon_3 \|\boldsymbol{u}\|_{L^2(0,t;\mathbb{U})}^2. \tag{5.54}$$

This inequality shows that the mild solution can be extended to the interval $[0, \bar{T}]$ [104, Theorem 6.1.4]. Since $\bar{T}$ was arbitrary, the mild solution exists globally. □

For inputs in $PC(\mathbb{R}^+;\mathbb{U})$, a similar density argument can be applied to prove the ISS of railway track IVP.

Corollary 5.1.9. *The railway track IVP is input-to-state stable (ISS) in the sense of Definition 5.1.3.*

Proof. For every $T > 0$, $PC([0,T];\mathbb{U}) \subset L^2_{loc}(0,\infty;\mathbb{U})$; thus, Theorem 5.1.8 ensures that a unique mild solution $\boldsymbol{x}(t)$ exists for all inputs $\boldsymbol{u}$ in $PC([0,T];\mathbb{U})$. Consider a sequence of initial conditions $\boldsymbol{x}_0^n \in D(\mathcal{A})$ converging to $\boldsymbol{x}_0$ in $\mathbb{X}$, and also a sequence of inputs $\boldsymbol{u}_n \in C^1([0,T];\mathbb{U})$ converging uniformly to $\boldsymbol{u} \in PC([0,T];\mathbb{U})$. Let $\boldsymbol{x}_n(t)$ be the classical solution to railway track IVP with initial condition $\boldsymbol{x}_0^n$ and input $\boldsymbol{u}_n(t)$. This solution also satisfies (2.32) which ensures that

$$\begin{aligned} &\boldsymbol{x}_n \to \boldsymbol{x} \text{ in } C([0,T];\mathbb{X}) \\ \text{as } &\boldsymbol{x}_0^n \to \boldsymbol{x}_0 \text{ in } \mathbb{X} \text{ and } \boldsymbol{u}_n(t) \to \boldsymbol{u}(t) \text{ uniformly.} \end{aligned} \tag{5.55}$$

See [104, Theorem 6.1.2] for Lipschitz continuity with respect to initial conditions, and [37, Proposition 5.2] for Lipschitz continuity with respect to inputs. Use Theorem 5.1.7 to obtain

$$\begin{aligned} \|\boldsymbol{x}_n(t)\|^2 \le e^{-\omega t}\left(\frac{c_u}{c_l}\|\boldsymbol{x}_0^n\|^2 + \frac{c_h}{c_l}\|\boldsymbol{x}_0^n\|^4\right) \\ + \frac{\epsilon_3}{\omega}\max_{t\in[0,T]}\|\boldsymbol{u}_n(t)\|^2_{\mathbb{U}}. \end{aligned} \tag{5.56}$$

This inequality continuously depends on the norm of initial conditions and inputs. Taking the limit yields a similar inequality for $\boldsymbol{x}(t)$ with $\boldsymbol{x}_0$ and $\boldsymbol{u}(t)$ replaced. Knowing that $\sup_{t\ge 0}\|\boldsymbol{u}(t)\|_{\mathbb{U}} < \infty$, the ISS property in Definition 5.1.3 follows immediately. □

5.2 Optimal Control and Actuator Location

In this section, we apply the results of previous sections to compute an optimal control and actuator location for the vibration suppression of the track. As discussed in chapter 3, the problem of finding the best control and actuator location is the optimization problem

$$\begin{cases} \min & J(\boldsymbol{u},\boldsymbol{r};\boldsymbol{x}_0) \\ \text{s.t.} & \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad \forall t \in (0,T], \\ & \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases} \tag{P}$$

The fitst order optimality conditions were derived in chapter 3. The optimality conditions use the derivative of the cost function with respect to the input and the actuator location. In that, the adjoint of the IVP needs to be calculated. The adjoint equation is the following final value problem (FVP):

$$\dot{\boldsymbol{p}}(t) = -(\mathcal{A}^* + \mathcal{F}'^*_{\boldsymbol{x}(t)})\boldsymbol{p}(t) - \mathcal{Q}\boldsymbol{x}(t), \quad \boldsymbol{p}(T) = 0 \tag{FVP}$$

For every $\boldsymbol{x}_0 \in \mathbb{X}$, the derivatives of the cost function with respect to $\boldsymbol{u}$ and $\boldsymbol{r}$ evaluated at $\boldsymbol{u} \in \text{int}(U_{ad})$, $\boldsymbol{r} \in \text{int}(K_{ad})$ are linear operators $D_{\boldsymbol{u}}J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ and $D_{\boldsymbol{r}}J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$, respectively. Identifying each operator with an element of its underlying space, they are derived as

$$D_{\boldsymbol{u}}J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) = \mathcal{B}^*(\boldsymbol{r})\boldsymbol{p}(t) + \mathcal{R}\boldsymbol{u}(t), \tag{5.57a}$$

$$D_{\boldsymbol{r}}J(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_0) = \int_0^T (\mathcal{B}'_{\boldsymbol{r}}\boldsymbol{u}(s))^*\boldsymbol{p}(s)\, ds. \tag{5.57b}$$

5.2.1 Optimization Algorithms

Several optimization algorithms were suggested in the literature for solving PDE-constrained optimization problems, see [59]. In this section, two common optimization algorithms for solving P will be discussed. These are projected gradient method and nonlinear conjugate gradient method. In projected gradient (or steepest descent) method, the negative of the gradient is considered as the search direction. This algorithm reads as follows:

Algorithm 1. *(Projected Gradient Method)*

1. ***input*** *initial guesses* $\boldsymbol{u}_1 \in U_{ad}$ *and* $\boldsymbol{r}_1 \in K_{ad}$
2. *Set* $n = 1$
3. ***while*** *a criteria is met* ***do***
4. *Solve the IVP with* $\boldsymbol{u}_n$ *and* $\boldsymbol{r}_n$*, and find* $\boldsymbol{x}_n$
5. *Solve the FVP with* $\boldsymbol{x}_n$*, and find* $\boldsymbol{p}_n$
6. *Evaluate* $d_n^{\boldsymbol{u}} := -D_{\boldsymbol{u}}J(\boldsymbol{u}_n, \boldsymbol{r}_n, \boldsymbol{x}_0)$ *and* $d_n^{\boldsymbol{r}} := -D_{\boldsymbol{r}}J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0)$ *from (5.67) and (5.68)*
7. *Obtain step lengths* $s_n^{\boldsymbol{u}}$ *and* $s_n^{\boldsymbol{r}}$ *by one of the line search methods discussed below*

8. *Set* $\boldsymbol{u}_{n+1} = \boldsymbol{u}_n + s_n^{\boldsymbol{u}} d_n^{\boldsymbol{u}}$ *and* $\boldsymbol{r}_{n+1} = \boldsymbol{r}_n + s_n^r d_n^{\boldsymbol{r}}$

9. *increase* n

10. ***end while***

Projected gradient method is typically converging to an optimizer slowly, whereas the nonlinear conjugate gradient method promises faster convergence [101]. The nonlinear conjugate gradient method reads as follows:

Algorithm 2. *(Nonlinear Conjugate Gradient Method)*

1. ***input*** *initial guesses* $\boldsymbol{u}_1 \in U_{ad}$ *and* $\boldsymbol{r}_1 \in K_{ad}$

2. *Set* $n = 1$

3. *Set* $d_n^{\boldsymbol{u}} = h_n^{\boldsymbol{u}} := -D_{\boldsymbol{u}} J(\boldsymbol{u}_n, \boldsymbol{r}_n, \boldsymbol{x}_0)$ *and* $d_n^{\boldsymbol{r}} = h_n^{\boldsymbol{r}} := -D_{\boldsymbol{r}} J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0)$

4. ***while*** *a criteria is met* ***do***

5. *Solve the IVP for* $\boldsymbol{u}_n$ *and* $\boldsymbol{r}_n$*, and find* $\boldsymbol{x}_n$

6. *Solve the FVP for* $\boldsymbol{x}_n$*, and find* $\boldsymbol{p}_n$

7. *Obtain step lengths* $s_n^{\boldsymbol{u}}$ *and* $s_n^{\boldsymbol{r}}$ *using, e.g., secant method in (5.64a)*

8. *Set* $\boldsymbol{u}_{n+1} = \boldsymbol{u}_n + s_n^{\boldsymbol{u}} d_n^{\boldsymbol{u}}$ *and* $\boldsymbol{r}_{n+1} = \boldsymbol{r}_n + s_n^{\boldsymbol{r}} d_n^{\boldsymbol{r}}$

9. *Solve the IVP for* $\boldsymbol{u}_{n+1}$ *and* $\boldsymbol{r}_{n+1}$*, and find* $\boldsymbol{x}_{n+1}$

10. *Solve the FVP for* $\boldsymbol{x}_{n+1}$*, and find* $\boldsymbol{p}_{n+1}$

11. *Evaluate* $h_{n+1}^{\boldsymbol{u}} := -D_{\boldsymbol{u}} J(\boldsymbol{u}_{n+1}, \boldsymbol{r}_{n+1}, \boldsymbol{x}_0)$ *and* $h_{n+1}^{\boldsymbol{r}} := -D_{\boldsymbol{r}} J(\boldsymbol{u}_{n+1}, \boldsymbol{r}_{n+1}; \boldsymbol{x}_0)$ *from (5.67) and (5.68)*

12. *Determine step lengths* $\beta_{n+1}^{\boldsymbol{u}}$ *and* $\beta_{n+1}^{\boldsymbol{r}}$ *using, e.g., Fletcher-Reeves or Polak-Ribière formula [101, Section 5.2]*

13. *Set* $d_{n+1}^{\boldsymbol{u}} := h_{n+1}^{\boldsymbol{u}} + \beta_{n+1}^{\boldsymbol{u}} d_n^{\boldsymbol{u}}$ *and* $d_{n+1}^{\boldsymbol{r}} := h_{n+1}^{\boldsymbol{r}} + \beta_{n+1}^{\boldsymbol{r}} d_n^{\boldsymbol{r}}$

14. ***if*** $\left\langle d_{n+1}^{\boldsymbol{u}}, h_{n+1}^{\boldsymbol{u}} \right\rangle_{L^2(0,T;\mathbb{U})} \leq 0$ ***then***

15. *Set* $d_{n+1}^{\boldsymbol{u}} = h_{n+1}^{\boldsymbol{u}}$

end if

if $\left\langle d^{\boldsymbol{r}}_{n+1}, h^{\boldsymbol{r}}_{n+1} \right\rangle_{\mathbb{K}} \leq 0$ ***then***

