A Collage-Based Approach to Inverse Optimal Control Problems with Unique Solutions

by

John J. Dewhurst

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Mathematics and Statistics

Guelph, Ontario, Canada

ABSTRACT

A COLLAGE-BASED APPROACH TO INVERSE OPTIMAL CONTROL PROBLEMS WITH UNIQUE SOLUTIONS

John J. Dewhurst
University of Guelph, 2021

Advisor:
Dr. K. Levere

Optimal control problems entail finding a control function that optimizes a given objective functional, subject to a set of constraints that include an ordinary differential equation. Conversely, inverse optimal control problems entail seeking objective functionals that are optimized by a given control system. In this thesis we develop a Collage-Based Approach to solving inverse optimal control problems with unique solutions based on the Collage method for ODE inverse problems and Pontryagin's Maximum Principle. Inverse problems are often formulated as the minimization of an approximation error over a set of parameters. Collage-type methods bound the approximation error by a quantity that is computationally preferable to optimize. We demonstrate this method through a variety of example scenarios and discuss its efficacy and robustness.

This thesis is dedicated to the memory of my father,

John Anthony Dewhurst

1946–2017

“Memories last forever”

ACKNOWLEDGEMENTS

Throughout my journey to this point I have crossed paths with many people, and I am a firm believer that each person we meet touches our lives in some way. Indeed, I would not be writing this today were it not for the individuals who, by chance or by circumstance, were present on this journey. The complete list of those to whom I owe thanks would fill many pages. I do however wish to articulate my gratitude to the following people:

To Marlene Belson, for opening your home to a young student in need. Your kindness and generosity will never be forgotten.

To Susan McCormick, for always ensuring that important deadlines were not missed, forms were not forgotten, and for overall ensuring that my time as a Graduate Student went as smoothly as possible.

To Dr. Herb Kunze, for offering your advice and guidance to a young graduate student.

To Dr. Kimberly Levere, for taking a chance on an undergraduate student you had never met before. You gave me the opportunity to pursue this degree, and throughout my time as your student, your support has never wavered and for this, I will be forever grateful.

To Doug and Sue Pearson, for your kindness and generosity. Thank you for accepting me into your family.

To my family, for always being supportive of the path that I have chosen and for believing that I would succeed.

To my Uncle Reg, for your support, financial and otherwise, and also for your assistance when Dad passed away. I really appreciate it.

To my sister Jenn. You have always been in my corner, offering love and support, and your wisdom as I have grown to be who I am today. I am thankful that despite the physical distance separating us, I have never felt that you were very far away at all.

To my mom, Pat. I really can not express the extent of your love and support for me always. Thank you for everything.

To Emily. We have really been on this journey together, for many years now. Thank you for your love and support, despite my flaws, and for being behind me, no matter what. I look forward to calling you my wife, very soon, and taking the next steps in our lives, together.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Modelling of dynamic physical systems is often an objective in applied mathematics. A popular approach to doing so is to use systems of differential equations. In some cases, we are interested in controlling or influencing a system with an input that is chosen to satisfy certain conditions. Optimal Control is an approach to this problem where a control function is chosen to optimize some measure of cost or performance. Regarded as an extension of the Calculus of Variations, Optimal Control was first studied systematically in the 1950s when Lev Pontryagin [9] and Richard Bellman [7] independently made important advances. We will consider one of these, now referred to as Pontryagin's Maximum Principle, in this thesis. Following these advances and the establishment of Optimal Control as a field of study, a differing view emerged. If we consider a given controlled system, the Inverse Optimal Control Problem involves identifying all objective functionals for which the given or observed control is optimal. Kalman [29] first stated and solved this problem for a specific class of problems in 1964.

In this thesis we develop a Collage-Based Approach for solving an Inverse Optimal Control Problem (IOCP). We utilize a common formulation for inverse problems where unknown parameter values, $\lambda \in \Lambda$, are to be determined so that the observed data x_{true} is close, in an appropriate norm, to the solution data produced using the recovered parameters, x_λ. That is, we wish to minimize the approximation error, or solve

$$\min_{\lambda \in \Lambda} \left\| x_{\text{true}} - x_\lambda \right\| .$$

Expressing x_λ in terms of the unknown parameters is generally difficult or impossible. The Collage-Based Approach to inverse problems instead bounds the approximation error above by a quantity that is more convenient to minimize. When dealing with ordinary differential equations (ODEs), this so-called collage distance is obtained using the Picard operator associated with the problem. By applying a modification of this idea to the system of ODEs provided by Pontryagin's Maximum Principle, we aim to handle an IOCP in a similar way.

The thesis is organized as follows. In Chapter 2 we present the requisite ODE theory. Existence and uniqueness of solutions is emphasized here, and the Picard operator is introduced which we return to later in the discussion of the Collage Method. We then formally introduce controlled ODE systems and present additional considerations to existence and uniqueness theory for these systems. We also briefly outline solution techniques for ODEs. In Chapter 3, the Collage-Based Approach to inverse ODE problems is introduced. The inverse problem is formally defined and we mention several other solution methods present in the literature. By drawing on the discussion of existence and uniqueness in Chapter 2, we present the Collage Theorem and formulate the inverse problem for ODEs using the Collage framework. This Chapter is closed by demonstrating the Collage Method with examples. Chapter 4 introduces Optimal Control Problems governed by ODEs. Pontryagin's Maximum Principle is stated and derived by an informal but accessible argument, and we then present an existence theorem for a particular class of Optimal Control Problems. A commonly used numerical solution procedure, the Forward Backward Sweep Method, is then introduced, its convergence properties are stated, and an example problem is solved. We also introduce the Linear Quadratic (LQ) Optimal Control Problem, a specific class of problems with desirable properties that we shall focus on. The culmination of the theory contained in Chapters 2–4 is presented in Chapter 5, where we state the IOCP governed by ODEs and present a Collage-Based Approach to its solution, assuming the corresponding forward problem has a unique solution. A variety of examples are included here to illustrate the capabilities and limitations of the method. We conclude in Chapter 6 with anticipated future work resulting from the research contained within this thesis.

Chapter 2

Systems of First Order ODEs: Classical Versus Controlled

The theory of ordinary differential equations (ODEs) underpins the majority of this thesis. In this Chapter we establish necessary background theory and, in particular, develop existence and uniqueness results for solutions of ODEs. These results are especially important to the theory of inverse problems, which we will see in Chapters 3 and 5. Following the discussion of "classical" ODEs, we introduce the idea of "controlled" ODEs and discuss the changes to the existing theory required to account for the additional term present in the controlled setting.

2.1 Classical First Order ODE Systems

In what follows, unless otherwise indicated, we let $I = [t_0 - a, t_0 + a], a > 0$ be a real-valued interval and we denote by $C^n(I)$ the space of all n-times continuously differentiable functions on I. We will use boldface notation to denote vectors and vector-valued functions, $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. Similarly, by $\mathbf{C}(I)$ we mean $C(I) \times C(I) \times \cdots \times C(I)$; the Cartesian product of n function spaces. We denote component-wise differentiation by $\frac{d}{dt}\mathbf{x}(t) = \dot{\mathbf{x}}(t)$. By $\|\mathbf{x}\|_\infty$, we denote the norm on $\mathbf{C}(I)$ given by,

$$\|\mathbf{x}\|_\infty \equiv \sum_{i=0}^{n} \|x_i(t)\|_\infty \equiv \sum_{i=0}^{n} \left\{ \sup_{t \in I} |x_i(t)| \right\}.$$

2.1.1 Existence and Uniqueness

We will consider first-order systems of ODEs of the form

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}) \tag{2.1}$$

$$\mathbf{x}(t_0) = \mathbf{x_0}, \tag{2.2}$$

where $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$, $\mathbf{g} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, and $\mathbf{x_0} \in \mathbb{R}^n$. Equation (2.1) is assumed to hold on some region $\mathcal{D}$ which is a subset of (t, x)-space, for example, $\mathcal{D} = I \times \Omega$, where $\Omega \subset \mathbb{R}^n$. The function $\mathbf{g}$ describes how the system variables $\mathbf{x}$ change in response to time t and their location in the space Ω. For now we will assume that $\mathbf{g}$ is at least continuous on $\mathcal{D}$. We also define the norm $\|\cdot\|_{\mathcal{D}}$,

$$\|\mathbf{g}(t, \mathbf{x})\|_{\mathcal{D}} \equiv \sum_{i=0}^{n} \left\{ \sup_{(t, \mathbf{x}) \in \mathcal{D}} |g_i(t, \mathbf{x})| \right\}.$$

Differential equations are the basis for models of physical systems which undergo smooth change. ODEs model systems in a single dimension, whereas partial differential equations (PDEs) model systems in many dimensions. We focus on ODE systems in this thesis.

When considering the *direct* or *forward* problem, one is given the vector field $\mathbf{g}$ and the initial condition $\mathbf{x}_0 \in \Omega$ and seeks a function $\mathbf{x}(t)$ such that

1. $\mathbf{x}(t) \in \mathbf{C^1}(I)$, and

2. $\mathbf{x}(t)$ satisfies (2.1) for all $t \in I$ and $\mathbf{x}(t) \in \Omega$.

When studying ODEs it is often of interest to determine when a solution to a given problem exists, and when it is unique. These attributes will also be of particular interest when we discuss inverse problems later on in this thesis. Banach's Fixed Point Theorem establishes conditions under which we can guarantee the existence and uniqueness of a solution to the IVP (2.1)–(2.2). We will first present a few definitions that are needed to understand this important theorem.

As we will see, Banach's Fixed Point Theorem is stated in terms of an operator $\mathbf{T}$ with particular properties. When working with ODEs, a natural choice for this operator is the *Picard Operator*, which we obtain by integrating both sides of the ODE (component-wise) in Equation (2.1)

$$\int_{t_0}^{t} \dot{\mathbf{x}}(s)ds = \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s))ds.$$

Applying the Fundamental Theorem of Calculus on the left-hand side, rearranging and using Equation (2.2) yields

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s))ds, \tag{2.3}$$

Equation (2.3) provides an implicit expression for the solution $\mathbf{x}(t)$, and suggests that we define the Picard operator $\mathbf{T}$ by

$$\mathbf{T}\mathbf{x}(t) = \mathbf{x_0} + \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s))ds. \tag{2.4}$$

Expressing the ODE problem in terms of the Picard operator permits a method to approximate solutions, called *Picard iteration* or the *Method of Successive Approximations.* One can define an iteration scheme beginning with an initial "guess" for the solution, and because the operator is *contractive*, the sequence of iterates converges to a *fixed point.*

Definition 2.1. *A fixed point of an operator* $\mathbf{T} : \mathbf{X} \to \mathbf{X}$ *is any point* $\bar{\mathbf{x}} \in \mathbf{X}$ *such that* $\mathbf{T}\bar{\mathbf{x}} = \bar{\mathbf{x}}$.

Definition 2.2. *Let* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *be a normed space. A mapping* $\mathbf{T} : \mathbf{X} \to \mathbf{X}$ *is called a contraction (or is contractive) if for any* $\mathbf{x}, \mathbf{y} \in \mathbf{X}$*, there exists* $c \in [0, 1)$ *such that*

$$\|\mathbf{T}\mathbf{x} - \mathbf{T}\mathbf{y}\|_{\mathbf{X}} \leq c \, \|\mathbf{x} - \mathbf{y}\|_{\mathbf{X}}$$

Mappings $\mathbf{T} : \mathbf{X} \to \mathbf{X}$, where the domain and codomain are the same, are said to be *space-preserving.* We denote the space of all contractive, space-preserving operators on $\mathbf{X}$ by $Con(\mathbf{X})$. For a full discussion of Picard iteration and the convergence of this method, see [17].

Fixed points of the Picard operator do in fact correspond to solutions of (2.1)–(2.2), as the following theorem states.

Theorem 2.1. *If* $\mathbf{x} \in \mathbf{C}(I)$ *is a fixed point of* (2.3) *then* $\mathbf{x} \in \mathbf{C}^1(I)$ *and is a solution of* (2.1)*–*(2.2)*.*

Proof. First, since $\mathbf{x}$ is a fixed point of Equation (2.3) we have that $\mathbf{x}(t_0) = \mathbf{x}_0$. Now, $\mathbf{x} \in \mathbf{C}(I)$ implies that $\mathbf{g}(t, \mathbf{x}(t))$ is a continuous function of t and so

$$\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s))ds \in \mathbf{C^1}(I) \implies \mathbf{x}(t) \in \mathbf{C^1}(I).$$

By the Fundamental Theorem of Calculus,

$$\frac{d}{dt}\left[\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s))ds\right] = \mathbf{g}(t, \mathbf{x}(t)).$$

Thus, $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x})$. □

Using the properties of an operator $\mathbf{T}$ that we have defined in Equation (2.4), we are now able to state Banach's Fixed Point Theorem, a classic result for proving the existence of a unique solution to the IVP (2.1)–(2.2).

Theorem 2.2. *(Banach's Fixed Point Theorem [31]) If* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *is a Banach space and* $\mathbf{T} \in Con(\mathbf{X})$*, then there exists a unique fixed point* $\bar{\mathbf{x}} \in \mathbf{X}$ *such that* $\mathbf{T}\bar{\mathbf{x}} = \bar{\mathbf{x}}$*.*

Proof. Let $\mathbf{x}_0 \in \mathbf{X}$ and define the sequence $\{\mathbf{x}_m\}_{m=0}^{\infty} \subset \mathbf{X}$ by

$$\mathbf{x}_{m+1} = \mathbf{T}\mathbf{x}_m$$

We will show that $\{\mathbf{x}_m\}$ is a Cauchy sequence.

$$\begin{aligned} \left\|\mathbf{x}_{m+1} - \mathbf{x}_m\right\|_{\mathbf{X}} &= \left\|\mathbf{T}\mathbf{x}_m - \mathbf{T}\mathbf{x}_{m-1}\right\|_{\mathbf{X}} \qquad m > 0 \\ &\leq c\left\|\mathbf{x}_m - \mathbf{x}_{m-1}\right\|_{\mathbf{X}} \\ &= c\left\|\mathbf{T}\mathbf{x}_{m-1} - \mathbf{T}\mathbf{x}_{m-2}\right\|_{\mathbf{X}} \\ &\leq c^2\left\|\mathbf{x}_{m-1} - \mathbf{x}_{m-2}\right\|_{\mathbf{X}} \end{aligned}$$

we can repeat this process to arrive at

$$\left\|\mathbf{x}_{m+1} - \mathbf{x}_m\right\|_{\mathbf{X}} \leq c^m\left\|\mathbf{x}_1 - \mathbf{x}_0\right\|_{\mathbf{X}}.$$

By the triangle inequality, for $n > m$,

$$\begin{aligned} \left\|\mathbf{x}_m - \mathbf{x}_n\right\|_{\mathbf{X}} &\leq \left\|\mathbf{x}_m - \mathbf{x}_{m+1}\right\|_{\mathbf{X}} + \left\|\mathbf{x}_{m+1} - \mathbf{x}_{m+2}\right\|_{\mathbf{X}} + \cdots + \left\|\mathbf{x}_{n-1} - \mathbf{x}_n\right\|_{\mathbf{X}} \\ &\leq \left(c^m + c^{m+1} + \cdots + c^{n-1}\right)\left\|\mathbf{x}_1 - \mathbf{x}_0\right\|_{\mathbf{X}}. \end{aligned}$$

Applying the formula for the sum of a geometric series yields the closed form

$$= c^m \frac{1 - c^{n-m}}{1 - c}\left\|\mathbf{x}_0 - \mathbf{x}_1\right\|_{\mathbf{X}}$$

Since $0 \leq c < 1$, we have in the numerator that $1 - c^{n-m} \leq 1$. So,

$$\left\|\mathbf{x}_m - \mathbf{x}_n\right\|_{\mathbf{X}} \leq \frac{c^m}{1 - c}\left\|\mathbf{x}_0 - \mathbf{x}_1\right\|_{\mathbf{X}}, \quad n > m. \tag{2.5}$$

The right-hand side of Equation (2.5) can be made arbitrarily small for large enough m. Therefore, $\{\mathbf{x}_m\}$ is a Cauchy sequence. Since $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is complete, there is some $\bar{\mathbf{x}} \in \mathbf{X}$ so that $\mathbf{x}_m \to \bar{\mathbf{x}}$. Now, we must show that $\bar{\mathbf{x}}$ is a fixed point of $\mathbf{T}$.

Using the triangle inequality and the contractivity of T,

$$\begin{aligned} \|\bar{\mathbf{x}} - \mathbf{T}\bar{\mathbf{x}}\|_{\mathbf{X}} &\leq \|\bar{\mathbf{x}} - \mathbf{x}_m\|_{\mathbf{X}} + \|\mathbf{x}_m - \mathbf{T}\bar{\mathbf{x}}\|_{\mathbf{X}} \\ &= \|\bar{\mathbf{x}} - \mathbf{x}_m\|_{\mathbf{X}} + \|\mathbf{T}\mathbf{x}_{m-1} - \mathbf{T}\bar{\mathbf{x}}\|_{\mathbf{X}} \\ &\leq \|\bar{\mathbf{x}} - \mathbf{x}_m\|_{\mathbf{X}} + c\,\|\mathbf{x}_{m-1} - \bar{\mathbf{x}}\|_{\mathbf{X}}\,. \end{aligned} \tag{2.6}$$

Since $\{\mathbf{x}_m\}$ converges to $\bar{\mathbf{x}}$, the right-hand side of (2.6) approaches zero as $m \to \infty$ so $\|\bar{\mathbf{x}} - \mathbf{T}\bar{\mathbf{x}}\| = 0$, hence $\bar{\mathbf{x}}$ is a fixed point of $\mathbf{T}$. All that remains is to demonstrate that $\bar{\mathbf{x}}$ is the only fixed point. Assume that there exists another fixed point $\tilde{x} \in \mathbf{X}$. Then,

$$\begin{aligned} \|\bar{\mathbf{x}} - \tilde{\mathbf{x}}\|_{\mathbf{X}} &= \|\mathbf{T}\bar{\mathbf{x}} - \mathbf{T}\tilde{\mathbf{x}}\|_{\mathbf{X}} \\ &\leq c\,\|\bar{\mathbf{x}} - \tilde{\mathbf{x}}\|_{\mathbf{X}} \\ &\implies \|\bar{\mathbf{x}} - \tilde{\mathbf{x}}\|_{\mathbf{X}} = 0 \end{aligned}$$

Thus $\bar{\mathbf{x}} = \tilde{\mathbf{x}}$. This completes the proof. □

We complete this section by formally stating and proving the Picard-Lindelöf existence and uniqueness theorem. This theorem states the conditions that the IVP (2.1)–(2.2) must meet in order to construct the Picard operator that satisfies the hypothesis of Banach's Fixed Point Theorem. First, we define the condition required on the right-hand side of Equation (2.1).

