OPTIMAL TRAJECTORY GENERATION IN A DYNAMIC MULTI-BODY ENVIRONMENT USING A PSEUDOSPECTRAL METHOD

THESIS

Jacob A. Dahlke, 2d Lt, USAF

AFIT-ENY-MS-18-M-248

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENY-MS-18-M-248

OPTIMAL TRAJECTORY GENERATION IN A DYNAMIC MULTI-BODY ENVIRONMENT USING A PSEUDOSPECTRAL METHOD

THESIS

Presented to the Faculty

Department of Astronautical Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Astronautical Engineering

Jacob A. Dahlke, B.S.M.E.

2d Lt, USAF

March 2, 2018

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENY-MS-18-M-248

OPTIMAL TRAJECTORY GENERATION IN A DYNAMIC MULTI-BODY ENVIRONMENT USING A PSEUDOSPECTRAL METHOD

Jacob A. Dahlke, B.S.M.E.
2d Lt, USAF

Committee Membership:

Capt Joshuah Hess, Ph.D.
Chair

Lt Col Kirk Johnson, Ph.D.
Member

Lt Col Stuart Stanton, Ph.D.
Member

Abstract

High-altitude parking orbits could provide resiliency to the military space infrastructure by providing redundancy in key assets, allowing for rapid reconstitution of underperforming satellites. When analyzing trajectories in a high-altitude regime, two-body models of Keplerian motion become less accurate since the gravitational effects of other bodies are no longer negligible. To provide a higher fidelity model of the dynamics in a high-altitude regime, a multiple-body model can be used. In the Earth-Moon system, a spacecraft operating in the high-altitude regime can be modeled with three-body dynamics. With certain simplifying assumptions, the model is called the circular-restricted three-body problem (CR3BP). The CR3BP provides unique dynamics that could be exploited to provide beneficial trajectories unavailable and unobservable in a lower-order model. The tradeoff for using this higher-order model is there is no closed-form analytical solution and the dynamics are chaotic. Methods to search for optimal trajectories within the CR3BP are analyzed to determine viability in rapid mission development. A direct orthogonal collocation pseudospectral method is utilized to generate fuel- and time- optimal trajectories within the CR3BP. These results are compared to benchmarks from two-body dynamics, such as Hohmann transfers. Numerical approaches to finding optimal solutions are highly dependent on initial guesses to converge on candidate optimal solutions. To compound this issue, the chaotic dynamics in the CR3BP mean small variations in the initial conditions could lead to wildly varying trajectories. The results from the current research provide a methodology to establish a framework for rapid mission development in a dynamical environment, which may be essential to maintain space superiority and responsiveness.

Acknowledgments

First and foremost, I would like to thank my wife for her love and encouragement throughout the writing of this thesis. Without her unwavering support, I could not have completed this Master's program.

I would also like to thank my advisor. Despite requiring frequent assistance, Capt Hess was always available to provide mentorship and feedback and I will always be grateful that he was my advisor.

My committee members, Lt Col Johnson and Lt Col Stanton, provided me invaluable feedback throughout the thesis writing process and I'm thankful that they were willing to take so much time to support me.

Finally, I would like to thank my family. They are always a source of support and I am grateful for them. I'd like to thank my parents for giving me the ambition to pursue an education. I'd specifically like to thank my father for transferring his passion of all things space to me from a young age.

Thank you to everyone that provided your support to me. I would not have been able to do it without you.

Jacob A. Dahlke

Table of Contents

Page

Abstract . . . iv
Acknowledgments . . . v
List of Figures . . . viii
List of Tables . . . xi
List of Acronyms . . . xii

I. Introduction . . . 1

1.1 Motivation . . . 1
1.1.1 Current Issues . . . 1
1.1.2 Benefits of the Circular-Restricted Three-Body Problem (CR3BP) . . . 4
1.2 Problem Statement . . . 5
1.3 Thesis Overview . . . 6

II. Background . . . 7

2.1 Chapter Overview . . . 7
2.2 N-body Problem . . . 7
2.2.1 Perturbation Methods . . . 9
2.2.2 Two-Body Problem . . . 10
2.3 CR3BP . . . 14
2.3.1 Nondimensionalization . . . 16
2.3.2 Derivation of the CR3BP equations of motion (EOMs) . . . 18
2.3.3 Transformation for Barycentric Synodic Reference Frame to Earth-Centered Inertial (ECI) . . . 22
2.4 Characteristics of the CR3BP . . . 25
2.4.1 Symmetries in the CR3BP . . . 25
2.4.2 Jacobi Constant . . . 27
2.5 Equilibrium Solutions . . . 30
2.6 Periodic Orbits . . . 34
2.6.1 Equations of Variation . . . 35
2.6.2 Differential Corrections Applied to the CR3BP . . . 38
2.7 Invariant Manifolds . . . 44
2.8 Heteroclinic Connections . . . 47
2.9 Static Optimization . . . 50
2.10 Dynamic Optimization Discretization . . . 55
2.10.1 Indirect Optimization . . . 55
2.10.2 Direct Optimization . . . 58
2.10.3 Characteristics of Direct and Indirect Methods . . . 59

Page

2.10.4 Pseudospectral Optimization . . . 60
2.11 General Purpose Optimal Control Software (GPOPS) . . . 65
2.12 Literature Review of Relevant Works . . . 66

III. Research Methodology . . . 70

3.1 Chapter Overview . . . 70
3.2 Test Plan Overview . . . 70
3.3 Selection of Parameters . . . 72
3.4 Equations of Motion . . . 73
3.5 Initial Guess in GPOPS . . . 79
3.6 Continuation Method . . . 88
3.7 Objective Functions . . . 89
3.8 Two-Body Problem (2BP) Benchmark . . . 90
3.9 Summary . . . 92

IV. Results . . . 93

4.1 Chapter Overview . . . 93
4.2 Heteroclinic Trajectory in GPOPS . . . 93
4.3 Nearby Periodic Orbits . . . 95
4.4 Compound Objective Function Initial Guesses . . . 98
4.4.1 Compound Objective Functional: Original Periodic Orbit . . . 98
4.4.2 Compound Objective Functional: Nearby Periodic Orbit . . . 103
4.5 Evaluation of Methods used in Current Research . . . 109
4.5.1 Evaluation of Inclusion of Mass Loss . . . 109
4.5.2 Evaluation of Utilizing Initial Guess . . . 110
4.5.3 Evaluation of Pseudospectral Method . . . 111

V. Conclusions and Recommendations . . . 114

5.1 Summary of Work . . . 114
5.2 Future Work . . . 115
5.3 Conclusion . . . 116
Bibliography . . . 117

List of Figures

Figure Page

1.1 Satellite Catalog Growth [1] . . . 2

2.1 The N-body Problem [2] . . . 8

2.2 The Two-Body Problem [2] . . . 11

2.3 The CR3BP in its Rotating Reference Frame . . . 16

2.4 Synodic and ECI reference frames . . . 23

2.5 Symmetry of CR3BP across the x-z plane and time . . . 26

2.6 Symmetry of CR3BP across the x-y plane . . . 27

2.7 Valid Trajectory: Jacobi Constant used to check validity of numerical integration . . . 29

2.8 Invalid trajectory (through the secondary body): Jacobi Constant demonstrates the invalidity of the numerical integration . . . 30

2.9 Lagrange Points in Earth-Moon System . . . 32

2.10 Lagrange Points as a function of μ. The current location of the points is at $\mu \approx 0.01215$ to represent the Earth-Moon System (Adapted from Doedel et al. [3]) . . . 34

2.11 Boundary Value Problem Example . . . 36

2.12 Crossing of the x-z plane for an initial guess and then again after differential corrections, such that the crossing is perpendicular to the x-z plane (Adapted from Grebow [4]) . . . 41

2.13 L1 Lyapunov Family in the Earth-Moon System (Adapted from Grebow [4]) . . . 44

2.14 Stable (left) and Unstable (right) Manifold in the Earth-Moon System (Reproduced with Permission from Elsevier) [5] . . . 47

Figure Page

2.15 Surface of Section placed between two Lyapunov periodic orbits in the Earth-Moon System (Reproduced with Permission from Elsevier) [5] 48

2.16 Lower: Poincare Map of the manifolds intersecting the Surface of Section. Upper: The Heteroclinic Trajectories found from the two points of intersection on the Poincare Map (Reproduced with Permission from Elsevier) [5] 50

2.17 Global and Local Maxima and Minima (Adapted from Hess [6]) 52

2.18 Difference between Guassian Quadrature points methods (Adapted from Rao [7]) 62

2.19 Linkage between Phases in Multi-Phase Collocation (Adapted from Rao et al. [8]) 65

3.1 Form of the Control as an Angle α 74

3.2 Form of the Control in the x- and y-direction 76

3.3 Periodic Orbits in the CR3BP 81

3.4 Periodic Orbits in the Entire CR3BP Synodic Reference Frame 81

3.5 Periodic Orbits in the Inertial Reference Frame 82

3.6 L1 Periodic Orbit Unstable Manifold from Perturbations in both the Positive and Negative Direction 83

3.7 Unstable and Stable Manifolds Propagated until Intersection with the Surface of Section 84

3.8 Poincaré Map Showing the Intersection Point between the stable and Unstable Manifold 85

3.9 The Heteroclinic Trajectory Found between the L1 Periodic Orbit and the Distant Prograde Orbit (DPO) 86

3.10 The Family of Orbits Produced from the Continuation Method 89

Figure Page

4.1 Heteroclinic Connection Found by Pseudospectral Method in GPOPS 94

4.2 Control and Hamiltonian for the Heteroclinic Connection Found by GPOPS 95

4.3 Min-Fuel Optimal Trajectories Found by GPOPS for "Nearby" Periodic Orbits of different Jacobi Constants 96

4.4 Optimal Δv Found by GPOPS for "Nearby" Periodic Orbits of different Jacobi Constants 97

4.5 Optimal Trajectories Found in GPOPS by varying the balance of α between Min-Fuel and Min-Time with an initial guess of the heteroclinic trajectory 99

4.6 Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of Minimum-Time 101

4.7 Trade-off between Δv and Transfer Time for Optimal Trajectories Found in GPOPS 102

4.8 (1 of 2) Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Heteroclinic Trajectory for a "Nearby" Orbit 104

4.9 (2 of 2) Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Heteroclinic Trajectory for a "Nearby" Orbit 105

4.10 Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Min-Time Solution for a "Nearby" Orbit 107

4.11 Trade-off between Δv and Transfer Time for Optimal Trajectories Found in GPOPS for a "Nearby" Periodic Orbit 108

4.12 Increasing Collocation Points Results in a New Trajectory for $\alpha = 0.95$ 113

List of Tables

Table		Page
2.1	Characteristic Quantities for Nondimensionalization in the CR3BP for the Earth-Moon System [9]	17
2.2	Location of the Lagrange Points in the Earth-Moon system [10]	33
3.1	Test Plan Overview	72
3.2	Selected Parameter Values	73
3.3	Initial Conditions for L1 Periodic Orbit and DPO	80
4.1	Resulting Time and Δv for Constant Mass and Mass Loss	110

List of Acronyms

2BP Two-Body Problem

3BP Three-Body Problem

ASAT anti-satellite

BVP boundary value problem

CR3BP Circular-Restricted Three-Body Problem

DoD Department of Defense

DPO Distant Prograde Orbit

ECI Earth-Centered Inertial

EOMs equations of motion

ER3BP Elliptic-Restricted Three-Body Problem

GPOPS General Purpose Optimal Control Software

GPS Global Positioning Satellite

IPOPT Interior Point Optimizer

ISR intelligence, surveillance and reconnaissance

KKT Karush-Kuhn-Tucker

LEO Low-Earth Orbits

LG Legendre-Gauss

LGL Legendre-Gauss-Lobatto

LGR Legendre-Gauss-Radau

MATLAB matrix laboratory

NLP nonlinear-programming

SNOPT Sparse Nonlinear Optimizer

SQP Sequential Quadratic Programming

TPBVP Two-Point Boundary Value Problem

OPTIMAL TRAJECTORY GENERATION IN A DYNAMIC MULTI-BODY ENVIRONMENT USING A PSEUDOSPECTRAL METHOD

I. Introduction

1.1 Motivation

1.1.1 Current Issues

Space is an increasingly relevant aspect of both civilian and military sectors. The National Security Space Strategy notes that the space domain has "worldwide services upon which the civil and commercial sectors depend" [1]. There are numerous assets utilized by the civilian sector including the Global Positioning Satellite (GPS), satellite television, and weather satellites. The military sector further employs the space architecture with assets ranging from global communication to space-based intelligence, surveillance and reconnaissance (ISR) [11]. The United States National Security Space Strategy notes that the space environment is becoming "increasingly congested, contested, and competitive [1]".

The number of satellites and amount of debris in orbit demonstrates the congested environment. Satellites and debris are catalogued and tracked by the Department of Defense (DoD). In 2010, the DoD tracked "approximately 22,000 man-made objects" [1]. The vast majority of tracked objects are debris, and it is estimated that there could be as many as "hundreds of thousands of additional pieces of debris" [1] that cannot be tracked due to their size. Figure 1.1 shows the growth of the number of objects cataloged by the DoD. Two major incidents led to the large collection of debris: China's anti-satellite (ASAT) system test and the Iridium-COSMOS colli-

sion. China's ASAT system test occurred on January 2007 and involved destroying a weather satellite [12]. The ASAT system test led to approximately 3,000 new pieces of space debris [1]. The Iridium-COSMOS collision occurred in February 2009 between the communication satellite Iridium and the non-operational Russian COSMOS satellite [13], leading to approximately 1,500 new pieces of debris [1]. These incidents demonstrate the congestion faced by space users and the imminent need for solutions to dealing with the overcrowded environment.

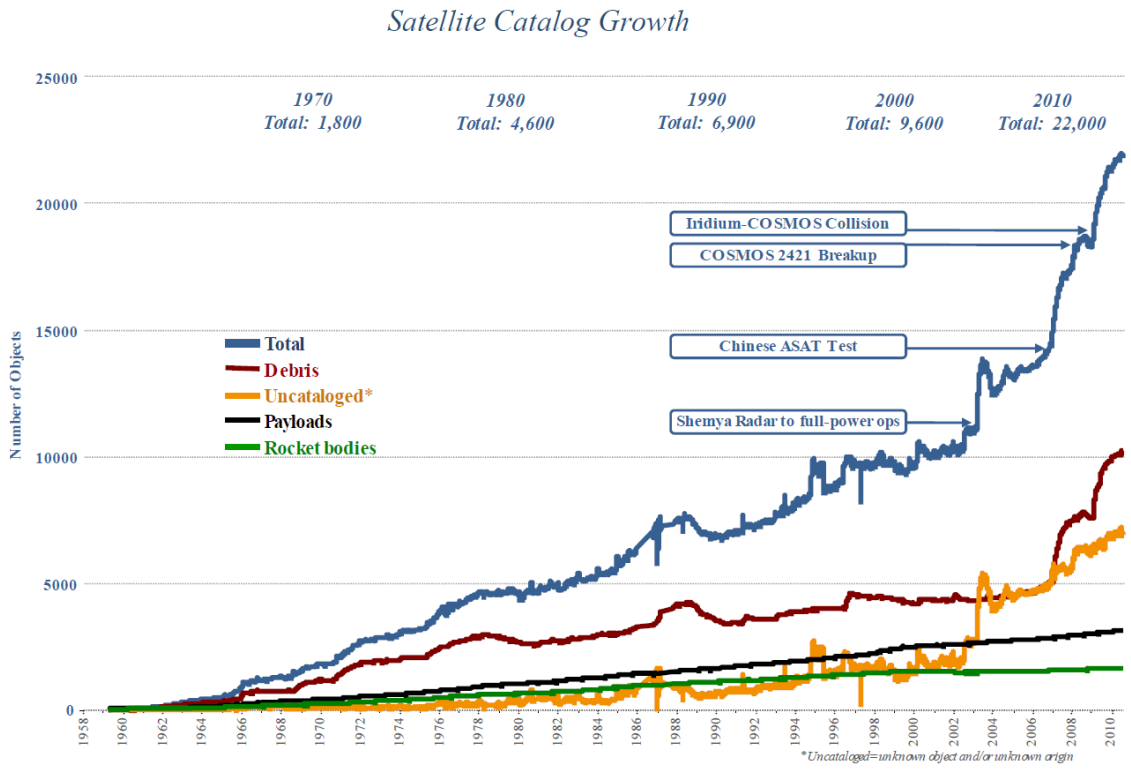


Figure 1.1. Satellite Catalog Growth [1]

The contested environment expressed in the National Security Space Strategy is concerned with potential adversaries threatening the "stability and security" [1] of the space domain. In the Pentagon's 2016 Annual Report to Congress, it is noted that other countries are developing anti-satellite capabilities and further measures to "prevent the use of space-based assets" [14]. As of November 2015, over 70 countries

have operated satellites in space and there are more than 1380 operational satellites [15]. The space domain is heavily contested and is only becoming increasingly utilized by nations and the private sector.

Competition in space can be seen by the demand for highly desirable orbits. One of the desirable orbits is a geostationary orbit. A satellite positioned in a geostationary orbit will orbit the Earth at approximately the same rate as the Earth's rotation. The satellite then appears to be stationary from the Earth. Remaining over a single point on Earth's surface is advantageous because it allows for constant visibility and communication. There is high demand for geostationary orbits and they are utilized by both the private and public sector, demonstrating the high competitiveness of the space domain.

The expense of space missions is another area of importance to commercial enterprises and nations as they try to be as cost-effective as possible. The space domain is inherently expensive. The environment prevents maintenance and repair, so high reliability and survivability is required of space assets, which raises costs. Launching is also an expensive endeavor as extra mass raises the cost of launch, so additional mass is essentially extra cost. For satellites that require orbit maintenance or maneuvers, extra fuel must be brought with the satellite. Extra fuel ultimately results in increased launch cost due to extra mass. Efficient orbit transfers and maintenance are desirable to prevent additional costs from increasing mass.

The United State military has unique priorities and needs beyond those of a civilian space-user. The priorities and goals of the United States military are derived from the National Space Policy. These priorities are to increase mission "assurance and resilience" and "ensure cost-effective survivability of space capabilities" [1]. One method to ensure space domain mission assurance is through reconstitution [16]. Reconstitution is the process of restoring "lost or diminished space capabilities",

such as by "reconfiguring surviving assets" [11].

1.1.2 Benefits of the Circular-Restricted Three-Body Problem (CR3BP)

The current research looks into the unique environment of the CR3BP to address the issues faced by civilian and military space-users from Section 1.1.1. The CR3BP is a model that accounts for the gravitational effects of two primary bodies on a third body. This model is generally used when the gravitational effects of the smaller body are no longer negligible in a two-body approximation. The benefit of using a higher order model is the accuracy of the model can be improved and unique trajectories can be determined. However, the inclusion of a third body complicates the dynamics preventing a closed-form, analytical solution. Instead, other methods can be used to gain insight into the problem, such as applying simplifying assumptions and linearizing the dynamics around equilibrium points. Motion in the CR3BP is chaotic, meaning "small changes in the initial state" can lead to very different final states [10]. However, this chaotic nature produces the unique dynamics that can allow for unconventional trajectories, such as low-cost transfers.