Set $d^{\boldsymbol{r}}_{n+1} = h^{\boldsymbol{r}}_{n+1}$

end if

increase n

end while

Several choices exist for selecting the step length $\beta^{\boldsymbol{u}}_{n+1}$ (similarly $\beta^{\boldsymbol{r}}_{n+1}$) of the previous algorithm [57]. Letting $\gamma^{\boldsymbol{u}}_{n+1} = h^{\boldsymbol{u}}_{n+1} - h^{\boldsymbol{u}}_{n}$, the following are for selecting the step length $\beta^{\boldsymbol{u}}_{n+1}$ (similarly $\beta^{\boldsymbol{r}}_{n+1}$) considered in this paper

$$\text{Fletcher-Reeves:} \quad \beta^{\boldsymbol{u}}_{n+1} = \frac{\left\| h^{\boldsymbol{u}}_{n+1} \right\|_{\mathbb{U}}}{\left\| h^{\boldsymbol{u}}_{n} \right\|_{\mathbb{U}}}, \tag{5.58a}$$

$$\text{Polan-Ribière:} \quad \beta^{\boldsymbol{u}}_{n+1} = \frac{\left\langle h^{\boldsymbol{u}}_{n+1}, \gamma^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}}{\left\| h^{\boldsymbol{u}}_{n} \right\|_{\mathbb{U}}}, \tag{5.58b}$$

$$\text{Hestenes-Stiefel:} \quad \beta^{\boldsymbol{u}}_{n+1} = \frac{\left\langle h^{\boldsymbol{u}}_{n+1}, \gamma^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}}{\left\langle d^{\boldsymbol{u}}_{n}, \gamma^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}}. \tag{5.58c}$$

A new formula was also proposed by Hager and Zhang [56]. Define $\bar{\beta}^{\boldsymbol{u}}_{n+1}$ and $\eta^{\boldsymbol{u}}_{n+1}$ as

$$\begin{aligned} \bar{\beta}^{\boldsymbol{u}}_{n+1} &= \frac{\left\langle \gamma^{\boldsymbol{u}}_{n+1} - 2 \frac{\left\| \gamma^{\boldsymbol{u}}_{n+1} \right\|^2_{\mathbb{U}}}{\left\langle d^{\boldsymbol{u}}_{n}, \gamma^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}} d^{\boldsymbol{u}}_{n}, h^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}}{\left\langle d^{\boldsymbol{u}}_{n+1}, \gamma^{\boldsymbol{u}}_{n+1} \right\rangle_{\mathbb{U}}}, \\ \eta^{\boldsymbol{u}}_{n+1} &= -\frac{1}{\left\| d^{\boldsymbol{u}}_{n} \right\|_{\mathbb{U}}} \min \left\{ 0.01, \left\| h^{\boldsymbol{u}}_{n} \right\|_{\mathbb{U}} \right\}. \end{aligned}$$

Then, the formula is

$$\text{Hager-Zhang:} \quad \beta^{\boldsymbol{u}}_{n+1} = \max \left\{ \bar{\beta}^{\boldsymbol{u}}_{n+1}, \eta^{\boldsymbol{u}}_{n+1} \right\} \tag{5.59}$$

Furthermore, several schemes have been proposed to choose the step length $s^{\boldsymbol{u}}_{n}$ (similarly $s^{\boldsymbol{r}}_{n}$) in each iteration of previous algorithms including bisection, (strong) Wolfe conditions, Secant method.

1. **Bisection [14]:** In each iteration n of the algorithms, initialize $s^{\boldsymbol{u}}_{n,1}$ and $s^{\boldsymbol{r}}_{n,1}$. Set $m = 1$. Compute $\boldsymbol{u}_{n,m} = \boldsymbol{u}_n + s^{\boldsymbol{u}}_{n,m} d^{\boldsymbol{u}}_n$, $\boldsymbol{r}_{n,m} = \boldsymbol{r}_n + s^{\boldsymbol{r}}_{n,m} d^{\boldsymbol{r}}_n$, and $J(\boldsymbol{u}_{n,m}, \boldsymbol{r}_{n,m}; \boldsymbol{x}_0)$. If
$$J(\boldsymbol{u}_{n,m}, \boldsymbol{r}_{n,m}; \boldsymbol{x}_0) \leq J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0), \tag{5.60}$$
then accept the step size; otherwise, set $s^{\boldsymbol{u}}_{n,m+1} = \frac{1}{2} s^{\boldsymbol{u}}_{n,m}$ and $s^{\boldsymbol{r}}_{n,m+1} = \frac{1}{2} s^{\boldsymbol{r}}_{n,m}$ and repeat the process.

2. **Wolfe conditions [101, Section 3.1]:** In each iteration n of the algorithms, initialize $s^{\boldsymbol{u}}_{n,1}$ and $s^{\boldsymbol{r}}_{n,1}$. Set $m = 1$. Pick constants c_1 and c_2 in the interval $(0, 1)$. Compute $\boldsymbol{u}_{n,m} = \boldsymbol{u}_n + s^{\boldsymbol{u}}_{n,m} d^{\boldsymbol{u}}_n$ and $\boldsymbol{r}_{n,m} = \boldsymbol{r}_n + s^{\boldsymbol{r}}_{n,m} d^{\boldsymbol{r}}_n$ together with $h^{\boldsymbol{u}}_{n,m} = -D_{\boldsymbol{u}} J(\boldsymbol{u}_{n,m}, \boldsymbol{r}_{n,m}; \boldsymbol{x}_0)$ and $h^{\boldsymbol{r}}_{n,m} = -D_{\boldsymbol{r}} J(\boldsymbol{r}_{n,m}, \boldsymbol{r}_{n,m}; \boldsymbol{x}_0)$. Iterate the step size $s^{\boldsymbol{u}}_{n,m}$ and $s^{\boldsymbol{r}}_{n,m}$ until the following conditions are met
$$J(\boldsymbol{u}_{n,m}, \boldsymbol{r}_n; \boldsymbol{x}_0) \leq J(\boldsymbol{u}_n, \boldsymbol{r}_n; \boldsymbol{x}_0) + c_1 s^{\boldsymbol{u}}_{n,m} \left\langle h^{\boldsymbol{u}}_{n,m}, d^{\boldsymbol{u}}_n \right\rangle_{\mathbb{U}} + c_1 s^{\boldsymbol{r}}_{n,m} \left\langle h^{\boldsymbol{r}}_{n,m}, d^{\boldsymbol{r}}_n \right\rangle_{\mathbb{K}}, \tag{5.61a}$$
$$\left\langle h^{\boldsymbol{u}}_{n,m}, d^{\boldsymbol{u}}_n \right\rangle_{\mathbb{U}} \geq c_2 \left\langle h^{\boldsymbol{u}}_n, d^{\boldsymbol{u}}_n \right\rangle_{\mathbb{U}}, \tag{5.61b}$$
$$\left\langle h^{\boldsymbol{r}}_{n,m}, d^{\boldsymbol{r}}_n \right\rangle_{\mathbb{K}} \geq c_2 \left\langle h^{\boldsymbol{r}}_n, d^{\boldsymbol{r}}_n \right\rangle_{\mathbb{K}}. \tag{5.61c}$$

3. **Strong Wolfe conditions [101, Section 3.1]:** Similar to Wolfe conditions except that condition (5.61) is replaced with
$$\left| \left\langle h^{\boldsymbol{u}}_{n,m+1}, d^{\boldsymbol{u}}_{n,m} \right\rangle_{\mathbb{U}} \right| \leq c_2 \left| \left\langle h^{\boldsymbol{u}}_{n,m}, d^{\boldsymbol{u}}_{n,m} \right\rangle_{\mathbb{U}} \right|, \tag{5.62a}$$
$$\left| \left\langle h^{\boldsymbol{r}}_{n,m+1}, d^{\boldsymbol{r}}_{n,m} \right\rangle_{\mathbb{K}} \right| \leq c_2 \left| \left\langle h^{\boldsymbol{r}}_{n,m}, d^{\boldsymbol{r}}_{n,m} \right\rangle_{\mathbb{K}} \right|. \tag{5.62b}$$

4. **Secant method:** The step lengths can be approximate minimizers of the function $\theta(s^{\boldsymbol{u}}, s^{\boldsymbol{r}}) := J(\boldsymbol{u}_n + s^{\boldsymbol{u}} d^{\boldsymbol{u}}_n, \boldsymbol{r}_n + s^{r} d^{\boldsymbol{r}}_n; \boldsymbol{x}_0)$. For instance, letting $\sigma^{\boldsymbol{u}}$ and $\sigma^{\boldsymbol{r}}$ be some positive constants, an approximate minimizer of $\theta(s^{\boldsymbol{u}}, s^{\boldsymbol{r}})$ can be derived by using secant formula as
$$s^{\boldsymbol{u}} = \frac{\theta_{s^{\boldsymbol{u}}}(0,0)}{\theta_{s^{\boldsymbol{u}}}(\sigma^{\boldsymbol{u}}, 0) - \theta_{s^{\boldsymbol{u}}}(0,0)}, \; s^{\boldsymbol{r}} = \frac{\theta_{s^{\boldsymbol{r}}}(0,0)}{\theta_{s^{\boldsymbol{r}}}(0, \sigma^{\boldsymbol{r}}) - \theta_{s^{\boldsymbol{r}}}(0,0)}, \tag{5.63}$$
where the subscripts indicate partial derivatives. In the first iteration, the constants $\sigma^{\boldsymbol{u}}$ and $\sigma^{\boldsymbol{r}}$ are chosen arbitrary; in next iterations, they are set to the values of $s^{\boldsymbol{u}}$ and $s^{\boldsymbol{r}}$ found in the previous iteration [59]. Accordingly, from the definition of $\theta(s^{\boldsymbol{u}}, s^{\boldsymbol{r}})$, and by arbitrary initializing $s^{\boldsymbol{u}}_0$ and $s^{\boldsymbol{r}}_0$, it follow that
$$s^{\boldsymbol{u}}_n = -\frac{\langle h^{\boldsymbol{u}}_n, d^{\boldsymbol{u}}_n \rangle_{\mathbb{U}}}{\left\langle h^{\boldsymbol{u}}_n + D_{\boldsymbol{u}} J(\boldsymbol{u}_n + s^{\boldsymbol{u}}_{n-1} d^{\boldsymbol{u}}_n, \boldsymbol{r}_n; \boldsymbol{x}_0), d^{\boldsymbol{u}}_n \right\rangle_{\mathbb{U}}}, \tag{5.64a}$$
$$s^{\boldsymbol{r}}_n = -\frac{\langle h^{\boldsymbol{r}}_n, d^{\boldsymbol{r}}_n \rangle_{\mathbb{K}}}{\left\langle h^{\boldsymbol{r}}_n + D_{\boldsymbol{r}} J(\boldsymbol{u}_n, \boldsymbol{r}_n + s^{\boldsymbol{r}}_{n-1} d^{\boldsymbol{r}}_n; \boldsymbol{x}_0), d^{\boldsymbol{u}}_n \right\rangle_{\mathbb{K}}}. \tag{5.64b}$$

5.2.2 Approximation Scheme

In each iteration, the IVP and FVP are solved numerically. To find a numerical solution to the IVP and FVP, a finite-dimensional approximation is needed. Let n_z, n_u, and n_r indicate the dimension of finite-dimensional subspaces of $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$, respectively. To avoid multiple subscripts, let $n = [n_z, n_u, n_r]$, and denote the subspaces by $\mathbb{X}_n$, $\mathbb{U}_n$, and $\mathbb{K}_n$. Also, let $\mathcal{P}_z : \mathbb{X} \to \mathbb{X}_n$, $\mathcal{P}_u : \mathbb{U} \to \mathbb{U}_n$, and $\mathcal{P}_r : \mathbb{K} \to \mathbb{K}_n$ be the projection of $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$ onto $\mathbb{X}_n$, $\mathbb{U}_n$, and $\mathbb{K}_n$, respectively. Define sets $K_{ad,n}$ and $U_{ad,n}$ in a similar way. For every $\boldsymbol{r} \in K_{ad,n}$, consider the finite-dimensional linear operators $\mathcal{A}_n \in \mathcal{L}(\mathbb{X}_n)$ and $\mathcal{B}_n(r) \in \mathcal{L}(\mathbb{U}_n, \mathbb{X}_n)$, and $\mathcal{Q}_n \coloneqq \mathcal{Q}|_{\mathbb{X}_n}$.