Definition 2.3. *Let $\mathcal{D} = I \times \Omega$. We say that $\mathbf{g} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is uniformly Lipschitz continuous in $\mathbf{x}$ on $\mathcal{D}$ if there exists a constant $L > 0$ such that for every $(t, \mathbf{x_1})$ and*

$(t, \mathbf{x_2}) \in \mathcal{D}$,

$$\|\mathbf{g}(t, \mathbf{x_1}) - \mathbf{g}(t, \mathbf{x_2})\|_{\mathcal{D}} \leq L \|\mathbf{x_1} - \mathbf{x_2}\|_{\infty} \tag{2.7}$$

Alternatively, $\mathbf{g}$ *is said to be uniformly Lipschitz in* $\mathbf{x}$.

Theorem 2.3. *(Picard-Lindelöf [46]) Consider the IVP* (2.1)–(2.2)*. Suppose* $\mathbf{g}$ *is continuous in* t *and uniformly Lipschitz in* $\mathbf{x}$ *with Lipschitz constant* L *on the domain* $\mathcal{D} = I \times \Omega$*, where* $I = [t_0 - a, t_0 + a]$*,* $\Omega = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x} - \mathbf{x}_0\|_2 \leq b, b > 0\}$*, and* $\mathbf{x}_0 \in \Omega$*. Then there exists a unique solution* $\mathbf{x}(t)$ *on* $\mathcal{D}$ *to* (2.1)–(2.2)*, provided that* $a < \dfrac{1}{L}$ *and* $\|\mathbf{g}(t, \mathbf{x})\|_{\mathcal{D}} \leq \dfrac{b}{a}$.

Proof. Define $\overline{\mathbf{C}(I)} = \{\mathbf{x}(t) \in \mathbf{C}(I) | \|\mathbf{x} - \mathbf{x}_0\|_{\infty} \leq b, \quad t \in I\}$. By hypothesis, $b > 0$ is chosen such that $\mathbf{x}(t_0) = \mathbf{x}_0 \in \Omega$. By choosing $a > 0$ small enough, we are guaranteed that $\mathbf{x}(t) \in \Omega$ for $t \in I$. We will show that the Picard operator defined in (2.4) is space-preserving and contractive, and thus by Banach's Fixed Point Theorem, has a unique fixed point.

For $\mathbf{x} \in \overline{\mathbf{C}(I)}$, continuity of $(\mathbf{Tx})(t)$ follows directly from the continuity of $\mathbf{g}$. Since $\mathbf{g}$ is a continuous function of $t \in I$ and Ω is a closed subset, by the Extreme Value Theorem $\mathbf{g}$ is bounded on Ω and the upper bound $\dfrac{b}{a}$ exists. Consider $t \in [t_0, t_0 + a]$. Then

$$\begin{aligned}
\|\mathbf{Tx} - \mathbf{x}_0\|_{\infty} &= \left\| \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s)) ds \right\|_{\infty} \\
&\leq \int_{t_0}^{t} \|\mathbf{g}(s, \mathbf{x}(s))\|_{\mathcal{D}} \, ds \\
&\leq \frac{b}{a} (t - t_0)
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{b}{a}a \\
&\le b.
\end{aligned}
$$

This result holds also for $t \in [t_0 - a, t_0]$. Consider now $\mathbf{x}, \mathbf{y} \in \overline{\mathbf{C}(I)}$. Because $\mathbf{g}$ is uniformly Lipschitz in $\mathbf{x}$ on $\mathcal{D}$,

$$
\begin{aligned}
\|\mathbf{Tx} - \mathbf{Ty}\|_\infty &= \left\| \int_{t_0}^{t} \mathbf{g}(s, \mathbf{x}(s)) - \mathbf{g}(s, \mathbf{y}(s)) ds \right\|_\infty \\
&\le \int_{t_0}^{t} \|\mathbf{g}(s, \mathbf{x}(s)) - \mathbf{g}(s, \mathbf{y}(s))\|_{\mathcal{D}} \, ds \\
&\le L \int_{t_0}^{t} \|\mathbf{x}(s) - \mathbf{y}(s)\|_\infty \, ds \\
&\le La \|\mathbf{x} - \mathbf{y}\|_\infty \\
&= c \|\mathbf{x} - \mathbf{y}\|_\infty .
\end{aligned}
$$

So $\mathbf{T}$ is a contraction with contraction factor $c = La < 1$, provided that $a < \frac{1}{L}$. Hence, by Banach's Fixed Point Theorem, $\mathbf{T}$ has a unique fixed point since $a < \frac{1}{L}$. By Theorem 2.1, this fixed point of $\mathbf{T}$ is the unique solution to (2.1)–(2.2) on I. □

One additional result which we do not prove here is the continuity of fixed points of Picard operators, and thus of solutions to IVPs given by Equations (2.1)–(2.2). See Proposition 1 in [39].

Remark: The preceding theory may be extended beyond the IVPs that have been presented here. For instance, some problems with conditions at a terminal time may also have unique solutions under similar conditions.

2.1.2 Solution Methods

There is a wealth of theory surrounding solution techniques for ODEs. Closed-form solutions can be found for a variety of ODEs, using methods such as integrating factors, separation of variables, and variation of parameters, among others. Many books can be found on this topic; see for instance [10].

While closed-form solutions exist for some ODE systems, in general they may be difficult to find. There are many numerical techniques for finding solutions that are often used in practice. The *Runge-Kutta* methods are a well-known family of so-called single-step methods used extensively for numerically solving many ODE systems, and include the Euler, Heun, and Runge-Kutta-Fehlberg methods. Such a method will be utilized later, so we will briefly introduce them here.

The Runge-Kutta methods are motivated by *quadrature rules*, or approximations of the definite integral of a function. Suppose that we want to numerically solve the IVP (2.1)–(2.2) on an interval $[t_0, t_f]$, using an N-point temporal discretization and fixed step size $h = \dfrac{t_f - t_0}{N+1}$. In order to solve Equation (2.1) from $t_n = t_0 + nh$ to t_{n+1}, one could integrate to obtain

$$\begin{aligned}\mathbf{x}(t_{n+1}) &= \mathbf{x}(t_n) + \int_{t_n}^{t_{n+1}} \mathbf{g}(s, \mathbf{x}(s))ds \\ &= \mathbf{x}(t_n) + h\int_0^1 \mathbf{g}(t_n + hs, \mathbf{x}(t_n + hs))ds, \qquad (2.8)\end{aligned}$$

and approximate the integral on the right-hand side of Equation (2.8) using quadrature. For instance,

$$\mathbf{x}(t_{n+1}) = \mathbf{x}(t_n) + h\sum_{j=1}^{\nu} b_j \mathbf{g}(t_n + c_j h, \mathbf{x}(t_n + c_j h)), \quad n = 0, 1, \ldots$$

Since the quantity $\mathbf{x}(t_n + c_j h)$ is unknown, we must use an approximation. Typically we denote the approximation of $\mathbf{x}(t_n + c_j h)$ by $\mathbf{k}_j, j = 0, \ldots, \nu$. In all *explicit* Runge-Kutta methods, these approximations are defined as

$$\begin{aligned}
\mathbf{k}_1 &= \mathbf{x}_n, \\
\mathbf{k}_2 &= \mathbf{x}_n + ha_{2,1}\mathbf{g}(t_n, \mathbf{k}_1), \\
\mathbf{k}_3 &= \mathbf{x}_n + ha_{3,1}\mathbf{g}(t_n, \mathbf{k}_1) + ha_{3,2}\mathbf{g}(t_n + c_2 h, \mathbf{k}_2), \\
&\vdots \\
\mathbf{k}_\nu &= \mathbf{x}_n + h\sum_{i=1}^{\nu-1} a_{\nu,i}\mathbf{g}(t_n + c_i h, \mathbf{k}_i),
\end{aligned}$$

which yields the following formula for the next step in the solution,

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\sum_{j=1}^{\nu} b_j \mathbf{g}(t_n + c_j h, \mathbf{k}_j).$$

Any Runge-Kutta method is defined by the vector of *weights*, $\mathbf{b} = [b_1, \ldots, b_\nu]^T$, the vector of *nodes*, $\mathbf{c} = [c_1, \ldots, c_\nu]^T$, and the Runge-Kutta matrix,

$$A = [a_{i,j}]_{i,j=1..\nu},$$

where any undefined entries are zero.

One particular Runge-Kutta method that will be utilized later in this thesis is the Runge-Kutta fourth order method, sometimes called the RK4 or RK41 method.

This method is an explicit, constant time-stepping method that is straightforward to implement while having convergence properties that are attractive in practice. Consider the interval $[t_0, t_f]$ and suppose that we wish to solve the IVP (2.1)–(2.2) on this interval using $N+1$ mesh points. Then we can state the RK4 method as follows.

```
Set $h \leftarrow \frac{t_f - t_0}{N+1}$
for $n = 0 \to N$ do
    $\mathbf{k}_1 \leftarrow \mathbf{g}(t_n, \mathbf{x}_n)$
    $\mathbf{k}_2 \leftarrow \mathbf{g}\left(t_n + \frac{h}{2}, \mathbf{x}_n + h\frac{\mathbf{k}_1}{2}\right)$
    $\mathbf{k}_3 \leftarrow \mathbf{g}\left(t_n + \frac{h}{2}, \mathbf{x}_n + h\frac{\mathbf{k}_2}{2}\right)$
    $\mathbf{k}_4 \leftarrow \mathbf{g}(t_n + h, \mathbf{x}_n + h\mathbf{k}_3)$
    $\mathbf{x}_{n+1} \leftarrow \mathbf{x}_n + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$
    $t_{n+1} \leftarrow t_n + h$
end for
```

A full discussion of these methods is beyond the scope of this thesis; see [13, 14, 25, 54] for more details.

2.2 Controlled First Order ODE Systems

So far, the ODE systems we have considered have not accounted for the effects of any *external* influences. The particular type of influence that we are interested in is a deliberate, calculated input that we can use to drive the system in a desirable way or to achieve a particular outcome. This idea is called *control*, and a *control system* uses external input to direct the behaviour of some process, which we will model with ODE systems.

Consider, for example, a model of cancer tumour growth. An uncontrolled system could model the dynamics of tumour growth in the absence of treatment. A controlled system might incorporate the use of chemotherapy, with the goal of diminishing tumour growth. In this case, the dosage and schedule of chemotherapy drugs are what constitute the control input to the system.

The control systems we will work with are similar to the form of the first-order ODE in (2.1), but include dependence on an external input,

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \tag{2.9}$$

where $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ is called the *state* or *process* variable and $\mathbf{u} : \mathbb{R} \to \mathbb{R}^m$ is the *control input* to the system. This equation, in the context of control problems, is often called the *state equation.*

Example 2.1. *The double integrator*

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u, \end{aligned}$$

is a first-order control system modelling the dynamics of a point mass under the influence of a time-varying control input force $u(t) \in \mathbb{R}$.

We can use control systems to model physical systems in a similar fashion as with classical ODEs, but the inclusion of the control input allows for the formal consideration of forces external to the underlying system. *Control theory* is the study of this type of system, and specifically considers the design and analysis of the control inputs and their effect on the physical system. Control theory is a vast field of study that is highly significant to many branches of engineering and and beyond. See [48] for an introductory text. We investigate one area of control theory, called *optimal control*, in this thesis.

The existence and uniqueness theory from the previous section does not address the added control term in the right-hand side of Equation (2.9). However, we can still utilize theory from Section 2.1.1 with some additional assumptions in order to establish an existence and uniqueness result for controlled systems. We note that the control input depends only on time, so we will consider the right-hand side of Equation (2.9) as

$$\bar{\mathbf{g}}(t, \mathbf{x}) \equiv \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)). \tag{2.10}$$

In order to apply Theorem 2.3 we know that $\mathbf{g}$ must be uniformly Lipschitz in $\mathbf{x}$ for each fixed t. Using Equation (2.10), the Lipschitz condition (2.7) becomes

$$\|\bar{\mathbf{g}}(t, \mathbf{x}_1) - \bar{\mathbf{g}}(t, \mathbf{x}_2)\|_{\mathbf{X}} = \|\mathbf{g}(t, \mathbf{x}_1(t), \mathbf{u}(t)) - \mathbf{g}(t, \mathbf{x}_2(t), \mathbf{u}(t))\|_{\mathbf{X}} \leq L \|\mathbf{x_1}(t) - \mathbf{x_2}(t)\|_{\mathbf{X}},$$

for $L > 0$, $\mathbf{x}_1, \mathbf{x}_2$ in a normed space $(\mathbf{X}, \|\cdot\|_X)$, and for any $t \in I$. Thus, as long as the right-hand side of Equation (2.9) satisfies this modified Lipschitz condition, we can apply Theorem 2.3. The regularity of the control $\mathbf{u}$ is then not required to be any stronger than piecewise continuous, and indeed may be relaxed, for instance, to locally integrable functions. See [40, 44, 60] for additional discussion.

Chapter 3

A Collage-Based Approach to Inverse Problems for Classical ODE Systems

With the theory of ODEs established, we now define and consider a method for solving an ODE inverse problem. In this Chapter, we will define an ODE inverse problem and develop the necessary background theory to present the Collage-Based Approach to these problems. An example is also presented to illustrate the method.

3.1 Inverse Problems

In Chapter 2 we defined and investigated the *direct* or *forward* (classical) ODE system problem. Now, we consider an associated inverse problem to the ODE system

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}) \tag{3.1}$$

$$\mathbf{x}(0) = \mathbf{x_0}. \tag{3.2}$$

Recall that the forward problem provides the function $\mathbf{g}$ and the initial condition $\mathbf{x}_0$ and seeks a $\mathbf{C}^1$ function $\mathbf{x}(t)$ that satisfies (3.1)–(3.2). Conversely, in the related inverse problem we suppose that we are given a *target* solution $\mathbf{x}_{target}(t)$, possibly in the form of data points, and we seek a vector field $\mathbf{g}(t, \mathbf{x})$ and possibly an initial condition $\mathbf{x}_0$ so that our target solution satisfies (3.1)–(3.2). In this case, the vector field $\mathbf{g}$ is defined in terms of a set of parameters $\lambda \in \Lambda$, that are to be chosen such that the approximation error is minimized,

$$\min_{\lambda \in \Lambda} \left\| \mathbf{x}_{target} - \mathbf{x}_\lambda \right\|, \tag{3.3}$$

where $\mathbf{x}_\lambda$ is the solution of (3.1)–(3.2) obtained using the recovered parameters, Λ is the space of parameters, and $\|\cdot\|$ is an appropriate norm.

Several solution techniques exist in the literature. For instance, regularization methods, such as Tikhonov regularization, approximate an ill-posed problem with a well-posed one, and then account for the discrepancy using penalization terms. Iterative methods, such as Landweber-Fridman iteration, instead look to directly minimize the approximation error by way of a convergent iterative scheme. An overview and comparison of these methods can be found in [41], although a full discussion of these methods is beyond the scope of this thesis. We instead focus on a Collage-Based Approach for solving such an inverse problem.

3.2 The Collage-Based Approach to ODE Inverse Problems

In Chapter 2 we defined the set $Con(\mathbf{X})$ to be the set of all contractive, space-preserving operators on $\mathbf{X}$. Here, we consider a subset of these operators that are defined by parameters λ, which we denote $Con_\lambda(\mathbf{X})$. One can view the inverse problem as searching for an operator $\mathbf{T}_\lambda \in Con_\lambda(\mathbf{X})$ with fixed point $\bar{\mathbf{x}}_\lambda$ such that the approximation error (3.3) is minimized. Using this formulation, we can define the parameter space $\Lambda = \{\lambda \in \mathbb{R}^{dim(\lambda)} | \mathbf{T}_\lambda \in Con_\lambda(\mathbf{X})\}$.

Using Banach's Fixed Point Theorem, one can view minimization of the approximation error instead as minimization of $\|\mathbf{x}_{\text{target}} - \bar{\mathbf{x}}_\lambda\|$. However, expressing $\bar{\mathbf{x}}_\lambda$ in terms of the parameters that define $\mathbf{T}_\lambda$ is difficult or impossible in practice. The idea behind a Collage-Based Approach is to bound the approximation error above by a quantity that is more easily determined. A simple consequence of Banach's Fixed Point Theorem, the *Collage Theorem*, provides this quantity, called the *collage distance.*

Theorem 3.1. *(Collage Theorem) Let* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *be a Banach space and* $\mathbf{T}_\lambda : \Lambda \times \mathbf{X} \to \mathbf{X}$ *be a contractive operator with contraction factor* c_λ *and unique fixed point* $\bar{\mathbf{x}}_\lambda \in \mathbf{X}$*. Then*

$$\|\mathbf{x} - \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} \leq \frac{1}{1 - c_\lambda} \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} .$$

Proof.

$$\begin{aligned} \|\mathbf{x} - \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} &= \|\mathbf{x} - \mathbf{T}_\lambda \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} \\ &\leq \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} + \|\mathbf{T}_\lambda \mathbf{x} - \mathbf{T}_\lambda \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} \\ &\leq \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} + c_\lambda \|\mathbf{x} - \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} \\ \implies (1 - c_\lambda) \|\mathbf{x} - \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} &\leq \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} \\ \implies \|\mathbf{x} - \bar{\mathbf{x}}_\lambda\|_{\mathbf{X}} &\leq \frac{1}{1 - c_\lambda} \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} \end{aligned}$$

□

The Collage Theorem ensures that we can control the approximation error, as long as $c_\lambda \ll 1$. The name of the theorem is derived from its origin in Barnsley's work on fractal image compression [5]. We define the collage distance to be the quantity $\|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}}$, and thus the inverse problem requires that we find the parameters $\lambda \in \Lambda$ which minimize the collage distance. Mathematically,

$$\min_{\lambda \in \Lambda} \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}} .$$

This bounding of the approximation error by the collage distance is the essence of the Collage-Based Approach. The advantage of this approach lies in the computation of the collage distance, as it requires only the operations needed for a single application of the operator $\mathbf{T}_\lambda$ to the target function $\mathbf{x}_{target}$. In contrast, direct computation of the approximation error by iteration to find the fixed point of $\mathbf{T}_\lambda$ will in general require many iterations to achieve a desired tolerance.