The unique dynamics in the CR3BP could provide solutions to the problems faced in Section 1.1.1. This environment is relatively uncongested, uncontested, and uncompetitive since Low-Earth Orbits (LEO) is where most human space activity occurs [17]. Other than occasional scientific missions, the equilibrium points are not utilized by "spacefaring nations at present" [17].

This domain could provide the ideal location to hold satellites in parking orbits until rapid reconstitution of degraded satellites is needed. The low-cost transfers in the CR3BP would potentially allow a satellite to perform reconstitution in a more rapid and cost-effective manner than a satellite launched from the ground. In addition, the trajectories designed in the CR3BP would not be the conventional trajectories

from a two-body approximation, resulting in unexpected trajectories. Ultimately, the CR3BP provides the decision-maker with more options and freedom to choose between those options. Rather than being limited to choices in the two-body domain, the design space can be expanded to allow for unique options in the CR3BP. The CR3BP domain can increase the space mission assurance and resilience desired by the United States military.

Performing analysis in the CR3BP provides the benefits of the unique dynamics associated with the CR3BP environment. To fully benefit from performing analysis in the CR3BP domain, optimization techniques can be utilized. Generating optimal trajectories in the CR3BP would allow for the unique dynamics and cost-saving in this environment to be exploited.

1.2 Problem Statement

The current research investigates the efficacy of the Legendre-Gauss-Rao pseudospectral method employed in General Purpose Optimal Control Software Version II (GPOPS-II, or GPOPS) to determine viability of the pseudospectral method for optimal low-thrust trajectory design in the CR3BP. The efficacy of the selected pseudospectral method is analyzed by testing in a variety of different settings and evaluating the resulting optimal trajectories. The settings evaluated include: providing different initial guesses, applying a compound objective function between minimum-time and minimum-fuel, and searching at orbits "nearby" to the starting orbit. Testing different scenarios with the pseudospectral method used within GPOPS will help determine the robustness of the optimization method in the CR3BP.

1.3 Thesis Overview

The current research analyzes the efficacy of a pseudospectral method in developing optimal low-thrust trajectories within the CR3BP. The CR3BP can provide unique trajectories that could be beneficial in mission planning, such as high-altitude parking orbits that allow for rapid reconstitution. By developing a method to analyze optimal trajectories in the CR3BP, the unique dynamics can be exploited. By analyzing the efficacy of the pseudospectral method, future work can develop a methodology for optimal mission design within the CR3BP. The background, methodology, results, analysis, and conclusions are organized as follows:

- Chapter 2 provides a background discussion on the fundamentals of the CR3BP and the unique dynamics that can be exploited in the CR3BP environment; optimization fundamentals and a description of the pseudospectral method utilized in the current research; and a literature review of relevant works on the topic.

- Chapter 3 describes a methodology for the research by providing the development of a test plan to analyze the efficacy of the pseudospectral method, a discussion of the problem set-up, a demonstration of the generation of an initial guess, and a development of a comparison metric to the Two-Body Problem (2BP).

- Chapter 4 implements the test plan and evaluates the results. The test plan evaluates and analyzes the following test scenarios: trying different initial guesses, applying a compound objective function between minimum-time and minimum-fuel, and searching at orbits "nearby" to the starting orbit. An evaluation of the resulting trajectories will also be discussed.

- Chapter 5 provides a brief summary of the main conclusions reached in the current research and recommendations for future work will be suggested.

II. Background

2.1 Chapter Overview

A background on Circular-Restricted Three-Body Problem (CR3BP) and optimization methods, specifically pseudospectral methods, is provided to lay the groundwork for the current research. The chapter begins with the derivation of the CR3BP starting from the N-body problem and then discuss features of the CR3BP, such as equilibrium solutions, periodic orbits, and "free"-transfer trajectories. A background on optimization methods is provided and discussion on the specifics of the pseudospectral method used with General Purpose Optimal Control Software (GPOPS) is given. The remainder of the chapter provides a literature review of relevant works.

2.2 N-body Problem

The N-body problem is a general method to calculate the motion of N objects only under the gravitational influence of each of their gravitational fields [2]. The N-body problem assumes each mass is a point mass and then describes the motion of these bodies. Figure 2.1 shows a notional example of the N-body problem, where there are N masses in an inertial reference frame.

The N-body problem can be derived as described in [2]. Begin with Newton's second law

$$\sum_{i=1}^{N} \vec{F}_i = m_i \ddot{\vec{r}}_i, \tag{2.1}$$

where $\vec{F}_i$ is the forces experienced by the i^{th} body, m is the masses of the bodies, and $\ddot{\vec{r}}$ is the inertial acceleration.

The only force is the gravitational attractions between each of the bodies, so for

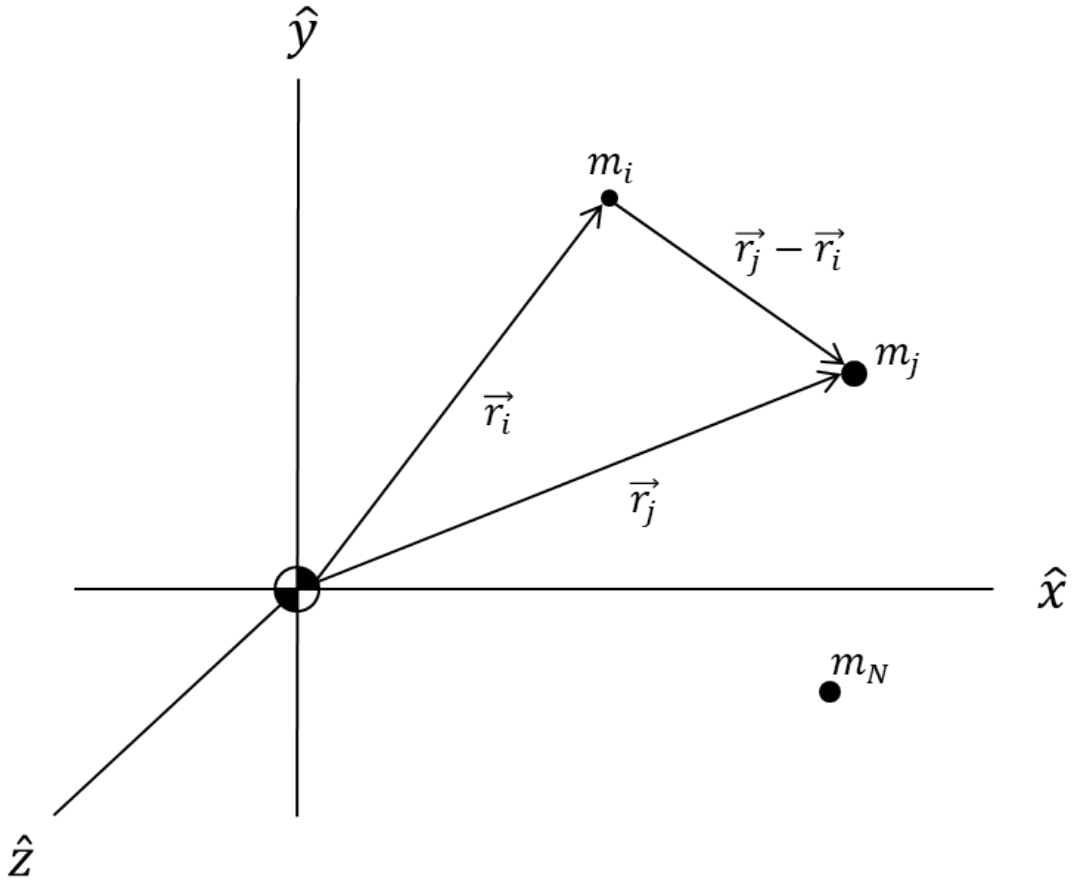


Figure 2.1. The N-body Problem [2]

the i^{th} body, Newton's second law can be written as

$$m_i \ddot{\vec{r}}_i = \sum_{j \neq i}^{N} \frac{G m_i m_j}{r_{ij}^2} \frac{\vec{r}_j - \vec{r}_i}{r_{ij}}. \tag{2.2}$$

Equation (2.2) represents the equations of motion (EOMs) for the N-body problem. Each body has an associated position and velocity $\vec{X} = [\vec{r}, \vec{v}]^T$. Since there are 6 states associated with each body, there are n unknowns in the N-body problem [18], where $n = 6N$.

To solve this problem analytically, all n quantities must be known for all time. However, the N-body problem has some conserved quantities that reduce the required number of values that need to be known to solve the problem. These conserved

quantities are called integrals of the motion, or constants of the motion. If there exists a full set of integrals of the motion, then the problem is completely solved [18]. The N-body problem possesses 10 integrals of the motion: six of these integrals of the motion are found from conservation of linear momentum, three are from conservation of total angular momentum, and one is found from the total energy [18]. With only ten integrals of the motion, there are not enough integrals to even solve a case where $N = 2$ because this would require twelve integrals of the motion [18].

2.2.1 Perturbation Methods

In the N-body problem, the gravitational effects of N bodies are taken into account, but only as point masses, so no other forces are accounted for outside of the point-mass gravitational effects. To improve on this model, other effects would need to be incorporated. These other effects could include drag, the oblateness of the body (i.e., J2), etc., and are incorporated into the model as perturbations. The *fundamental assumption of perturbation theory* is that the perturbing forces are small, so their effects are also small [18]. Incorporating these small perturbing forces helps to more accurately represent the real-world motion of an object.

There are two ways to model perturbations: general and special. General perturbations find an approximate analytical solution in the form of equations of motion by using an infinite series expansion [18, 19]. Special perturbations do not derive an analytical solution, but instead numerically integrate the equations of motion [18]. General perturbations have the benefit of giving insight into the problem because the effects of perturbations can be characterized to determine the corresponding change in the motion. General perturbation methods allow the desired time to be calculated directly without going through intermediate times, such as in numerical integration, which makes answers quick to find [18]. However, general perturbation methods have

the added difficulty of finding an infinite series expansion to approximate the analytical solution, which can be numerically cumbersome. Special perturbations have the drawback of having to use numerical integration, which requires calculating the state of the object at many intermediate time steps between the initial and final time [18]. However, this numerical integration is less difficult to employ than general perturbations [19]. With the improvements in computing power, special perturbations is becoming more feasible and more common to implement [19].

In the current research, third-body effects will be analyzed. However, the *fundamental assumption of perturbation theory* will be violated because the effects of the secondary body cannot be considered small. Thus, perturbation methods cannot be used. Instead, the third-body effects will be directly incorporated into the equations of motion. Since no closed-form analytical solution is available when $N = 3$, numerical integration, the same technique used to solve special perturbations, will be utilized. Before discussing the $N = 3$ case, the two-body problem where $N = 2$ is discussed to provide a simpler introduction to the N-body problem.

2.2.2 Two-Body Problem

The Two-Body Problem (2BP) is the simplest N-body problem where two bodies are only under the gravitational influence of each other. An in-depth derivation of the equations of motion in the 2BP an be found in [2]. By beginning in an inertial frame similar to the N-body problem, the position vectors $\vec{r}_1$ and $\vec{r}_2$ can be defined as the distance from the inertial frame to m_1 and m_2, respectively as seen in Fig. 2.2.

From Newton's second law, the gravitational force for each of these bodies can be written as

$$m_1 \ddot{\vec{r}}_1 = -\frac{G m_1 m_2}{||\vec{r}_1 - \vec{r}_2|\,|^3} (\vec{r}_1 - \vec{r}_2)\,, \tag{2.3}$$

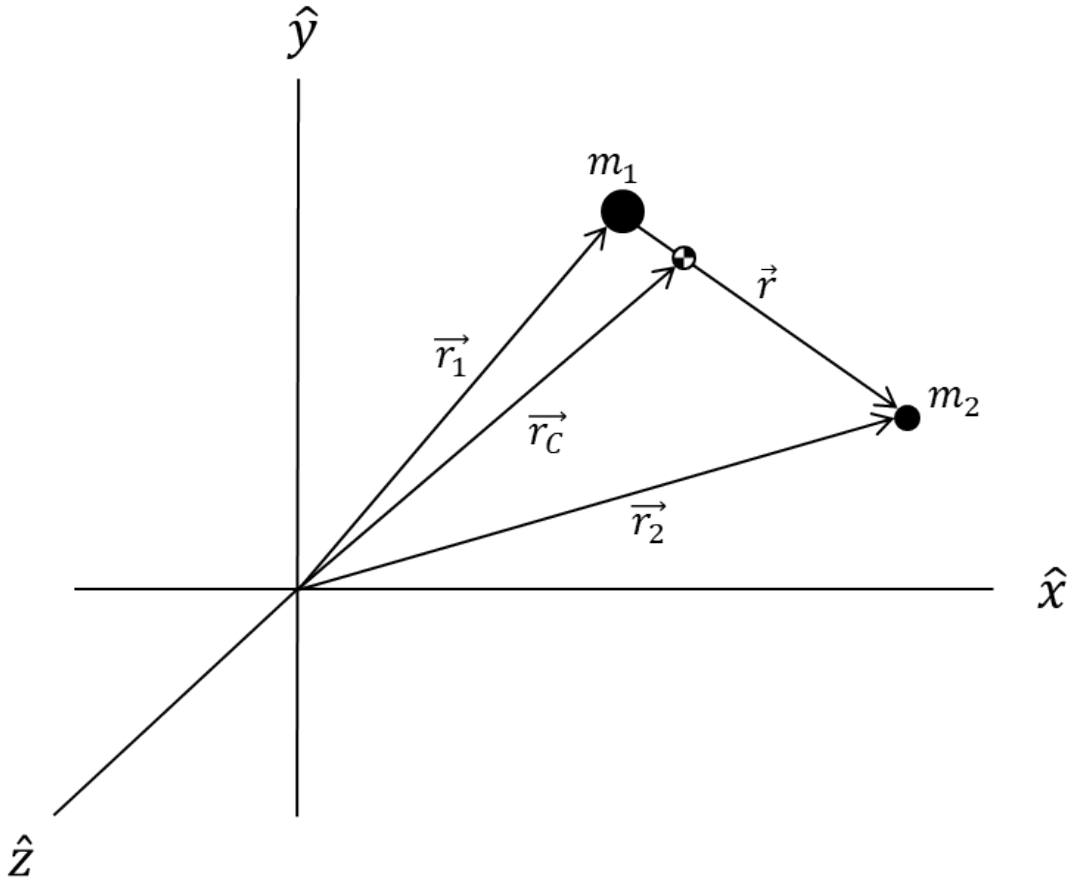


Figure 2.2. The Two-Body Problem [2]

$$m_2\ddot{\vec{r}}_2 = -\frac{Gm_1m_2}{||\vec{r}_1 - \vec{r}_2|\,|^3}\left(\vec{r}_2 - \vec{r}_1\right), \tag{2.4}$$

where G is the universal gravitational constant and is approximately equal to $6.67408*10^{-11}\ \frac{m^3}{kg*s^2}$.

Adding Eq. (2.3) and Eq. (2.4) together gives

$$m_1\ddot{\vec{r}}_1 + m_2\ddot{\vec{r}}_2 = 0. \tag{2.5}$$

Define the vector from the center of the inertial frame to the barycenter of the two masses as $\vec{r}_C$ as seen in Fig. 2.2. $\vec{r}_C$ is written as

$$\vec{r}_C = \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}. \tag{2.6}$$

The second inertial derivative of $\vec{r}_C$ gives

$$\ddot{\vec{r}}_C = \vec{0}. \tag{2.7}$$

Now define the vector from m_1 to m_2 as $\vec{r}$ as seen in Fig. 2.2

$$\vec{r} = \vec{r}_2 - \vec{r}_1. \tag{2.8}$$

Take the second inertial derivative of $\vec{r}$

$$\ddot{\vec{r}} = \ddot{\vec{r}}_2 - \ddot{\vec{r}}_1. \tag{2.9}$$

Next substitute Eq. (2.3) and Eq. (2.4) into Eq. (2.9 resulting in

$$\ddot{\vec{r}} = -\frac{G\left(m_1 + m_2\right)}{r^3}\vec{r}. \tag{2.10}$$

By defining μ as $\mu = G\left(m_1 + m_2\right)$, Eq. (2.10) becomes

$$\ddot{\vec{r}} = -\frac{\mu\vec{r}}{r^3}. \tag{2.11}$$

Equation (2.11) is the EOMs for the 2BP. In the case where the second body's mass is much smaller than the first body's mass, $m_2 \ll m_1$, μ simply becomes $\mu \approx Gm_1$. When this occurs, the larger mass can be treated as an inertial point because $\vec{r}_C \approx \vec{r}$. The EOMs from Eq. (2.11) then simply describe the motion of the second body around the first. Now that the first body is essentially an inertial point, the problem has been simplified from an $N = 2$ problem to an $N = 1$ problem. Therefore, only 6 integrals of the motion are required to solve the problem.

To find the first integral of the motion, take the dot product of Eq. (2.11) with $\dot{\vec{r}}$. After simplification, this results in the "vis-viva" equation [2]

$$\varepsilon = \frac{1}{2}v^2 - \frac{\mu}{r}, \tag{2.12}$$

where ε is the total energy of the system and v is the velocity of the second mass.

Equation (2.12) is the law of conservation of energy. It is the total energy per unit mass of the satellite and is one of the integrals of the motion [2].

To find the next integral of the motion, take a cross product of each side of Eq. (2.11) with $\vec{r}$. After several simplifying steps and an integration, this leads to specific angular momentum $\vec{h}$

$$\vec{r} \times \vec{v} = \vec{h}. \tag{2.13}$$

The specific angular momentum is a vector with three components. Since angular momentum is conserved, this leads to an additional three constants of the motion. To find the final two integrals of the motion, take the cross product of Eq. (2.11) and Eq. (2.13), which leads to

$$\dot{\vec{r}} \times \vec{H} - \mu\frac{\vec{r}}{r} = \mu\vec{e}, \tag{2.14}$$

where $\vec{e}$ is the eccentricity vector and is also a constant of the motion. Although, $\vec{e}$ is also a vector with three components like angular momentum $\vec{H}$, it only accounts for two more integrals of the motion. This is due to the fact that eccentricity must lie perpendicular to the angular momentum vector, which makes the eccentricity vector in-plane [2]. Thus, one of the eccentricity vector's integrals of motions is dependent on a previous integral of the motion.

There are now enough integrals of motion to guarantee that there is a closed-form

solution to the 2BP. The integrals of the motion can be combined into a closed form solution by taking the dot product of Eq. (2.14) and solving for $\vec{r}$ to yield

$$r = \frac{\frac{H^2}{\mu}}{1 + e\cos(\nu)}, \tag{2.15}$$

where ν is called the true anomaly and is the angle between $\vec{r}$ and $\vec{e}$.