There are different techniques to handle the nonlinear operators $\mathcal{F}(\cdot)$, $\mathcal{F}'_{\cdot}$, $\mathcal{B}(\cdot)$, and $\mathcal{B}'_{\cdot}$. A common way is to approximate the nonlinear operator $\mathcal{F}(\cdot)$ with an operator $\mathcal{F}_n(\cdot)$ that coincides with $\mathcal{F}(\cdot)$ on $\mathbb{X}_n$. The operators $\mathcal{F}'_{\cdot}$, $\mathcal{B}(\cdot)$, and $\mathcal{B}'_{\cdot}$ can be approximated in similar ways.

Then, the approximated IVP and FVP are governed by

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \mathcal{A}_n \boldsymbol{x}(t) + \mathcal{F}_n(\boldsymbol{x}(t)) + \mathcal{B}_n(\boldsymbol{r})\boldsymbol{u}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_{n0} \coloneqq \mathcal{P}_n \boldsymbol{x}_0, \\ \dot{\boldsymbol{p}}(t) &= -(\mathcal{A}_n^* + \mathcal{F}'^*_{n,\boldsymbol{x}(t)})\boldsymbol{p}(t) - \mathcal{Q}_n \boldsymbol{x}(t), \qquad \boldsymbol{p}(T) = 0. \end{aligned} \tag{5.65}$$

For the optimality conditions, the operator $(\mathcal{B}'_{n,\boldsymbol{r}}\boldsymbol{u})^* : \mathbb{X}_n \to \mathbb{K}_n$ is defined by a sesquilinear form

$$\left\langle (\mathcal{B}'_{n,\boldsymbol{r}}\boldsymbol{u})^* \boldsymbol{p}, \boldsymbol{r} \right\rangle_{\mathbb{K}} = \left\langle \boldsymbol{p}, (\mathcal{B}'_{n,\boldsymbol{r}}\boldsymbol{r})\boldsymbol{u} \right\rangle, \quad \forall (\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{r}) \in \mathbb{U}_n \times \mathbb{X}_n \times \mathbb{K}_n. \tag{5.66}$$

Then, letting $\mathcal{R}_n \coloneqq \mathcal{R}|_{\mathbb{X}_n}$, the approximated optimality conditions are

$$D_{\boldsymbol{u}} J_n(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_{n0}) = \mathcal{B}_n^*(\boldsymbol{r})\boldsymbol{p}(t) + \mathcal{R}_n \boldsymbol{u}(t), \tag{5.67}$$

$$D_{\boldsymbol{r}} J_n(\boldsymbol{u}, \boldsymbol{r}, \boldsymbol{x}_{n0}) = \int_0^T (\mathcal{B}'_{n,\boldsymbol{r}}\boldsymbol{u}(s))^* \boldsymbol{p}(s)\, ds. \tag{5.68}$$

These operators should satisfy a set of assumption to be qualified as an approximation of the operators in the original system. Assumptions A1-A3 in [95] ensures this for a linear system with infinite horizon cost function.

By means of a basis for the underlying spaces, the approximation can be fully realized. This yields an approximation that satisfies assumptions A1-A3 in [95]. Using $i \in \mathbb{N}_n$ to enumerate the bases, let $\boldsymbol{e}_i^{\mathbb{X}}$, $\boldsymbol{e}_i^{\mathbb{U}}$, and $\boldsymbol{e}_i^{\mathbb{K}}$ denote orthonormal bases of $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{K}$, respectively. It is assumed that $\boldsymbol{e}_i^{\mathbb{X}} \in D(\mathcal{A})$ for all $i \in \mathbb{N}_n$. Denoted by x_i and p_i are projections of the state $\boldsymbol{x}$ and adjoint state $\boldsymbol{p}$ onto the one-dimensional subspaces spanned by $\boldsymbol{e}_i^{\mathbb{X}}$. Define u_i and r_i in a similar way. Consider the vectors $\underset{\sim}{x}$, $\underset{\sim}{u}$, $\underset{\sim}{r}$, and $F(\underset{\sim}{x})$ as

$$\underset{\sim}{x} \coloneqq \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \underset{\sim}{p} \coloneqq \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}, \quad \underset{\sim}{u} \coloneqq \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \quad \underset{\sim}{r} \coloneqq \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \tag{5.69}$$

$$\underset{\sim}{x}_0 \coloneqq \begin{bmatrix} \langle \boldsymbol{x}_0, \boldsymbol{e}_1^{\mathbb{X}} \rangle \\ \vdots \\ \langle \boldsymbol{x}_0, \boldsymbol{e}_n^{\mathbb{X}} \rangle \end{bmatrix}, \quad F(\underset{\sim}{x}) \coloneqq \begin{bmatrix} \langle \mathcal{F}(\boldsymbol{x}), \boldsymbol{e}_1^{\mathbb{X}} \rangle \\ \vdots \\ \langle \mathcal{F}(\boldsymbol{x}), \boldsymbol{e}_n^{\mathbb{X}} \rangle \end{bmatrix}, \tag{5.70}$$

and matrices A_n, Q_n, and $dF_n(\underset{\sim}{x})$ with ith row and jth column, $i, j = 1, ..., n_z$, as

$$A_{ij} \coloneqq \langle \mathcal{A}\boldsymbol{e}_j^{\mathbb{X}}, \boldsymbol{e}_i^{\mathbb{X}} \rangle, \quad Q_{ij} \coloneqq \langle \mathcal{Q}\boldsymbol{e}_j^{\mathbb{X}}, \boldsymbol{e}_i^{\mathbb{X}} \rangle, \quad dF_{ij}(\underset{\sim}{x}) \coloneqq \langle \mathcal{F}'_{\boldsymbol{x}}\boldsymbol{e}_j^{\mathbb{X}}, \boldsymbol{e}_i^{\mathbb{X}} \rangle, \tag{5.71}$$

matrix $B(\underset{\sim}{r})$ with $i = 1, ..., n_z$ and $j = 1, ..., n_u$ as

$$B_{ij}(\underset{\sim}{r}) \coloneqq \langle \mathcal{B}(\boldsymbol{r})\boldsymbol{e}_j^{\mathbb{U}}, \boldsymbol{e}_i^{\mathbb{X}} \rangle. \tag{5.72}$$

The superscript * will denote conjugate transpose, $A^* = \bar{A}^T$. A finite-dimensional state-space representation of the approximated IVP and FVP is

$$\begin{aligned} \dot{\underset{\sim}{x}}(t) &= A_n \underset{\sim}{x}(t) + F_n(\underset{\sim}{x}(t)) + B_n(\underset{\sim}{r})\underset{\sim}{u}(t), & \underset{\sim}{x}(0) &= \underset{\sim}{x}_0, \\ \dot{\underset{\sim}{p}}(t) &= -(A_n^* + dF_n^*(\underset{\sim}{x}(t)))\underset{\sim}{p}(t) - Q_n \underset{\sim}{x}(t), & \underset{\sim}{p}(T) &= 0. \end{aligned} \tag{5.73}$$

Also, define the matrix R_n with $i, j = 1, ..., n_u$ as

$$R_{ij} \coloneqq \langle \mathcal{R}\boldsymbol{e}_j^{\mathbb{U}}, \boldsymbol{e}_i^{\mathbb{U}} \rangle, \tag{5.74}$$

The optimality condition (5.67) becomes

$$D_{\boldsymbol{u}} J_n(\underset{\sim}{u}, \underset{\sim}{r}; \underset{\sim}{x}_0) = B^*(\underset{\sim}{r})\underset{\sim}{p}(t) + R\underset{\sim}{u}(t). \tag{5.75}$$

To write the optimality condition (5.68) in a vector form, use the sesquilinear form (5.66) together with Corollary 5.8 in [37], let $i = 1, ..., n_u$, $j = 1, ..., n_r$, and $k = 1, ..., n_z$, define the array $dB_n(\underset{\sim}{r})$ as

$$dB_{ijk}(\underset{\sim}{r}) = \langle b'_{i,\boldsymbol{r}} e_j^{\mathbb{K}}, e_k^{\mathbb{Z}} \rangle, \tag{5.76}$$

see [37] for the definition of $b'_{j,\boldsymbol{r}}$. This optimality condition becomes

$$D_{\boldsymbol{r}} J_{n,j}(\underset{\sim}{u}, \underset{\sim}{r}, \underset{\sim}{x}_0) = \int_0^T \sum_{i=1}^{n_u} \sum_{k=1}^{n_z} \bar{u}_i(s) dB_{ijk}(\underset{\sim}{r}) p_k(s)\, ds. \tag{5.77}$$

Knowing the components of the above optimality condition, this equation is compactly written as

$$D_{\boldsymbol{r}} J_n(\underset{\sim}{u}, \underset{\sim}{r}, \underset{\sim}{x}_0) = \int_0^T \underset{\sim}{u}^*(s) dB_n(\underset{\sim}{r}) \underset{\sim}{p}(s)\, ds. \tag{5.78}$$

5.2.3 Simulation Results

A basis is chosen for the numerical simulation. Let

$$c_i = \sqrt{\frac{2\ell^3 \pi^4 i^4}{EI\pi^4 i^4 + k\ell^4 + \rho a \ell^4 \pi^4 i^4}}. \tag{5.79}$$

It is straightforward to show that the following sequence forms an orthonormal basis for $\mathbb{X}$:

$$\boldsymbol{e}^{\mathbb{X}}_{2i-1} = \left(\frac{c_i}{\pi^2 i^2} \sin(\frac{\pi i}{\ell}\xi), c_i \sin(\frac{\pi i}{\ell}\xi) \right), \quad i \in \mathbb{N}, \tag{5.80a}$$

$$\boldsymbol{e}^{\mathbb{X}}_{2i} = \left(-\frac{c_i}{\pi^2 i^2} \sin(\frac{\pi i}{\ell}\xi), c_i \sin(\frac{\pi i}{\ell}\xi) \right), \quad i \in \mathbb{N}. \tag{5.80b}$$