Although the Collage-Based Approach saves significantly on computational resources, it is suboptimal as the following theorem states.

Theorem 3.2. *(Suboptimality of the Collage Theorem)[33] Let* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *be a Banach space, and* $\mathbf{x} \in \mathbf{X}$ *be a target function. Further, let* $\lambda_{min} = \arg\min_{\lambda \in \Lambda} \|\mathbf{x} - \mathbf{T}_\lambda \mathbf{x}\|_{\mathbf{X}}$ *be the parameter values that minimize the collage distance, with corresponding fixed point* $\bar{\mathbf{x}}_{\lambda_{min}}$ *of the operator* $\mathbf{T}_{\lambda_{min}} \in Con_\lambda(\mathbf{X})$*. Let* $\bar{\mathbf{x}}_{\lambda_{opt}}$ *be the optimal fixed point that minimizes the approximation error* $\|\bar{\mathbf{x}}_\lambda - \mathbf{x}\|_{\mathbf{X}}$*; that is,* $\bar{\mathbf{x}}_{\lambda_{opt}}$ *satisfies*

$$\left\|\bar{\mathbf{x}}_{\lambda_{opt}} - \mathbf{x}\right\|_{\mathbf{X}} \le \|\mathbf{y} - \mathbf{x}\|_{\mathbf{X}} \tag{3.4}$$

for all $\mathbf{y}$ *satisfying* $\mathbf{T}_\lambda \mathbf{y} = \mathbf{y}$ *and some* $\lambda \in \Lambda$*. Then*

$$\left\|\bar{\mathbf{x}}_{\lambda_{opt}} - \bar{\mathbf{x}}_{\lambda_{min}}\right\|_{\mathbf{X}} \le \frac{2}{1 - c_{\lambda_{min}}} \|\mathbf{x} - \mathbf{T}_{\lambda_{min}} \mathbf{x}\|_{\mathbf{X}},$$

where $c_{\lambda_{min}}$ *is the contraction factor of* $\mathbf{T}_{\lambda_{min}}$*.*

Proof. By the triangle inequality,

$$\begin{aligned}\left\|\bar{\mathbf{x}}_{\lambda_{opt}} - \bar{\mathbf{x}}_{\lambda_{min}}\right\|_{\mathbf{X}} &\le \left\|\bar{\mathbf{x}}_{\lambda_{opt}} - \mathbf{x}\right\|_{\mathbf{X}} + \|\mathbf{x} - \bar{\mathbf{x}}_{\lambda_{min}}\|_{\mathbf{X}} \\ &\le 2\|\mathbf{x} - \bar{\mathbf{x}}_{\lambda_{min}}\|_{\mathbf{X}}\end{aligned}$$

where we have taken $\mathbf{y} = \bar{\mathbf{x}}_{\lambda_{min}}$ in Equation (3.4). Applying the Collage Theorem to the right-hand side yields

$$\left\|\bar{\mathbf{x}}_{\lambda_{opt}} - \bar{\mathbf{x}}_{\lambda_{min}}\right\|_{\mathbf{X}} \leq \frac{2}{1 - c_{\lambda_{min}}} \left\|\mathbf{x} - \mathbf{T}_{\lambda_{min}}\mathbf{x}\right\|_{\mathbf{X}}.$$

□

The Collage-Based Approach for solving inverse problems depends on the following continuity result, which ensures that contractive, space-preserving operators that are "nearby" in the sense of $Con_\lambda(\mathbf{X})$ have fixed points that are also close together.

Theorem 3.3. *(Continuity Theorem for Fixed Points) [39] Let* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *be a Banach space and* $T_\lambda^1 : \Lambda \times \mathbf{X} \to \mathbf{X}$ *and* $T_\lambda^2 : \Lambda \times \mathbf{X} \to \mathbf{X}$ *be contractive operators with contraction factors* c_λ^1 *and* c_λ^2*, respectively. If* $\bar{\mathbf{x}}_\lambda^1$ *and* $\bar{\mathbf{x}}_\lambda^2$ *are the unique fixed points of* $\mathbf{T}_\lambda^1$ *and* $\mathbf{T}_\lambda^2$ *respectively then*

$$\left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \leq \frac{1}{1 - \min\{c_\lambda^1, c_\lambda^2\}} \left\|\mathbf{T}_\lambda^1 - \mathbf{T}_\lambda^2\right\|_{Con_\lambda(\mathbf{X})}$$

where $\left\|\mathbf{T}_\lambda^1 - \mathbf{T}_\lambda^2\right\|_{Con_\lambda(\mathbf{X})} = \sup_{\mathbf{x}\in\mathbf{X}} \left\|\mathbf{T}_\lambda^1\mathbf{x} - \mathbf{T}_\lambda^2\mathbf{x}\right\|_{\mathbf{X}}$

Proof.

$$\begin{aligned} \left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} &= \left\|\mathbf{T}_\lambda^1\bar{\mathbf{x}}_\lambda^1 - \mathbf{T}_\lambda^2\bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \\ &\leq \left\|\mathbf{T}_\lambda^1\bar{\mathbf{x}}_\lambda^1 - \mathbf{T}_\lambda^2\bar{\mathbf{x}}_\lambda^1\right\|_{\mathbf{X}} + \left\|\mathbf{T}_\lambda^2\bar{\mathbf{x}}_\lambda^1 - \mathbf{T}_\lambda^2\bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \\ &\leq \left\|\mathbf{T}_\lambda^1\bar{\mathbf{x}}_\lambda^1 - \mathbf{T}_\lambda^2\bar{\mathbf{x}}_\lambda^1\right\|_{\mathbf{X}} + c_\lambda^2 \left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \\ \implies (1 - c_\lambda^2) \left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} &\leq \sup_{\mathbf{x}\in\mathbf{X}} \left\|\mathbf{T}_\lambda^1\mathbf{x} - \mathbf{T}_\lambda^2\right\|_{\mathbf{X}} \end{aligned}$$

$$\implies \left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \leq \frac{1}{1-c_\lambda^2} \left\|\mathbf{T}_\lambda^1 - \mathbf{T}_\lambda^2\right\|_{Con_\lambda(\mathbf{X})}.$$

By a similar argument we also have that

$$\left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \leq \frac{1}{1-c_\lambda^1} \left\|\mathbf{T}_\lambda^1 - \mathbf{T}_\lambda^2\right\|_{Con_\lambda(\mathbf{X})}.$$

Therefore,

$$\left\|\bar{\mathbf{x}}_\lambda^1 - \bar{\mathbf{x}}_\lambda^2\right\|_{\mathbf{X}} \leq \frac{1}{1-\min\{c_\lambda^1, c_\lambda^2\}} \left\|\mathbf{T}_\lambda^1 - \mathbf{T}_\lambda^2\right\|_{Con_\lambda(\mathbf{X})}$$

□

In Chapter 2 in the context of the ODE problem (3.1)–(3.2), we defined the Picard operator $\mathbf{T} : \mathbf{C}(I) \to \mathbf{C}(I)$ on the Banach space of continuous functions on I, $(\mathbf{C}(I), \|\cdot\|_\infty)$. In practice, however, working with the sup norm $\|\cdot\|_\infty$ is computationally undesirable. We instead work with the more convenient L^2 norm, $\|\cdot\|_2$. It is important to note that although $(\mathbf{C}(I), \|\cdot\|_2)$ is not a Banach space, there is no issue with our theory because $\mathbf{C}(I) \subset \mathbf{L}^2(I)$.

The Collage-Based Approach to inverse problems has been extended beyond ODE inverse problems, to boundary value problems [32, 42], variational equations and PDEs [34, 37, 38, 43], and random differential equations [15, 35, 37], for instance.

We conclude this Chapter by illustrating the Collage Method with two examples.

Example 3.1. *Consider the following first-order ODE*

$$\begin{cases} \dot{x}(t) = g(t)x(t), \\ x(0) = 1. \end{cases} \tag{3.5}$$

Suppose that we wish to solve the inverse problem of identifying $g(t)$ given $x(t)$ (possibly in the form of observed data). We choose $x(t) = exp(-\frac{1}{3}t(t^2 - 3t - 3))$ and simulate observational data by sampling at N uniformly distributed data points on $I = [0, 1]$. To account for experimental error, we add normally distributed noise with amplitude ϵ to these data points. We then fit a third degree polynomial target function, $x_{target}(t)$ to this data using a least-squares procedure. Note that $g_{true}(t) = 1 + 2t - t^2$ is the exact function that we seek. We use this for comparison purposes.

Assuming a second degree polynomial representation of $g_{collage}(t)$,

$$g_{collage}(t) = b_0 + b_1 t + b_2 t^2,$$

we determine the coefficients b_i, $i = 0, 1, 2$, such that the collage distance is minimized. We then solve the IVP (3.5) defined by the recovered coefficients, which we denote by $x_{collage}(t)$, for comparison purposes. The results for various numbers of sample points N and amplitudes of noise ϵ are presented in Table 3.1. Recovered parameters in $g_{collage}(t)$ are presented in Table 3.2.

As expected, we see that the approximation error decreases as the number of sample points N is increased, and this error increases as we increase the amplitude of the noise.

Example 3.1 is greatly simplified by the fact that a polynomial function is being approximated by a polynomial function. This serves as a useful example to illustrate the Collage Method as the parameters may be directly compared, but this is an

ϵ	N	$\Vert x_{true} - x_{collage}\Vert_{L^2(I)}$	$\Vert g_{true} - g_{collage}\Vert_{L^2(I)}$	$\Vert x - Tx\Vert_{L^2(I)}$
0	5	3.686×10^{-3}	8.436×10^{-3}	5.756×10^{-5}
	10	2.945×10^{-3}	7.762×10^{-3}	4.694×10^{-5}
	25	2.520×10^{-3}	7.906×10^{-3}	4.028×10^{-5}
0.01	5	9.107×10^{-3}	4.685×10^{-2}	3.452×10^{-5}
	10	5.554×10^{-3}	3.769×10^{-2}	2.511×10^{-5}
	25	4.797×10^{-3}	3.368×10^{-2}	3.228×10^{-5}
0.05	5	3.389×10^{-2}	2.011×10^{-1}	1.475×10^{-5}
	10	2.192×10^{-2}	1.590×10^{-1}	6.825×10^{-5}
	25	2.073×10^{-2}	1.413×10^{-1}	2.153×10^{-4}

Table 3.1: Results for the Collage Method applied to Equation (3.5). The exact function is $g_{true}(t) = 1 + 2t - t^2$.

ϵ	N	$g_{collage}(t)$
0	5	$0.9770 + 2.0949t - 1.0725t^2$
	10	$0.9777 + 2.1022t - 1.0879t^2$
	25	$0.9786 + 2.1081t - 1.1019t^2$
0.01	5	$0.8600 + 2.5552t - 1.4516t^2$
	10	$0.8873 + 2.4464t - 1.3762t^2$
	25	$0.8993 + 2.3898t - 1.3240t^2$
0.05	5	$0.3971 + 4.3793t - 2.9551t^2$
	10	$0.5273 + 3.8179t - 2.5252t^2$
	25	$0.5819 + 3.5166t - 2.2122t^2$

Table 3.2: Function $g(t)$ recovered using the Collage Method applied to Equation (3.5). The exact function is $g_{true}(t) = 1 + 2t - t^2$.

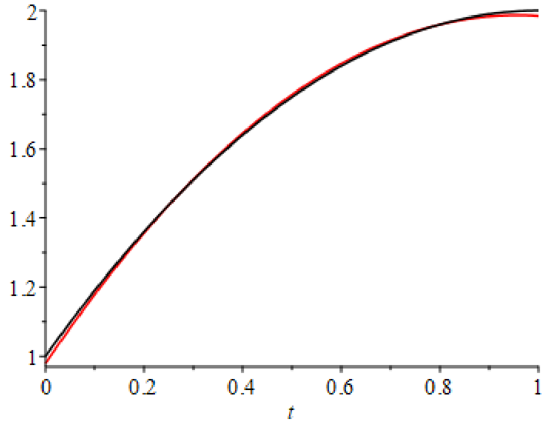

Figure 3.1: Plot of $g_{true}(t)$ (black) versus $g_{collage}(t)$ (red) for $N = 25$ and $\epsilon = 0$.

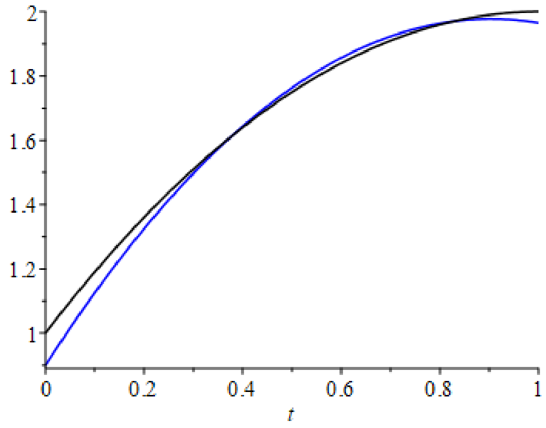

Figure 3.2: Plot of $g_{true}(t)$ (black) versus $g_{collage}(t)$ (blue) for $N = 25$ and $\epsilon = 0.01$.

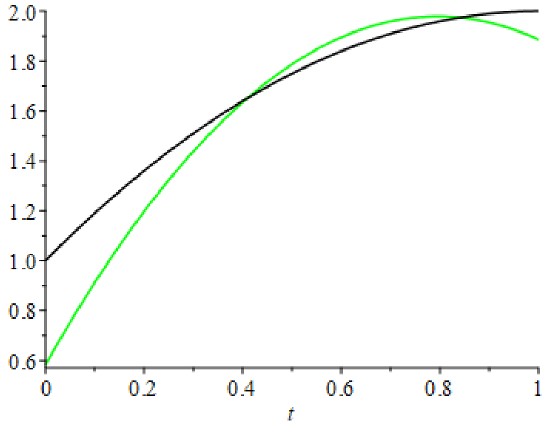

Figure 3.3: Plot of $g_{true}(t)$ (black) versus $g_{collage}(t)$ (green) for $N = 25$ and $\epsilon = 0.05$.

ideal scenario. We now present a second example where the unknown function in the inverse problem is not a polynomial.

Example 3.2. *Consider the ODE problem*

$$\begin{cases} \dot{x}(t) = e^{1-x(t)}, \\ x(0) = 1. \end{cases} \tag{3.6}$$

The exact solution of this problem is $x_{true}(t) = 1 + \ln(t+1)$. We again sample from the true solution and add normally distributed noise with amplitude ϵ to simulate experimental error before fitting a third degree polynomial, $x_{target}(t)$ to the data via a least squares procedure. We consider the inverse problem of finding a polynomial function $p(x)$ that admits $x_{target}(t)$ as an approximate solution to the ODE

$$\begin{cases} \dot{x}(t) = p(x), \\ x(0) = 1. \end{cases}$$

The results for various numbers of sample points N and amplitudes of noise ϵ are presented in Table 3.3. Recovered polynomial representations of degree M for $p_{collage}(x)$ may be found in Table 3.4. For comparison, the second degree Taylor polynomial for $p_{true}(x) = e^{1-x}$ is $p_{taylor}(x) = 2.7183 - 2.7183x + 1.3591x^2$. The approximation errors tend to decrease with a greater number of sample points and increase with greater amounts of error, as expected. We also see an improvement in the approximation error between the true solution and the collage solution when we move from recovering a linear function to a quadratic one. Although the recovered $p_{collage}(x)$ may not appear to closely match $p_{taylor}(x)$ based solely on the parameter values, the error *reported for even these relatively small values of N and the resulting approximation error in the solution demonstrate that the true problem and the recovered problem are nearby. That being said, Figures 3.4–3.6 illustrate the necessity for a greater number of sample points, N.*

ϵ	N	M	$\|x_{true} - x_{collage}\|_{L^2(I)}$	$\|p_{true} - p_{collage}\|_{L^2(I)}$	$\|x - Tx\|_{L^2(I)}$
0	5	1	1.421×10^{-3}	1.105×10^{-2}	1.640×10^{-6}
	10	1	1.421×10^{-3}	1.102×10^{-2}	1.783×10^{-6}
	25	1	1.420×10^{-3}	1.099×10^{-2}	1.917×10^{-6}
0.01	5	1	5.788×10^{-3}	4.216×10^{-2}	4.798×10^{-6}
	10	1	3.820×10^{-3}	2.939×10^{-2}	7.978×10^{-6}
	25	1	3.516×10^{-3}	2.995×10^{-2}	7.715×10^{-6}
0.05	5	1	2.922×10^{-2}	1.980×10^{-1}	3.546×10^{-4}
	10	1	2.140×10^{-2}	1.337×10^{-1}	3.901×10^{-4}
	25	1	2.034×10^{-2}	1.382×10^{-1}	3.136×10^{-4}
0	5	2	2.171×10^{-4}	1.578×10^{-3}	5.700×10^{-8}
	10	2	1.502×10^{-4}	1.149×10^{-3}	4.921×10^{-8}
0.01	5	2	5.681×10^{-3}	4.820×10^{-2}	5.682×10^{-7}
	10	2	3.719×10^{-3}	4.190×10^{-2}	9.678×10^{-7}
0.05	5	2	3.123×10^{-2}	2.582×10^{-1}	3.562×10^{-6}
	10	2	1.718×10^{-2}	2.188×10^{-1}	3.482×10^{-5}

Table 3.3: Results for the Collage Method applied to (3.6).

Remark: In both examples we minimized the collage distance by solving the systems generated by first-order conditions. The minimization scheme may need to be chosen strategically to compute the minimum for a particular problem. For example, a gradient descent scheme may be employed, or a more exotic method, such as Particle Swarm Ant Colony Optimization (PSACO) may be used to minimize the collage distance [12].

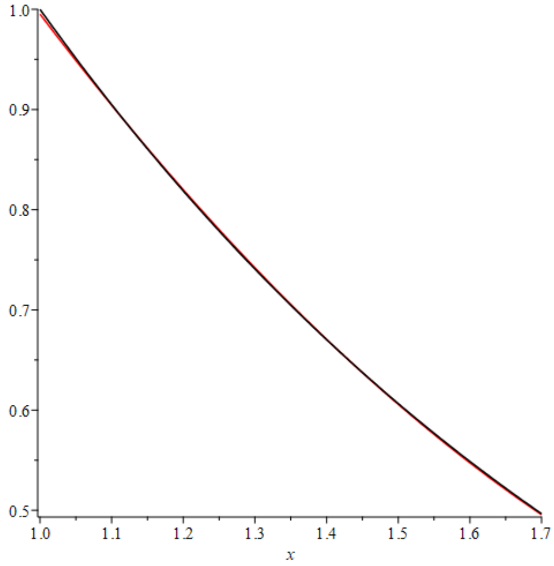

Figure 3.4: Plot of $p_{true}(t)$ (black) versus $p_{collage}(t)$ (red) for $N = 10$ and $\epsilon = 0$.