Equation (2.15) gives the behavior of $\vec{r}$ in terms of the integrals of motion and ν. It is the "polar form of a conic section with the origin at one focus" [2]. Conic sections include circles, ellipses, parabolas, and hyperbolas. Therefore, the second mass is in an orbit with a conic section shape with the first mass located at one of the foci.

2.3 CR3BP

Another instance of the N-body problem is when $N = 3$, known as the Three-Body Problem (3BP). In the restricted case of the 2BP, the motion of a body of negligible mass is analyzed near a primary body. This analysis is effective only when near a single body. In the vicinity of a single body, such as the Earth, the effects of other bodies can be assumed to be negligible. However, if other bodies' masses are large or if the motion being analyzed is farther away from the primary mass, other bodies' gravitational effects may not be able to be considered negligible. As mentioned earlier, these effects are frequently analyzed by using the 2BP as the baseline model and then incorporating the gravitational effects of other bodies as perturbations. If these effects get too large, though, they cannot be considered perturbations because they violate the fundamental assumption of perturbations. In these cases, a higher fidelity model is needed. In the 3BP, the third body's effects are accounted for directly in the model. An example of this is in the Earth-Moon system. It is assumed that the Earth and the Moon are the only two contributors to the motion of a third body. When a third

body is sufficiently far from Earth (typically considered to be outside geostationary orbit), then the Moon's effects are too large to be modeled as perturbations.

Unlike two-body motion, no closed-form analytical solution has been found. There are only the ten regular constants of the motion and 18 are needed in the 3BP. Other techniques must be used to gain insight into motion in the 3BP. One of these techniques is making simplifying assumptions to reduce the required number of integrals of motion. The simplifying assumptions of the CR3BP as found in Grebow and Szebehely [4, 20] are

1. The third mass is negligible relative to the primary masses (called the primary and secondary mass). Furthermore, in many systems, such as the Earth-Moon system, the secondary mass is less massive than the primary mass (known as the Copenhagen Problem).

$$m_1 > m_2 \gg m_3 \tag{2.16}$$

2. The primary and secondary masses are in a circular orbit about their common barycenter.

The first assumption implies that the third mass has no effect on the motion of the other two masses. Therefore, the motion of the other two masses is Keplerian [4]. This allows for the motion of the orbits to be any of the shapes of the conic section, such as circular or elliptical. However, the second assumption restricts the problem to the circular case. There are other well-known problems, such as the Elliptic-Restricted Three-Body Problem (ER3BP) that allow the orbit to be elliptical instead of circular. However, the ER3BP is outside the scope of the current research. Figure 2.3 shows the CR3BP in its rotating reference frame.

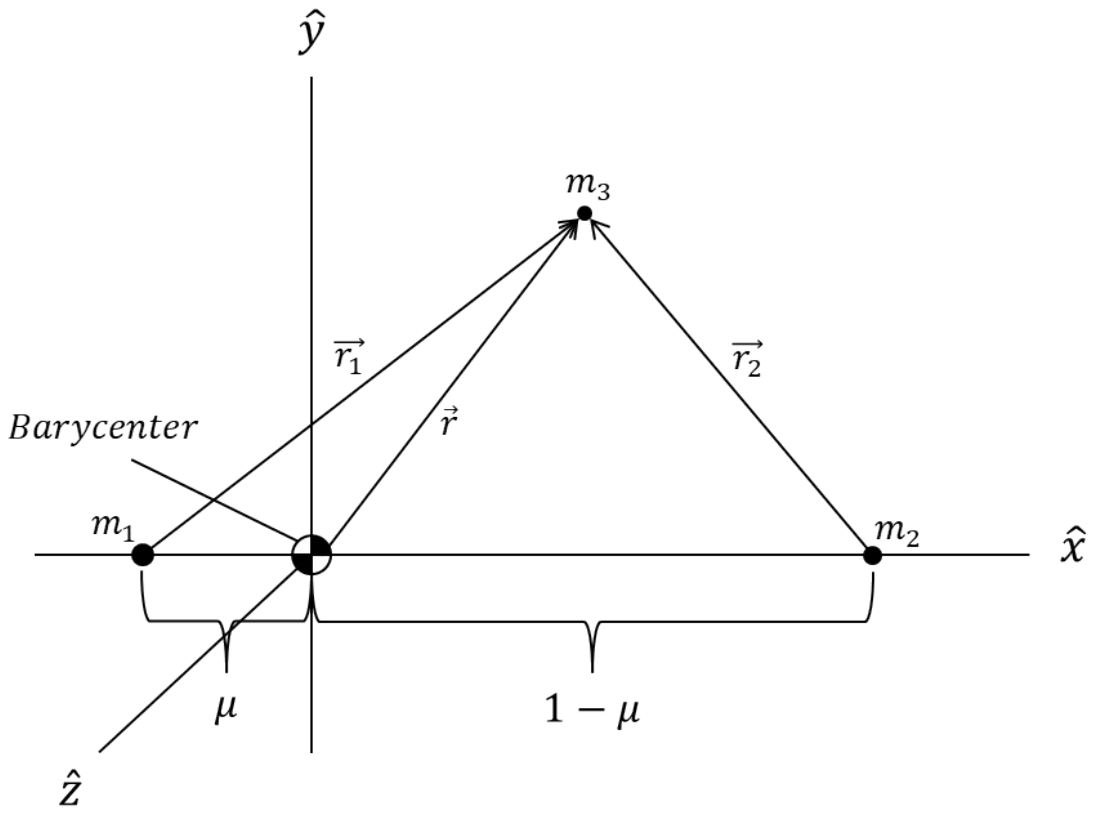


Figure 2.3. The CR3BP in its Rotating Reference Frame

2.3.1 Nondimensionalization

Although not specifically an assumption in the CR3BP, an important generalization is the nondimensionalization of the CR3BP EOMs. Nondimensionalization allows systems of different masses and distances to be directly compared. It also helps improve numerical solution methods, such as numerical integration [10]. By keeping all units the same order of magnitude, the problem avoids numerical scaling issues. The terms that are nondimensionalized are time, mass, and length. These can be best described by introducing characteristic quantities length, mass, and time. A table summarizing the characteristic quantities and their dimensionalized value are given in Table 2.1.

Table 2.1. Characteristic Quantities for Nondimensionalization in the CR3BP for the Earth-Moon System [9]

Symbol	Unit	Definition	Dimensional Value
l^*	Length	Distance between the primaries	384,400 km
m^*	Mass	Total system mass, $m_1 + m_2$	$6.046 \cdot 10^{24}$ kg
t^*	Time	$\sqrt{\frac{{l^*}^3}{Gm^*}}$	4.342 days

The characteristic length, l^*, will be defined as the distance between the two primaries. The length l^* is constant since the motion is assumed to be circular [4].

Characteristic mass, m^*, is defined as the sum of the masses of two primaries,

$$m^* = m_1 + m_2. \tag{2.17}$$

The characteristic time, t^*, is computed as the inverse of the primaries orbital angular velocity [21]

$$t^* = \sqrt{\frac{{l^*}^3}{Gm^*}}. \tag{2.18}$$

This is done so that the gravitational constant, G, equals 1 [2]. The characteristic time also makes it such that one period, T, of the moon around the earth is equal to 2π, shown as

$$T = 2\pi \left(\frac{l^*}{Gm^*}\right)^{\frac{1}{2}} = 2\pi. \tag{2.19}$$

To convert the dimensionalized units to nondimensional units:

$$r_{nondimensional} = \frac{r_{dimensional}}{l^*}, \tag{2.20}$$

$$\mu = \frac{m_2}{m^*}, \tag{2.21}$$

$$t_{nondimensional} = \frac{t_{dimensional}}{t^*}. \tag{2.22}$$

The value μ is a unique variable in Eq. (2.21) because it can describe the nondimensional mass of either primary with the following relation [2]:

$$m_1 = 1 - \mu \tag{2.23}$$

$$m_2 = \mu \tag{2.24}$$

The unique aspect of this variable is, since nondimensional units are being utilized, "anything can be measured in terms of μ" [2]. Even the locations of the primaries can be defined by μ. The more massive primary is located at μ from the barycenter of the system, while the secondary is located at $1 - \mu$, shown in Fig. 2.3.

2.3.2 Derivation of the CR3BP EOMs

The derivation of the CR3BP EOMs follows the derivation laid out by Wiesel [2]. However, Szebehely [20] provides an alternative approach to the derivations by first deriving the EOMs in the sidereal (fixed) reference frame and then converting them to the synodic (rotating) reference frame. Since the mass of the third body is considered negligible, the motion of the two primaries is the same EOMs as the 2BP. Thus, only the motion of the third body will be evaluated. The synodic reference frame used in this derivation is shown in Fig. 2.3. Similar to the 2BP we begin with Newton's second law

$$\sum_{i=1}^{2} \vec{F}_i = m_3 \ddot{\vec{r}}, \tag{2.25}$$

where $\vec{F}_i$ are the forces acting on the third body, m_3 is the mass of the third body, and $\ddot{\vec{r}}$ is the acceleration of the third body. The summation only goes to N=2 because the two primaries are the only two bodies that have any effect on the motion of the third-body.

To solve for $\ddot{\vec{r}}$, the position of the third body in the synodic reference frame is given as

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}. \tag{2.26}$$

The second derivative of this equation in time with respect to the inertial frames gives the acceleration, which is required to apply Newton's second Law. However, $\vec{r}$ is given in the synodic frame, so the transport theorem must be applied to take the inertial derivative of $\vec{r}$ in the synodic reference frame. The transport theorem is a way to take the derivative with respect to the inertial frame of a vector given in non-inertial reference frame coordinates. Details of the transport theorem are described by Kunz [22]. The first derivative of $\vec{r}$ is given as

$${}^{i}\frac{d}{dt}\vec{r} = {}^{s}\frac{d}{dt}\vec{r} + \vec{\omega}^{si} \times \vec{r}. \tag{2.27}$$

The second derivative is

$${}^{i}\frac{d^2}{dt^2}\vec{r} = {}^{s}\frac{d^2}{dt^2}\vec{r} + 2\vec{\omega}^{si} \times {}^{s}\frac{d}{dt}\vec{r} + \vec{\omega}^{si} \times \left(\vec{\omega}^{si} \times \vec{r}\right), \tag{2.28}$$

where the superscript before the derivative is the reference frame the derivative is taken with respect to (i being the inertial frame and s being the synodic reference frame) and $\vec{\omega}^{si}$ is the angular velocity of the synodic reference frame with respect to

the inertial frame. Since the rotational period of the synodic reference frame is 2π TU, $\vec{\omega}^{si} = 1$ [2]. Thus, the only unknown components in Eq. (2.28) are ${}^{s}\frac{d}{dt}\vec{r}$ and ${}^{s}\frac{d^2}{dt^2}\vec{r}$, which can be simply calculated by taking the derivative in time with respect to the synodic reference frame of $\vec{r}$ as follows

$$ {}^{s}\frac{d}{dt}\vec{r} = \dot{x}\hat{x} + \dot{y}\hat{y} + \dot{z}\hat{z}, \tag{2.29} $$

$$ {}^{s}\frac{d^2}{dt^2}\vec{r} = \ddot{x}\hat{x} + \ddot{y}\hat{y} + \ddot{z}\hat{z}. \tag{2.30} $$

After simplifying Eq. (2.28), it becomes

$$ {}^{i}\frac{d^2}{dt^2}\vec{r} = (\ddot{x} - 2\dot{y} - x)\,\hat{x} + (\ddot{y} - 2\dot{x} - y)\,\hat{y} + \ddot{z}\hat{z}. \tag{2.31} $$

Now that $\ddot{\vec{r}_3}$ is found, the gravitational force $\vec{F}_i$ acting on the third body needs to be calculated. The force of gravity can be written as

$$ \vec{F}_g = -\frac{Gm_a m_b}{r^2}\frac{\vec{r}}{r}. \tag{2.32} $$

In the CR3BP, $G = 1$, so Eq. (2.32) can be simplified by substituting $G = 1$. As can be seen in Eq. (2.32), gravitational force is an inverse square law, so the distance between the primary bodies and the third body must be determine. The distance between the primary bodies and the third body is shown in Fig. 2.3 and can be written as

$$ r_1 = \left[(x+\mu)^2 + y^2 + z^2\right]^{1/2} \tag{2.33} $$

and

$$ r_2 = \left[(x-1+\mu)^2 + y^2 + z^2\right]^{1/2}, \tag{2.34} $$

where r_1 is the distance between the primary body and the third body and r_2 is the distance between the secondary body and the third body.

With the distance defined, the sum of the forces acting on the third body can be calculated

$$\vec{F} = -m_3 \left(\frac{(1-\mu)\vec{r_1}}{r_1^3} + \frac{\mu \vec{r_2}}{r_2^3} \right). \tag{2.35}$$

Now that the force and acceleration are known, they can be equated from Newton's second law in Eq. (2.25). The assumption that $m_3 \approx 0$ is intended to be representative of a limit as $m_3 \to 0$ [2]. The assumption is more relevant in ensuring the objects of the two primary bodies are unaffected by the third body's mass, so it is still acceptable to divide the m_3 from both sides of the equation [2]. Breaking Eq. (2.25) into its three components gives

$$\ddot{x} - 2\dot{y} - x = -\frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+mu)}{r_2^3} \tag{2.36}$$

$$\ddot{y} + 2\dot{x} - y = -\frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} \tag{2.37}$$

$$\ddot{z} = -\frac{(1-\mu)z}{r_1^3} - \frac{\mu z}{r_2^3}. \tag{2.38}$$

Equations (2.36) - (2.38) are the EOMs for the nondimensionalized CR3BP. Interestingly, while these equations are three-dimensional, if the third body does not have any components or motion in the $\hat{z}$-direction, then the motion of the object will remain in the $\hat{x}$-$\hat{y}$ plane for all time, since $\ddot{z} = 0$. This situation is called the planar CR3BP. If there is motion in the $\hat{z}$-direction, then the problem is called the spatial CR3BP.

These equations of motion are frequently written in an alternative form involving a pseudopotential U^* [10], where

$$U^* = \frac{1}{2}\left(x^2 + y^2\right) + \frac{1-\mu}{r_1} + \frac{\mu}{r_2}, \tag{2.39}$$

and thus, the equations of motion become

$$\ddot{x} = 2\dot{y} + U_x^* \tag{2.40}$$

$$\ddot{y} = -2\dot{x} + U_y^* \tag{2.41}$$

$$\ddot{z} = U_z^*, \tag{2.42}$$

where $U_x^* = \frac{\partial U^*}{\partial x}$, $U_y^* = \frac{\partial U^*}{\partial y}$, and $U_z^* = \frac{\partial U^*}{\partial z}$.

2.3.3 Transformation for Barycentric Synodic Reference Frame to Earth-Centered Inertial (ECI)

Although the synodic frame is useful for analyzing trajectories within the CR3BP, it is frequently beneficial to view those trajectories in an inertial reference frame. A common inertial reference frame is the ECI frame. The ECI frame has the center of the reference frame at the center of the Earth and the reference frame is nonrotating [23]. The coordinate frame is defined by the $\hat{x}$-axis being pointed toward the vernal equinox, the $\hat{z}$-axis is about Earth's rotation axis (perpendicular to the equatorial plane), and the $\hat{y}$-axis simply completes the right-handed reference frame [23]. The vernal equinox is the line that meets at the intersection of the Earth's equatorial planes and the plane of the Earth-Sun [23]. The vernal equinox is a line that is assumed to be fixed or close enough to fixed to have the coordinate system be considered inertial. Figure 2.4 demonstrates the relationship between the synodic reference frame and the ECI. The subscript I represent the inertial frame, and the subscript s represents the synodic frame. The z component of both reference frames is out of the page. The variable θ corresponds to dimensionless time in the CR3BP [24].

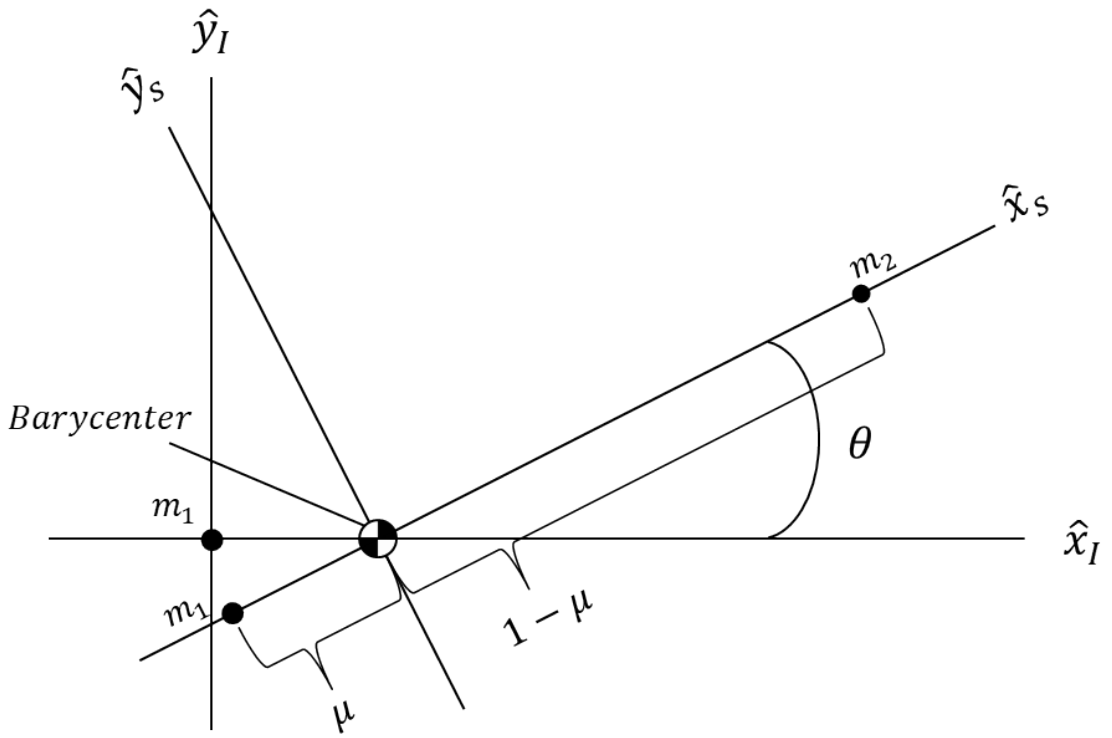


Figure 2.4. Synodic and ECI reference frames

The description of the transformation follows the steps in [24]. Begin the transformation by defining the state of the system to be transformed

$$\vec{X}_s = \begin{bmatrix} \vec{R}_s \\ \vec{V}_s \end{bmatrix}, \tag{2.43}$$

where $\vec{R}_s$ and $\vec{V}_s$ are the synodic frame position and velocity of the point to be transformed. In component form, they are

$$\vec{R}_s = [x_s, y_s, z_s]^T \tag{2.44}$$

$$\vec{V}_s = [\dot{x}_s, \dot{y}_s, \dot{z}_s]^T . \tag{2.45}$$

As seen in Fig. 2.4, the $\hat{z}$-axis is in-line in both reference frames. However, the

reference frames are offset, since one is focused at the barycenter of the system and the other is centered at the primary body. The x_s coordinates should be translated before converting reference frames by altering them in the $\hat{x}$ direction [24]

$$x_{s'} = x_s + \mu, \tag{2.46}$$

where $x_{s'}$ is the x_s coordinate translated to have the Earth at the center of the reference frame. The other two position coordinates,y_s and z_s, are unaffected by this translation, so the new position $\vec{R_{s'}}$ is

$$\vec{R_{s'}} = [x_{s'}, y_s, z_s]^T . \tag{2.47}$$

The equation to rotate from the synodic frame to the inertial frame is given as

$$\vec{R_I} = Q^{IS}\vec{R}_{s'} \tag{2.48}$$

$$\vec{V_I} = Q^{IS}\vec{V_s} + \dot{Q}^{IS}\vec{R}_{s'}, \tag{2.49}$$

where Q^{IS} is the rotation matrix from the synodic frame to the inertial frame. The rotation matrix is defined in [22] and allows for the transformation from one coordinate system to another. The matrix Q^{IS} is given as

$$Q^{IS} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{2.50}$$

Taking the time derivative gives

$$\dot{Q}^{IS} = \dot{\theta} \begin{bmatrix} -\sin(\theta) & -\cos(\theta) & 0 \\ \cos(\theta) & -\sin(\theta) & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.51}$$

In the CR3BP, $\dot{\theta} = 1$ and θ may be treated as dimensionless time [24]. The assumption will be made that $\theta_0 = 0$, so that there does not need to be an initial angular offset.