Set $\mathcal{Q}(w, v) = (c_w w, c_v v)$ and $\mathcal{R} = c_u$ in the cost function for some positive constants c_w, c_v, and c_u. Then, the FVP and the optimality conditions (5.67) and (5.68) follow immediately. Letting $i = 1, ..., 2n$, and

$$\begin{aligned} a_i^1 &= -\tfrac{\mu l^4 + C_d \pi^4 i^4}{2l^3} c_i^2, & a_i^2 &= \tfrac{kl^4 + EI\pi^4 i^4}{l^3 i^2 \pi^2} c_i^2 \\ q_i^1 &= c_i(c_v - c_w)\rho a \ell^4 n^4 \pi^4 + c_w/c_i^2, & q_i^2 &= c_i(c_v + c_w)\rho a \ell^4 n^4 \pi^4 - c_w/c_i^2, \end{aligned}$$

the matrices A_n, Q_n, and B_n are calculated as

$$\begin{aligned} A_n =& \text{blockdiag} \begin{bmatrix} a_i^1 & a_i^1 + a_i^2 \\ a_i^1 - a_i^2 & a_i^1 \end{bmatrix}, \\ Q_n =& \text{blockdiag} \begin{bmatrix} q_i^1 & q_i^2 \\ q_i^2 & q_i^1 \end{bmatrix}, \end{aligned}$$

$$B_{(2i-1)1} = B_{(2i)1} = c_i \int_0^\ell b(\xi; r) \sin(\frac{\pi i}{\ell}\xi)\, d\xi.$$

The components of the nonlinear operator $F(\cdot)$ are calculated as

$$\begin{aligned} &F_{(2i-1)1} = F_{(2i)1} \\ &= -\alpha c_i \int_0^\ell \left(\sum_{j=1}^n \frac{c_j(x_{2j-1} - x_{2j})}{\pi^2 j^2} \sin(\frac{\pi j}{\ell}\xi) \right)^3 \sin(\frac{\pi i}{\ell}\xi)\, d\xi \end{aligned} \tag{5.81}$$

$$\begin{aligned} &dF_{(2i)(2j)} = dF_{(2i-1)(2j)} = dF_{(2i)(2j-1)} = dF_{(2i-1)(2j-1)} \\ &= -3\alpha c_j c_i \int_0^\ell \left(\sum_{k=1}^n \frac{x_k c_k}{\pi^2 k^2} \sin(\frac{\pi k}{\ell}\xi) \right)^2 \sin(\frac{\pi j}{\ell}\xi) \sin(\frac{\pi i}{\ell}\xi)\, d\xi \end{aligned} \tag{5.82}$$

For the derivative of the input operator with respect to the actuator location, let $b_r(\xi; r)$ denote the derivative of $b(\xi; r)$ with respect to r, then

$$dB_n = \text{row}\left(c_i \int_0^\ell b_r(\xi; r) \sin(\frac{\pi i}{\ell}\xi)\, d\xi \right). \tag{5.83}$$

The function $b(\xi; r)$ is chosen as follows:

$$b(\xi; r) = \begin{cases} (\xi - r + \delta)^2 (\xi - r - \delta)^2, & |\xi - r| \le \delta \\ 0 \quad \text{otherwise} \end{cases} \tag{5.84}$$

Also, it follows that $R_n = c_u$. Now, the approximate IVP and FVP is governed by (5.73); the optimality conditions are given by (5.75) and (5.77).

In the railway track model, if the variables w, t, and ξ are appropriately substituted with dimensionless variables, a dimensionless PDE model for the railway track can be derived. Let c_1, c_2, c_3, and c_4 be some generic constants, a dimensionless railway track is formatted as

$$\begin{cases} \frac{\partial^2 w}{\partial t^2} + \frac{\partial}{\partial \xi^2}(\frac{\partial^2 w}{\partial \xi^2} + c_1 \frac{\partial^3 w}{\partial \xi^2 \partial t}) + c_2 \frac{\partial w}{\partial t} + c_3 w + c_4 w^3 = u(\xi, t), \\ w(\xi, 0) = w_0(\xi), \quad \frac{\partial w}{\partial t}(\xi, 0) = v_0(\xi), \\ w(0, t) = w(1, t) = 0, \\ \frac{\partial^2 w}{\partial \xi^2}(0, t) + c_1 \frac{\partial^3 w}{\partial \xi^2 \partial t}(0, t) = 0, \\ \frac{\partial^2 w}{\partial \xi^2}(1, t) + c_1 \frac{\partial^3 w}{\partial \xi^2 \partial t}(1, t) = 0. \end{cases} \tag{5.85}$$

The relative value of the coefficients is important. The coefficients are chosen with nominal values

$$c_1 = 10^{-4}, \quad c_2 = 0.1, \quad c_3 = 1, \quad c_4 = 10. \tag{5.86}$$

The coefficients c_1, c_2 and c_4 will be changed in simulations to observe their influence. Moreover, the final interval time is set to $\tau = 10$. This gives the state of the controlled system enough time to settle. In addition, we choose $\delta = 5 \times 10^{-5}$ so that the input force is concentrated on a relatively small region on the track. In the cost function, we choose the same weights for the deflection, rate of deflection, and input; so $c_w = c_v = c_u = 1$ is selected for the simulations.

Given an order of approximation, the initial conditions are chosen such that all modes are excited. The initial conditions are chosen from

$$\boldsymbol{x}_0 = \left(2,\, 3,\, \frac{2}{2},\, \frac{3}{2},\, \frac{2}{4},\, \frac{3}{4},\, \frac{2}{8},\, \frac{3}{8},\, \frac{2}{16},\, \frac{3}{16}\right). \tag{5.87}$$

The order of approximation is equal to the dimension of an initial condition. For example, if the order of approximation is 4, the initial condition is

$$\boldsymbol{x}_0 = \left(2,\, 3,\, \frac{2}{2},\, \frac{3}{2}\right). \tag{5.88}$$

The initial condition is illustrated in Figure 5.1 for the 10th order approximation.

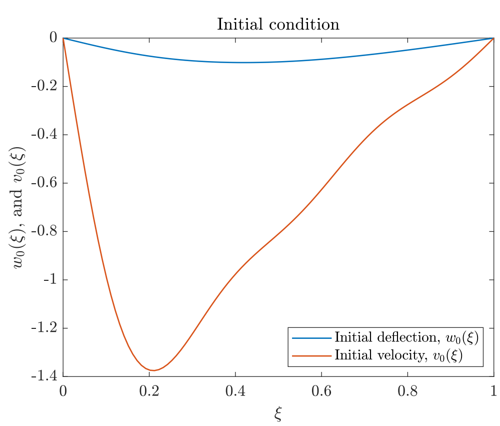

Figure 5.1: Graph of the the initial condition for the simulations.

Simulations were conducted using the software MATLAB. The ODE solver `ode15s` was used to solve the finite-dimensional approximation of the system. MATLAB optimization routine `fmincon` was also used as the optimization algorithm. The convergence of the approximation method is illustrated in Figures 5.2 and 5.3. It is observed that beyond 6th order approximation, increasing the approximation order will not make a noticeable difference. Figures 5.4 and 5.5 compare the cost and optimal input for the linear and nonlinear model in the presence and absence of damping. These figures indicate a significant change in the cost of control and in the optimal input. Figure 5.6 shows how the cost and optimal location of actuators change when the coefficient of nonlinearity, α, is increased. As a general rule of thumb, increasing α increases the cost of control. Moreover, Figures 5.7 and 5.8 show how the cost and location of actuators change when the coefficient of viscous and Kelvin-Voigt damping are decreased. Simulations show that the optimal location of actuators moves away from the center as the damping is decreased. Also, an interesting observation is made in Figure 5.8 where local optimizers appear by decreasing the coefficient of Kelvin-Voigt damping. Lastly, Figure 5.9 shows the improvement in the performance of the control system when the optimal location is chosen for the actuator over the center location.

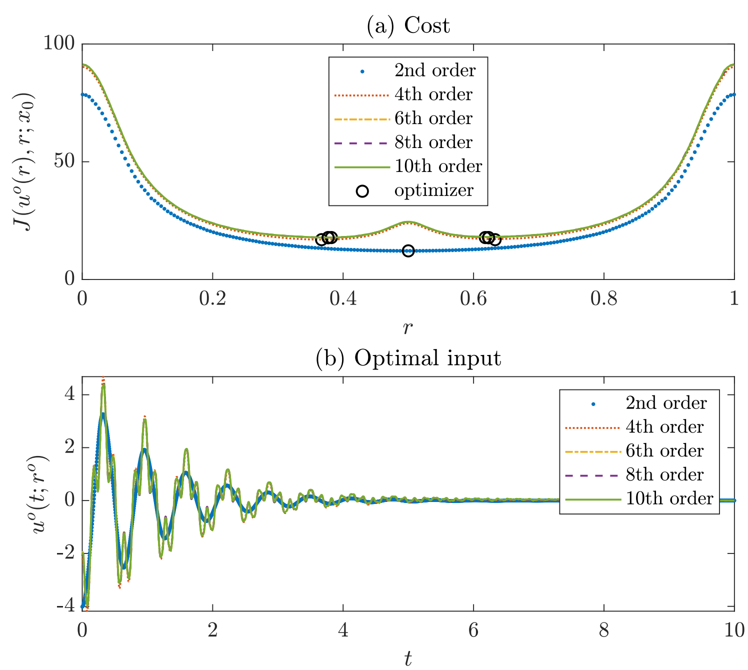

Figure 5.2: Convergence of the numerical scheme for different orders of approximation in undamped beam. No significant improvement is observed for 4th order approximation or higher.

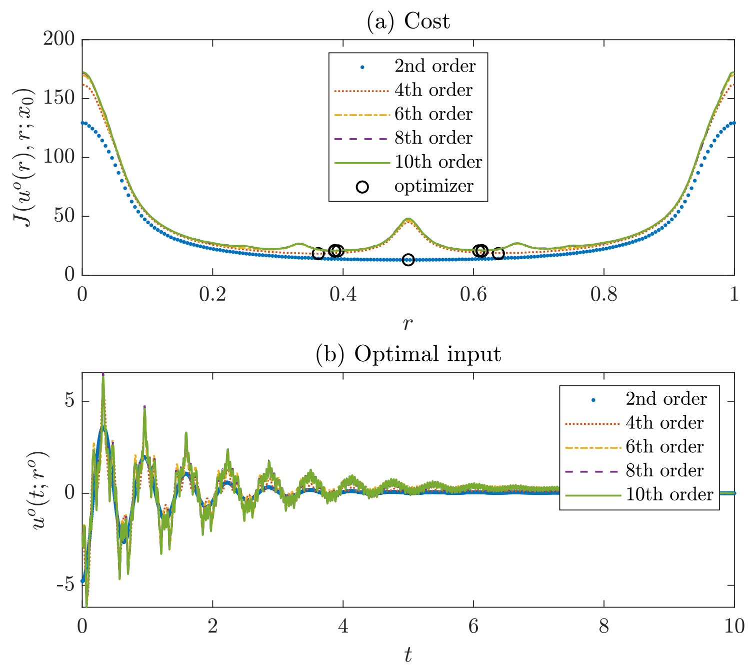

Figure 5.3: Convergence of the numerical scheme for different orders of approximation in damped beam. No significant improvement is observed for 6th order approximation or higher.