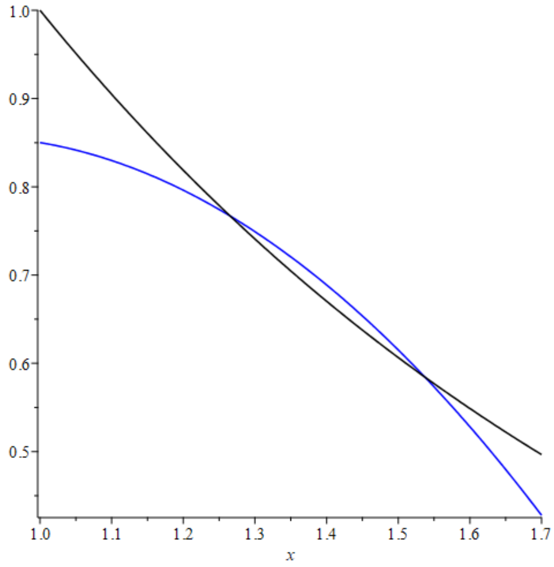

Figure 3.5: Plot of $p_{true}(t)$ (black) versus $p_{collage}(t)$ (blue) for $N = 10$ and $\epsilon = 0.01$.

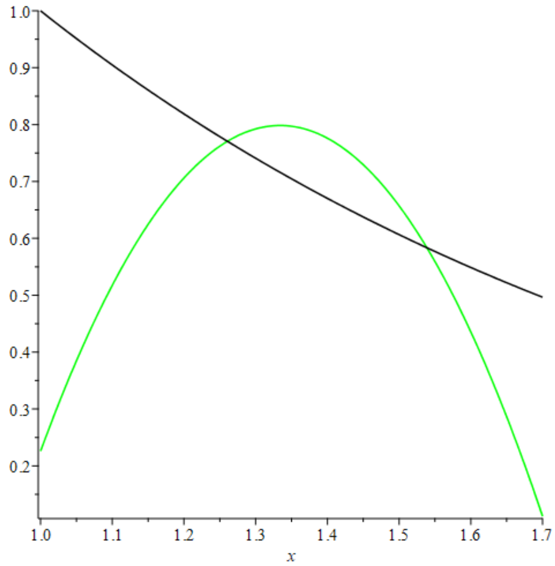

Figure 3.6: Plot of $p_{true}(t)$ (black) versus $p_{collage}(t)$ (green) for $N = 10$ and $\epsilon = 0.05$.

ϵ	N	$p_{collage}(x)$
0	5	$1.6933 - 0.7241x$
	10	$1.6926 - 0.7235x$
	25	$1.6918 - 0.7229x$
0.01	5	$1.3568 - 0.4836x$
	10	$1.4881 - 0.5840x$
	25	$1.4727 - 0.5706x$
0.05	5	$0.0731 + 0.4426x$
	10	$0.6812 - 0.0320x$
	25	$0.6058 + 0.0336x$
0	5	$2.2376 - 1.5544x + 0.3110x^2$
	10	$2.2625 - 1.5928x + 0.3256x^2$
0.01	5	$0.4550 + 0.8895x - 0.5133x^2$
	10	$0.3196 + 1.1974x - 0.6669x^2$
0.05	5	$-8.6467 + 13.6398x - 4.9026x^2$
	10	$-8.3297 + 13.6848x - 5.1291x^2$

Table 3.4: Function $p(x)$ recovered using the Collage Method applied to (3.5). For comparison, $p_{taylor}(x) = 2.7183 - 2.7183x + 1.3591x^2$.

Chapter 4

Optimal Control of ODE Systems

We briefly introduced the concept of a control system in Chapter 2. Now, we look at one strategy for choosing control inputs, called *optimal control.* Optimal control problems are a type of optimization problem, where one of the constraints on the quantity being optimized is given by a control system. In other words, the objective functional in an ODE optimal control problem is subject to a controlled ODE system.

In this Chapter, we explore the theory of optimal control applied to ODE systems. This theory was largely developed in the 1950s by Lev Pontryagin and Richard Bellman. We begin by introducing the optimal control problems that we'll be considering. We then derive a set of necessary conditions that optimal controls must satisfy (called Pontryagin's Maximum Principle). The existence of optimal controls is then examined, and an existence theorem is stated. We conclude this Chapter with a discussion on the Forward Backward Sweep method, a numerical approach to finding optimal controls.

Some definitions are important for what follows.

Definition 4.1. *Let X be a Banach space and $J : X \to \mathbb{R}$ be a functional. The Gâteaux derivative $\delta J(x;\eta)$ of J at $x \in X$ in the direction $\eta \in X$ is defined as*

$$\delta J(x;\eta) = \lim_{\alpha \to 0} \frac{J(x+\alpha\eta) - J(x)}{\alpha} = \frac{d}{d\alpha} J(x+\alpha\eta)\bigg|_{\alpha=0}$$

Definition 4.2. *Let X and Y be normed spaces and $T : X \to Y$ be a bounded linear operator. Then the adjoint operator $T^* : Y^{-1} \to X^{-1}$ is defined by*

$$f(x) = (T^*g)(x) = g(Tx)$$

where f, g are bounded linear functionals on X and Y, respectively, and X^{-1}, Y^{-1} are the dual spaces of X, Y.

Definition 4.3. *Let U be a closed, convex subset of a real Hilbert space V with inner product $\langle \cdot, \cdot \rangle$. The set*

$$N_U(u) = \{w \in V : \langle v - u, w \rangle \leq 0 \quad \forall v \in U\}$$

is called the normal cone to U at u.

Example 4.1. *Let $V = L^2([t_0, t_f])$.*

1. *If $U = V$, then $N_U(u) = \{0\}$.*

2. *If $U = \{w \in L^2([t_0, t_f]) : L_1 \leq w(t) \leq L_2 \text{ a.e. } t \in (t_0, t_f), L_1, L_2 \in \mathbb{R}\}$, then*

$$N_U(u) = \{w \in L^2([t_0, t_f]) : w(t) \geq 0 \text{ if } u(t) = L_2, w(t) \leq 0 \text{ if } u(t) = L_1,$$

$$w(t) = 0 \text{ if } L_1 < u(t) < L_2 \text{ a.e. } t \in (t_0, t_f)\}.$$

3. *In general, if $u \in int(U)$, then $N_U(u) = \{0\}$.*

Remark: *In this example, we use a.e. (almost everywhere) to mean that the statement holds on the entire interval except possibly at a finite number of points.*

4.1 An Optimal Control Problem

ODE optimal control problems take the form:

$$\sup_{\mathbf{u} \in \mathcal{U}} \int_{t_0}^{t_f} f(t, \mathbf{x}(t), \mathbf{u}(t))dt + \psi(t_f, \mathbf{x}_f), \tag{4.1}$$

subject to the controlled ODE problem,

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \tag{4.2}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{4.3}$$

where $\mathcal{U} \subset \mathbf{L}^2([t_0, t_f])$, called the set of *admissible controls*, is a closed, convex set of all functions $\mathbf{u} : [t_0, t_f] \to U \subset \mathbb{R}^m$ that are considered as control inputs. The set U is the *control space*, which specifies the number of control inputs, through its dimension, and constraints on the controls. The time t_0 is the *initial time* and t_f is the *terminal time*, which we define below. Equation (4.1) maps $\mathcal{U} \to \mathbb{R}$ and is called the *objective functional* (or alternatively, the *cost functional* or *performance index*), which is the quantity being optimized. The function $f : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}$ is called the *running cost*, which models costs that are continuously accumulated in the problem, and $\psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is the *terminal cost*, which represents a cost in the problem that is incurred at the conclusion of the process being investigated.

Remarks:

1. Optimal control problems with the cost functional given by Equation (4.1) are said to be in the *Bolza* form. Bolza problems with no terminal cost are said to be in the *Lagrange* form, and Bolza problems with no running cost are in the *Mayer* form.

2. We can transform problems in the Lagrange form into the Mayer form and vice-versa by applying a transformation.

Equation (4.2) is called the *state equation*, and $\mathbf{x}(t)$ is called the *state*. A pair $(\mathbf{x}^{\text{opt}}, \mathbf{u}^{\text{opt}})$ that solves Equations (4.1)-(4.3) is called an *optimal pair*, where $\mathbf{u}^{\text{opt}}$ is the *optimal control* and $\mathbf{x}^{\text{opt}}$ is the corresponding *optimal state*.

The final component of the optimal control problem is called the *target set*. A target set $S \subset [t_0, \infty) \times \mathbb{R}^n$ is an additional constraint on the optimal control problem that specifies what values are permitted for the terminal time and corresponding *terminal state*, $\mathbf{x}(t_f) = \mathbf{x}_f$. The most general target set is $S = [t_0, \infty) \times \mathbb{R}^n$. This choice of target set results in a so-called *free time, free endpoint* optimal control problem, since the terminal time and state are free to take any value in the optimization of the cost functional. Conversely, the most restrictive target set is $S = \{t_f\} \times \{\mathbf{x}_f\}$, resulting in a *fixed time, fixed endpoint* problem, since both the terminal time and state are specified by S. We define t_f to be the smallest time such that the pair $(t_f, \mathbf{x}_f) \in S$. The target set must be a closed set to ensure that the limit point of any trajectory $(t, \mathbf{x}(t))$ is also in S.

4.2 Pontryagin's Maximum Principle

In this section, we will establish a set of necessary conditions for the optimality of controls, known as Pontryagin's Maximum Principle. These conditions were first established by Lev Pontryagin in [9]. In particular, we focus on optimal control problems in which the state and control are real-valued for $t \in [t_0, t_f]$, $t_f < \infty$, f and g are continuously differentiable, and the terminal cost is absent:

$$\sup_{u \in U} \int_{t_0}^{t_f} f(t, x(t), u(t))dt, \tag{4.4}$$

subject to the state equation

$$\dot{x}(t) = g(t, x(t), u(t)) \tag{4.5}$$

$$x(t_0) = x_0. \tag{4.6}$$

We consider a target set $S = \{t_f\} \times \mathbb{R}$. Necessary conditions are those which are satisfied by optimal pairs $(x^{\mathrm{opt}}, u^{\mathrm{opt}})$. However, these conditions are not sufficient for optimality; they do not guarantee that a pair (x, u) will be optimal. When looking for optimal controls, we must also ensure that the conditions for their existence are met. We investigate existence in the next Section. For now, we begin by stating Pontryagin's Maximum Principle.

Theorem 4.1. *(Pontryagin's Maximum Principle) Let (x^{opt}, u^{opt}) be an optimal pair for the problem described above. Then there exists a piecewise differentiable function $p(t)$ such that:*

1. *For all controls $u \in \mathcal{U}$ at each time $t \in [t_0, t_f]$,*

$$\begin{aligned} &f(t, x^{opt}(t), u(t)) + p(t)g(t, x^{opt}(t), u(t)) \\ &\qquad \le f(t, x^{opt}(t), u^{opt}(t)) + p(t)g(t, x^{opt}(t), u^{opt}(t)) \end{aligned}$$

2. *x^{opt} and p satisfy the equations*

$$\dot{x}^{opt}(t) = g(t, x^{opt}(t), u^{opt}(t)) \tag{4.7}$$

$$\dot{p}(t) = -g_x(t, x^{opt}(t), u^{opt}(t))p(t) - f_x(t, x^{opt}(t), u^{opt}(t)) \tag{4.8}$$

with initial condition $x^{opt}(t_0) = x_0$ and final time condition $p(t_f) = 0$.

3. *The optimality condition*

$$g_u(\cdot, x^{opt}, u^{opt})p(\cdot) + f_u(\cdot, x^{opt}, u^{opt}) \in N_{\mathcal{U}}(u^{opt}) \tag{4.9}$$

is satisfied for all $t \in [t_0, t_f]$

A proof of Pontryagin's Maximum Principle can be found in [44]. Here, we derive the optimality condition (4.9) by considering a small perturbation of an optimal control and the resulting perturbation of the optimal state trajectory. It is through the properties of this perturbation that the optimality condition is realized.

Let $u^{\text{opt}} \in \mathcal{U}$ be a control that maximizes (4.4). We will consider the effect of perturbing this optimal control.

$$\tilde{u}(t, \alpha) = u^{\text{opt}}(t) + \alpha v(t) \tag{4.10}$$

where $v : [t_0, t_f] \to \mathbb{R}$ and $\alpha > 0$ are chosen such that u is an admissible control. In other words, $v \in V$ where

$$V = \{v \in L^2([t_0, t_f]), \alpha > 0 : u^{\text{opt}} + \alpha v \in \mathcal{U}\}$$

We know that this perturbed control will result in a perturbed state trajectory,

$$\tilde{x}(t, \alpha) = x^{\text{opt}}(t) + \alpha z(t) + o(\alpha), \tag{4.11}$$

where the function $z : [t_0, t_f] \to \mathbb{R}$ is the corresponding perturbation to the state. We use the notation $o(\alpha)$ to mean higher-order terms depending on α. Clearly, $z(t_0) = 0$ (up to terms of order $o(\alpha)$) so as to satisfy the initial condition (4.3). Observe that

$$\tilde{x}_\alpha(t, 0) = z(t). \tag{4.12}$$

We now use Equation (4.12) to develop a differential equation, that, together with the initial condition $z(t_0) = 0$, will be used to establish a condition required for $(x^{\text{opt}}, u^{\text{opt}})$

to be optimal. Differentiating Equation (4.12) with respect to time, and assuming sufficient smoothness of the first partial derivatives of $\tilde{x}$,

$$\begin{aligned}\dot{z}(t) &= \frac{\partial}{\partial t}\tilde{x}_\alpha(t,0)\\ &= \frac{\partial}{\partial \alpha}\dot{\tilde{x}}(t,\alpha)\bigg|_{\alpha=0}\\ &= \frac{\partial}{\partial \alpha}g(t,\tilde{x}(t,\alpha),\tilde{u}(t,\alpha))\bigg|_{\alpha=0}.\end{aligned}$$

Recalling Equations (4.10) and (4.11),

$$= g_x(t,x^{\text{opt}}(t),u^{\text{opt}}(t))\tilde{x}_\alpha(t,0) + g_u(t,x^{\text{opt}}(t),u^{\text{opt}}(t))v(t)$$

$$\Longrightarrow \dot{z}(t) = g_x(t,x^{\text{opt}}(t),u^{\text{opt}}(t))z(t) + g_u(t,x^{\text{opt}}(t),u^{\text{opt}}(t))v(t) \tag{4.13}$$

So, for a fixed perturbation of the optimal control, $v \in V$, Equation (4.13) describes how the resulting perturbation to the state propagates in time. This system is the linearization of the state equation (4.2) about the optimal pair $(x^{\text{opt}}, u^{\text{opt}})$.

We will return to Equation (4.13) shortly, but we first need to turn our attention to the so-called *adjoint equation*,

$$\dot{p}(t) = -g_x(t,x^{\text{opt}}(t),u^{\text{opt}}(t))p(t) - f_x(t,x^{\text{opt}}(t),u^{\text{opt}}(t)), \quad t \in [t_0,t_f] \tag{4.14}$$

$$p(t_f) = 0. \tag{4.15}$$

A solution $p(t)$ to (4.14)–(4.15) is called the *adjoint state* or the *costate.*

Remark: The adjoint equation gets its name because it is adjoint to the equation

$$\dot{y} = \begin{pmatrix} 0 & f_x(t,x^{\text{opt}},u^{\text{opt}}) \\ 0 & g_x(t,x^{\text{opt}},u^{\text{opt}}) \end{pmatrix} y$$

which arises in the formal proof of Pontryagin's Maximum Principle. This equation plays a similar role as our DE for $z(t)$ in that it is a linearization of the state equation about the optimal control u^{opt}, however it is found by considering perturbations to u^{opt} differently than we do in our derivation.

Multiplying our DE for $z(t)$ in Equation (4.13) by $p(t)$ and integrating over $[t_0, t_f]$,

$$\int_{t_0}^{t_f} \dot{z}(t)p(t)dt = \int_{t_0}^{t_f} [g_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))z(t) + g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))v(t)]p(t)dt \tag{4.16}$$

Integrating by parts on the left-hand side, we find that

$$\int_{t_0}^{t_f} \dot{z}(t)p(t)dt = z(t)p(t)\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} z(t)\dot{p}(t)dt = -\int_{t_0}^{t_f} z(t)\dot{p}(t)dt.$$

Substituting in Equation (4.16) gives

$$-\int_{t_0}^{t_f} z(t)\dot{p}(t)dt = \int_{t_0}^{t_f} \left[g_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))z(t) + g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))v(t)\right] p(t)dt$$

for any $v \in V$. We can replace $\dot{p}(t)$ on the left-hand side using Equation (4.14),

$$\begin{aligned}&\int_{t_0}^{t_f} z(t)[g_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))p(t) + f_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))]dt \\ &= \int_{t_0}^{t_f} [g_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))z(t) + g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))v(t)]p(t)dt.\end{aligned}$$

Simplifying gives

$$\int_{t_0}^{t_f} v(t)g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))p(t)dt = \int_{t_0}^{t_f} z(t)f_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))dt \tag{4.17}$$

for any $v \in V$.