Now that the dimensionless position and velocity have been converted to the inertial reference frame, they now need to be changed to dimensional units. This can be done by scaling by the characteristic units as follows

$$\vec{R}_{I'} = \vec{R}_I l^* \tag{2.52}$$

$$\vec{V}_{I'} = \vec{V}_I \frac{l^*}{t^*}, \tag{2.53}$$

where $\vec{R}_{I'}$ and $\vec{V}_{I'}$ are the dimensionalized position and velocity, $\vec{R}_I$ and $\vec{V}_I$ are the nondimensional position and velocity, and l^* and t^* are the characteristic length and characteristic time, respectively.

2.4 Characteristics of the CR3BP

2.4.1 Symmetries in the CR3BP

The EOMs in the CR3BP exhibit two symmetries. The first symmetry is with the x-z plane and time [20]. In simple terms, reflections over the x-z plane are backwards in time. When the state $[x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t)]^T$ is reflected over the x-z plane, another solution is found, $[x(-t), -y(-t), z(-t), -\dot{x}(-t), \dot{y}(-t), -\dot{z}(-t)]^T$ [10, 3]. This symmetry is useful in finding periodic orbits because an orbit that intersects

the x-z plane will have a mirrored orbit that flows in the opposite direction [20]. While this is not enough to find a periodic orbit it is an essential aspect of the process and will be discussed further in Section 2.6. An example of this symmetry can be seen in Fig. 2.5. In this image, the trajectory is reflected over the x-z plane. It is important to note the direction of travel of both orbits: the initial point on the "original" trajectory is the final point in the "symmetric trajectory". This demonstrates the symmetry in time, as well as symmetry in the x-z plane.

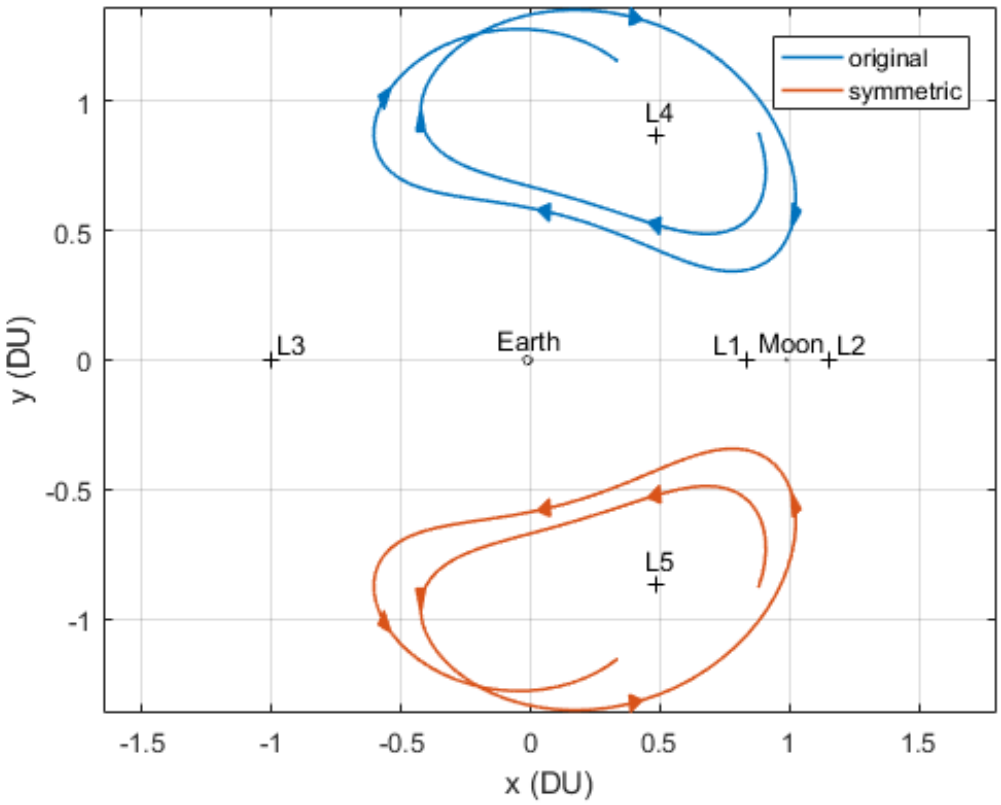


Figure 2.5. Symmetry of CR3BP across the x-z plane and time

The other symmetry in the CR3BP EOMs is over the x-y plane [3]. This symmetry comes from the motion in the $\hat{z}$-direction being decoupled from the motion in the $\hat{x}$- and $\hat{y}$-directions in the EOMs. This symmetry is described as $[x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t)]^T$ is mirrored with $[x(t), y(t), -z(t), \dot{x}(t), \dot{y}(t), -\dot{z}(t)]^T$ [10, 3]. This symmetry is out-of-

plane, so it is not the plane of the primaries' rotation. This symmetry is shown in Fig. 2.6. Since the symmetry is across the x-y plane, the y-z plane is shown. There is symmetry as the trajectory moves toward and away from the plane of the primaries.

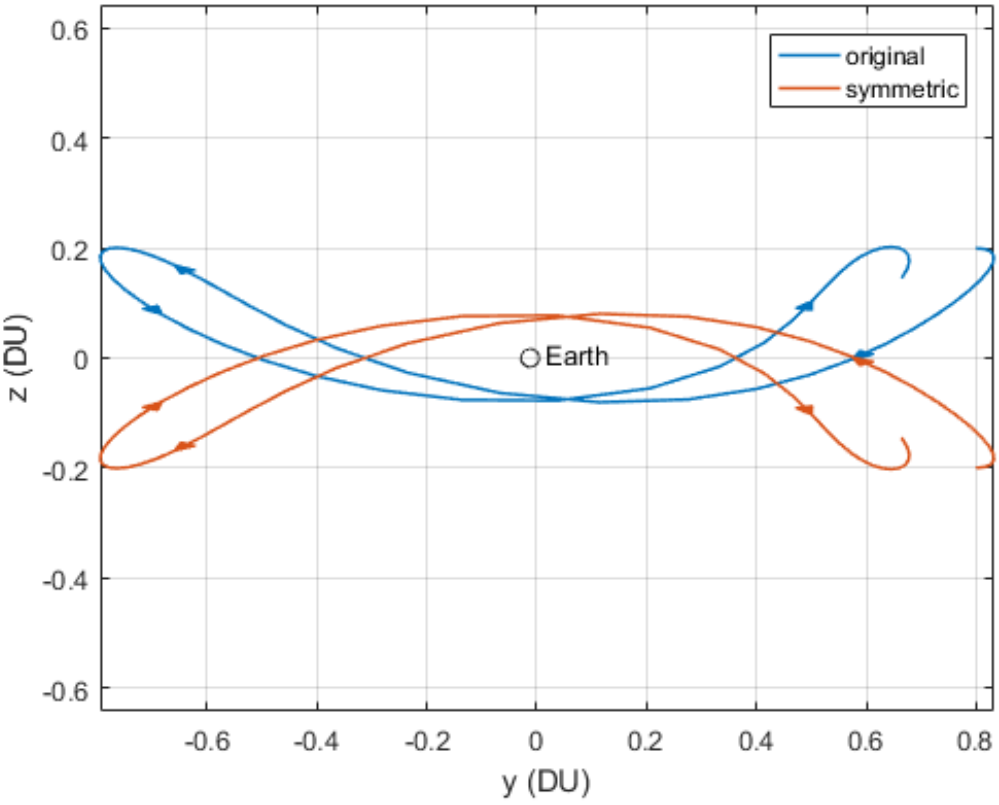


Figure 2.6. Symmetry of CR3BP across the x-y plane

2.4.2 Jacobi Constant

Since the CR3BP is a specific case of the N-body problem, it still has the standard integrals of motion: total linear momentum, total energy, and total angular momentum. However, these conserved quantities all deal with the two primaries, since the third body's mass is considered negligible [2]. In order to solve the CR3BP, six integrals of the motion would need to exist, two for each second-order EOMs [2]. As discussed earlier, this problem has no closed form solution, so six integrals of the

motion have not been found, but one does exist. The one constant of the motion in the CR3BP is called the Jacobi Constant, or Jacobi's Integral.

To find the Jacobi Constant, the derivation from [2] will be followed. The process is done by looking for energy conservation, but in a rotating frame [2]. Take the dot product of the EOMs from Eq. (2.36), (2.37), and (2.38) in vector form with $\dot{\vec{r}}$, which can be written as $\ddot{\vec{r}} \cdot \dot{\vec{r}}$. This results in

$$\ddot{x}\dot{x}+\ddot{y}\dot{y}+\ddot{z}\dot{z}-x\dot{x}-y\dot{y} = -\frac{(1-\mu)}{r_1^3}\left((x+\mu)\dot{x}+y\dot{y}+z\dot{z}\right)-\frac{\mu}{r_2^3}\left((x-1+mu)\dot{x}+y\dot{y}+z\dot{z}\right). \tag{2.54}$$

It can be recognized now that each side is a perfect time derivative of two different quantities as follows

$$\frac{d}{dt}\left(\frac{1}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)\right)-\frac{d}{dt}\left(\frac{1}{2}\left(x^2+y^2\right)\right)=\frac{d}{dt}\left(\frac{1-\mu}{r_1}\right)+\frac{d}{dt}\left(\frac{\mu}{r_2}\right). \tag{2.55}$$

Thus, Jacob's Integral is

$$JC=\frac{1}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)-\frac{1}{2}\left(x^2+y^2\right)-\frac{1-\mu}{r_1}-\frac{\mu}{r_2}. \tag{2.56}$$

Using Eq. (2.39) and simplifying $\dot{x}^2+\dot{y}^2+\dot{z}^2$ to v^2 this can be written in pseudopotential form as

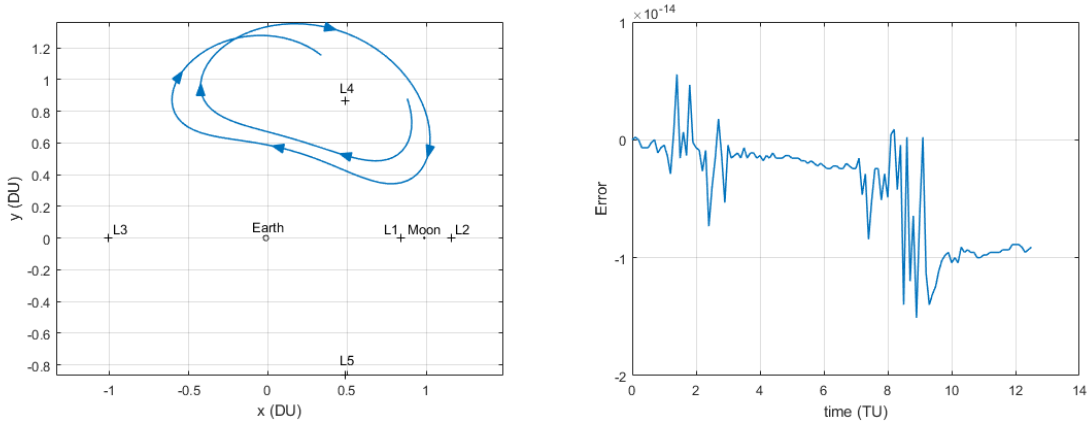
$$JC=U^*-\frac{1}{2}v^2, \tag{2.57}$$

where U^* is the pseudopotential defined in Eq. (2.39).

The Jacobi Integral has units of energy, but this energy is not necessarily "the total energy of the system measured relative to an inertial frame" [25]. Instead, the Jacobi constant is an energy-like quantity [10]. The velocity term is not in the

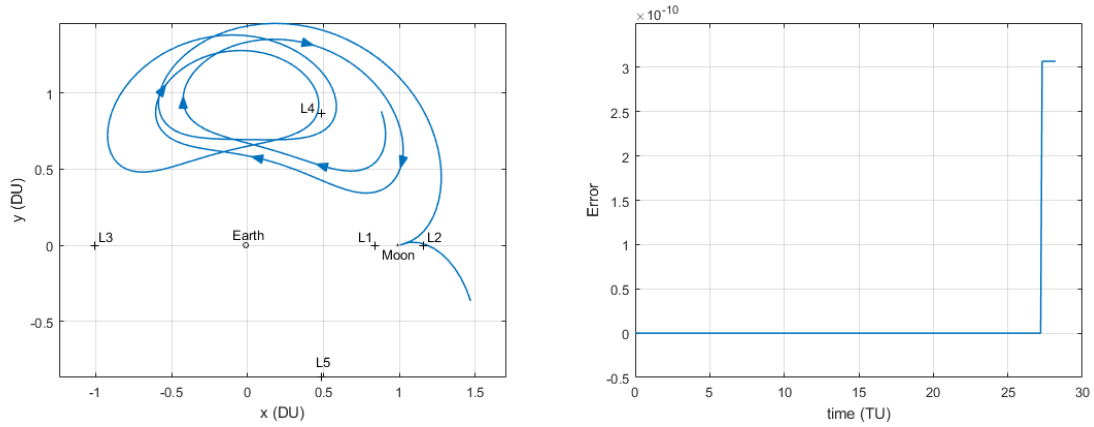
inertial frame, but rather the rotating frame. If converted back to the inertial frame, the Jacobi constant would be a "combination of the total energy of the third mass and its total angular momentum" [2].

Since an analytical solution cannot be produced, the Jacobi constant is frequently used to check the accuracy of the numerical integration [10]. Since it is a constant of the motion, it should remain constant throughout a trajectory. Therefore, when numerically integrating a trajectory, the Jacobi Constant can be evaluated at each time step to ensure that it is constant. Since numerical integration inherently has some error, the Jacobi constant can never remain perfectly constant, but Fig. 2.7 demonstrates how constant Jacobi's Integral remains in an acceptable numerical integration procedure. Note in Fig. 2.7b that the error is of the order 10^{-14}, which is the same order as the tolerance used in the numerical integration. Figure 2.8 demonstrates what happens to Jacobi's Integral when a trajectory follows an invalid trajectory, such as going through a numerically unstable point at the second body. When the trajectory passes through the secondary body, the Jacobi constant spikes as seen in Fig. 2.8b.



(a) Numerically Integrated Trajectory **(b) Jacobi Constant Error**

Figure 2.7. Valid Trajectory: Jacobi Constant used to check validity of numerical integration



(a) Numerically Integrated Trajectory **(b) Jacobi Constant Error**

Figure 2.8. Invalid trajectory (through the secondary body): Jacobi Constant demonstrates the invalidity of the numerical integration

2.5 Equilibrium Solutions

While no closed-from analytical solution exists for the CR3BP, five equilibrium solutions are present. An equilibrium point is a stationary location where the forces in the system balance each other. This results in no velocity or acceleration on a particle located at an equilibrium point [2]. To solve for the equilibrium points, the accelerations $\ddot{\vec{r}}$ and velocities $\dot{\vec{r}}$ in the EOMs from Eq. (2.36), (2.37), and (2.38) must be set equal to zero [2],

$$x = \frac{(1-\mu)(x+\mu)}{r_1^3} + \frac{\mu(x-1+\mu)}{r_2^3} \tag{2.58}$$

$$y = \frac{(1-\mu)y}{r_1^3} + \frac{\mu y}{r_2^3} \tag{2.59}$$

$$0 = -\frac{(1-\mu)z}{r_1^3} - \frac{\mu z}{r_2^3}. \tag{2.60}$$

Equation (2.60) quickly demonstrates that $z = 0$ to find an equilibrium solution.

This simplifies the equilibrium points to entirely in the plane. Now x and y must be solved for from Eq. (2.58) and (2.59). Leonhard Euler discovered the first three equilibrium points in 1765 [10]. They can be found using a root-finding technique, such a the Newton-Raphson method. These three roots are all collinear lying on the $\hat{x}$-axis. The first equilibrium point lies between the primary and secondary body; the second equilibrium point lies beyond the secondary body; and the third equilibrium point lies on the far side the primary body, opposite the secondary mass [3] called L_1, L_2, and L_3, respectively.

The final two equilibrium points were discovered by Lagrange in 1772 [20]. This is where the equilibrium points get one of their alternative names: Lagrange points. The final two Lagrange points, called L_4 and L_5, are found by setting $r_1 = r_2 = 1$. While this appears to just cancel out all the variables in Eq. (2.58) and (2.59), it actually contains all the necessary information to find the final Lagrange points [2]. This is because the points lie at points on an equilateral triangles, with two vertices being the primary and secondary body and the final being the Lagrange point. The location of the final two points is at $x = \mu - \frac{1}{2}$ and $y = \pm\frac{\sqrt{3}}{2}$, where the positive sign corresponds to the fourth Lagrange point and the negative sign corresponds to the fifth Lagrange point [20, 3]. Figure 2.9 shows all five Lagrange points location in the Earth-Moon system.