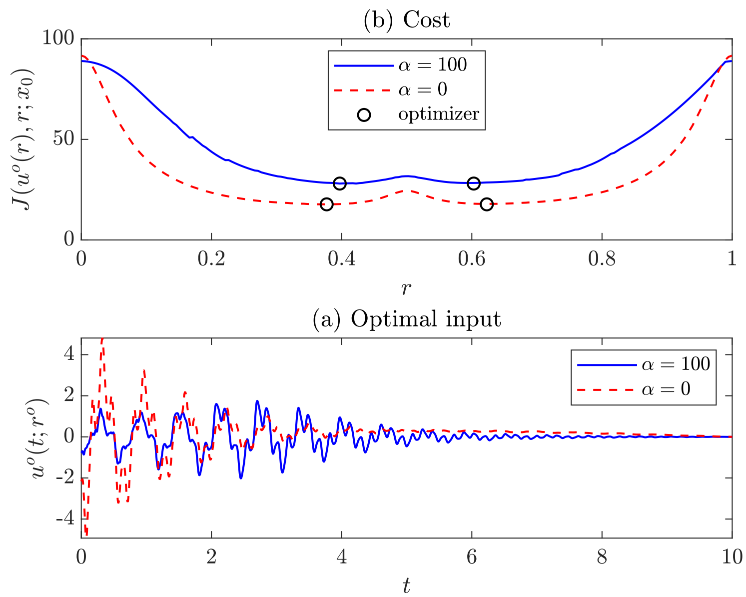

Figure 5.4: Comparison of the optimal input and cost function in linear and nonlinear damped beam. The cost of control increases by increasing the nonlinearity.

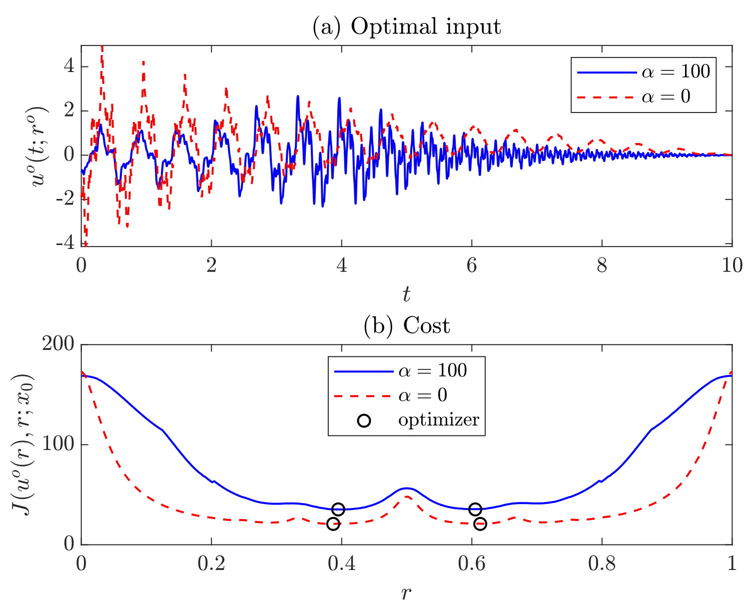

Figure 5.5: Comparison of the optimal input and cost function in linear and nonlinear undamped beam. The cost of control increases by increasing the nonlinearity.

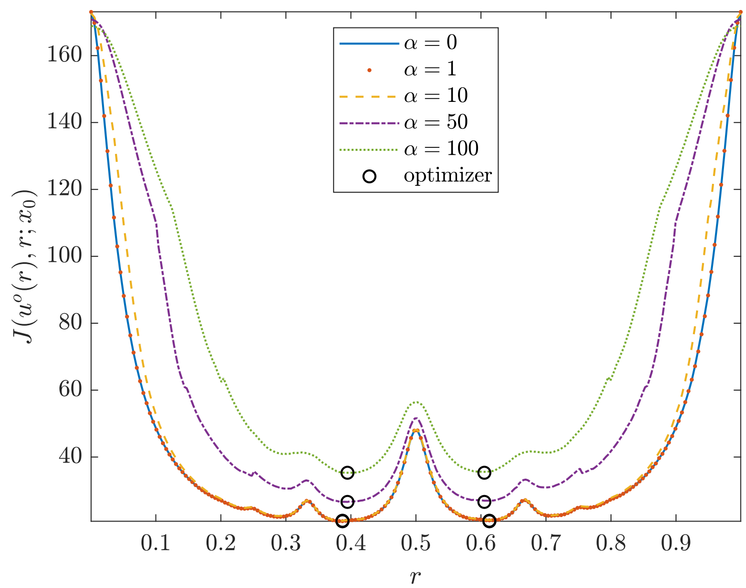

Figure 5.6: Effect of nonlinearity on the cost function. The optimal actuator locations do not change significantly despite the change in the cost.

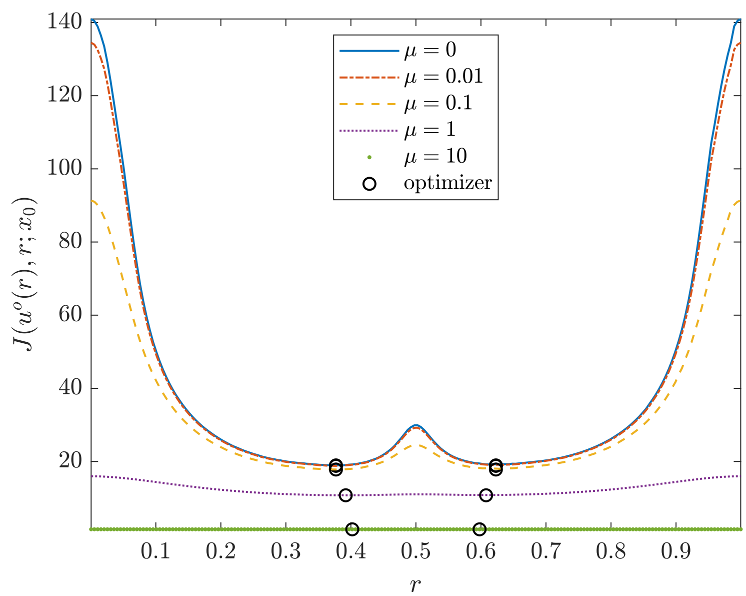

Figure 5.7: Effect of viscous damping on the cost function. The optimal actuator locations move away from center as the damping is decreased.

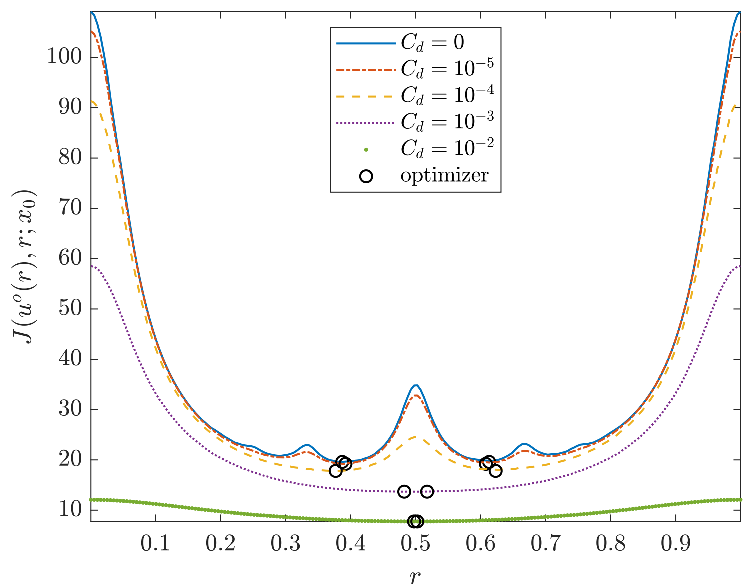

Figure 5.8: Effect of Kelvin-Voigt damping on the cost function. If $C_d = 0$, the beam models is hyperbolic. The optimal actuator locations move away from center as the damping is decreased.

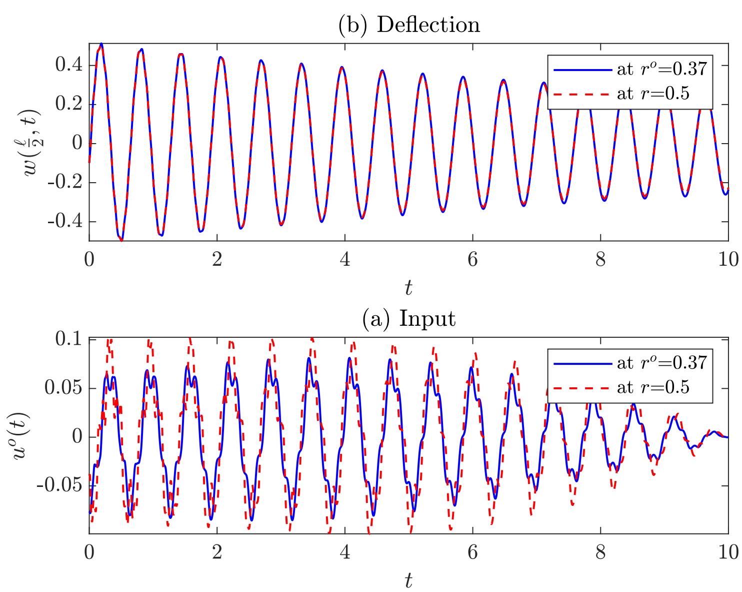

Figure 5.9: Comparison of optimal inputs: optimal location vs center. Actuators on optimal locations improve the control input.

5.3 Concluding Remarks

The stability and well-posedness of a nonlinear railway track model was studied in this chapter. Using a suitable Lyapunov function, it was proved that the model admits a global (in time) classical solution for a continuously differentiable input. The solution is also exponentially stable. For less regular inputs, belonging only to $L^2_{loc}(0,\infty;\mathbb{U})$ or $PC(\mathbb{R}^+;\mathbb{U})$, existence and stability of a mild solution as well as input-to-state stability (ISS) of the model were established.

Furthermore, the optimal control and actuator design which were derived in chapter 3 were numerically calculated in this chapter. Numerical simulations were conducted for various physical parameters. The simulations show that the optimal actuator location is off-center, and improves the performance of the control system significantly.

Future work is concerned with optimal shape design of actuators and development of suitable numerical schemes. Shape of actuators is particularly important if the model is in two or three space dimension.