By bounding the left-hand side of Equation (4.17), we will soon arrive at the optimality condition we have been working towards. Recall Equations (4.10) and (4.11). To obtain this bound, notice that, for a fixed $v \in V$,

$$\int_{t_0}^{t_f} f(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha))dt \leq \int_{t_0}^{t_f} f(t, x^{\text{opt}}(t), u^{\text{opt}}(t))dt$$

Dividing both sides of this inequality by α, and bringing all terms to the left-hand side,

$$\int_{t_0}^{t_f} \frac{1}{\alpha}[f(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) - f(t, x^{\text{opt}}(t), u^{\text{opt}}(t))]dt \leq 0.$$

Taking the limit as $\alpha \to 0^+$, and bringing the limit inside the integral,

$$\begin{aligned} &\lim_{\alpha \to 0^+} \int_{t_0}^{t_f} \frac{1}{\alpha}[f(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) - f(t, x^{\text{opt}}(t), u^{\text{opt}}(t))]dt \\ &= \int_{t_0}^{t_f} \lim_{\alpha \to 0^+} \left(\frac{1}{\alpha}[f(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) - f(t, x^{\text{opt}}(t), u^{\text{opt}}(t))] \right) dt \end{aligned} \tag{4.18}$$

The integrand is now the definition of the Gâteaux derivative of f at v. We evaluate this derivative,

$$\begin{aligned} &\frac{d}{d\alpha} f(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) \bigg|_{\alpha=0} \\ &\quad = \left[f_x(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) \frac{\partial \tilde{x}(t, \alpha)}{\partial \alpha} + f_u(t, \tilde{x}(t, \alpha), \tilde{u}(t, \alpha)) \frac{\partial \tilde{u}(t, \alpha)}{\partial \alpha} \right]_{\alpha=0} \\ &\quad = f_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))z(t) + f_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))v(t). \end{aligned}$$

Replacing this in Equation (4.18), we arrive at

$$\int_{t_0}^{t_f} f_x(t, x^{\text{opt}}(t), u^{\text{opt}}(t))z(t) + f_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))v(t)dt \leq 0. \tag{4.19}$$

Turning back now to Equation (4.17), we can use Equation (4.19) to bound the left-hand side of Equation (4.17),

$$\int_{t_0}^{t_f} v(t) g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t)) p(t) dt \leq -\int_{t_0}^{t_f} v(t) f_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t)) dt$$
$$\Longrightarrow \int_{t_0}^{t_f} v(t) \left[g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t)) p(t) + f_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t)) \right] dt \leq 0$$

for any fixed $v \in V$. This implies the condition

$$g_u(\cdot, x^{\text{opt}}, u^{\text{opt}}) p + f_u(\cdot, x^{\text{opt}}, u^{\text{opt}}) \in N_{\mathcal{U}}(u^{\text{opt}}) \tag{4.9}$$

where $N_{\mathcal{U}}(u^{\text{opt}})$ is the normal cone to $\mathcal{U}$ at u^{opt}.

Remark: When $\mathcal{U}$ is open and the left-hand side is concave in u, we may write the optimality condition (4.9) as

$$g_u(\cdot, x^{\text{opt}}, u^{\text{opt}}) p + f_u(\cdot, x^{\text{opt}}, u^{\text{opt}}) = 0.$$

The state equation (4.5), adjoint equation (4.14), and optimality condition (4.9) taken together constitute Pontryagin's Maximum Principle.

Example 4.2. *Consider the optimal control problem*

$$\max_u \int_0^1 2x(t) + \frac{3}{2} u(t)^2 dt$$
$$\textit{subject to} \begin{cases} \dot{x}(t) = 7x(t) + 5u(t), \\ x(0) = \dfrac{25(2 - e^7)}{147}. \end{cases}$$

We can solve this problem using Pontryagin's Maximum Principle. To begin, we write the adjoint equation provided by (4.14),

$$\begin{cases} \dot{p}(t) = -2 - 7p(t), \\ p(1) = 0. \end{cases}$$

Solving this DE yields

$$p(t) = \frac{2(e^{7(1-t)} - 1)}{7}.$$

The optimality condition (4.9) *for this problem is given by*

$$3u^{opt}(t) + 5p(t) = 0,$$

so we have that $u^{opt}(t) = \dfrac{10(1 - e^{7(1-t)})}{21}$. *We can now substitute this optimal control into the state equation,*

$$\dot{x}^{opt}(t) = 7x^{opt}(t) + \frac{50(1 - e^{7(1-t)})}{21}, \quad x(0) = \frac{25(2 - e^7)}{147}.$$

Solving for the optimal state yields $x^{opt}(t) = \dfrac{25(e^{7(1-t)} - 2)}{147}$. *Thus, the optimal pair that solves this optimal control problem is*

$$\begin{aligned} x^{opt}(t) &= \frac{25(e^{7(1-t)} - 2)}{147} \\ u^{opt}(t) &= \frac{10(1 - e^{7(1-t)})}{21}. \end{aligned}$$

This optimal pair yields a maximum value of the objective functional of 29164.189.

Remark: The maximum principle is often stated in terms of the *Hamiltonian*,

$$H(t, x, u, p) = f(t, x(t), u(t)) + p(t)g(t, x(t), u(t)).$$

Using the Hamiltonian we can write the state equation as

$$\dot{x}(t) = H_p(t, x, u, p) = g(t, x(t), u(t)),$$

the adjoint equation as

$$\dot{p}(t) = -H_x(t, x, u, p) = -f_x(t, x(t), u(t)) - p(t)g_x(t, x(t), u(t)),$$

and finally, we can write the optimality condition, in general, as

$$H_u(t, x, u, p) \in N_{\mathcal{U}}(u^{\text{opt}}).$$

To reiterate, Pontryagin's Maximum Principle provides, in general, necessary conditions for an optimal control. There are circumstances, however, where the Maximum Principle is also sufficient for optimality. We briefly discuss two such scenarios here, see [11] for more details.

Firstly, certain convexity assumptions on the set of admissible controls $\mathcal{U}$ and on the objective functional will guarantee that any pair (x, u) satisfying the Maximum Principle is optimal for the problem.

Secondly, if we can prove that an optimal control and corresponding optimal state exist for the given problem, and if we assume that $u_1, \ldots, u_k$ are admissible controls that satisfy the Maximum Principle, then the $u_i, i \in \{1, \ldots, k\}$ that yields the greatest value of the objective functional is optimal. We investigate the existence of optimal controls in the next section.

4.3 Existence of Optimal Controls

The maximum principle does not address whether an admissible optimal pair $(x^{\mathrm{opt}}, u^{\mathrm{opt}})$ actually exists for a given problem. We now consider an existence theorem for the optimal control problem discussed in Section 4.2.

There are some properties of control systems that we need to define in order to understand the existence of optimal controls. The first property, called *controllability*, must hold for the control system in question.

Definition 4.4. *The system*

$$\dot{x}(t) = g(t, x(t), u(t))$$

is called controllable if for any initial condition $x(t_0) = x_0$ and final condition $x(t_f) = x_f$, there exists a control $u(t)$ and a time $t_f < \infty$ such that $x(t_f) = x_f$.

Controllability is thus an intrinsic property of a control system, without which the existence of optimal controls is meaningless. Next, we define the *reachable sets* of a control system.

Definition 4.5. *The reachable sets $R(t)$ of the control system* (4.2)–(4.3) *is the set of all points reachable from $x(t_0) = x_0$ at time t when the system is driven by admissible controls,*

$$R(t) = \{x(t) : \dot{x}(t) = g(t, x(t), u(t)), u(t) \in \mathcal{U}\}.$$

We will show that existence of optimal controls holds when the reachable sets of the control system being considered are compact. The conditions for compactness are given by Filippov's Lemma [40], which we now state.

Lemma 4.2. *(Filippov) Consider the control system*

$$\dot{x}(t) = g(t, x(t), u(t)) \tag{4.2}$$

$$x(t_0) = x_0, \tag{4.3}$$

whose solutions exist on the time interval $[t_0, t_f]$*,* $t_f < \infty$*, for all* $u \in \mathcal{U}$*. Assume that*

1. *there exists a uniform bound*

   $$|x(t)| \leq b, \quad t \in [t_0, t_f], \quad b \in \mathbb{R}$$

   for all state trajectories $x(t)$ *in response to controls* $u(t)$*, and*

2. *for every pair* (t, x)*, the set* $Q(t, x) = \{g(t, x, u) : u \in \mathcal{U}\}$ *is compact and convex.*

Then the reachable sets $R(t)$ *are compact for each* $t \in [t_0, t_f]$*.*

The proof of Filippov's Lemma is beyond the scope of this thesis. Making use of these required conditions for compact reachable sets, we can now state a result regarding existence of optimal controls.

Theorem 4.3. *(Existence of optimal controls [40]) Consider the optimal control problem*

$$\sup_{u \in \mathcal{U}} \int_{t_0}^{t_f} f(t, x(t), u(t)) dt$$

subject to

$$\dot{x}(t) = g(t, x(t), u(t)) \tag{4.2}$$

$$x(t_0) = x_0, \tag{4.3}$$

with target set $S = \{t_f\} \times \mathbb{R}$. *If the running cost* f *is a continuous function on* $\mathbb{R} \times \mathbb{R} \times U$, *and we assume that*

1. $R(t)$ *is compact; and*

2. *the system is controllable.*

Then there exists an optimal control $u^{opt}(t)$ *on* $[t_0, t_f]$ *such that the objective function is maximized.*

Proof. We define a new state variable, $x^0 \in \mathbb{R}$, to be the solution of

$$\dot{x}^0(t) = f(t, x(t), u(t))$$

$$x^0(t_0) = 0,$$

and arrive at the augmented system

$$\dot{x}^0(t) = f(t, x(t), u(t))$$

$$\dot{x}(t) = g(t, x(t), u(t)),$$

with the initial condition $[0, x_0]^T$. Using this new variable, the cost functional can be rewritten as

$$\int_{t_0}^{t_f} \dot{x}^0(t)dt = x^0(t_f) \tag{4.20}$$

Since the reachable sets $R(t)$ of (4.2)–(4.3) are compact, $R(t_f)$ is the compact set of all possible state trajectory endpoints $x(t_f) = x_f$. Equation (4.20) is the transformation of the problem from Lagrange type to Mayer type, so the right-hand side of (4.20) may be thought of as a function of the new state, $\psi([x^0, x]^T) = x^0$. Since ψ is continuous, by the Extreme Value Theorem it attains its maximum at some point x^{opt} on $R(t_f)$. By definition of reachable sets, there must be at least one admissible control $u^{\text{opt}}(t)$ with corresponding state trajectory $x^{\text{opt}}(t)$ such that $x^{\text{opt}}(t_f) = x_f^{\text{opt}}$, and every such control is optimal. □

4.4 Linear Quadratic Optimal Control Problems

Conditions for the existence of optimal controls have now been established. Uniqueness of optimal controls is, unfortunately, less straightforward to consider. Rather than attempt a general discussion of uniqueness, we instead present a uniqueness result for a particular class of problems, called Linear Quadratic (LQ) problems, which will be the focus of our later examples.

LQ problems are named for the property that they are linear in the state equation and quadratic in the objective functional. In other words, for $x(t) \in \mathbb{R}$, $u(t) \in \mathbb{R}$, and the target set $\{t_f\} \times \mathbb{R}$, they have the form,

$$\sup_{u \in \mathcal{U}} \int_{t_0}^{t_f} qx(t)^2 + ru(t)^2 dt, \tag{4.21}$$

subject to

$$\dot{x}(t) = ax(t) + bu(t) \tag{4.22}$$

$$x(t_0) = x_0, \tag{4.23}$$

where the coefficients $a, b, q, r \in \mathbb{R}$.

These problems can be shown to have unique optimal controls, as the following theorem states.

Theorem 4.4. *Consider the LQ problem given by Equations* (4.21)–(4.23)*. If $q \geq 0$ and $r > 0$, the for any initial condition x_0, there is an optimal control and corresponding optimal state that uniquely optimizes Equation* (4.21)*.*

A proof can be found in [48]. See also Example 4.3 in Section 4.5.

4.5 The Forward Backward Sweep Method

Pontryagin's Maximum Principle provides a convenient and straightforward avenue to search for solutions to ODE optimal control problems. However, the differential-algebraic system generated by the maximum principle does not, in general, have a closed-form solution. Therefore, in practice we often must turn to numerical methods. In this Section, we investigate one such technique, called the Forward Backward Sweep method.

Numerical methods for solving optimal control methods can be classified as either *direct* or *indirect*. With a direct method, one seeks to optimize the objective functional directly. In contrast, an indirect method uses the maximum principle equations to construct a sequence of controls and states that converges to an optimal pair. The Forward Backward Sweep method is indirect in this sense.

Recall that the optimality condition from Pontryagin's Maximum Principle is given by

$$g_u(\cdot, x^{\text{opt}}, u^{\text{opt}})p + f_u(\cdot, x^{\text{opt}}, u^{\text{opt}}) \in N_{\mathcal{U}}(u^{\text{opt}}). \tag{4.9}$$

The Forward Backward Sweep method is applicable in the case where $\mathcal{U}$ is open and the left-hand side is concave in the control, in which case the optimality condition becomes

$$g_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t))p(t) + f_u(t, x^{\text{opt}}(t), u^{\text{opt}}(t)) = 0. \tag{4.24}$$

Beginning with an initial estimate or guess for the optimal control, $u^{(0)} \in \mathcal{U}$, the Forward Backward Sweep method entails solving the state equation using $u^{(0)}$ to generate an initial iterate for the state, $x^{(0)}$. Next, the adjoint equation is solved backwards in time using $u^{(0)}$ and $x^{(0)}$, yielding $p^{(0)}$. The next iterate for the control is then found by solving the optimality equation (4.24)

$$g_u(t, x^{(0)}(t), u^{(1)}(t))p^{(0)}(t) + f_u(t, x^{(0)}(t), u^{(1)}(t)) = 0,$$

for $u^{(1)}(t)$. This process is repeated until reaching some stopping criterion. For instance,

$$\frac{\|u^{(k+1)} - u^{(k)}\|}{\|u^{(k)}\|} + \frac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k)}\|} + \frac{\|p^{(k+1)} - p^{(k)}\|}{\|p^{(k)}\|} \le \epsilon, \tag{4.25}$$

where $u^{(k)}$, $u^{(k+1)}$, $x^{(k)}$, $x^{(k+1)}$, $p^{(k)}$ and $p^{(k+1)}$ are the kth and $(k+1)$st iterates of u, x, and p, respectively, and $\|\cdot\|$ is an appropriate norm.

Remark: The choice of stopping criterion is not limited to Equation (4.25), and could be chosen to involve only the control or to be an absolute error quantity, for example.

The Forward Backward Sweep method is independent of the choice of solution method for the state and adjoint equations. In practice the solution method may be chosen to best accommodate the problem.

To guarantee that the method will converge for a given problem, certain conditions must be met as outlined in the following theorem.

Theorem 4.5. *Consider the optimal control problem* (4.4)–(4.6)

$$\sup_{u\in U}\int_{t_0}^{t_f} f(t,x(t),u(t))dt \tag{4.4}$$

subject to

$$\dot{x}(t) = g(t,x(t),u(t)) \tag{4.5}$$

$$x(t_0) = x_0. \tag{4.6}$$

Assume that the optimality equation (4.24) *may be solved for $u(t)$ and expressed as $u(t) = h(t,x(t),p(t))$. Suppose that the functions g, g_x, f_x and h each satisfy a uniform Lipschitz condition with respect to their second and third arguments with Lipschitz constants L_g, L_{g_x}, L_{f_x} and L_h. For example, $\forall t\in[t_0,t_f]$,*

$$|g(t,x_1,u_1) - g(t,x_2,u_2)| \leq L_g\left(|x_1-x_2| + |u_1-u_2|\right). \tag{4.26}$$

Suppose further that $\Lambda = \|p\|_\infty < \infty$ *and* $H = \|g_x\|_\infty < \infty$. *If*

$$L_h \left(\left(e^{(L_g(t_f - t_0))} - 1\right) + (L_{g_x} + \Lambda L_{f_x}) \frac{1}{H} \left(e^{H(t_f - t_0)} - 1\right) \left(e^{L_g(t_f - t_0)} + 1\right) \right) < 1,$$

then in the limit as $k \to \infty$,

$$\max_{t \in I} \left| x(t) - x^{(k)}(t) \right| + \max_{t \in I} \left| p(t) - p^{(k)}(t) \right| + \max_{t \in I} \left| u(t) - u^{(k)}(t) \right| \to 0.$$

That is, the solution obtained numerically converges, in the limit as $k \to \infty$, *to the exact solution.*

The proof of this Theorem is found in [45] for the interested reader, along with a thorough exploration of convergence of the Forward Backward Sweep method.

We shall consider the Forward Backward Sweep method utilizing the RK4 method for solving the state and adjoint equations; see Chapter 2, Section 2.1.2 for the algorithm. Solving the adjoint equation backwards in time requires that the initial point in the process is replaced with the final point, $p(t_f)$, and that the time step at each iteration is decreased, $t_{n+1} = t_n - h$, rather than increased.

We demonstrate the Forward Backward Sweep method with the following example from [45].

Example 4.3. *Consider the linear quadratic optimal control problem*

$$\max_{u \in \mathcal{U}} \frac{1}{2} \int_0^1 x(t)^2 + u(t)^2 dt \tag{4.27}$$

$$\textit{subject to} \begin{cases} \dot{x}(t) = -x(t) + u(t), \\ x(0) = 1. \end{cases} \tag{4.28}$$

Pontryagin's Maximum Principle can provide a closed-form solution for this problem. The adjoint equation, given by Equation (4.14) *is*

$$\dot{p}(t) = p(t) - x(t),$$

with the final time condition $p(1) = 0$. *The optimality condition given by* (4.9) *is*

$$u(t) = -p(t).$$

Together with the state equation, Pontryagin's Maximum Principle for this problem consists of the equations

$$\begin{cases} \dot{x}(t) = -x(t) + u(t), \qquad x(0) = 1 \\ \dot{p}(t) = p(t) - x(t), \qquad p(1) = 0 \\ u(t) = -p(t). \end{cases}$$

We have solved this problem numerically via the Forward Backward Sweep method using the RK4 method with a tolerance of $\epsilon = 10^{-3}$ *and varying numbers* N *of equidistantly spaced mesh points. The initial guess for the state was* $u^{(0)} = 0$ *and the stopping condition was that of Equation* (4.25). *For comparison, the exact solution of this problem, obtained algebraically using Maple 21, is*

$$\begin{cases} x(t) = \dfrac{\sqrt{2}\cosh(\sqrt{2}(t-1)) - \sinh(\sqrt{2}(t-1))}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})} \\ p(t) = -\dfrac{\sinh(\sqrt{2}(t-1))}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})} \\ u(t) = \dfrac{\sinh(\sqrt{2}(t-1))}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})}. \end{cases}$$

This solution admits a maximum value of the objective functional of 0.1929092981. The results, including the L^2-error between the true and numerical state and adjoint, in addition to the absolute error in the maximum objective values, are presented in Table 4.1. We can see a marked improvement in the accuracy of the numeric solution when the number of mesh points is increased.

N	$\|x_{true} - x_{FBS}\|_2$	$\|p_{true} - p_{FBS}\|_2$	$\lvert J_{true} - J_{FBS}\rvert$
30	0.1431	0.0733	9.813×10^{-3}
100	0.0782	0.0398	2.954×10^{-3}
300	0.0449	0.0228	2.349×10^{-3}
500	0.0347	0.0176	1.933×10^{-3}

Table 4.1: Results for the Forward Backward Sweep method applied to the problem (4.27)–(4.28).