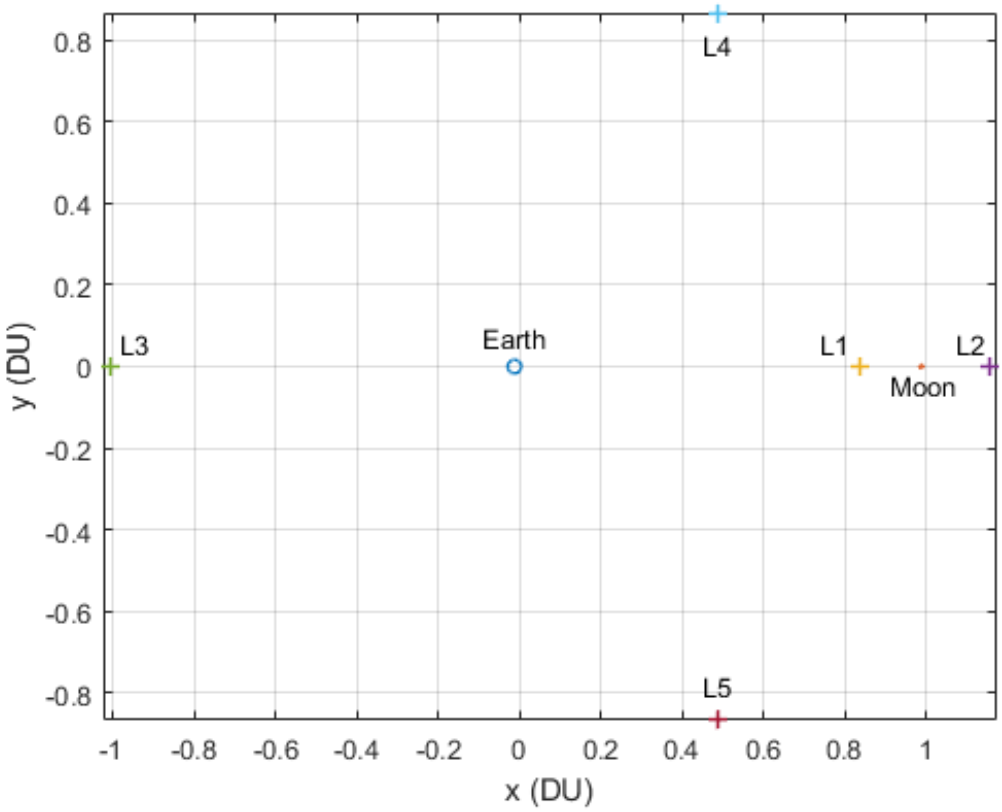


Figure 2.9. Lagrange Points in Earth-Moon System

The location of the Lagrange points in the Earth-Moon system can be seen in Table 2.2.

Table 2.2. Location of the Lagrange Points in the Earth-Moon system [10]

Lagrange Point	X	Y	Units
L_1	0.836915121142417	0	Nondimensional
	321710.172567150	0	Kilometers
L_2	1.155682169063842	0	Nondimensional
	444244.2257881407	0	Kilometers
L_3	-1.005062646202315	0	Nondimensional
	-386346.0812001700	0	Kilometers
L_4	0.487849413449431	0.866025403784439	Nondimensional
	187529.3145299614	332900.1652147382	Kilometers
L_5	0.487849413449431	-0.866025403784439	Nondimensional
	187529.3145299614	-332900.1652147382	Kilometers

The Earth-Moon system has a $\mu \approx 0.01215$ [3]. However, in a different system, the μ-value will change. For example, in the Sun-Jupiter system $\mu \approx 9.53e-4$ and in the Sun-Earth system $\mu \approx 3.0e-6$ [3]. The value of μ in each system affects the location of the Lagrange points. As the μ value changes, a distinct path is followed by the Lagrange points. Figure 2.10 demonstrates the location of the Lagrange points for different values of μ. Notice that the planar Lagrange points follow a distinct S-shaped curve with L_1 and L_2 coalescing as $\mu \to 0$ and L_1 and L_3 coalescing as $\mu \to 1$ [3]. L_4 and L_5 are in straight lines that connect at points that lie on the circle created by L_1 coalescing with L_2 and L_3 [3].

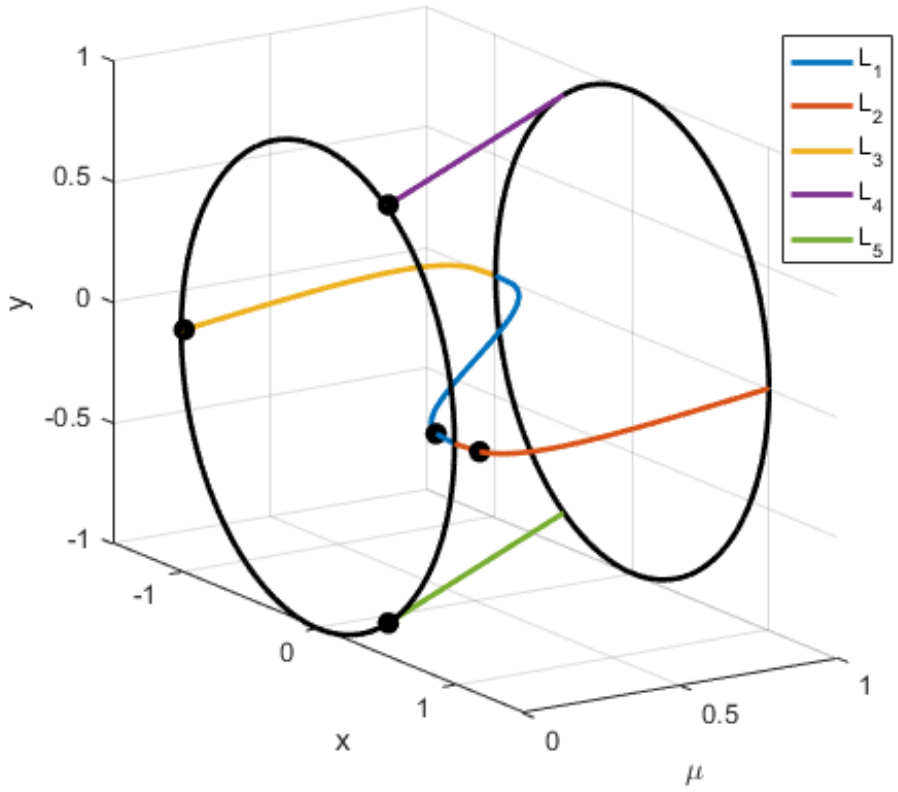


Figure 2.10. Lagrange Points as a function of μ. The current location of the points is at $\mu \approx 0.01215$ to represent the Earth-Moon System (Adapted from Doedel et al. [3])

2.6 Periodic Orbits

Determining how a trajectory behaves over a long period of time, such as $t \to \infty$, is not possible for many trajectories in problems that require numerical methods to solve [20]. However, several special cases allow insight into the trajectory as $t \to \infty$. These special cases are "asymptotic, periodic, or almost periodic" trajectories [20]. Periodic motion is when the "same configuration is repeated at regular intervals of time" [20]. When the trajectories are known for a long period of time, analysis can be done to determine general behavior of the problem. In the case of CR3BP, families of orbits can be generated with the periodic property [20].

While periodic orbits in the CR3BP are not currently being utilized by the mil-

itary, they could be exploited in the future. Periodic orbits could offer potential military applications as a location to be used as a parking orbit. Satellites could be held in these repeating trajectories until needed for a desired mission, such as reconstitution.

2.6.1 Equations of Variation

To generate periodic orbits, a boundary value problem (BVP) must be solved. In a BVP, a final position $\vec{r}_f$ is sought from some initial guess by numerically integrating the EOMs in the form

$$\dot{\vec{x}} = \vec{f}(\vec{x}, t). \tag{2.61}$$

Unless the desired trajectory is already known, the initial guess will not follow the desired trajectory exactly [18] as seen in Fig. 2.11. A differential corrections process can be used to determine the initial conditions. The differentials corrections process involves finding information about all nearby orbits and the influence of $\vec{x}_0$ on $\vec{x}_f$ [18], so that the initial guess can be updated in a logical manner.

The differential corrections process will be described following the explanation from [2]. Begin by selecting a set of initial conditions $\vec{x}_0(t_0)$ that, when numerically integrating Eq. (2.61), gives the trajectory $\vec{x}_0(t)$. A general nearby trajectory would be described as

$$\vec{x}(t) = \vec{x}_0(t) + \delta\vec{x}(t), \tag{2.62}$$

where the $\delta\vec{x}(t)$ is a small variation away from the orbit $\vec{x}_0(t)$.

Substituting Eq. (2.62) into Eq. (2.61) gives

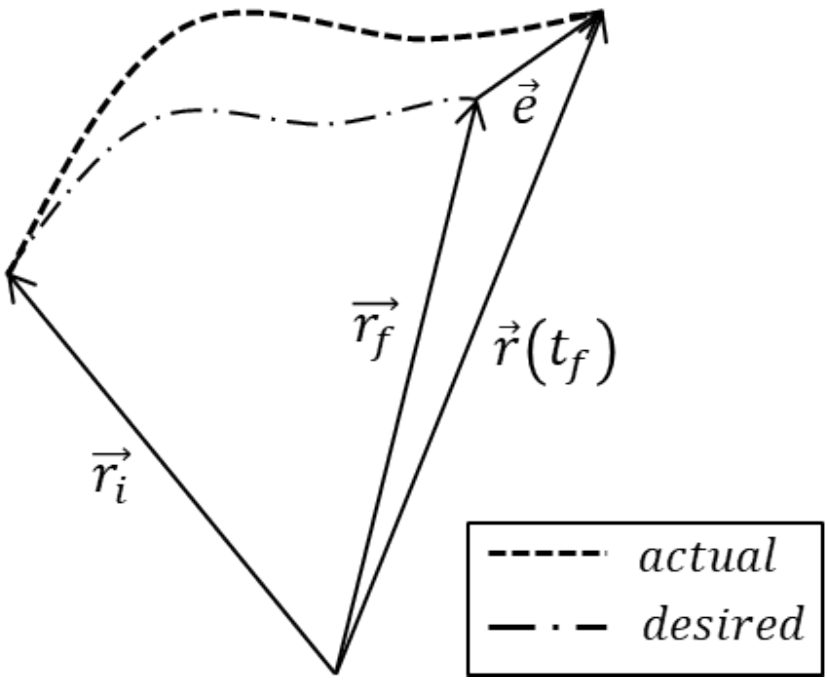


Figure 2.11. Boundary Value Problem Example

$$\dot{\vec{x}}_0 + \delta\dot{\vec{x}} = \vec{f}(\vec{x}_0 + \delta\vec{x}, t). \tag{2.63}$$

Expanding this equation in a first-order Taylor series about $\vec{x} = 0$ [3, 18] gives

$$\dot{\vec{x}}_0 + \delta\dot{\vec{x}} \approx \vec{f}(\vec{x}_0, t) + \left.\frac{\partial\vec{f}}{\partial\vec{x}}\right|_{\vec{x}_0} \delta\vec{x}. \tag{2.64}$$

After simplifying, the derivation evolves as

$$\delta\dot{\vec{x}} = \left.\frac{\partial\vec{f}}{\partial\vec{x}}\right|_{\vec{x}_0} \delta\vec{x} = A(t)\delta\vec{x}. \tag{2.65}$$

Equation (2.65) is known as the first-order equations of variation [18, 4]. $\frac{\partial\vec{f}}{\partial\vec{x}}$ is a square matrix of partial derivatives called $A(t)$. $A(t)$ is found by taking the partial derivative of the the EOMs with respect to the states.

To find the general solution to the equations of variation in Eq. (2.65), the state

transition matrix must be introduced. The state transition matrix is used to predict variations along a path [4]. The state transition matrix provides information about the sensitivity of the final conditions to changes in the initial conditions [10]. It must satisfy the the differential equation

$$\dot{\Phi}(t, t_0) = A(t)\Phi(t, t_0) \tag{2.66}$$

with initial conditions

$$\Phi(t_0, t_0) = I, \tag{2.67}$$

where I is the identity matrix. Equation (2.66) is an alternate form of the equations of variation [18].

The state transition matrix obeys the following identities [18]:

$$\Phi(t_2, t_0) = \Phi(t_2, t_1)\Phi(t_1, t_0) \tag{2.68}$$

$$\Phi(t_0, t_1) = \Phi^{-1}(t_1, t_0). \tag{2.69}$$

With the state transition defined, a general solution to the equations of variation in Eq. (2.65) can be found as

$$\delta\vec{x}(t) = \Phi(t, t_0)\delta\vec{x}(t_0). \tag{2.70}$$

Thus, by numerically integrating both the equations of motion in Eq. (2.61) and the equations of variation in Eq. (2.66) in parallel, everything about the trajectories nearby to our original trajectory x_0 is known [18]. This means changes to our initial conditions can be made through Eq. (2.70) to arrive at a desired trajectory. Numerically integrating the equations of motion and the equations of variation can

be computationally expensive. There are N^2 equations of variation in Eq. (2.66) and N EOMs, so $N^2 + N$ equations must be numerically integrated simultaneously [2]. For the CR3BP, there are 42 equations that must be numerically integrated simultaneously.

The problem is now a BVP, as seen in Fig. 2.11. Common methods of solving boundary value problems include single-shooting and multiple-shooting methods. In a shooting method, initial conditions are selected and propagated forward in time. A correction is made on the initial condition to allow the propagated trajectory to approach the desired trajectory within a specified tolerance. When the state transition matrix is used to correct the initial conditions, this process is known as differential corrections. Since Eq. (2.65) is a linear first-order approximation, more than one iteration may be required to converge on the desired trajectory.

2.6.2 Differential Corrections Applied to the CR3BP

A detailed derivation of the variational equations of motion specific to the CR3BP can be found from Szebehely and Brick [20, 10]. To find the equations of variation in the CR3BP, the $A(t)$ matrix discussed in Section 2.6.1 must be found. To find $A(t)$, N^2 partial derivatives of the EOMs must be found [18]. The equations of motion are given in pseudopotential form in Eq. (2.40)-(2.42). The partials of these equations leads to the following [4, 20]

$$\mathbf{A}(t) = \left[\begin{array}{c|c} \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \\ \hline \mathbf{U_{XX}} & \mathbf{\Omega} \end{array}\right], \tag{2.71}$$

where $\mathbf{0}_{3\times3}$ is a 3×3 submatrix of zeros, $\mathbf{I}_{3\times3}$ is a 3×3 identity submatrix, $\mathbf{U_{XX}}$ and $\mathbf{\Omega}$ are submatrices defined as follows

$$\mathbf{U_{XX}} = \begin{bmatrix} U^*_{xx} & U^*_{xy} & U^*_{xz} \\ U^*_{yx} & U^*_{yy} & U^*_{yz} \\ U^*_{zx} & U^*_{zy} & U^*_{zz} \end{bmatrix} \tag{2.72}$$

and

$$\mathbf{\Omega} = \begin{bmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.73}$$

The term $\mathbf{U_{XX}}$ is a submatrix of second partial derivatives where $U_{ij} = \frac{\partial^2 U}{\partial j \partial i}$.

With $A(t)$ defined, only $\Phi(t, t_0)$ needs to be derived to allow the differential equation dealing with the state transition matrix in Eq. (2.66) to be solved. The state transition matrix is the "Jacobian of the final state, $\vec{X}(t)$, with respect to the initial state, $\vec{X}_0(t)$" [10] as follows

$$\Phi(t, t_0) = \begin{bmatrix} \frac{\partial x(t)}{\partial x(t_0)} & \frac{\partial x(t)}{\partial y(t_0)} & \frac{\partial x(t)}{\partial z(t_0)} & \frac{\partial x(t)}{\partial \dot{x}(t_0)} & \frac{\partial x(t)}{\partial \dot{y}(t_0)} & \frac{\partial x(t)}{\partial \dot{z}(t_0)} \\ \frac{\partial y(t)}{\partial x(t_0)} & \frac{\partial y(t)}{\partial y(t_0)} & \frac{\partial y(t)}{\partial z(t_0)} & \frac{\partial y(t)}{\partial \dot{x}(t_0)} & \frac{\partial y(t)}{\partial \dot{y}(t_0)} & \frac{\partial y(t)}{\partial \dot{z}(t_0)} \\ \frac{\partial z(t)}{\partial x(t_0)} & \frac{\partial z(t)}{\partial y(t_0)} & \frac{\partial z(t)}{\partial z(t_0)} & \frac{\partial z(t)}{\partial \dot{x}(t_0)} & \frac{\partial z(t)}{\partial \dot{y}(t_0)} & \frac{\partial z(t)}{\partial \dot{z}(t_0)} \\ \frac{\partial \dot{x}(t)}{\partial x(t_0)} & \frac{\partial \dot{x}(t)}{\partial y(t_0)} & \frac{\partial \dot{x}(t)}{\partial z(t_0)} & \frac{\partial \dot{x}(t)}{\partial \dot{x}(t_0)} & \frac{\partial \dot{x}(t)}{\partial \dot{y}(t_0)} & \frac{\partial \dot{x}(t)}{\partial \dot{z}(t_0)} \\ \frac{\partial \dot{y}(t)}{\partial x(t_0)} & \frac{\partial \dot{y}(t)}{\partial y(t_0)} & \frac{\partial \dot{y}(t)}{\partial z(t_0)} & \frac{\partial \dot{y}(t)}{\partial \dot{x}(t_0)} & \frac{\partial \dot{y}(t)}{\partial \dot{y}(t_0)} & \frac{\partial \dot{y}(t)}{\partial \dot{z}(t_0)} \\ \frac{\partial \dot{z}(t)}{\partial x(t_0)} & \frac{\partial \dot{z}(t)}{\partial y(t_0)} & \frac{\partial \dot{z}(t)}{\partial z(t_0)} & \frac{\partial \dot{z}(t)}{\partial \dot{x}(t_0)} & \frac{\partial \dot{z}(t)}{\partial \dot{y}(t_0)} & \frac{\partial \dot{z}(t)}{\partial \dot{z}(t_0)} \end{bmatrix} \tag{2.74}$$

Now that all the required equations are defined relative to the CR3BP, the differential corrections process may be described. In the CR3BP, the differential corrections process takes advantage of the symmetry over the x-z plane and in time. This description of the differential corrections process in the CR3BP follows the description by [26, 4]. Begin by selecting an initial condition that is perpendicular to the x-z plane

$$\vec{X}_0 = [x_0, 0, z_0, 0, \dot{y}_0, 0]^T . \tag{2.75}$$

The initial condition in Eq. (2.75) will be propagated forward in time until it crosses the x-z plane. The solution will be found when this crossing of the x-z plane is also perpendicular to the plane [26]

$$\vec{X}(T/2) = [x, 0, z, 0, \dot{y}, 0]^T , \tag{2.76}$$

where T is the period.

By crossing perpendicular to the plane, the symmetry in time and over the x-z plane, guarantees the trajectory will return to the original starting point. Thus, the crossing at $y = 0$ is halfway through one period, shown in Fig. 2.12.