Chapter 6

Conclusion and Future Research

6.1 Summary

Optimal actuator design for nonlinear infinite-dimensional systems was studied in this thesis. The underlying state space was an abstract reflexive Banach space. The nonlinear system was described by

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \mathcal{A}\boldsymbol{x}(t) + \mathcal{F}(\boldsymbol{x}(t)) + \mathcal{B}(\boldsymbol{r})\boldsymbol{u}(t), \quad t > 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0, \end{cases} \tag{6.1}$$

where $\boldsymbol{x}(t)$ and $\boldsymbol{u}(t)$ are the state and input with values in reflexive Banach spaces $\mathbb{X}$ and $\mathbb{U}$, respectively. Also, $\boldsymbol{r}$ is the actuator design parameter that takes value in a topological space $\mathbb{K}$. It was assumed that the linear operator $\mathcal{A} : D(\mathcal{A}) \to \mathbb{X}$ is the generator of a strongly continuous semigroup $\mathcal{T}(t)$. The nonlinear operator $\mathcal{F}(\cdot) : \mathbb{V} \to \mathbb{X}$ was defined on a space $\mathbb{V} \hookrightarrow \mathbb{X}$ (the symbol $\hookrightarrow$ denotes continuous embedding). For each $\boldsymbol{r} \in \mathbb{K}$, the input operator $\mathcal{B}(\boldsymbol{r})$ is a linear bounded operator that maps the input space $\mathbb{U}$ into the state space $\mathbb{X}$. This family of input operators $\mathcal{B}(\cdot) : \mathbb{K} \to \mathcal{L}(\mathbb{U}, \mathbb{X})$ is continuous with respect to $\boldsymbol{r}$ in the operator norm topology. A cost function $J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0)$ was introduced to evaluate the cost of a control input and actuator design. In the cost function, two convex continuous functionals $\phi(\cdot)$ on $\mathbb{X}$ and $\psi(\cdot)$ on $\mathbb{U}$ are considered. That is

$$J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) = \int_0^\tau \phi(\boldsymbol{x}(t)) + \psi(\boldsymbol{u}(t)) \, dt$$

The admissible input set was defined as

$$U_{ad} = \{\boldsymbol{u} \in L^p(0, \tau; \mathbb{U}) : \ \|\boldsymbol{u}\|_p \le R, \ 1 < p < \infty\}. \tag{6.2}$$

Let K_{ad} be a compact and convex subset of $\mathbb{K}$. For a fixed initial condition $\boldsymbol{x}_0$, an optimal actuator design problem was defined as follows

$$\begin{cases} \min & J(\boldsymbol{u}, \boldsymbol{r}; \boldsymbol{x}_0) \\ \text{s.t.} & (6.1), \\ & (\boldsymbol{u}, \boldsymbol{r}) \in U_{ad} \times K_{ad}. \end{cases} \tag{P}$$

The existence of an optimal control together with an optimal actuator design for two general classes of systems was guaranteed under some assumptions.

In Chapter 3, the first class of systems, semilinear systems, was studied. In semilinear systems, the nonlinear operator $\mathcal{F}(\cdot)$ is defined on the whole state space so $\mathbb{V} = \mathbb{X}$, and it is locally Lipschitz continuous. That is, for every open set D in $\mathbb{X}$, there is $L_D > 0$ such that

$$\|\mathcal{F}(\boldsymbol{x}_2) - \mathcal{F}(\boldsymbol{x}_1)\| \leq L_D \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|, \; \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D. \tag{6.3}$$

Examples of such PDE's include the railway track model [39], Gordon family of equations [37], and some nonlinear plate equations [65]. The conditions under which the existence of a solution is guaranteed were derived in Section 2.4. If the nonlinear operator $\mathcal{F}$ also satisfies the following assumptions, then the existence of an optimal actuator design and control is guaranteed. If $\boldsymbol{x}_n(t)$ is bounded in $C(0, \tau; \mathbb{X})$ and weakly convergent to $\boldsymbol{x}(t)$ in $L^p(0, \tau; \mathbb{X})$ then $\mathcal{F}(\boldsymbol{x}_n(t))$ weakly converges to $\mathcal{F}(\boldsymbol{x}(t))$ in $L^p(0, \tau; \mathbb{X})$. This assumption is implied by weak continuity of $\mathcal{F}$. See [42, Chapter 1] for some examples concerning weak continuity. As an example, a system satisfying this condition is a semilinear wave equation on $[0, 1]$. For any $b(\cdot, r) \in L^2(0, 1)$ and for all $0 < r < 1$, consider

$$\begin{cases} w_{tt}(\xi, t) \\ \quad = w_{\xi\xi}(\xi, t) + w^3(\xi, t) + b(\xi, r)u(t), \; \xi \in (0, 1), \, t > 0, \\ w(0, t) = w(1, t) = 0, \; t \geq 0, \\ w(\xi, 0) = w_0(\xi), \; w_t(\xi, 0) = v_0(\xi), \; \xi \in [0, 1]. \end{cases}$$

A state space for this problem with state $\boldsymbol{x} := (w, w_t)$ is $\mathbb{X} = H_0^1(0, 1) \times L^2(0, 1)$. Because of the compact embedding of $H^1(0, 1)$ in $C(0, 1)$ and [110, Theorem 3], if a sequence w_n in $C(0, \tau; H_0^1(0, 1))$ weakly converges to w in $L^p(0, \tau; H_0^1(0, 1))$, then w_n^3 converges weakly to w^3 in $L^p(0, \tau; L^2(0, 1))$. The required condition on $\mathcal{F}$ follows immediately. A similar argument can be used to establish the required regularity for other second-order in time PDE's. The optimality conditions for such systems were derived in [37, Theorem. 5.7].

In Chapter 4, the second class of systems, nonlinear parabolic systems, was studied. In nonlinear parabolic systems, the nonlinear operator $\mathcal{F}(\cdot)$ is not continuous with respect to the norm on the state space. The conditions ensuring the existence of a solution were derived in Section 2.4. Extra assumptions on the linear part $\mathcal{A}$ are required to obtain existence of solutions to the PDE; see, for example [89]. Similar assumptions also imply a solution to the optimization problem (P). It is assumed that the operator $\mathcal{A}$ has maximal L^p regularity, see Definition 4.1.1. Every generator of an analytic semigroup on a Hilbert space $\mathbb{X}$ has maximal L^p regularity [35, Theorem 4.1]. Some examples of PDE's where the linear part generates an analytic semigroup are Burger's equation, Kuramoto-Sivashinsky equation, and Euler-Bernoulli beams with Kelvin-Voigt damping. As an example, consider Burger's equation with periodic boundary conditions

$$\begin{cases} w_t(\xi,t) + w(\xi,t)w_\xi(\xi,t) \\ \qquad = w_{\xi\xi}(\xi,t) + b(\xi,r)u(t),\ \xi \in (-\pi,\pi),\ t>0, \\ w(-\pi,t) = w(\pi,t),\ t \geq 0, \\ w(\xi,0) = w_0(\xi),\ \xi \in [-\pi,\pi]. \end{cases}$$

A suitable state space is $\mathbb{X} = L^2_{per}(-\pi,\pi)$. The nonlinear operator $\mathcal{F}(\boldsymbol{x}) = -ww_\xi$ is not strongly continuous from $\mathbb{X}$ to $\mathbb{X}$. However, defining $\mathbb{V} = H^1_{per}(-\pi,\pi)$, it is strongly continuous from $\mathbb{V}$ to $\mathbb{X}$. The linear part of the PDE

$$\mathcal{A}w = w_{\xi\xi}, \quad D(\mathcal{A}) = H^2_{per}(-\pi,\pi),$$

generates an analytic semigroup [109, Theorem 32.1]. It therefore has maximal L^p regularity [35, Theorem 4.1]. Assume that $\mathcal{F}(\cdot)$ is defined on $\mathbb{V}$, where $D(\mathcal{A}) \hookrightarrow \mathbb{V} \hookrightarrow \mathbb{X}$, and is locally Lipschitz continuous with respect to $\mathbb{V}$; that is, for every open set $D \subset \mathbb{V}$ there is a positive number L_D such that

$$\|\mathcal{F}(\boldsymbol{x}_2) - \mathcal{F}(\boldsymbol{x}_1)\| \leq L_D \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_{\mathbb{V}},\ \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D. \tag{6.4}$$

If also the domain $D(\mathcal{A})$ is compactly embedded in $\mathbb{X}$, then there is an optimal control and actuator design for (P). The optimality conditions for such systems were derived in [38, Theorem. 5.7].

In Chapter 5, an in-depth study of optimal control and actuator location for nonlinear railway track model was conducted. The study begins with investigating the well-posedness and stability of classical and mild solutions to a nonlinear railway track model. Further on, the optimality conditions characterizing the optimal control and actuator location were derived for this model. The optimal control and actuator location were computed for various parameters of the model. It was observed that the optimal actuator locations are off-center, contrary to intuition.

6.2 Future Research

This thesis could have an impact on the development of numerical methods and optimization algorithms for optimal control and actuator design. The numerical theory for control of nonlinear PDEs is still in early stages. The existing numerical results are only for specific nonlinear PDE models. The conditions derived in this thesis for the existence of an optimal control and actuator design can initiate the investigation of convergence and stability of suitable numerical schemes. Proving the convergence of a numerical scheme often follows similar steps as proving the existence of a solution. The abstract optimization setting in this thesis will also facilitate the investigation of convergence of optimization algorithms formulated in Banach spaces. Proving the convergence of an optimization algorithm is also linked to proving the existence of an optimizer. Computation of actuator shapes, as opposed to locations, for PDE models has also not yet been developed for nonlinear PDE models. This thesis also provides an abstract framework for numerical computation of actuator shapes.

This thesis also helps to better understand the difference between the finite-dimensional and infinite-dimensional models of a nonlinear system. In this thesis, the existence and optimality conditions were stated for an abstract reflexive Banach space. However, the numerical schemes for finite-dimensional systems sometimes fail to approximate the solution of an infinite-dimensional system. This is also true when computing the optimal control and actuator design for an infinite-dimensional system.

There are still many open problems in this field. Establishing that a given nonlinear PDE falls into one of the classes discussed here is not always straightforward, particularly for problems in several space dimensions. There are also nonlinear PDEs that do not fall into any of these classes. Also, the existence of a solution to the optimality conditions is important, yet unknown. In some systems, the actuators should be installed on the boundary. This yields unbounded input operators which violate one of the assumptions in this thesis. Extension of the results to unbounded input operators is interesting. Moreover, extension of the current results to optimal sensor design is interesting, yet challenging.

Bibliography

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Pure and Applied Mathematics, Elsevier Science, 2003.

[2] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, *Dissipation inequalities for the analysis of a class of PDEs*, Automatica, 66 (2016), pp. 163–171.

[3] R. Al Jamal and K. Morris, *Linearized stability of partial differential equations with application to stabilization of the Kuramoto–Sivashinsky equation*, SIAM Journal on Control and Optimization, 56 (2018), pp. 120–147.

[4] H. Amann, *Compact embeddings of vector valued Sobolev and Besov spaces*, Glasnik matematički, 35 (2000), pp. 161–177.