Remark: The Forward Backward Sweep method presented here is applicable to fixed-time, free-endpoint problems as described in Section 4.1. A modified version of this problem may be used to numerically solve fixed-endpoint problems by incorporating a shooting method to solve the state equation.

Chapter 5

A Collage-Based Approach to Inverse Optimal Control Problems with Unique Solutions

In Chapter 4, we considered the *forward* optimal control problem, in which given an objective functional and an ODE system, we endeavour to find a control function that optimizes the objective. While many problems can be formulated in this way (such as [2, 18, 20, 23, 26], for example), a criticism of optimal control is that it is not always clear what cost functional or performance index to employ as the objective [29].

A different approach to optimal control is to investigate control systems for which we can observe, or otherwise obtain knowledge of, both the state variables and control inputs, and reasoning that there may exist some objective functional that is optimized by the observed system. This line of inquiry leads to the *inverse optimal control problem (IOCP)*. In this Chapter, we discuss this problem and formulate a Collage-Based Approach for solving a class of these problems using Pontryagin's Maximum Principle.

5.1 Applying the Collage Method to an Inverse Optimal Control Problem

In this Chapter, we will be working on the assumption that the Optimal Control Problems being considered have unique solutions. Recall that LQ problems satisfy this requirement under certain conditions; we will consider this type of problem specifically in Section 5.2.

First, we restate the Optimal Control Problem from Chapter 4: Find a function $u(t)$ that maximizes the objective functional

$$\int_{t_0}^{t_f} f(t, x(t), u(t))dt, \tag{5.1}$$

subject to the controlled ODE constraint

$$\dot{x}(t) = g(t, x(t), u(t)), \tag{5.2}$$

$$x(t_0) = x_0, \tag{5.3}$$

with the target set $S = \{t_f\} \times \mathbb{R}$, $t_f < \infty$.

In the IOCP, we seek unknowns present in the function f. From this point forward, we will denote the integrand of Equation (5.1) by f^λ to emphasize the dependence on unknown parameters $\lambda \in \Lambda$.

Recall that Pontryagin's Maximum Principle stipulates that for this Optimal Control Problem and an unconstrained control set U, an optimal pair $(x^{\text{opt}}, u^{\text{opt}})$ must satisfy the necessary conditions given by

$$\dot{x}(t) = g(t, x, u), \tag{5.4}$$

$$\dot{p}(t) = -f_x^\lambda(t, x, u) - p(t) g_x(t, x, u), \tag{5.5}$$

$$g_u(t, x, u) p(t) + f_u^\lambda(t, x, u) = 0, \tag{5.6}$$

with initial condition $x(t_0) = x_0$ and terminal condition $p(t_f) = 0$. Equations (5.4)–(5.5) represent a system of controlled ODEs that incorporate information pertaining to the unknown parameters λ present in the function f^λ. We propose a solution technique for the IOCP facilitated by applying a Collage-Based Approach to this system of ODEs.

Remark: The system of ODEs (5.4)–(5.5) provided by the Maximum Principle do not constitute an IVP as we considered in Chapters 2 and 3. We can, however, apply the same ideas with some modifications.

We will require the following result in the sequel.

Theorem 5.1. *(Implicit Function Theorem) Let $F(t,x,p,u)$ be C^2 and non-zero. If $F_u \neq 0$, then there exists a neighbourhood of the point (t,x,p,u) and a continuously differentiable function h such that $h(t,x,p) = u$ and $F(t,x,p,h(t,x,p)) = 0$ in this neighbourhood.*

In what follows, we shall work on the interval $I = [t_0, t_f]$, the region $\Psi = \{(x,p) \in \mathbb{R}^2 | |x - x_0| \leq \beta, |p| \leq \gamma\}$, and on the Banach space $\overline{C(I)}$ defined as

$$\overline{C(I)} = C_\beta(I) \times C_\gamma(I),$$

where $C_\beta(I) = \{x \in C(I) : \|x - x_0\|_\infty \leq \beta\}$ and $C_\gamma(I) = \{x \in C(I) : \|x\|_\infty \leq \gamma\}$. That is, the elements of the space $\overline{C(I)}$ are of the form $\mathbf{y} = [x,p]^T$. The sup-norm associated with this space is

$$\|\mathbf{y}_1 - \mathbf{y}_2\|_\infty \equiv \|x_1 - x_2\|_\infty + \|p_1 - p_2\|_\infty ,$$

where $\mathbf{y}_1 = [x_1, p_1], \mathbf{y}_2 = [x_2, p_2] \in \overline{C(I)}$. A routine argument establishes completeness of this space. We also require the norm on $\mathcal{G} = I \times \Psi$,

$$\|g(t,x,p)\|_{\mathcal{G}} \equiv \sup_{(t,x,p)\in\mathcal{G}} |g(t,x,p)|.$$

Using the Implicit Function Theorem, we can write the optimality condition in Equation (5.6) as $u = h(t, x(t), p(t))$. Doing this allows us to write $\bar{g}(t,x,p) \equiv g(t,x,h(t,x,p))$ and, similarly, $\bar{f}^\lambda(t,x,p) \equiv f(t,x,h(t,x,p))$, and thus express Equations (5.4)–(5.6) as a system of ODEs in only x, p:

$$\dot{x}(t) = \bar{g}(t, x, p), \qquad x(t_0) = x_0 \tag{5.7}$$

$$\dot{p}(t) = -\bar{f}_x^\lambda(t, x, p) - p(t)\bar{g}_x(t, x, p), \qquad p(t_f) = 0 \tag{5.8}$$

In order to apply a Collage-Based Approach to the controlled ODE system (5.7)–(5.8), we build the "Picard-like" operator $\mathbf{Ty} = [T_1\mathbf{y}, T_2\mathbf{y}]^T$, whose components are

$$\mathbf{Ty} = \begin{pmatrix} T_1\mathbf{y} \\ \\ T_2\mathbf{y} \end{pmatrix} = \begin{pmatrix} x_0 + \int_{t_0}^t \bar{g}(s, x(s), p(s))\, ds \\ \\ \int_t^{t_f} (\bar{f}_x(s, x(s), p(s)) + \bar{g}_x(s, x(s), p(s))p(s))\, ds \end{pmatrix} \tag{5.9}$$

where $\mathbf{y} = [x, p]^T \in \overline{C(I)}$. The collage distance for this operator is

$$\Delta^2 = \|x - T_1\mathbf{y}\|_2^2 + \|p - T_2\mathbf{y}\|_2^2\,. \tag{5.10}$$

To apply the Collage Theorem, we require that $\mathbf{T}$ is space-preserving and contractive on $\overline{C(I)}$.

Theorem 5.2. *Let $I = [t_0, t_f]$ and consider the system of ODEs* (5.7)–(5.8) *provided by Pontryagin's Maximum Principle. Suppose that*

$$\|\bar{g}(t, x, p)\|_{\mathcal{G}} \leq \frac{\beta}{t_f - t_0}, \text{ and}$$
$$\left\|p\bar{g}_x(t, x, p) + \bar{f}_x^\lambda(t, x, p)\right\|_{\mathcal{G}} \leq \frac{\gamma}{t_f - t_0},$$

then the operator given in Equation (5.9) *preserves the space $\overline{C(I)}$. That is, the components of **Ty** satisfy $\|T_1\mathbf{y} - x_0\|_\infty \leq \beta$ and $\|T_2\mathbf{y}\|_\infty \leq \gamma$, for $\beta, \gamma \in \mathbb{R}^+$, and $\boldsymbol{y} = [x, p]^T$.*

Proof. We first show that $\|T_1\mathbf{y} - x_0\|_\infty \leq \beta$. For $t \in I$,

$$\begin{aligned}
|T_1\mathbf{y} - x_0| &= \left| \int_{t_0}^{t} \bar{g}(s, x(s), p(s)) ds \right| \\
&\leq \int_{t_0}^{t} |\bar{g}(s, x(s), p(s))| \, ds.
\end{aligned}$$

Assuming that (x, p) remain in Ψ for $t \in I$, we bound the integrand to obtain

$$\begin{aligned}
&\leq \|\bar{g}(t, x, p)\|_{\mathcal{G}} \int_{t_0}^{t} ds \\
&= \|\bar{g}(t, x, p)\|_{\mathcal{G}} (t - t_0) \\
&\leq \frac{\beta}{t_f - t_0}(t - t_0).
\end{aligned}$$

Taking the supremum of both sides over $t \in I$ establishes the result.

In a similar fashion, we now show that $\|T_2\mathbf{y}\|_\infty \leq \gamma$. For $t \in I$,

$$\begin{aligned}
|T_2\mathbf{y}| &= \left| \int_{t}^{t_f} \left[p(s)\bar{g}_x(s, x(s), p(s)) + \bar{f}_x^\lambda(s, x(s), p(s)) \right] ds \right| \\
&\leq \int_{t}^{t_f} \left| p(s)\bar{g}_x(s, x(s), p(s)) + \bar{f}_x^\lambda(s, x(s), p(s)) \right| ds.
\end{aligned}$$

Assuming that (x, p) remain in Ψ for $t \in I$, we again bound the integrand to obtain

$$\begin{aligned}
&\leq \left\| p\bar{g}_x(t, x, p) + \bar{f}_x^\lambda(t, x, p) \right\|_{\mathcal{G}} \int_{t}^{t_f} ds \\
&= \left\| p\bar{g}_x(t, x, p) + \bar{f}_x^\lambda(t, x, p) \right\|_{\mathcal{G}} (t_f - t) \\
&\leq \frac{\gamma}{t_f - t_0}(t_f - t).
\end{aligned}$$

Taking the supremum of both sides over $t \in I$ once again establishes the result. Thus, $\mathbf{T} : \overline{C(I)} \to \overline{C(I)}$. □

Theorem 5.3. *Let $I = [t_0, t_f]$ and recall the system of ODEs given in* (5.7)*–*(5.8)*. The associated operator* $\mathbf{T}$ *given in Equation* (5.9) *is contractive with respect to the norm* $\|\cdot\|_\infty$ *if* $\bar{g}$*,* $\bar{g}_x$*, and* $\bar{f}_x^\lambda$ *are each Lipschitz continuous in their second and third arguments (as defined in Equation* (4.26)*), continuous in their first argument and the following conditions hold:*

1. $L_g(t_f - t_0) < 1$*; and*

2. $\left(L_{f_x} + \gamma L_{g_x}\right)(t_f - t_0) < 1$.

Proof. Let $\mathbf{y}_1 = [x_1, p_1]^T$, $\mathbf{y}_2 = [x_2, p_2]^T \in \overline{C(I)}$. Then for the first component of the operator, we have, for $t \in I$,

$$\begin{aligned} |T_1\mathbf{y}_1 - T_1\mathbf{y}_2| &= \left| \int_{t_0}^{t} \bar{g}(s, x_1(s), p_1(s)) - \bar{g}(s, x_2(s), p_2(s)) ds \right| \\ &\leq \int_{t_0}^{t} \left| \bar{g}(s, x_1(s), p_1(s)) - \bar{g}(s, x_2(s), p_2(s)) \right| ds. \end{aligned}$$

Since $\bar{g}$ is uniformly L_g-Lipschitz, we have that

$$\begin{aligned} |T_1\mathbf{y}_1 - T_1\mathbf{y}_2| &\leq \int_{t_0}^{t} L_g \left(|x_1(s) - x_2(s)| + |p_1(s) - p_2(s)| \right) ds \\ &\leq L_g \sup_{t \in I} \left(|x_1(t) - x_2(t)| + |p_1(t) - p_2(t)| \right) \int_{t_0}^{t} ds \\ &\leq L_g \left\| \mathbf{y}_1 - \mathbf{y}_2 \right\|_\infty (t - t_0) \end{aligned}$$

Taking the supremum over $t \in I$, we arrive at

$$\left\| T_1\mathbf{y}_1 - T_1\mathbf{y}_2 \right\|_\infty \leq L_g(t_f - t_0) \left\| \mathbf{y}_1 - \mathbf{y}_2 \right\|_\infty \tag{5.11}$$

Thus, under the assumption that $L_g(t_f - t_0) < 1$, T_1 is contractive with respect to $\|\cdot\|_\infty$.

Similarly, for the second component of $\mathbf{T}$ we have, for $t \in I$,

$$|T_2\mathbf{y}_1 - T_2\mathbf{y}_2| = \left| \int_t^{t_f} p_1(s)\bar{g}_x(x, x_1(s), p_1(s)) - \bar{f}_x^\lambda(s, x_1(s), p_1(s)) \right.$$
$$\left. -p_2(s)\bar{g}_x(x, x_2(s), p_2(s)) + \bar{f}_x^\lambda(s, x_2(s), p_2(s)) ds \right|$$

$$\leq \int_t^{t_f} \left| p_1(s)\bar{g}_x(x, x_1(s), p_1(s)) - \bar{f}_x^\lambda(s, x_1(s), p_1(s)) \right.$$
$$\left. -p_2(s)\bar{g}_x(x, x_2(s), p_2(s)) + \bar{f}_x^\lambda(s, x_2(s), p_2(s)) \right| ds.$$

Applying the triangle inequality to the integrand yields

$$|T_2\mathbf{y}_1 - T_2\mathbf{y}_2| \leq \int_t^{t_f} \left[\left| \bar{f}_x^\lambda(s, x_1(s), p_1(s)) - \bar{f}_x^\lambda(s, x_2(s), p_2(s)) \right| \right.$$
$$\left. + \left| p_1(s)\bar{g}_x(x, x_1(s), p_1(s)) - p_2(s)\bar{g}_x(x, x_2(s), p_2(s)) \right| \right] ds$$

Since $\bar{f}_x^\lambda$ is uniformly L_{f_x}-Lipschitz, we have that

$$|T_2\mathbf{y}_1 - T_2\mathbf{y}_2| \leq \int_t^{t_f} \left[L_{f_x} \left(|x_1(t) - x_2(t)| + |p_1(t) - p_2(t)| \right) \right.$$
$$\left. + \left| p_1(s)\bar{g}_x(x, x_1(s), p_1(s)) - p_2(s)\bar{g}_x(x, x_2(s), p_2(s)) \right| \right] ds.$$

Since $p_i \in C_\gamma(I)$, we have that $|p_i| \leq \gamma$, for $i = 1, 2$ and $t \in I$. Thus,

$$|T_2\mathbf{y}_1 - T_2\mathbf{y}_2| \leq \int_t^{t_f} \left[L_{f_x} \left(|x_1(t) - x_2(t)| + |p_1(t) - p_2(t)| \right) \right.$$
$$\left. +\gamma \left| \bar{g}_x(x, x_1(s), p_1(s)) - \bar{g}_x(x, x_2(s), p_2(s)) \right| \right] ds.$$

Lastly, since $\bar{g}_x$ is uniformly L_{g_x}-Lipschitz,

$$\begin{aligned}|T_2\mathbf{y}_1 - T_2\mathbf{y}_2| &\leq \int_t^{t_f} \left[L_{f_x}\left(|x_1(t)-x_2(t)|+|p_1(t)-p_2(t)|\right)\right. \\ &\qquad \left.+\gamma L_{g_x}\left(|x_1(t)-x_2(t)|+|p_1(t)-p_2(t)|\right)\right] ds \\ &\leq \left(L_{f_x}+\gamma L_{g_x}\right)\int_t^{t_f}\left[|x_1(s)-x_2(s)|+|p_1(s)-p_2(s)|\right] ds \\ &\leq \left(L_{f_x}+\gamma L_{g_x}\right)\sup_{t\in I}\left(|x_1(t)-x_2(t)|+|p_1(t)-p_2(t)|\right)\int_t^{t_f} ds \\ &\leq \left(L_{f_x}+\gamma L_{g_x}\right)\|\mathbf{y}_1-\mathbf{y}_2\|_\infty (t_f - t).\end{aligned} \tag{5.12}$$

Taking the supremum over $t \in I$, we arrive at

$$\|T_2\mathbf{y}_1 - T_2\mathbf{y}_2\|_\infty \leq \left(L_{f_x}+\gamma L_{g_x}\right)\|\mathbf{y}_1-\mathbf{y}_2\|_\infty (t_f - t_0). \tag{5.13}$$

Thus by (5.11) and (5.13), we have shown that if $L_g(t_f - t_0) < 1$ and $\left(L_{f_x}+\gamma L_{g_x}\right)(t_f - t_0) < 1$, then $\mathbf{T}$ is contractive with respect to $\|\cdot\|_\infty$ with contraction factor $c = \max\left\{L_g(t_f-t_0), \left(L_{f_x}+\gamma L_{g_x}\right)(t_f-t_0)\right\}$. □

A continuity result for the fixed points of these operators similar to that for Picard operators mentioned in Chapter 2 may also be established.

Remark: It may be necessary to divide the interval I into smaller subintervals to ensure that $\mathbf{T}$ meets the conditions to be contractive. In this case, the collage distances obtained on each subinterval, Δ_i may be summed to obtain a collage distance over the entire interval of interest.

5.2 Examples

5.2.1 A LQ IOCP Subject to a Single Controlled ODE and a Single Control

We now investigate a variety of scenarios all drawn from the following problem. Consider a linear control system for which the state equation is given by

$$\dot{x}(t) = -x(t) + u(t). \tag{5.14}$$

If this system is subject to the control input

$$u(t) = \frac{\sinh(\sqrt{2}(t-1))}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})}, \tag{5.15}$$

then this system admits the state trajectory

$$x(t) = \frac{\sqrt{2}\cosh(\sqrt{2}(t-1)) - \sinh(\sqrt{2}(t-1))}{\sqrt{2}\cosh(\sqrt{2}) + \sinh(\sqrt{2})}. \tag{5.16}$$

This pair is optimal (see Chapter 4) for the quadratic objective functional

$$\sup_{u \in \mathcal{U}} \int_0^1 \frac{1}{2}x(t)^2 + \frac{1}{2}u(t)^2 dt. \tag{5.17}$$

We use this fact for comparative purposes in the following analysis. We denote the integrand of Equation (5.17) by $f_{\text{true}}(t, x(t), u(t))$.

Suppose that we wish to solve the inverse optimal control problem of identifying a performance index,

$$\int_{t_0}^{t_f} f^{\lambda}(t, x(t), u(t))dt, \tag{5.18}$$

that is optimized by (5.15)–(5.16). We select a form for $f^\lambda(t, x(t), u(t))$ dependent on the vector of unknown parameters λ and apply the Collage Method as outlined in Section 5.1.