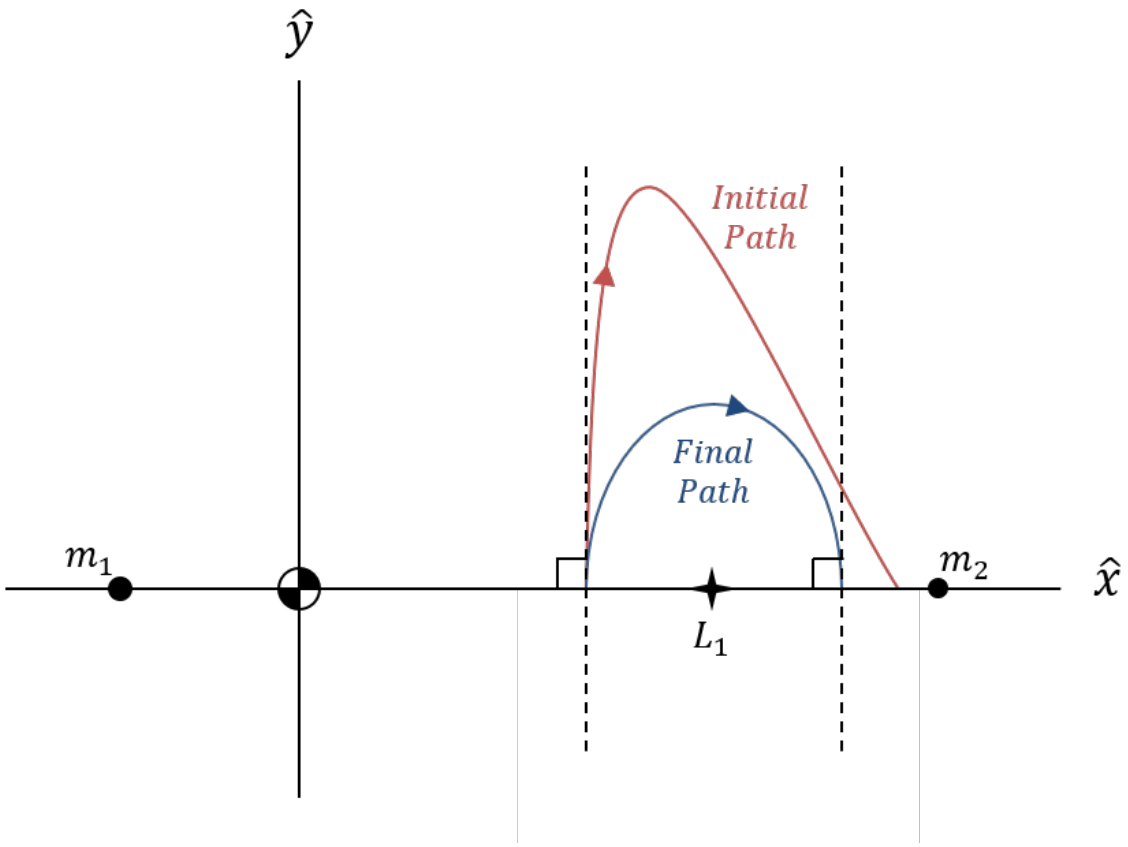


Figure 2.12. Crossing of the x-z plane for an initial guess and then again after differential corrections, such that the crossing is perpendicular to the x-z plane (Adapted from Grebow [4])

With the objective in mind of finding a perpendicular x-z crossing at $T/2$, the differential corrections process may be used to slightly vary the initial conditions to reach the desired end point. The differential corrections process begins by propagating the initial conditions and the state transition matrix forward in time until "y changes signs" [26], indicating an x-z plane crossing. If both $|\dot{x}|$ and $|\dot{z}|$ are less than a specified tolerance, then a periodic orbit has been found. [26] defines 10^{-8} as the required tolerance to declare an orbit as periodic. If the tolerance is not met, then the initial conditions need to be adapted using the differential corrections process in Eq. (2.70)

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \delta \dot{x} \\ \delta \dot{y} \\ \delta \dot{z} \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_{16} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & \Phi_{26} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & \Phi_{34} & \Phi_{35} & \Phi_{36} \\ \Phi_{41} & \Phi_{42} & \Phi_{43} & \Phi_{44} & \Phi_{45} & \Phi_{46} \\ \Phi_{51} & \Phi_{52} & \Phi_{53} & \Phi_{54} & \Phi_{55} & \Phi_{56} \\ \Phi_{61} & \Phi_{62} & \Phi_{63} & \Phi_{64} & \Phi_{65} & \Phi_{66} \end{bmatrix} \begin{bmatrix} \delta x_0 \\ \delta y_0 \\ \delta z_0 \\ \delta \dot{x}_0 \\ \delta \dot{y}_0 \\ \delta \dot{z}_0 \end{bmatrix} + \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} \delta\tau, \tag{2.77}$$

where $[\delta x_0, \delta y_0, \delta z_0, \delta \dot{z}_0, \delta \dot{y}_0, \delta \dot{z}_0]^T$ are the changes to the initial conditions, $[\delta x, \delta y, \delta z, \delta \dot{z}, \delta \dot{y}, \delta \dot{z}]^T$ are the desired changes to the final condition, and $\delta\tau$ is the change in time of the period of the orbit due to the changes in the initial conditions [27].

In essence, the differential corrections process in this case is adapting the three non-zero initial states $(x_0, z_0, \dot{y}_0)$ to target the two final states $(\dot{x}, \dot{z})$. Thus, many of the term in Eq. (2.77) are unneeded. There cannot be any change in the initial states y_0, $\dot{x}_0$, and $\dot{z}_0$ because the trajectory would not start perpendicular to the $x-z$ plane. Any terms in the state transition matrix associated with y_0, $\dot{x}_0$, and $\dot{z}_0$ are therefore unneeded[27]. Similarly, at the final state, the terms x, z, and $\dot{y}$ can take any value. Since their value is not being target in the differential corrections scheme, the terms in the state transition matrix associated with x, z, and $\dot{y}$ are also unneeded [27]. These changes are shown in Eq. (2.78) as follows

$$\begin{bmatrix} \cancel{\delta x} \\ \delta y \\ \cancel{\delta z} \\ \delta \dot{x} \\ \cancel{\delta \dot{y}} \\ \delta \dot{z} \end{bmatrix} = \begin{bmatrix} \cancel{\Phi_{11}} & \cancel{\Phi_{12}} & \Phi_{13} & \cancel{\Phi_{14}} & \Phi_{15} & \cancel{\Phi_{16}} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} & \Phi_{24} & \Phi_{25} & \Phi_{26} \\ \cancel{\Phi_{31}} & \cancel{\Phi_{32}} & \Phi_{33} & \cancel{\Phi_{34}} & \Phi_{35} & \cancel{\Phi_{36}} \\ \Phi_{41} & \Phi_{42} & \Phi_{43} & \Phi_{44} & \Phi_{45} & \Phi_{46} \\ \cancel{\Phi_{51}} & \cancel{\Phi_{52}} & \Phi_{53} & \cancel{\Phi_{54}} & \Phi_{55} & \cancel{\Phi_{56}} \\ \Phi_{61} & \Phi_{62} & \Phi_{63} & \Phi_{64} & \Phi_{65} & \Phi_{66} \end{bmatrix} \begin{bmatrix} \delta x_0 \\ 0 \\ \delta z_0 \\ 0 \\ \delta \dot{y}_0 \\ 0 \end{bmatrix} + \begin{bmatrix} \cancel{\dot{x}} \\ \dot{y} \\ \cancel{\dot{z}} \\ \ddot{x} \\ \cancel{\ddot{y}} \\ \ddot{z} \end{bmatrix} \delta\tau. \tag{2.78}$$

Since it is known the $\delta y = 0$, since y must equal 0 at the x-z plane crossings, δy can be used to find $\delta\tau$ [27]

$$\delta y = 0 = \Phi_{21}\delta x_0 + \Phi_{23}\delta z_0 + \Phi_{25}\delta\dot{y}_0 + \dot{y}\delta\tau. \tag{2.79}$$

To solve for $\delta\dot{x}$ and $\delta\dot{z}$, it is convenient to constrain one of the initial variables, so that the inverted matrix is a 2×2, instead of a 2×3 [28]. Thus, one option is to only change z_0 and $\dot{y}_0$, and leave x_0 fixed [26]

$$\begin{bmatrix} \delta\dot{x} \\ \delta\dot{z} \end{bmatrix} = \left[\begin{pmatrix} \Phi_{43} & \Phi_{45} \\ \Phi_{63} & \Phi_{65} \end{pmatrix} - \frac{1}{\dot{y}} \begin{bmatrix} \ddot{x} \\ \ddot{z} \end{bmatrix} \begin{pmatrix} \Phi_{23} & \Phi_{25} \end{pmatrix} \right] \begin{pmatrix} \delta z_0 \\ \delta\dot{y}_0 \end{pmatrix}. \tag{2.80}$$

The other option is to only change x_0 and $\dot{y}_0$, and leave z_0 fixed [26]

$$\begin{bmatrix} \delta\dot{x} \\ \delta\dot{z} \end{bmatrix} = \left[\begin{pmatrix} \Phi_{41} & \Phi_{45} \\ \Phi_{61} & \Phi_{65} \end{pmatrix} - \frac{1}{\dot{y}} \begin{bmatrix} \ddot{x} \\ \ddot{z} \end{bmatrix} \begin{pmatrix} \Phi_{21} & \Phi_{25} \end{pmatrix} \right] \begin{pmatrix} \delta x_0 \\ \delta\dot{y}_0 \end{pmatrix}. \tag{2.81}$$

Either of the methods can be used, so if one does not work, then the other can be attempted. Use Eq. (2.80) and Eq. (2.81) to solve for the changes in the initial conditions. Adapt the initial conditions, then perform the differential procedure over again. According to [26] and [28], the process should converge in less than 5 iterations. The second half of the orbit can be found using the symmetry about the x-z plane or by numerically integrating the initial conditions for an entire period, T [26].

Since this is a differential corrections process, the initial guess must be close to the true initial conditions. Grebow tabulates initial conditions for different periodic orbits that can be used as an initial guess in the differential corrections process [4].

There are many different periodic orbits in the CR3BP. Describing all of these periodic orbits is beyond the scope of the current research, but a detailed description of numerous periodic orbits can be found in Grebow and Doedel et al. [4, 3]. One

example periodic orbit is the Lyapunov family of orbits. A family of orbits are orbits that "posess similar behavior" [10]. The Lyapunov family of orbit are planar periodic orbits that are around the collinear libration points [3]. The L_1 Lyapunov family is shown in Fig. 2.13. It should be noted that bifurcations are locations where one family of orbits coincides with another family of orbits [3].

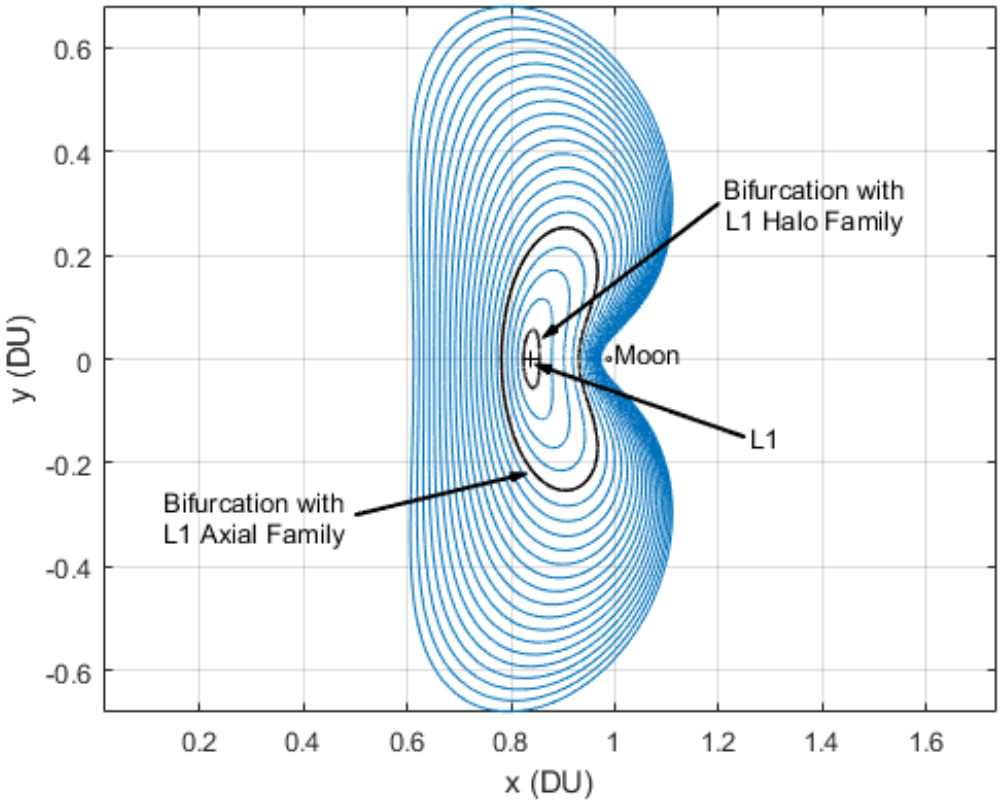


Figure 2.13. L1 Lyapunov Family in the Earth-Moon System (Adapted from Grebow [4])

2.7 Invariant Manifolds

Manifolds are a set of all trajectories that make up "a surface of lower dimension imbedded within the phase space" [18]. In other words, manifolds are the description of the motion an object can take if perturbed from its motion about some equilib-

rium point [29]. Manifolds are more precisely named invariant manifolds due to the CR3BP being "time-invariant" [18]. Every periodic orbit and equilibrium point has a set of invariant manifolds [29]. When on an invariant manifold, a spacecraft will stay on it for all time [18]. While on an invariant manifold, no fuel expenditure is required because manifolds share a constant energy [30]. An invariant manifold can be separated into two categories: stable and unstable. An unstable manifold is one that originates on a periodic orbit (or equilibrium point) and departs that orbit as $t \to \infty$ [18]. Similarly, a stable manifold will arrive on a periodic orbit as $t \to \infty$ [18].

To find a manifold, the periodic orbit must be perturbed in the unstable or stable direction in the positive or negative time, respectively [30]. To determine the direction of perturbation, the state transition matrix must be propagated for one full orbit. When the state transition matrix has been propagated for an entire orbit, it contains information about every point along the orbit and is called the monodromy matrix [29]. The monodromy matrix contains six eigenvalues that describe "stable, oscillating and/or unstable modes" [30]. In order to determine the dominant effects, the smallest and largest real eigenvalue must be determined [30, 29]. The smallest eigenvalue, λ_S, identifies the dominant stable perturbation, while the largest eigenvalue, λ_U, identifies the dominant unstable perturbation [30]. These eigenvalues have associated eigenvectors, $\vec{v}_S$ and $\vec{v}_U$, that describe the direction of stable and unstable perturbations [30]. The direction of perturbation can be calculated at each point in the periodic orbit by use of the monodromy matrix's stable and unstable eigenvectors.

After calculating the stable and unstable eigenvectors, the manifolds can be determined. At any point in the trajectory of the periodic orbit, the state is defined as $\vec{X}(t) = [x, y, z, \dot{x}, \dot{y}, \dot{z}]$. To find the initial condition of the manifold, it must be perturbed away from the initial state as follows [30]

$$\vec{X}_S(t) = \vec{X}(t) \pm \epsilon \hat{\vec{v}}_S \tag{2.82}$$

$$\vec{X}_U(t) = \vec{X}(t) \pm \epsilon \hat{\vec{v}}_U, \tag{2.83}$$

where $\vec{X}_S(t)$ is the stable manifold initial conditions, $\vec{X}_U(t)$ is the unstable manifold initial conditions, ϵ is the size of the perturbation, $\hat{\vec{v}}_S$ is the normalized stable eigenvector, and $\hat{\vec{v}}_U$ is the normalized unstable eigenvector.

In theory, the perturbation ϵ is infinitely small, leading to the infinite time required to depart or arrive at the periodic orbit [18]. However, finite perturbations can be used to approximate the manifold, but they must be carefully selected: too large results in inaccurate trajectories and too small results in large computation times [30]. In [30], Truesdale recommends a perturbation value of $1e-4$ in the Earth-Moon system.

With the initial conditions of the manifold found, symmetry allows the manifold to be numerically integrated forward or backward in time to find the unstable or stable manifold, respectively. Since there are an infinite number of points on a periodic orbit, there are also an infinite number of manifolds that a spacecraft could follow [30]. These manifolds form a tube [29]. An example manifold tube can be seen in Fig. 2.14.

Notice how in the stable manifold, motion on the trajectory moves toward the periodic orbit, while it moves away from the periodic orbit on the unstable manifold. The manifolds also have a symmetry about the x-z plane, which is expected for motion in the CR3BP.

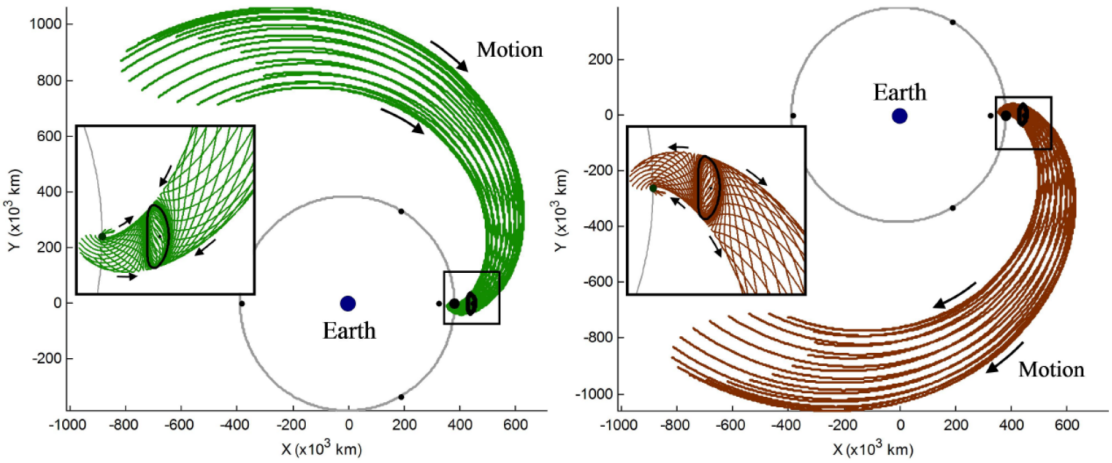


Figure 2.14. Stable (left) and Unstable (right) Manifold in the Earth-Moon System (Reproduced with Permission from Elsevier) [5]

2.8 Heteroclinic Connections

Periodic Orbits in the CR3BP contain homoclinic and heteroclinic connections. A homoclinic connection is when an unstable manifold connects with itself, allowing it to return to its original periodic orbit [5]. Homoclinic connections are not the focus of the current research, so they will not be discussed any further. A heteroclinic connection is found when a stable and unstable manifold intersect [31]. Thus, a spacecraft would theoretically be able to make a "free" transfer from one periodic orbit and arrive on a secondary orbit [5]. These "free" transfers make heteroclinic trajectories appealing in the CR3BP. In reality, a small cost would be associated with heteroclinic trajectories to ensure the spacecraft transferred between manifolds. However, the small cost would still offer immense cost savings over conventional transfers in the 2BP.