[5] M. Ansari, E. Esmailzadeh, and D. Younesian, *Frequency analysis of finite beams on nonlinear Kelvin-Voight foundation under moving loads*, Journal of Sound and Vibration, 330 (2011), pp. 1455–1471.

[6] C. Antoniades and P. D. Christofides, *Integrating nonlinear output feedback control and optimal actuator/sensor placement for transport-reaction processes*, Chemical Engineering Science, 56 (2001), pp. 4517–4535.

[7] A. Armaou and M. A. Demetriou, *Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes*, AIChE journal, 54 (2008), pp. 2651–2662.

[8] H. T. Banks and K. Ito, *A unified framework for approximation in inverse problems for distributed parameter systems*, Control, Theory and Advanced Technology, 4 (1988), pp. 73–90.

[9] V. Barbu and T. Precupanu, *Convexity and optimization in Banch spaces*, Springer Science & Business Media, 2012.

[10] I. Bena, C. Misbah, and A. Valance, *Nonlinear evolution of a terrace edge during step-flow growth*, Physical Review B, 47 (1993), pp. 7408–7419.

[11] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*, Springer Science & Business Media, 2007.

[12] M. Bergounioux and K. Kunisch, *On the structure of Lagrange multipliers for state-constrained optimal control problems*, Systems & Control Letters, 48 (2003), pp. 169–176.

[13] J. L. Boldrini, B. M. C. Caretta, and E. Fernández-Cara, *Some optimal control problems for a two-phase field model of solidification*, Revista Matemática Complutense, 23 (2009), pp. 49–75.

[14] R. Buchholz, H. Engel, E. Kammann, and F. Tröltzsch, *On the optimal control of the Schlögl-model*, Computational Optimization and Applications, 56 (2013), pp. 153–185.

[15] E. Casas, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM Journal on Control and Optimization, 35 (1997), pp. 1297–1327.

[16] E. Casas, J. C. De Los Reyes, and F. Tröltzsch, *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM Journal on Optimization, 19 (2008), pp. 616–643.

[17] E. Casas, C. Ryll, and F. Tröltzsch, *Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems*, Computational Methods in Applied Mathematics, 13 (2013), pp. 415–442.

[18] E. Cerpa, *Null controllability and stabilization of the linear Kuramoto-Sivashinsky equation*, Communications on Pure and Applied Analysis, 9 (2010), pp. 91–102.

[19] E. Cerpa and A. Mercado, *Local exact controllability to the trajectories of the 1-D Kuramoto-Sivashinsky equation*, Journal of Differential Equations, 250 (2011), pp. 2024–2044.

[20] S. N. Chandler-Wilde, D. P. Hewett, and A. Moiola, *Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples*, Mathematika, 61 (2015), pp. 414–443.

[21] S. P. Chen and R. Triggiani, *Proof of extensions of two conjectures on structural damping for elastic systems*, Pacific Journal of Mathematics, 136 (1989), pp. 15–55.

[22] P. D. Christofides and A. Armaou, *Global stabilization of the Kuramoto-Sivashinsky equation via distributed output feedback control*, Systems and Control Letters, 39 (2000), pp. 283–294.

[23] G. Ciaramella and A. Borzi, *Quantum optimal control problems with a sparsity cost functional*, Numerical Functional Analysis and Optimization, 37 (2016), pp. 938–965.

[24] P. Clément and S. Li, *Abstract parabolic quasilinear equations and applications to a groundwater flow problem*, Advances in mathematical sciences and applications, 3 (1993), pp. 17–32.

[25] R. V. Craster and O. K. Matar, *Dynamics and stability of thin liquid films*, Reviews of Modern Physics, 81 (2009), pp. 1131–1198.

[26] R. F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Texts in Applied Mathematics, Springer New York, 2012.

[27] T. Dahlberg, *Dynamic interaction between train and nonlinear railway track model*, in Proc. Fifth Euro. Conf. Struct. Dyn., Munich, Germany, 2002, pp. 1155–1160.

[28] N. Darivandi, K. Morris, and A. Khajepour, *An algorithm for LQ optimal actuator location.* Smart Materials and Structures, 22 (2013), 035001 (10pp).

[29] J. C. de los Reyes, R. Herzog, and C. Meyer, *Optimal Control of Static Elastoplasticity in Primal Formulation*, SIAM Journal on Control and Optimization, 54 (2016), pp. 3016–3039.

[30] M. S. de Queiroz, D. M. Dawson, S. P. Nagarkatti, and F. Zhang, *Lyapunov-based Control of Mechanical Systems*, Springer Science & Business Media, 2012.

[31] M. A. Demetriou, *A numerical algorithm for the optimal placement of actuators and sensors for flexible structures*, in American Control Conference, 2000. Proceedings of the 2000, vol. 4, IEEE, 2000, pp. 2290–2294.

[32] D. Dier, *Non-autonomous maximal regularity for forms of bounded variation*, Journal of Mathematical Analysis and Applications, 425 (2015), pp. 33–54.

[33] H. Ding, L.-Q. Chen, and S.-P. Yang, *Convergence of Galerkin truncation for dynamic response of finite beams on nonlinear foundations under a moving load*, Journal of Sound and Vibration, 331 (2012), pp. 2426–2442.

[34] H. Ding, K.-L. Shi, L.-Q. Chen, and S.-P. Yang, *Dynamic response of an infinite Timoshenko beam on a nonlinear viscoelastic foundation to a moving load*, Nonlinear Dynamics, 73 (2013), pp. 285–298.

[35] G. Dore, *L_p-regularity for abstract differential equations*, in Functional Analysis and Related Topics, 1991, Springer, (1993), pp. 25–38.

[36] M. S. Edalatzadeh and A. Alasty, *Boundary exponential stabilization of non-classical micro/nano beams subjected to nonlinear distributed forces*, Applied Mathematical Modelling, 40 (2016), pp. 2223–2241.

[37] M. S. Edalatzadeh and K. A. Morris, *Optimal actuator design for semilinear systems*, SIAM Journal on Control and Optimization, 57 (2019), pp. 2992–3020.

[38] M. S. Edalatzadeh and K. A. Morris, *Optimal controller and actuator design for nonlinear parabolic systems*, To be submitted, (2019).

[39] M. S. Edalatzadeh and K. A. Morris, *Stability and well-posedness of a nonlinear railway track model*, IEEE Control Systems Letters, 3 (2019), pp. 162–167.

[40] M. S. Edalatzadeh, R. Vatankhah, and A. Alasty, *Suppression of dynamic pull-in instability in electrostatically actuated strain gradient beams*, in 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), oct 2014, pp. 155–160.

[41] K.-J. Engel and R. Nagel, *One-parameter Semigroups for Linear Evolution Equations*, Springer-Verlag New York, 2000.

[42] L. C. Evans, *Weak Convergence Methods for Nonlinear Partial Differential Equations*, no. 74 in CBMS Regional Conference Series, Conference Board of the Mathematical Sciences, 1990.

[43] F. Fahroo and M. A. Demetriou, *Optimal actuator/sensor location for active noise regulator and tracking control problems*, Journal of Computational and Applied Mathematics, 114 (2000), pp. 137–158.

[44] G. Fantuzzi and A. Wynn, *Semidefinite relaxation of a class of quadratic integral inequalities*, in Decision and Control (CDC), 2016 IEEE 55th Conference on, IEEE, 2016, pp. 6192–6197.

[45] E. Fašangová and J. Prüss, *Asymptotic behaviour of a semilinear viscoelastic beam model*, Archiv der Mathematik, 77 (2001), pp. 488–497.

[46] H. O. Fattorini, *Infinite Dimensional Optimization and Control Theory*, vol. 54, Cambridge University Press, 1999.

[47] A. Fleig and R. Guglielmi, *Optimal control of the Fokker–Planck equation with space-dependent controls*, Journal of Optimization Theory and Applications, 174 (2017), pp. 408–427.

[48] M. I. Frecker, *Recent advances in optimization of smart structures and actuators*, Journal of Intelligent Material Systems and Structures, 14 (2003), pp. 207–216.

[49] T. Fu Ma, *Boundary stabilization for a nonlinear beam on elastic bearings*, Mathematical Methods in the Applied Sciences, 24 (2001), pp. 583–594.

[50] P. Gao, *Optimal distributed control of the Kuramoto-Sivashinsky equation with pointwise state and mixed control-state constraints*, IMA Journal of Mathematical Control and Information, 33 (2016), pp. 791–811.

[51] F. Gazzola and M. Squassina, *Global solutions and finite time blow up for damped semilinear wave equations*, in Annales de l'Institut Henri Poincare (C) Non Linear Analysis, 23 (2006), pp. 185–207.

[52] S. Gerbi and B. Said-Houari, *Local existence and exponential growth for a semilinear damped wave equation with dynamic boundary conditions*, Advances in Differential Equations, 13 (2008), pp. 1051–1074.

[53] S. Gerbi and B. Said-Houari, *Asymptotic stability and blow up for a semilinear damped wave equation with dynamic boundary conditions*, Nonlinear Analysis: Theory, Methods & Applications, 74 (2011), pp. 7137–7150.

[54] S. N. Gomes, D. T. Papageorgiou, and G. A. Pavliotis, *Stabilizing nontrivial solutions of the generalized Kuramoto-Sivashinsky equation using feedback and optimal control*, IMA Journal of Applied Mathematics (Institute of Mathematics and Its Applications), 82 (2017), pp. 158–194.

[55] M. Grasselli, G. Schimperna, and S. Zelik, *On the 2D Cahn–Hilliard equation with inertial term*, Communications in Partial Differential Equations, 34 (2009), pp. 137–170.

[56] W. W. Hager and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM Journal on optimization, 16 (2005), pp. 170–192.

[57] W. W. Hager and H. Zhang, *A survey of nonlinear conjugate gradient methods*, Pacific journal of Optimization, 2 (2006), pp. 35–58.

[58] P. Hebrard and A. Henrot, *A spillover phenomenon in the optimal location of actuators*, SIAM Journal of Control and Optimization, 44 (2005), pp. 349–366.

[59] R. Herzog and K. Kunisch, *Algorithms for PDE-constrained optimization*, GAMM Mitteilungen, 33 (2010), pp. 163–176.

[60] M. Hintermüller, T. Keil, and D. Wegner, *Optimal control of a semidiscrete Cahn–Hilliard–Navier–Stokes system with nonmatched fluid densities*, SIAM Journal on Control and Optimization, 55 (2017), pp. 1954–1989.

[61] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, vol. 23, Springer Science & Business Media, 2008.

[62] D. Hömberg, C. Meyer, J. Rehberg, W. Ring, and D. H. Omberg, *Optimal control for the thermistor problem*, SIAM Journal on Control and Optimization, 48 (2010), pp. 3449–3481.

[63] B. Jacob, R. Nabiullin, J. R. Partington, and F. L. Schwenninger, *Infinite-dimensional input-to-state stability and Orlicz spaces*, SIAM Journal on Control and Optimization, 56 (2018), pp. 868–889.