Recall that the IOCP starts with only a given control u_{target} and associated state x_{target}. These target functions are obtained by sampling from the known control (5.15) and state (5.16) at N equidistant points in I and to simulate observation error, normally distributed noise of amplitude ϵ is added to the data. The resulting datasets for the state and control have $M_x = 2$ and $M_u = 3$ degree polynomials respectively fit to them via least squares.

There is a practical problem that we must address. The collage distance Δ, which we wish to minimize, depends on the known control and state, the parameters that define f^λ, and also the adjoint $p(t)$. In practice, we cannot directly observe the adjoint $p(t)$ as it is a mathematical construct. We address this issue by rearranging the optimality condition in Equation (5.6) to express $p(t)$ in terms of the known control and state and the parameters λ:

$$p(t) = \frac{f_u^\lambda(t, x_{target}(t), u_{target}(t))}{g_u(t, x_{target}(t), u_{target}(t))}. \tag{5.19}$$

Thus, if $g_u \neq 0$, which we can check a priori, we may replace all instances of $p(t)$ in the collage distance with this equivalent quantity.

We apply the Collage-Based Approach to this problem by considering a variety of scenarios. To showcase the performance of the method, we report the parameters recovered and the collage distance. We omit the approximations error here as it will remain the same from trial to trial as the state equation (5.14) does not contain any parameters.

Example 5.1. *First we consider the case where the form of the unknown function* f^λ *matches the exact form of the true objective:* $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 u^2$*. These results are presented in Table 5.1.*

ϵ	N	λ_0	λ_1	Δ^2
0	10	0.4993	0.4995	2.257×10^{-5}
	25	0.4997	0.4999	1.972×10^{-5}
	50	0.4998	0.5000	1.919×10^{-5}
	100	0.4999	0.5001	1.903×10^{-5}
0.01	10	0.4998	0.4985	2.292×10^{-4}
	25	0.5004	0.4995	2.096×10^{-4}
	50	0.5002	0.5001	2.016×10^{-4}
	100	0.5007	0.4996	2.276×10^{-4}
0.05	10	0.5022	0.4945	5.404×10^{-3}
	25	0.5034	0.4979	4.922×10^{-3}
	50	0.5017	0.5004	4.657×10^{-3}
	100	0.5041	0.4978	5.270×10^{-3}
0.1	10	0.5051	0.4895	2.166×10^{-2}
	25	0.5072	0.4959	1.971×10^{-2}
	50	0.5036	0.5007	1.861×10^{-2}
	100	0.5083	0.4954	2.046×10^{-2}

Table 5.1: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$ and $\lambda_1^{\text{true}} = 0.5$.

Recovery of the parameters is robust, remaining within 0.0105 *or* 2.1% *of the exact values even with error of amplitude* $\epsilon = 0.1$.

With this “ideal” scenario producing strong results, we now investigate the performance of the method when we include extraneous terms in f^λ not present in f_{true}.

Example 5.2. *Next we consider several forms of f^λ containing extraneous terms. Tables 5.2–5.4 contain the results for trials with a single additional polynomial term present in f^λ. In each of these scenarios, values the Collage Method recovers for the extraneous parameter λ_1 appear to approach* 0 *as the number of sample points increases, while the parameters λ_0 and λ_2 remain close to the expected value of* 0.5.

N	λ_0	λ_1	λ_2	Δ^2
10	0.4912	0.0105	0.5010	2.231×10^{-5}
25	0.4929	0.0087	0.5011	1.954×10^{-5}
50	0.4936	0.0081	0.5011	1.903×10^{-5}
100	0.4939	0.0078	0.5012	1.889×10^{-5}

Table 5.2: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t, x, u) = \lambda_0 x^2 + \lambda_1 x + \lambda_2 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = 0$ and $\lambda_2^{\text{true}} = 0.5$.

N	λ_0	λ_1	λ_2	Δ^2
10	0.4942	-0.0172	0.4772	2.256×10^{-5}
25	0.4957	-0.0136	0.4823	1.972×10^{-5}
50	0.4962	-0.0123	0.4841	1.919×10^{-5}
100	0.4965	-0.0116	0.4850	1.902×10^{-5}

Table 5.3: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t, x, u) = \lambda_0 x^2 + \lambda_1 x u + \lambda_2 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = 0$ and $\lambda_2^{\text{true}} = 0.5$.

Example 5.3. *Extending, we select forms for f^λ with multiple extraneous polynomial terms. We include two particular choices of f^λ; we first consider $f^\lambda(t, x, u) = \lambda_0 x^2 + \lambda_1 x^3 + \lambda_2 u^3 + \lambda_3 x + \lambda_4 u^2$, expecting that the recovered values for the three additional parameters λ_1, λ_2, and λ_3 will be small. These results are found in Table 5.5, and while the recovered parameters appear, at least superficially, to define a nearby problem, only*

N	λ_0	λ_1	λ_2	Δ^2
10	0.4997	-0.0136	0.4934	2.233×10^{-5}
25	0.5000	-0.0112	0.4948	1.956×10^{-5}
50	0.5001	-0.0104	0.4953	1.905×10^{-5}
100	0.5002	-0.0099	0.4956	1.890×10^{-5}

Table 5.4: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^{\lambda}(t,x,u) = \lambda_0 x^2 + \lambda_1 u^3 + \lambda_2 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = 0$ and $\lambda_2^{\text{true}} = 0.5$.

λ_3 appears to be approaching zero as the number of sample points increases. In an effort to improve the results and reduce error, the same scenario was considered but with higher degree polynomial approximations for x_{target} and u_{target} of degree $M_x = M_u = 4$. This illustrates that the quality of the approximation of the system and control data become more significant as we attempt to identify more unknown values. As Table 5.6 shows, with this configuration of the collage method, we now recover parameters close to their expected values, even with as few as $N = 10$ sample points, and these parameters appear to approach their expected values as the number of sample points increases.

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	0.5865	-0.1357	0.2407	0.0641	0.6394	2.229×10^{-5}
25	0.5923	-0.1401	0.2454	0.0625	0.6422	1.952×10^{-5}
50	0.5945	-0.1417	0.2473	0.0620	0.6433	6.433×10^{-5}
100	0.5956	-0.1426	0.2482	0.0617	0.6439	1.887×10^{-5}

Table 5.5: Parameters recovered by the Inverse Optimal Control Collage Method with polynomial approximations of degree $M_x = 2$ and $M_u = 3$ for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^{\lambda}(t,x,u) = \lambda_0 x^2 + \lambda_1 x^3 + \lambda_2 u^3 + \lambda_3 x + \lambda_4 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = \lambda_2^{\text{true}} = \lambda_3^{\text{true}} = 0$ and $\lambda_4^{\text{true}} = 0.5$.

However, for some choices of f^{λ}, improving the quality of the approximation still

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	0.4900	0.0125	-0.0197	-0.0034	0.4886	1.322×10^{-9}
25	0.4932	0.0087	-0.0141	-0.0026	0.4919	1.043×10^{-9}
50	0.4947	0.0069	-0.0114	-0.0022	0.4934	7.182×10^{-11}
100	0.4956	0.0059	-0.0098	-0.0020	0.4943	9.493×10^{-10}

Table 5.6: Parameters recovered by the Inverse Optimal Control Collage Method with polynomial approximations of degree $M_x = M_u = 4$ for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t, x, u) = \lambda_0 x^2 + \lambda_1 x^3 + \lambda_2 u^3 + \lambda_3 x + \lambda_4 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = \lambda_2^{\text{true}} = \lambda_3^{\text{true}} = 0$ and $\lambda_4^{\text{true}} = 0.5$.

does not yield a parameter set corresponding to the expected objective function. Table 5.7 contains the parameters recovered for $f^\lambda(t, x, u) = \lambda_0 x^2 + \lambda_1 x + \lambda_2 u + \lambda_3 u^2$, *again using* $M_x = 2$ *and* $M_u = 3$ *degree polynomial approximations. For this choice of objective function, the minimizing parameter values for the collage distance are quite clear, as there is little change for increasing values of* N. *Increasing the degree of the approximating polynomials, and even using the true state and control functions in the collage method yield similar results.*

The consistency of the results in Table 5.7 suggest that the recovered objective function may also be optimized by the state (5.16) *and control* (5.15). *This possibility from a practical viewpoint is one that can not be discounted, and since Pontryagin's Maximum Principle is only a necessary condition, one must verify that an objective recovered in this way does in fact admit the given data as a solution when subject to the known state dynamics, described by the state equation.*

So far, we have only considered extraneous terms that are polynomial. The next example demonstrates that there is potential for the method to handle other classes

N	λ_0	λ_1	λ_2	λ_3	Δ^2
10	7.8242×10^{-24}	0.3858	-0.3858	7.9802×10^{-24}	2.229×10^{-5}
25	6.9540×10^{-24}	0.3858	-0.3858	7.0687×10^{-24}	1.952×10^{-5}
50	2.1511×10^{-23}	0.3858	-0.3858	2.1840×10^{-23}	1.901×10^{-5}
100	1.3676×10^{-23}	0.3858	-0.3858	1.3876×10^{-23}	1.887×10^{-5}

Table 5.7: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 x + \lambda_2 u + \lambda_3 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = \lambda_2^{\text{true}} = 0$ and $\lambda_3^{\text{true}} = 0.5$.

of functions as well.

Example 5.4. *In this trial, we use an objective function of the form* $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 e^x + \lambda_2 u + \lambda_3 u^2$ *is implemented. The results are found in Table 5.8. For this choice of* f^λ*, the collage method identifies values for the extraneous term parameters that appear to tend towards zero, and the parameters that match the true objective function appear to tend toward their true values. As expected, improving the approximations used for the state and control also resulted in the recovery of parameters closer to the expected values. Improvement was also found by increasing the degree of polynomial approximations from* $M_x = 2, M_u = 3$.

N	λ_0	λ_1	λ_2	λ_3	Δ^2
10	0.4411	0.0333	-0.0157	0.4802	2.230×10^{-5}
25	0.4448	0.0310	-0.0156	0.4804	1.954×10^{-5}
50	0.4463	0.0302	-0.0156	0.4805	1.903×10^{-5}
100	0.4471	0.0297	-0.0156	0.4806	1.889×10^{-5}

Table 5.8: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 \mathrm{e}^x + \lambda_2 u + \lambda_3 u^2$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = \lambda_2^{\text{true}} = 0$ and $\lambda_3^{\text{true}} = 0.5$.

We now turn to a variation of the inverse problem studied so far in this Section. To this point, it has been assumed that the structure of the system is known through the state equation, but we now discard this assumption, seeking unknown parameters not only in the objective function, but also in the state equation.

Example 5.5. *Consider again $f^{\lambda}(t, x, u) = \lambda_0 x^2 + \lambda_1 u^2$, and also a parameterized form of the state equation, $g^{\lambda}(t, x, u) = \lambda_2 x + \lambda_3 u$. We proceed in the same way as for the prequel examples. Results are presented in Table 5.9. The approximation error is included for the state in this case. The true state found in Equation* (5.16) *is denoted by x_{true} and the state found by solving the state equation using the recovered parameters in denoted by $x_{collage}$.*

The effect of added noise is much more significant when compared to only recovering the unknown objective function in Example 5.1. Furthermore, a greater quantity of data is required before recovering parameters within a similar deviation from the exact values that was obtained for only unknowns in the objective. We include $N = 1000$ sample points in Table 5.9 for noise levels of $\epsilon = 0$ and $\epsilon = 0.01$ to illustrate this. Beyond this level of noise, there are significant differences between recovered and true values of the parameters.

Example 5.6. *We complete our analysis of this example problem by including the results obtained when, rather than sampling from the true state and control functions, the sample data is drawn from a numerical solution of the forward problem, obtained via the Forward Backward Sweep Method. The purpose of this example is to illus-*

ϵ	N	λ_0	λ_1	λ_2	λ_3	$\lVert x_{\text{true}} - x_{\text{collage}} \rVert$	Δ^2
0	10	0.5110	0.3966	-1.0698	0.7924	6.298×10^{-3}	1.994×10^{-5}
	25	0.5058	0.4461	-1.0364	0.8915	7.289×10^{-3}	1.900×10^{-5}
	50	0.5039	0.4645	-1.0240	0.9283	7.724×10^{-3}	1.887×10^{-5}
	100	0.5028	0.4741	-1.0175	0.9475	7.962×10^{-3}	1.885×10^{-5}
	1000	0.5019	0.4830	-1.0115	0.9653	8.187×10^{-3}	1.887×10^{-5}
0.01	10	0.4986	0.5668	-0.9925	1.1732	1.656×10^{-2}	2.587×10^{-5}
	25	0.4816	0.7152	-0.8881	1.4366	1.774×10^{-2}	3.323×10^{-5}
	50	0.4880	0.6566	-0.9276	1.3158	1.929×10^{-2}	4.223×10^{-5}
	100	0.4836	0.7025	-0.8985	1.4102	1.907×10^{-2}	3.729×10^{-5}
	1000	0.4870	0.6597	-0.9235	1.3219	1.951×10^{-2}	4.249×10^{-5}
0.05	10	0.4448	1.2710	-0.6545	2.5962	7.067×10^{-2}	2.521×10^{-4}
	25	0.3850	1.7713	-0.2853	3.6329	7.544×10^{-2}	3.886×10^{-4}
	50	0.4233	1.4282	-0.5303	2.8958	7.870×10^{-2}	5.901×10^{-4}
	100	0.4049	1.6168	-0.4038	3.3051	7.836×10^{-2}	4.343×10^{-4}
0.1	10	0.3672	2.2190	-0.1570	4.6266	0.1490	1.006×10^{-3}
	25	0.2652	3.0436	0.4874	6.4042	0.1821	1.544×10^{-3}
	50	0.3398	2.3974	-0.0071	4.9392	0.1749	2.277×10^{-3}
	100	0.3017	2.7624	0.2594	5.7767	0.1780	1.677×10^{-3}

Table 5.9: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 u^2$ and for $g^\lambda(t,x,u) = \lambda_2 x + \lambda_3 u$. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$, $\lambda_1^{\text{true}} = 0.5$, $\lambda_2^{\text{true}} = -1$ and $\lambda_3^{\text{true}} = 1$. The $L^2(I)$ norm is used to calculate the error in the states.

trate the effectiveness of the Collage-Based Approach to IOCPs, even when additional uncertainty is introduced. An N-element partition of the time interval $[0,1]$ is used along with a combined relative error tolerance of $\delta = 0.001$.

Comparing Tables 5.1 and 5.10, we see that this added error introduced by the Forward Backward Sweep method does not prevent the Collage Method from identifying the underlying objective. We see very similar results between this scenario and sampling directly from the true state and control functions.

ϵ	N	λ_0	λ_1	Δ^2
0	10	0.5000	0.4994	2.340×10^{-5}
	25	0.4998	0.4999	1.980×10^{-5}
	50	0.4998	0.5000	1.921×10^{-5}
	100	0.4999	0.5001	1.903×10^{-5}
0.01	10	0.5004	0.4985	2.326×10^{-4}
	25	0.5005	0.4995	2.193×10^{-4}
	50	0.5002	0.5001	2.003×10^{-4}
	100	0.5007	0.4996	2.319×10^{-4}
0.05	10	0.5018	0.4949	5.323×10^{-3}
	25	0.5033	0.4979	5.112×10^{-3}
	50	0.5017	0.5004	4.602×10^{-3}
	100	0.5040	0.4978	2.700×10^{-3}
0.1	10	0.5036	0.4903	2.126×10^{-2}
	25	0.5069	0.4959	2.044×10^{-2}
	50	0.5036	0.5008	1.838×10^{-2}
	100	0.5082	0.4955	2.135×10^{-2}

Table 5.10: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.15)–(5.16), (5.18) for $f^\lambda(t,x,u) = \lambda_0 x^2 + \lambda_1 u^2$. The state and control functions were obtained by a Forward Backward Sweep scheme. For comparison, the exact values are $\lambda_0^{\text{true}} = 0.5$ and $\lambda_1^{\text{true}} = 0.5$.

5.2.2 A LQ IOCP Subject to a System of Two Controlled ODEs and Two Controls

We now explore an extension of the theory presented in Section 5.1 to handle an IOCP with a system of state equations and multiple controls. This extension is not presented here in full; the required modifications to Theorems 5.2 and 5.3 follow naturally for a state $\mathbf{x} \in \mathbb{R}^n$ and control $\mathbf{u} \in \mathbb{R}^m$, and details regarding Pontryagin's Maximum Principle for systems may be found in [3]. We are considering optimal control problems of the form

$$\sup_{\mathbf{u}\in\mathcal{U}} \int_{t_0}^{t_f} f(t, \mathbf{x}(t), \mathbf{u}(t))dt$$

subject to the controlled system

$$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)),$$

$$\mathbf{x}(t_0) = \mathbf{x}_0.$$

The key relation needed to solve the inverse problem using the Collage Method was obtained from the optimality condition (5.6). In the system setting, this condition is given by

$$\mathbf{g}^*_{\mathbf{u}}(t, \mathbf{x}, \mathbf{u})\mathbf{p}(t) + f_{\mathbf{u}}(t, \mathbf{x}, \mathbf{u}) = 0, \tag{5.20}$$

where $\mathbf{g}^*_{\mathbf{u}}$ is the adjoint of the Jacobian matrix of $\mathbf{g}$ with respect to $\mathbf{u}$ and $f_{\mathbf{u}}$ is the gradient vector of f with respect to $\mathbf{u}$. Equation (5.20) cannot, in general, be solved for $\mathbf{p}$. However, for $m = n$, if $\mathbf{g}^*_{\mathbf{u}}(t, \mathbf{x}, \mathbf{u})$ is invertible, we may write

$$\mathbf{p}(t) = -\left(\mathbf{g}^*_{\mathbf{u}}(t, \mathbf{x}, \mathbf{u})\right)^{-1} f_{\mathbf{u}}(t, \mathbf{x}, \mathbf{u}).$$

This invertibility condition on $\mathbf{g}^*_{\mathbf{u}}$ is the natural extension of the condition $g_u \neq 0$ in the $n = m = 1$ case.

We complete our demonstration of the Collage-Based Approach to solving IOCPs with the following example, where the state and control both belong to $\mathbb{R}^2$.