Although there are heteroclinic connections in the spatial CR3BP, it is easier to gain an understanding of these free transfers in the planar CR3BP. To start, a planar manifold must be characterized by $\vec{X} = [x, y, \dot{x}, \dot{y}]$ making it four-dimensional [5]. Before going further, an assumption is made that both periodic orbits must have the same Jacobi Constant. This implies that the stable and unstable manifolds

propagating from these periodic orbits also have the same Jacobi Constant [5]. Thus, the dimensionality of the problem is reduced to 3.

Next, both the unstable manifold of the initial periodic orbit and the stable manifold of the final periodic orbit are propagated until they cross a "surface of section" selected at some x-position [5]. A "surface of section" was proposed by Henri Poincaré and is used to reduce the dimensionality of the space from 3-dimensions to 2-dimensions [18]. Figure 3.7 demonstrates the placement of a surface of section.

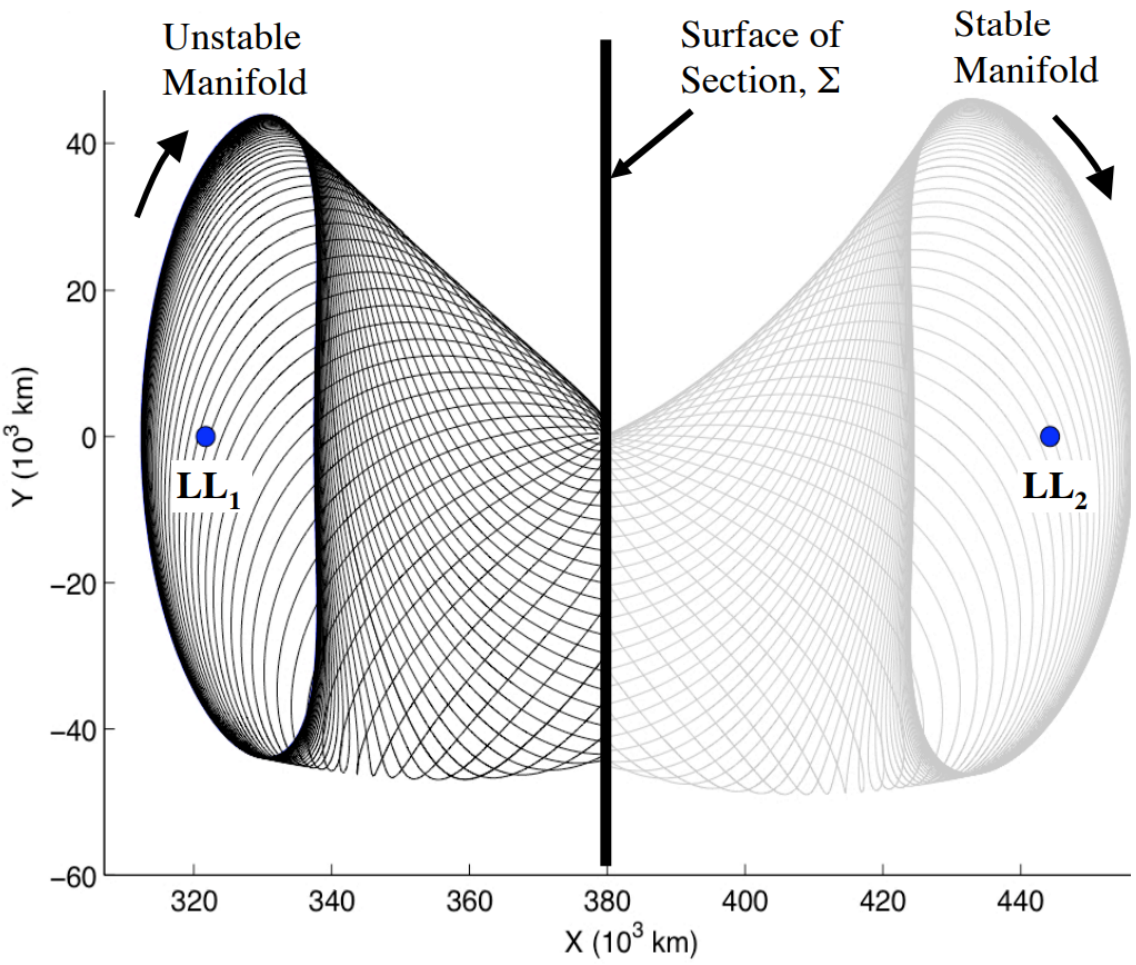


Figure 2.15. Surface of Section placed between two Lyapunov periodic orbits in the Earth-Moon System (Reproduced with Permission from Elsevier) [5]

A point is plotted every time the trajectory intersects the surface of section [18]. The surface of section then results in a two-dimensional Poincaré map [5]. There are

three different types of Poincaré maps that allow for sign ambiguity in the Jacobi Constant to be accounted for [5]:

- One-sided Maps:
  - P_+: Only trajectories with positive intersection with the surface of section are plotted (for example, positive $\dot{x}$ values) [5]
  - P_-: Only trajectories with negative intersection with the surface of section are plotted (for example, negative $\dot{x}$ values) [5]
- Two-sided Maps:
  - $P_\pm$: all trajectories with intersection with the surface of section are plotted [5]

Ignoring sign ambiguities in the Jacobi constant, any intersection of points on the Poincaré map guarantees position and velocity are aligned at that point [31]. Since all the states are aligned, then a heteroclinic connection is present. At this point a spacecraft could transfer from the unstable manifold of the starting orbit to the stable manifold of the final orbit at no cost [5]. Figure 3.8 demonstrates a Poincare Map and the associated heteroclinic trajectories.

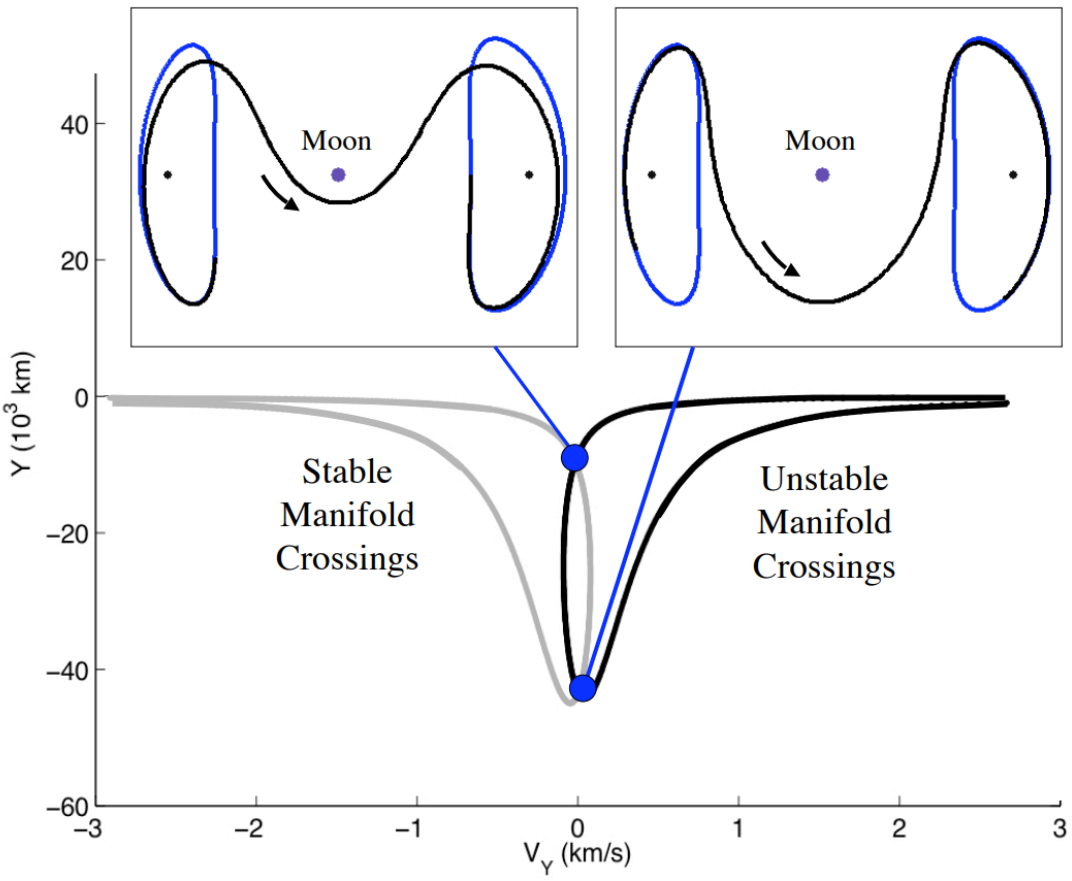


Figure 2.16. Lower: Poincare Map of the manifolds intersecting the Surface of Section. Upper: The Heteroclinic Trajectories found from the two points of intersection on the Poincare Map (Reproduced with Permission from Elsevier) [5]

2.9 Static Optimization

There are many ways to classify optimal control problems, but an important distinction is between static and dynamic optimization. Static optimization takes place at a single point in time, while dynamic optimization problems deals with a changing problem. The dynamics of the problem describe how it changes over time, but these dynamics can be discretized to solve the dynamic optimization problem. The fundamentals used in static optimization can be used in dynamic optimization after discretization. Methods for discretization of a dynamic problem will be discussed in Section 2.10, but the basics of static optimization will be demonstrated to give a

foundation to apply to dynamic optimization.

The derivation of static optimization fundamentals is taken mostly from [32], [33], [34], and [35].

Static Optimization in its unconstrained form is given as

$$\underset{\vec{x}}{\text{minimize}} \quad f(\vec{x}). \tag{2.84}$$

Equation (2.84) shows that the variable $\vec{x}$ is to be found, while minimizing the cost function J, where $J = f(\vec{x})$.

In a situation where the maximum needs to be found instead of the minimum, the cost function can simply be made negative so that the function remains in standard form. It should be noted that the problem can be constrained or unconstrained. The form shown in 2.84 is for the unconstrained case.

In the unconstrained case, a critical point of a function can be determined by setting the first derivative equal to zero, which can also be stated as setting the Jacobian equal to zero if $\vec{x}$ is a vector. The critical points can be found as

$$\nabla f(\vec{x}^*) = \vec{0}. \tag{2.85}$$

This is called the first-order necessary condition because it is necessary to be a minimum, but it is not sufficient to declare the point a minimum because it could also be a maximum or an inflection point. In order to determine if the critical point is truly a minimum, second-order conditions must be checked. Second-order conditions can be found by taking the Hessian and determining if it is positive definite or positive semi-definite. The Hessian is a matrix of second-partial derivatives. Positive semi-definite is still a second-order necessary condition because it still does not guarantee that the critical point is a minimum

$$H(\vec{x}^*) = \nabla^2 f(\vec{x}^*) \geq \vec{0}. \tag{2.86}$$

However, if the Hessian is positive-definite, then the second-order sufficient condition is satisfied and the critical point is a minimum

$$H(\vec{x}^*) = \nabla^2 f(\vec{x}^*) > \vec{0}. \tag{2.87}$$

A point that satisfies both the necessary and sufficient conditions is guaranteed to be at least a locally optimal solution. However, it is unknown if it is globally optimal. To be globally optimal there must not be any other values of x that are smaller than it within the boundary. Figure 2.17 is an example of global and local points.

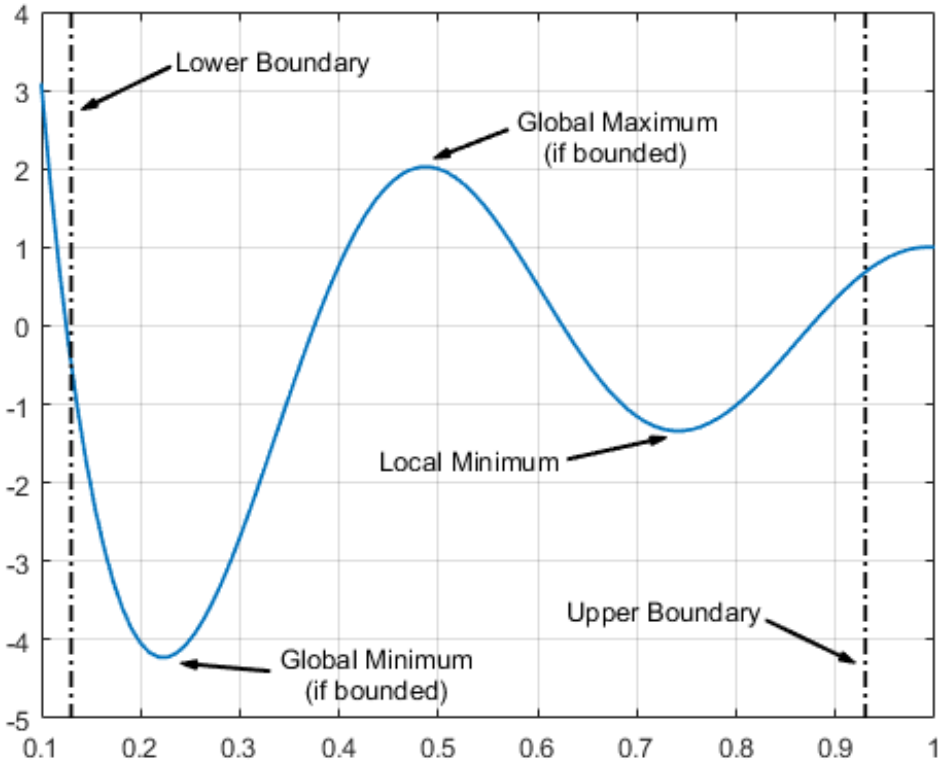


Figure 2.17. Global and Local Maxima and Minima (Adapted from Hess [6])

Static Optimization in the constrained case is typically given in the standard form:

$$\begin{aligned} \underset{\vec{x}}{\text{minimize}} \quad & J = f(\vec{x}) \\ \text{subject to} \quad & \vec{h}(\vec{x}) = \vec{0}, \\ & \vec{g}(\vec{x}) \leq \vec{0}. \end{aligned} \tag{2.88}$$

In the standard form, $\vec{x}$ is the variable that needs to be found, J is the cost function that is being minimized, $\vec{h}(\vec{x})$ are the equality constraints, and $\vec{g}(\vec{x})$ are the inequality constraints. In order to solve the unconstrained static optimization problem, the constraints must be appended onto the cost function to form the Lagrangian. To add on the constraints without changing the cost function, they are added on in "convenient forms of zero" [33]

$$\begin{aligned} L(\vec{x}, \vec{v}, \vec{u}) = f(\vec{x}) + \sum_{i=1}^{p} v_i h_i(\vec{x}) + \sum_{j=1}^{m} u_i \left(g_i(\vec{x}) + s_j^2 \right) \\ = f(\vec{x}) + \vec{v}^T \vec{h}(\vec{x}) + \vec{u}^T \vec{g}(\vec{x}) \end{aligned}, \tag{2.89}$$

where $\vec{v}$ are the Lagrange Multipliers associated with the equality constraints and $\vec{u}$ are the Lagrange Multipliers associated with the inequality constraints. In order to append the inequality constraints, a "slack variable" s^2 is for each inequality constraint is introduced to convert the inequality constraints to equality constraints. Since the equality constraint and the converted inequality constraint are both equal to zero, adding this on is the same as appending zeros to the cost function.

To find the minimum of the Lagrangian, Karush-Kuhn-Tucker (KKT) conditions must be applied. Like in the unconstrained case, there are necessary and sufficient conditions. The first-order necessary conditions include the following 4 conditions [34]:

1. Gradient check - ensure the Lagrangian is stationary with respect to the design variables, the Lagrange multipliers, and the slack variables

$$\begin{aligned} \frac{\partial L}{\partial \vec{x}} &= \vec{0} \\ \frac{\partial L}{\partial \vec{u}} &= \vec{0} \\ \frac{\partial L}{\partial \vec{v}} &= \vec{0} \\ \frac{\partial L}{\partial \vec{s}} &= \vec{0} \end{aligned} \tag{2.90}$$

2. Feasibility Check for Inequality constraints - ensure the slack variables are meeting their requirement

$$s_j^2 \geq 0 \quad \text{for } j = 1, 2, \ldots, n \tag{2.91}$$

3. Non-negativity of Lagrange Multipliers for inequality constraints

$$u_j^* \geq 0 \quad \text{for } j = 1, 2, \ldots, n \tag{2.92}$$

4. Regularity check - the gradients of the active constraints are linearly independent. An active constraints is when the slack variable is equal to zero, meaning that the associated inequality constraint is essentially an equality constraint.

The second-order sufficient conditions are test to guarantee the point is a local minimum. This is done by taking the Hessian of the Lagrange function. If the Hessian of the Lagrangian is positive definite, then the point is a minimum point

$$H(\vec{x}^*) = \nabla^2 L(\vec{x}^*) > \vec{0}. \tag{2.93}$$

The KKT conditions are the fundamentals behind solving a constrained nonlinear-programming (NLP) problem [35]. An NLP problem is optimization of a nonlinear

scenario. Computers are well-equipped to hand NLP algorithms and there are many algorithms available to solve them in different scenarios, including Interior Point Optimizer (IPOPT) and Sparse Nonlinear Optimizer (SNOPT).

IPOPT and SNOPT are both sparse NLP solvers and are capable of solving problems with a large number of variables and constraints [7]. The difference lies in what method the two solvers use. SNOPT uses a Sequential Quadratic Programming (SQP) quasi-Newton Method, while IPOPT uses an interior-point method [7].

2.10 Dynamic Optimization Discretization

Dynamical optimal control problems must be reformulated to allow an NLP solver to be used. There are two categories of methods to convert the problem from continuous to discrete: direct and indirect. The major distinction between the two methods is indirect optimization utilizes the Euler-Lagrange equations in Optimal Control Theory, while direct optimization does not utilize these equations explicitly [36].