[64] B. Jayawardhana, H. Logemann, and E. P. Ryan, *Infinite-dimensional feedback systems: the circle criterion and input-to-state stability*, Communications in Information and Systems, 8 (2008), pp. 413–444.

[65] G. Ji and I. Lasiecka, *Nonlinear boundary feedback stabilization for a semilinear Kirchhoff plate with dissipation acting only via moments-limiting behavior*, Journal of Mathematical Analysis and Applications, 229 (1999), pp. 452–479.

[66] D. Kalise, K. Kunisch, and K. Sturm, *Optimal actuator design based on shape calculus*, Mathematical Models and Methods in Applied Sciences, 28 (2018), pp. 2667–2717.

[67] I. Karafyllis and M. Krstic, *ISS with respect to boundary disturbances for 1-D parabolic PDEs*, IEEE Transactions on Automatic Control, 61 (2016), pp. 3712–3724.

[68] D. Kasinathan and K. Morris, $\mathbb{H}_\infty$*-optimal actuator location*, IEEE Transactions on Automatic Control, 58 (2013), pp. 2522–2535.

[69] J. U. Kim and Y. Renardy, *Boundary control of the Timoshenko beam*, SIAM Journal on Control and Optimization, 25 (1987), pp. 1417–1429.

[70] S.-J. Kimmerle, M. Gerdts, and R. Herzog, *Optimal control of an elastic crane-trolley-load system-a case study for optimal control of coupled ODE-PDE systems*, Mathematical and Computer Modelling of Dynamical Systems, 24 (2018), pp. 182–206.

[71] V. Komornik, *Exact Controllability and Stabilization: the Multiplier Method*, vol. 36, Masson, 1994.

[72] G. Kouroussis, C. Caucheteur, D. Kinet, G. Alexandrou, O. Verlinden, and V. Moeyaert, *Review of trackside monitoring solutions: from strain gages to optical fibre sensors*, Sensors, 15 (2015), pp. 20115–20139.

[73] Y. Kuramoto and T. Tsuzuki, *On the formation of dissipative structures in reaction-diffusion systems: Reductive perturbation approach*, Progress of Theoretical Physics, 54 (1975), pp. 687–699.

[74] J. E. Lagnese and G. Leugering, *Uniform stabilization of a nonlinear beam by nonlinear boundary feedback*, Journal of Differential Equations, 91 (1991), pp. 355–388.

[75] S. Lang, *Real and functional analysis*, vol. 142, Springer Science & Business Media, 2012.

[76] R. E. LaQuey, S. M. Mahajan, P. H. Rutherford, and W. M. Tang, *Nonlinear saturation of the trapped-ion mode*, Physical Review Letters, 34 (1975), pp. 391–394.

[77] I. Lasiecka and R. Triggiani, *Differential and algebraic Riccati equations with application to boundary/point control problems: continuous theory and approximation theory*, vol. 164, Springer, 1991.

[78] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: continuous and approximation theories*, vol. 1, Cambridge University Press Cambridge, 2000.

[79] G. Leugering, S. Engell, A. Griewank, M. Hinze, R. Rannacher, V. Schulz, M. Ulbrich, and S. Ulbrich, *Constrained optimization and optimal control for partial differential equations*, vol. 160, Springer Science & Business Media, 2012.

[80] C. Li, E. Feng, and J. Liu, *Optimal control of systems of parabolic PDEs in exploitation of oil*, Journal of Applied Mathematics and Computing, 13 (2003), pp. 247–259.

[81] W. Liu, K. Chen, and J. Yu, *Asymptotic stability for a non-autonomous full von Kármán beam with thermo-viscoelastic damping*, Applicable Analysis, 97 (2018), pp. 400–414.

[82] W. J. Liu and M. Krstic, *Stability enhancement by boundary control in the Kuramoto–Sivashinsky equation*, Nonlinear Analysis: Theory, Methods and Applications, 46 (2001), pp. 485 – 507.

[83] Y. Lou and P. D. Christofides, *Optimal actuator/sensor placement for nonlinear control of the Kuramoto-Sivashinsky equation*, IEEE Transactions on Control Systems Technology, 11 (2003), pp. 737–745.

[84] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Springer Science & Business Media, 2012.

[85] A. Martínez, C. Rodríguez, and M. E. Vázquez-Méndez, *Theoretical and numerical analysis of an optimal control problem related to wastewater treatment*, SIAM Journal on Control and Optimization, 38 (2000), pp. 1534–1553.

[86] F. Mazenc and C. Prieur, *Strict Lyapunov functions for semilinear parabolic partial differential equations*, Mathematical Control and Related Fields, 1 (2011), pp. 231–250.

[87] A. R. Mehrabian and A. Yousefi-Koma, *Optimal positioning of piezoelectric actuators on a smart fin using bio-inspired algorithms*, Aerospace Science and technology, 11 (2007), pp. 174–182.

[88] J. Merger, A. Borzi, and R. Herzog, *Optimal control of a system of reaction–diffusion equations modeling the wine fermentation process*, Optimal Control Applications and Methods, 38 (2017), pp. 112–132.

[89] C. Meyer and L. M. Susu, *Optimal control of nonsmooth, semilinear parabolic equations*, SIAM Journal on Control and Optimization, 55 (2017), pp. 2206–2234.

[90] A. Mironchenko, *Input-to-state stability of infinite-dimensional control systems*, PhD thesis, University of Bremen, 2012.

[91] A. Mironchenko, *Local input-to-state stability: characterizations and counterexamples*, Systems and Control Letters, 87 (2016), pp. 23–28.

[92] A. Mironchenko and H. Ito, *Characterizations of integral input-to-state stability for bilinear systems in infinite dimensions.*, Mathematical Control and Related Fields, 6 (2016), pp. 447–466.

[93] A. Mironchenko and F. Wirth, *Characterizations of input-to-state stability for infinite-dimensional systems*, IEEE Transactions on Automatic Control, 63 (2018), pp. 1602–1617.

[94] S. H. Moon, *Finite element analysis and design of control system with feedback output using piezoelectric sensor/actuator for panel flutter suppression*, Finite Elements in Analysis and Design, 42 (2006), pp. 1071–1078.

[95] K. Morris, *Linear-quadratic optimal actuator location*, IEEE Transactions on Automatic Control, 56 (2011), pp. 113–124.

[96] K. Morris, M. A. Demetriou, and S. D. Yang, *Using $\mathbb{H}_2$-control performance metrics for the optimal actuator location of distributed parameter systems*, IEEE Transactions on Automatic Control, 60 (2015), pp. 450–462.

[97] K. Morris and S. Yang, *Comparison of actuator placement criteria for control of structures*, Journal of Sound and Vibration, 353 (2015), pp. 1–18.

[98] K. A. Morris, *Noise reduction achievable by point control*, Journal of dynamic systems, measurement, and control, 120 (1998), pp. 216–223.

[99] K. A. Morris and A. Vest, *Design of damping for optimal energy dissipation of vibrations*, in 2016 IEEE 55th Conference on Decision and Control (CDC), IEEE, 2016, pp. 532–536.

[100] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*, Applied Mathematical Sciences, Springer-Verlag New York, 2000.

[101] J. Nocedal and S. Wright, *Numerical Optimization*, Springer-Verlag New York, 2006.

[102] A. Novick-Cohen, *Interfacial instabilities in directional solidification of dilute binary alloys: The Kuramoto-Sivashinsky equation*, Physica D: Nonlinear Phenomena, 26 (1987), pp. 403–410.

[103] Y. Orlov, *On general properties of eigenvalues and eigenfunctions of a Sturm–Liouville operator: comments on "ISS with respect to boundary disturbances for 1-D parabolic PDEs"*, IEEE Transactions on Automatic Control, 62 (2017), pp. 5970–5973.

[104] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, vol. 44, Springer Science & Business Media, 2012.

[105] A. Pisano and Y. Orlov, *On the ISS properties of a class of parabolic DPS' with discontinuous control using sampled-in-space sensing and actuation*, Automatica, 81 (2017), pp. 447–454.

[106] Y. Privat, E. Trélat, and E. Zuazua, *Optimal location of controllers for the one-dimensional wave equation*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 30 (2013), pp. 1097–1126.

[107] J. P. Raymond and H. Zidani, *Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations*, Applied mathematics & optimization, 39 (1999), pp. 143–177.

[108] M. R. Saviz, *An optimal approach to active damping of nonlinear vibrations in composite plates using piezoelectric patches*, Smart Materials and Structures, 24 (2015), p. 115024.

[109] G. R. Sell and Y. You, *Dynamics of Evolutionary Equations*, vol. 143, Springer Science & Business Media, 2013.

[110] J. Simon, *Compact sets in the space $L^p(0,T;B)$*, Annali di Matematica pura ed applicata, 146 (1986), pp. 65–96.

[111] G. I. Sivashinsky, *Nonlinear analysis of hydrodynamic instability in laminar flames—I. Derivation of basic equations*, Acta astronautica, 4 (1977), 1177–1206.

[112] M. Sprengel, G. Ciaramella, and A. Borzì, *Investigation of optimal control problems governed by a time-dependent Kohn-Sham model*, Journal of Dynamical and Control Systems, 24 (2018), pp. 1–23.

[113] B. Sun, *Maximum principle for optimal boundary control of the Kuramoto-Sivashinsky equation*, Journal of the Franklin Institute, 347 (2010), pp. 467–482.

[114] H. Takeda and S. Yoshikawa, *On the initial value problem of the semilinear beam equation with weak damping II: Asymptotic profiles*, Journal of Differential Equations, 253 (2012), pp. 3061–3080.

[115] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, Graduate studies in mathematics, American Mathematical Society, 2010.

[116] A. Unger and F. Tröltzsch, *Fast solution of optimal control problems in the selective cooling of steel*, ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik, 81 (2001), pp. 447–456.

[117] M. Van De Wal and B. De Jager, *A review of methods for input/output selection*, Automatica, 37 (2001), pp. 487–510.

[118] B. Van Keulen, *H_∞-control for distributed parameter systems: A state-space approach*, Springer Science & Business Media, 2012.

[119] A. Wouk, *A Course of Applied Functional Analysis*, Wiley, 1979.

[120] I. Yousept, *Optimal control of non-smooth hyperbolic evolution Maxwell equations in type-II superconductivity*, SIAM Journal on Control and Optimization, 55 (2017), pp. 2305–2332.

[121] A. Zettl, *Sturm-Liouville Theory*, Mathematical Surveys and Monographs, American Mathematical Society, 2005.

[122] K. Zhou, J. C. Doyle, K. Glover, et al., *Robust and Optimal Control*, vol. 40, Prentice hall New Jersey, 1996.

[123] S. Zhu, J. Yang, H. Yan, L. Zhang, and C. Cai, *Low-frequency vibration control of floating slab tracks using dynamic vibration absorbers*, Vehicle System Dynamics, 53 (2015), pp. 1296–1314.