Example 5.7. *Consider an OC problem governed by the linear state system*

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t). \tag{5.21}$$

The matrices A and B are given as

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

This problem admits the following solution (obtained algebraically using Maple 21 and expressed only approximately for brevity),

$$\begin{aligned} x_1(t) = -0.0500e^t(&53.6647\cos(2.4495t)t - 112.5998305\sin(2.4495t)t \\ &-20.0000\cos(2.4495t) + 72.6937\sin(2.4495t)) \end{aligned} \tag{5.22}$$

$$\begin{aligned} x_2(t) = -0.1000e^t(&-9.9999\cos(2.4495t) + 45.9687\cos(2.4495t)t \\ &-49.4516\sin(2.4495t) + 21.9085\sin(2.4495t)t) \end{aligned} \tag{5.23}$$

$$\begin{aligned} u_1(t) = 0.0311e^t(&140.7318\sin(2.4495t)t + 295.2847\cos(2.4495t)t \\ &-28.3386\sin(2.4495t) - 202.1245\cos(2.4495t)) \end{aligned} \tag{5.24}$$

$$\begin{aligned} u_2(t) = -0.0078e^t(&344.7212\cos(2.4495t)t - 723.2969\sin(2.4495t)t \\ &-837.1555\cos(2.4495t) + 129.1992\sin(2.4495t)), \end{aligned} \tag{5.25}$$

when the quadratic objective functional is given by

$$\sup_{u \in U} \int_0^1 3x_1^2(t) + 2x_2^2(t) + u_1^2(t) + u_2^2(t)dt. \tag{5.26}$$

As we did in the previous section, we attempt to solve the IOCP of identifying the function $f^\lambda(t, \mathbf{x}, \mathbf{u})$ that defines an unknown objective function,

$$\int_0^1 f^\lambda(t, \mathbf{x}, \mathbf{u})dt, \tag{5.27}$$

which when subject to the state system (5.21), *will admit the given state* (5.22)–(5.23) *and control* (5.24)–(5.25) *data as a solution. The given solutions are sampled at N equidistant points and fourth-degree polynomials are fit to the data via a least squares procedure to simulate observation. We use the true objective* (5.26) *for comparison with the results of the Collage Method applied to the IOCP.*

The form of the unknown objective that we focus on is

$$f^{\lambda}(t, \mathbf{x}, \mathbf{u}) = \lambda_0 x_1^2 + \lambda_1 x_1 x_2 + \lambda_2 x_2^2 + \lambda_3 u_1^2 + \lambda_4 u_1 u_2 + \lambda_5 u_2^2.$$

A minimizing set of parameters for the collage distance could not be obtained for this full set of unknowns, and further exploration is necessary. Instead, we provided one out of the six unknown parameters and were able to effectively recover the remainder in the four cases where the supplied parameter corresponded to one of the nonzero parameters in the true objective (5.26)*. These results are presented in Tables 5.11–5.14. In each case, the parameter which was provided is indicated and the remaining five are labelled as λ_0–λ_4. Each recovered parameter was driven closer to its expected value as the number of sample points N was increased. The squared collage distances are also consistent across each choice of f^{λ}.*

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	3.154×10^{-2}	2.0374	1.0055	5.236×10^{-3}	1.0201	2.097×10^{-3}
25	1.934×10^{-2}	2.0231	1.0036	3.235×10^{-3}	1.0125	1.572×10^{-3}
50	1.305×10^{-2}	2.0157	1.0026	2.203×10^{-3}	1.0085	1.446×10^{-3}
100	9.393×10^{-3}	2.0114	1.0020	1.603×10^{-3}	1.0062	1.409×10^{-3}

Table 5.11: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.24)–(5.25), (5.27) for $f^\lambda(t, \mathbf{x}, \mathbf{u}) = q_{11}x_1^2 + \lambda_0 x_1 x_2 + \lambda_1 x_2^2 + \lambda_2 u_1^2 + \lambda_3 u_1 u_2 + \lambda_4 u_2^2$. For comparison, the exact values are $\lambda_0 = 0, \lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 0, \lambda_4 = 1$.

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	2.9057	5.167×10^{-2}	0.9791	8.622×10^{-3}	1.0027	2.022×10^{-3}
25	2.9355	3.505×10^{-2}	0.9860	5.878×10^{-3}	1.0019	1.535×10^{-3}
50	2.9485	2.776×10^{-2}	0.9891	4.679×10^{-3}	1.0016	1.420×10^{-3}
100	2.9555	2.386×10^{-2}	0.9907	4.038×10^{-3}	1.0014	1.388×10^{-3}

Table 5.12: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.24)–(5.25), (5.27) for $f^\lambda(t, \mathbf{x}, \mathbf{u}) = \lambda_0 x_1^2 + \lambda_1 x_1 x_2 + q_{22} x_2^2 + \lambda_2 u_1^2 + \lambda_3 u_1 u_2 + \lambda_4 u_2^2$. For comparison, the exact values are $\lambda_0 = 3, \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 0, \lambda_4 = 1$.

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	2.9772	3.981×10^{-2}	2.0327	6.610×10^{-3}	1.0183	2.077×10^{-3}
25	2.9842	2.566×10^{-2}	2.0206	4.284×10^{-3}	1.0117	1.562×10^{-3}
50	2.9877	1.891×10^{-2}	2.0150	3.177×10^{-3}	1.0085	1.439×10^{-3}
100	2.9897	1.513×10^{-2}	2.0118	2.556×10^{-3}	1.0068	1.403×10^{-3}

Table 5.13: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.24)–(5.25), (5.27) for $f^\lambda(t, \mathbf{x}, \mathbf{u}) = \lambda_0 x_1^2 + \lambda_1 x_1 x_2 + \lambda_2 x_2^2 + r_{11} u_1^2 + \lambda_3 u_1 u_2 + \lambda_4 u_2^2$. For comparison, the exact values are $\lambda_0 = 3, \lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 0, \lambda_4 = 1$.

N	λ_0	λ_1	λ_2	λ_3	λ_4	Δ^2
10	2.8937	5.351×10^{-2}	1.9943	0.9755	8.934×10^{-3}	2.014×10^{-3}
25	2.9266	3.653×10^{-2}	1.9959	0.9834	6.128×10^{-3}	1.530×10^{-3}
50	2.9408	2.917×10^{-2}	1.9966	0.9868	4.917×10^{-3}	1.416×10^{-3}
100	2.9483	2.525×10^{-2}	1.9970	0.9887	4.273×10^{-3}	1.385×10^{-3}

Table 5.14: Parameters recovered by the Inverse Optimal Control Collage Method for the optimal control problem given by (5.24)–(5.25), (5.27) for $f^\lambda(t, \mathbf{x}, \mathbf{u}) = \lambda_0 x_1^2 + \lambda_1 x_1 x_2 + \lambda_2 x_2^2 + \lambda_3 u_1^2 + \lambda_4 u_1 u_2 + r_{22} u_2^2$. For comparison, the exact values are $\lambda_0 = 3, \lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 1, \lambda_4 = 0$.

Chapter 6

Conclusions and Future Work

In this thesis we have presented a Collage-Based Approach for solving IOCPs governed by ODEs with unique solutions. This method is motivated by Pontryagin's Maximum Principle, a necessary (and sufficient, for problems with unique solutions) condition for a control to be optimal. This formulation retains the fundamental idea behind Collage-Based methods, where we bound the approximation error by a quantity that is minimized more readily.

The results of this Collage-Based Approach to IOCPs with unique solutions demonstrated that in a variety of scenarios, the method is robust to observation error and can effectively identify both necessary and extraneous terms in the objective function. The results even suggest that this approach may have success in identifying unknown parameters in both the objective and the state equation(s). We have also shown that this method can extend to problems in higher dimensions, although additional investigation and rigour is still needed to fully understand the complexities of this extension.

There are a number of considerations that arose in the course of developing this approach. Optimal Control Problems are, in general, plagued by a lack of existence and uniqueness, which is central to the Collage framework as presented here. Thus, significant limitations of this work are that we require that the problem has a unique solution and that the matrix $\mathbf{g_u}^*$ must be invertible (or equivalently, $g_u \neq 0$). While this excludes many classes of problems, the method remains useful for others, such as the LQ problems on which we have demonstrated the method. The conditions on the operator required to implement the Collage Method may be restrictive, and a partition of the time interval of interest may be required to satisfy these conditions.

One possible way to deal with these limitations is to extend our theory to use a *generalized* collage distance, introduced in [36]. This approach is free from many of the trappings of the Collage Method based on Banach's Fixed Point Theorem presented in this thesis. Additional future work in this area may include

- publication of these results;
- a complete and thorough investigation of the extension to IOCPs with systems of state equations and multiple controls;
- extending to other classes of ODE IOCPs, for example free-time or fixed-endpoint problems;

- investigating more complex existence criteria;
- investigating the application of the Collage-Based Approach to dynamic programming and the Hamilton-Jacobi-Bellman equation to solve an IOCP; and
- developing a Collage-Based Approach to IOCPs governed by partial differential equations.

Bibliography

[1] N. Ab Azar, A. Shahmansoorian, and M. Davoudi. From inverse optimal control to inverse reinforcement learning: A historical review. *Annual Reviews in Control*, 50:119–138, 2020.

[2] M. Alamir and S. Chareyron. State-constrained optimal control applied to cell-cycle-specific cancer chemotherapy. *Control Appl. Meth*, 28:175–190, 2007.

[3] S. Aniţa, V. Arnăutu, and V. Capasso. *An Introduction to Optimal Control Problems in Life Sciences and Economics.* Springer Science+Business Media, 2011.

[4] V. Arnautu and P. Neittaanmaki. *Optimal Control from Theory to Computer Programs.* Kluwer Academic Publishers, 2003.

[5] M.F. Barnsley, V. Ervin, D Hardin, and J. Lancaster. Solution of an inverse problem for fractals and other sets. *Proceedings of the National Academy of Sciences*, 83:1975–1977, 1985.

[6] Y. Batmani and H. Khaloozadeh. Optimal chemotherapy in cancer treatment: State dependent riccati equation control and extended kalman filter. *Optimal Control Applications and Methods*, 34:562–577, 2013.

[7] R. Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

[8] A. Bensoussan. *Estimation and Control of Dynamical Systems*, volume 48. Springer International Publishing AG, 2018.

[9] V.G. Boltyanski, R.V. Gamkrelidze, E.F. Mishchenko, and L.S. Pontryagin. The maximum principle in the theory of optimal processes of control. pages 464–469, 1960.

[10] W.E. Boyce, R.C. DiPrima, and D.B. Meade. *Elementary Differential Equations and Boundary Value Problems.* John Wiley & Sons Inc., 2018.

[11] A. Bressan and B. Piccoli. *Introduction to the mathematical theory of control*, volume 1. American Institute of Mathematical Sciences Springfield, 2007.

[12] V. Brott. A collage-based approach to a sturm-liouville boundary value inverse problem with boundary data, 2018.

[13] R.L. Burden and J.D. Faires. *Numerical Analysis.* Nelson Education Ltd., ninth edition, 2010.

[14] J.C. Butcher. *Numerical Methods for Ordinary Differential Equations.* 2016.

[15] V. Capasso, H.E. Kunze, D. La Torre, and E.R. Vrscay. Solving inverse problems for differential equations by a "generalized collage" method and application to a mean field stochastic model. *Nonlinear Analysis: Real World Applications*, 15:276–289, 2014.

[16] J. Cea. *Lectures on Optimization - Theory and Algorithms.* Springer-Verlag, 1978.

[17] E.A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company Inc., 1955.

[18] L.G. De Pillis and A. Radunskaya. The dynamics of an optimally controlled tumor model: A case study. *COMPUTER MODELLING Mathematical and Computer Modelling*, 37:1221–1244, 2003.

[19] K.D. Do. Global inverse optimal control of vertical take-off and landing aircraft. *IFAC Journal of Systems and Control*, 15:100132, 2021.

[20] M. Engelhart, D. Lebiedz, and S. Sager. Optimal control for selected cancer chemotherapy ode models: A view on the potential of optimal schedules and choice of objective function. 2011.

[21] H.P. Geering. *Optimal Control with Engineering Applications.* Springer-Verlag.

[22] M. Gostimirovic, P. Kovac, and M. Sekulic. An inverse optimal control prob-

lem in the electrical discharge machining. *Sadhana - Academy Proceedings in Engineering Sciences*, 43:1–10, 2018.

[23] D. Grass, J.P. Caulkins, G. Feichtinger, G. Tragler, and D.A. Behrens. *Optimal Control of Nonlinear Processes.* 2008.

[24] K. Hatz, J.P. Schloder, and H.G. Bock. Estimating parameters in optimal control problems. *SIAM Journal on Scientific Computing*, 34:1707–1728, 2012.

[25] I. Iserles. *A First Course in the Numerical Analysis of Differential Equations.* Cambridge University Press, 1996.

[26] M. Itik, M.U. Salamci, and S.P. Banks. Optimal control of drug therapy in cancer treatment. *Nonlinear Analysis*, 71:e1473–e1486, 2009.

[27] M. Johnson, N. Aghasadeghi, and T. Bretl. Inverse optimal control for deterministic continuous-time nonlinear systems. 2013.

[28] W. Jun, J. Ping, C. Zhong, and L. Wei. Adaptive inverse optimal control for a class of nonlinear systems. pages 1925–1928, 2011.

[29] R.E. Kalman. When is a Linear Control System Optimal? *Journal of Basic Engineering*, pages 51–60, 1964.

[30] G. Knowles. *An Introduction to Applied Optimal Control.* Academic Press, Inc., New York, 1981.

[31] E. Kreyzig. *Introductory Functional Analysis.* John Wiley & Sons, 1978.

[32] H. Kunze and S. Murdock. Solving inverse two-point boundary value problems using collage coding. *Inverse Problems*, 22, 2006.

[33] H.E. Kunze, J.E. Hicken, and E.R. Vrscay. Inverse problems for odes using contraction maps and suboptimality of the "collage method". *Inverse Problems*, 20:977–991, 2004.

[34] H.E. Kunze and D. La Torre. A collage-type approach to inverse problems for elliptic pdes on perforated domains, 2015.

[35] H.E. Kunze, D. La Torre, and E.R. Vrscay. Random fixed point equations and inverse problems using the collage method for contraction mappings. *J. Math. Anal. Appl*, 334:1116–1129, 2007.

[36] H.E. Kunze, D. La Torre, and E.R. Vrscay. A generalized collage method based upon the lax-milgram functional for solving boundary value inverse problems. *Nonlinear Analysis*, 71:1337–1343, 2009.

[37] H.E. Kunze, D. La Torre, and E.R. Vrscay. Inverse problems for random differential equations using the collage method for random contraction mappings. *Journal of Computational and Applied Mathematics*, 223:853–861, 2009.

[38] H.E. Kunze, D. La Torre, and E.R. Vrscay. Solving inverse problems for variational equations using "generalized collage methods," with applications to bound-

ary value problems. *Nonlinear Analysis: Real World Applications*, 11:3734–3743, 2010.

[39] H.E. Kunze and E.R. Vrscay. Solving Inverse Problems for ODEs Using the Picard Contraction Mapping. *Inverse Problems*, 15, 1999.

[40] E.B. Lee and L. Markus. *Foundations of Optimal Control Theory.* John Wiley & Sons, New York, 1967.

[41] K.M. Levere, B. Boreland, and J. Dewhurst. A computational comparison of three methods for solving a 1d boundary value inverse problem, 2021. In M. D. Kilgour, H. Kunze, R. Makarov, R. Melnik, & S. X. Wang (Eds.), Recent Developments in Mathematical, Statistical and Computational Sciences. Springer International Publishing AG. https://doi.org/10.1007/978-3-030-63591-6.

[42] K.M. Levere and H.E. Kunze. Using the collage method to solve one-dimensional two-point boundary value problems at steady-state. *Nonlinear Analysis, Theory, Methods and Applications*, 71:1–15, 2009.

[43] K.M. Levere, H.E. Kunze, and D. La Torre. A collage-based approach to solving inverse problems for second-order nonlinear parabolic pdes. *J. Math. Anal. Appl*, 406:120–133, 2013.

[44] D. Liberzon. *Calculus of Variations and Optimal Control Theory A Concise Introduction.* Princeton University Press, Princeton, 2012.

[45] M. McAsey, L. Mou, and W. Han. Convergence of the Forward-Backward Sweep Method in Optimal Control. *Computational Optimization and Applications*, 53(1):207–226, 2012.

[46] J.D. Meiss. *Differential Dynamical Systems.* Society for Industrial and Applied Mathematics, revised edition, 2017.

[47] T.L. Molloy, J.J. Ford, and T. Perez. Finite-horizon inverse optimal control for discrete-time nonlinear systems. *Automatica*, 87:442–446, 2018.

[48] K. Morris. *Introduction to Feedback Control.* Harcourt Academic Press, San Diego, 2001.

[49] P.J. Moylan and B.D.O. Anderson. Nonlinear regulator theory and an inverse optimal control problem. *IEEE Transactions on Automatic Control*, AC-18, 1973.

[50] R.W. Obermayer and F.A. Muckler. On the inverse optimal control problem in manual control systems. page 29, 1965.

[51] E. Pauwels, D. Henrion, and J.B. Lasserre. Inverse optimal control with polynomial optimization. *Proceedings of the IEEE Conference on Decision and Control*, 2015-Febru:5581–5586, 2014.

[52] R. Pytlak. *Numerical Methods for Optimal Control Problems with State Constraints.* Springer-Verlag, 1999.

[53] R. Ruiz-Cruz, E.N. Sanchez, F. Ornelas-Tellez, A.G. Loukianov, and R.G. Harley. Particle swarm optimization for discrete-time inverse optimal control of a doubly fed induction generator. *IEEE TRANSACTIONS ON CYBERNETICS*, 43, 2013.

[54] T. Sauer. *Numerical Analysis*, volume 38. Pearson Education Inc., second edition, 2012.

[55] R. Sepulchre, M. Jankovic, and P. Kokotovic. *Constructive Nonlinear Control*. Springer-Verlag, 1997.

[56] V.V. Ternovskii and M.M. Khapaev. Inverse problem method in optimal control. *Doklady Mathematics*, 83:357–360, 2011.

[57] F.E. Thau. On the inverse optimum control problem for a class of nonlinear autonomous systems. *IEEE Transactions on Automatic Control*, AC-12:674–681, 1967. Ref [3] in Moylan 73.

[58] R. Vinter. *Optimal Control*. Birkhauser, 2000.

[59] R. Yokoyama and E. Kinnen. The inverse problem of the optimal regulator. *IEEE Trans. Automat. Contr. (Short. Papers)*, AC-17:497–504, 1972. Ref [4] in Moylan 73.

[60] L.C. Young. *Lectures on the Calculus of Variations and Optimal Control Theory*. AMS Chelsea Publishing, second edition, 1980.