2.10.1 Indirect Optimization

Indirect optimization utilizes optimal control theory, which uses the first-order differential Euler-Lagrange equations derived from calculus of variations to find the optimal solution. The following is a derivation of the Euler-Lagrange equations following the steps provided in [36] and [7]. A more detailed derivation can be found in [**?**] and [25]. An optimal control problem is given in the Bolza form

$$\begin{aligned} \underset{\vec{u}}{\text{minimize}} \quad & J = \Phi(t_0, \vec{x}(t_0), t_f, \vec{x}(t_f)) + \int_{t_0}^{t_f} L(t, \vec{x}(t), \vec{u}(t))dt \\ \text{subject to} \quad & \dot{\vec{x}} = \vec{f}(t, \vec{x}(t), \vec{u}(t)), \\ & \phi(t_0, \vec{x}(t_0), t_f, \vec{x}(t_f)) = \vec{0}, \\ & \vec{C}_{min} \leq \vec{C}(t, \vec{x}(t), \vec{u}(t)) \leq \vec{C}_{max}, \end{aligned} \tag{2.94}$$

where t_0 is the initial time, t_f is the final time, $\vec{x}$ is the state vector, and $\vec{u}$ is the control vector. In Eq. (2.94), J is the cost functional that is to be minimized, $\dot{\vec{x}}$ are the dynamic constraints in the form of differential equations, ϕ is the collection of boundary conditions, and $\vec{C}$ are the path constraints.

In order to solve for first-order optimality conditions, the constraints must be appended on to the cost function by multiplying each constraint by a Lagrange multiplier and subtracting it from the cost function.The augmented cost function is

$$J_a = \Phi - \vec{\upsilon}^T \vec{\phi} + \int_{t_0}^{t_f} \left[L + \vec{\lambda}^T (\vec{f} - \dot{\vec{x}}) \right] dt, \tag{2.95}$$

where υ is the Lagrange multiplier of the boundary conditions and λ is the is the Lagrange multiplier of the state (called the costate or adjoint of the differential equation).

Taking the variation of the cost function and setting it equal to zero gives the necessary conditions for optimality [37]. This is similar to the static case [37]. To perform this operation, the Hamiltonian must be introduced [37]

$$H(\vec{x}, \vec{\lambda}, \vec{u}) = L + \vec{\lambda}^T \vec{f}. \tag{2.96}$$

After setting the variation of the cost function equal to zero, the necessary conditions are found [37]. The first set of the necessary conditions in Eq. (2.97) are know as the control equations [37]

$$\begin{aligned}
\delta\vec{v} = \vec{0}; \delta\vec{x_0} = \vec{0}; \delta\vec{x_f} = \vec{0}; \delta\vec{t_0} = \vec{0}; \delta\vec{t_f} = \vec{0}; \\
\phi(t_0, \vec{x}(t_0), t_f, \vec{x}(t_f)) = \vec{0}; \\
\vec{\lambda}(t_0) = -\left[\frac{\partial\Phi}{\partial\vec{x}(t_0)}\right]^T + \left[\frac{\partial\phi}{\partial\vec{x}(t_0)}\right]^T \vec{v}; \vec{\lambda}(t_f) = \left[\frac{\partial\Phi}{\partial\vec{x}(t_f)}\right]^T - \left[\frac{\partial\phi}{\partial\vec{x}(t_f)}\right]^T \vec{v}; \\
H(t_0) = \frac{\partial\Phi}{\partial t_0} - \vec{v}^T\frac{\partial\phi}{\partial t_0}; H(t_f) = -\frac{\partial\Phi}{\partial t_f} + \vec{v}^T\frac{\partial\phi}{\partial t_f}
\end{aligned} \tag{2.97}$$

The next set of necessary conditions in Eq. 2.98is known as the Euler-Lagrange equations

$$\dot{\vec{x}} = \left[\frac{\partial H}{\partial\vec{\lambda}}\right]^T = \vec{f}(\vec{x}(t), \vec{u}(t), t) \tag{2.98a}$$

$$\dot{\vec{\lambda}} = -\left[\frac{\partial H}{\partial\vec{x}}\right]^T \tag{2.98b}$$

$$\left[\frac{\partial H}{\partial\vec{u}}\right]^T = \vec{0}. \tag{2.98c}$$

The Euler-Lagrange equations from Eq. (2.98) have three parts. Equation (2.98a) is the state dynamics, Eq.(2.98b) is the costate dynamics, and Eq. (2.98c) is the stationarity condition. Collectively, Eq. (2.97) and (2.98) are the first-order necessary conditions that, when satisfied, represent the optimality conditions. The problem is now a Hamiltonian Boundary Value Problem. Hamiltonian Boundary Value problems can be solved in a variety of ways, including shooting and multiple-shooting methods [7]. These solutions are generally only used to find analytical solutions to straightforward optimization problems. More complex problems generally use numerical optimization methods [36].

2.10.2 Direct Optimization

Unlike indirect optimization, direct optimization does not involve deriving optimality conditions and is a more "brute force" approach [36]. There are two classifications of direct optimization based on what is being parametrized. Parametrization is the process of selecting a function to approximate the optimized variables at discrete points. This type of direct optimization utilizes control parameterization, which requires solving by shooting or multiple shooting methods, similar to how the Hamiltonian Boundary Value Problem is solved in the indirect case. The second type of direct optimization uses both state and control parametrization. This method uses collocation to discretize the optimal control problem and allow it to be solved with the NLP solvers.

In collocation, dynamics are no longer propagated. Instead, the differential equations are converted to algebraic constraints [7]. This is done by discretizing the problem in time by dividing into segments. The following description of collocation is taken from Rao and Stanton [7, 36]. Time is divided into segments as

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N, \tag{2.99}$$

where $t_N = t_f$. The time where these mesh points are placed can be of variable-length.

The parameters x that are going to be optimized represent the value of the control and the states at each of those mesh points

$$\vec{x} = (x_1, u_1, \ldots, x_M, u_m), \tag{2.100}$$

where M is the number of states and m is the number of controls in the problem.

If initial and final time are free, then they must also be included as parameters

that need to be optimized gving

$$\vec{x} = (x_1, u_1, \ldots, x_M, u_m, t_0, t_f). \tag{2.101}$$

The cost, constraint functions, boundary conditions, and dynamics must be derived in terms of x . The continuous dynamics, called the continuity constraint, must be discretized to algebraic equations to create a NLP of the following form:

$$\begin{aligned} \min_{\vec{x}} \quad & J = f(\vec{y}) \\ \text{s.t.} \quad & \vec{g}(\vec{x}) = \vec{0}, \\ & \vec{h}_{min} \leq \vec{h}(\vec{x}) \leq \vec{h}_{max}. \end{aligned} \tag{2.102}$$

Since the dynamics are now approximated by algebraic equations there will be defects [7]. The goal is to make these defects as close to zero as possible, so that the dynamics are accurately represented.

2.10.3 Characteristics of Direct and Indirect Methods

Indirect methods and direct methods have distinct tradeoffs. Indirect methods have the difficulty of deriving optimality conditions. This can be cumbersome and is generally only used for simple problems [36]. To further exacerbate the issue, solving the Hamiltonian Boundary Value Problem can be very difficult and the problem has a small radii of convergence [7]. However, indirect methods have the benefit of producing not only the state and control, but also the costate. This allows greater insight into the problem, specifically by seeing how close a solution is to optimal.

Direct methods do not have the issue of deriving optimality conditions, so they can be very appealing as they do not require a deep understanding of optimization to apply. However, the tradeoff is a large number of parameters after discretizing the

problem. In direct methods, each node contains all the states, control, and possibly even t_0 and t_f. This means that a problem utilizing direct methods has n parameters [36]

$$n = (n_x + n_u))\, n_N + 2, \tag{2.103}$$

where n_x is the number of states, n_u is the number of controls, and n_N is the the number of nodes. The additional two parameters account for t_0 and t_f.

This demonstrates how large direct methods can become. This generally is not an issue however because using collocation leads to a large sparse NLP [7]. A sparse NLP is when there are a large number of zeros in the derivatives of the functions. Sparse NLP problems are easy to solve using known NLP solving algorithms, specifically SNOPT. This means that direct methods large size are not necessarily an issue and can still lead to effective solving of optimal control problems. NLP solvers are even easier to solve than the Hamiltonian Boundary Value Problem used in indirect methods leading to a larger radius of convergence.

The other major issue with direct collocation is costates are not found. As discussed earlier, this prevents gaining insight into the problem and determining closeness to optimality. If more accurate solutions are desired, the results for indirect optimization will be more accurate than those from direct optimization.

2.10.4 Pseudospectral Optimization

Interestingly, pseudospectral optimization methods allow for the convenience of direct methods, but also incorporate accurately approximating the costates. With an approximation of the costates, the closeness of the optimization can be determined, which also provides one of the major benefits of indirect methods. In essence, pseudospectral methods are direct methods that have approximations of the added

information found from indirect methods [7].

In short, pseudospectral optimization involves determining the location of collocation points and then discretizing the problem. The following description of pseudospectral optimization follows that provided by Rao [7]. Before selecting where to place the collocation points, the proper discretization scheme needs to be determined. If the points are placed improperly, accurately approximating the optimal trajectory might be difficult. The first value considered in the discretization is the cost function. To find the cost function, an integral must be approximated at discretized points. To do this accurately, a numerical quadrature is introduced. A numerical quadrature is the approximation of an integral of a function. To approximate the integral, the function is sampled at certain support points within the interval. To get an accurate representation, these points must be weighted. The discrete approximation is given as

$$\int_{-1}^{+1} f(\tau)d\tau = \sum_{i=1}^{N} w_i f(\tau_i), \tag{2.104}$$

where f is the function that is being approximated, N is the number of support points, w_i is the weight of each of these points, and τ_i is time where the support points are located.

The method to space these points has an effect on the convergence of the function to lower-order polynomials. For example, if evenly space points were used, there may be poor convergence of the integral and a polynomial of a high degree would have to be used for the approximation. Alternatively, if Gaussian quadrature points are used as the support points, there is in as exponential convergence for smooth functions [7], which is the "spectral" convergence in pseudospectral methods. The three most common Gaussian quadrature approximations are Legendre-Gauss (LG), Legendre-Gauss-Radau (LGR), and Legendre-Gauss-Lobatto (LGL) [8]. These methods are

distinguished by their weighting and their inclusion of endpoints. Each form of the Gaussian quadrature has a different weighting associated with it. For example, the weight of LGR has the following formula:

$$w_1 = \frac{2}{N^2} \tag{2.105}$$

$$w_i = \frac{1}{(1-\tau_i)\,\dot{P}^2_{N_1}(\tau_i)}, \quad i = 2 \ldots N \tag{2.106}$$

LG methods do not include either endpoint, LGR methods only include one endpoint, and LGL methods include both endpoints. This is best summarized in Fig. 2.18.

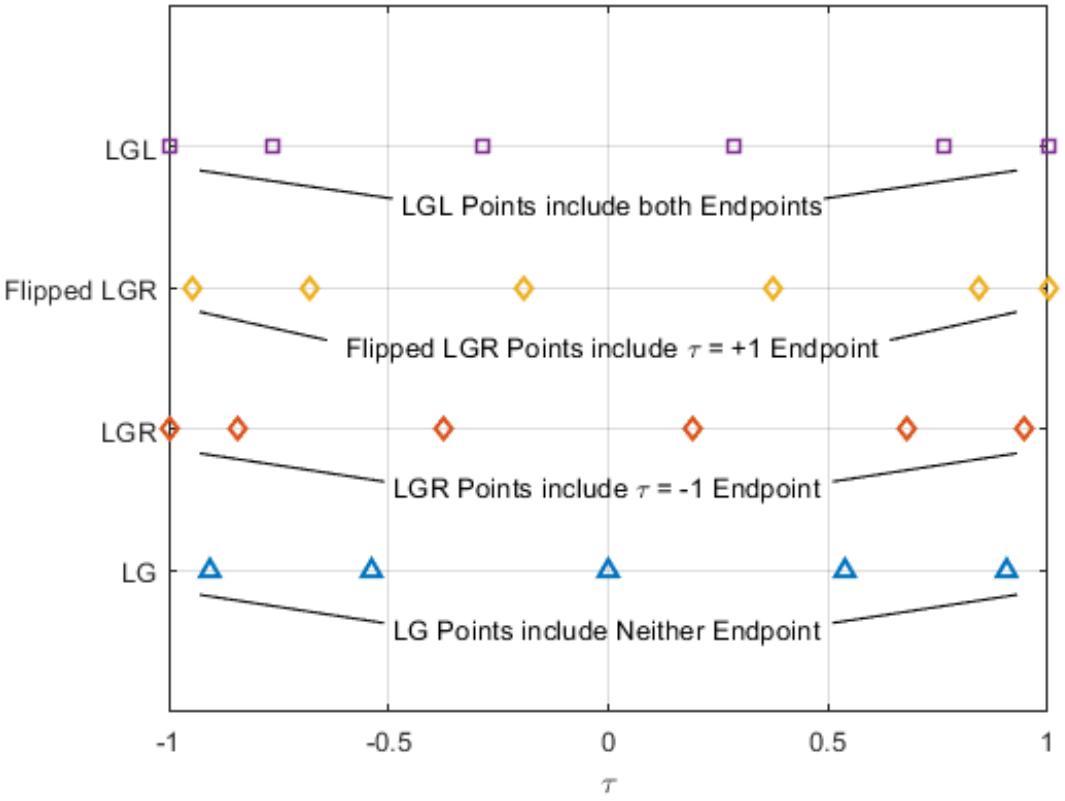


Figure 2.18. Difference between Guassian Quadrature points methods (Adapted from Rao [7])

The Gaussian quadrature points will be used as the spacing for the collocation points, since they will allow for the cost function to be approximated via a Gauss Quadrature. The Gaussian Quadrature points are determined from the roots of a Legendre polynomial or some linear combination of Legendre polynomials and its derivatives [8]. The roots of each method is listed below for an Nth degree Legendre polynomial given as [8]

$$P_N = \frac{1}{2^N N!} \frac{d^N}{d\tau^N} \left[\tau^2 - 1\right]^N \tag{2.107}$$

- LG: Roots of $P_N(\tau)$
- LGR: Roots of $P_N(\tau) + P_{N-1}(\tau)$
- LGL: Roots of $\dot{P}_{N-1}(\tau)$ together with $\tau = -1$ and $\tau = +1$.

This form of collocation is sometimes referred to as orthogonal collocation since Legendre polynomials are orthogonal [8].

Each method of selecting collocation becomes an exact approximation when certain degrees of polynomial are used. For example, LGR quadrature is exact when the polynomial used is of degree $2N - 2$ or less, where N is the number of support points. However, all the methods become exact within 2 degrees of polynomial of each other, so they are all extremely accurate [7].

Now that the location of the support points and how to approximate the cost function has been determined, collocation can occur. The state $\vec{x}(\tau)$ can be shown as

$$\vec{x}(\tau) \approx \vec{X}(\tau) = \sum_{i=0}^{N} \vec{X}(\tau_i) L_i(\tau) \tag{2.108}$$

Approximations are generally done with polynomials [7]. Although any polynomial can be used, Lagrange interpolating polynomials have beneficial characteristics.

A basis of N Lagrange polynomials will be constructed of degree $N-1$

$$L_i(t) = \prod_{k=1;k\neq i}^{N} \frac{t-t_k}{t_i-t_k} \text{ (for } = 1,\ldots,N). \tag{2.109}$$

Lagrange polynomials satisfy the isolation property [7],[8]

$$L_i(t_k) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}. \tag{2.110}$$

It should be noted that if global collocation is inadequate (this could be due to discontinuities, like impulsive burns, or rapidly changing dynamics [7]), then it can be divided into subintervals, called phases, and global collocation can be performed over each phase [8]. A "linkage constraint" will need to be included between each phase to ensure continuity [8]. Figure 2.19 gives an example of how the multi-phase problem is divided along with linkage constraints between each subinterval.

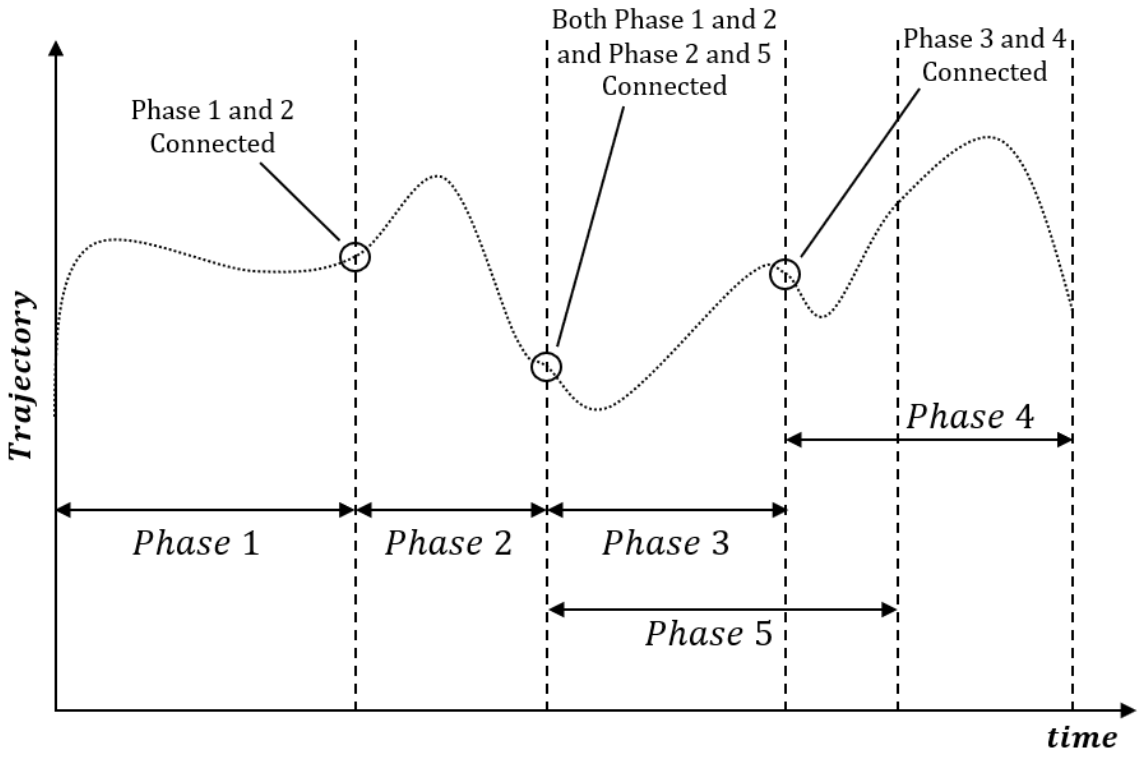


Figure 2.19. Linkage between Phases in Multi-Phase Collocation (Adapted from Rao et al. [8])

2.11 GPOPS

The optimization techniques described in Section 2.10.4 are implemented in GPOPS. GPOPS is software used in matrix laboratory (MATLAB) that is meant to solve general nonlinear optimal control problems [38]. Specifically, GPOPS utilizes pseudospectral direct collocation methods with a quadrature method of LGR [39]. After the problem has been discretized, GPOPS approximates the first and second derivatives and plugs them into the derivative matrix to be solved by the NLP solvers [39]. The software gives a choice between using SNOPT and IPOPT to solve the NLP problem. While the program has the option of dividing the problem into multiple-phases, this feature goes unused in this research since the problem is not set-up for multiple phases.

One of the interesting features of GPOPS is the way the mesh points are defined.

A mesh is the division of a segment into pieces. The number of mesh intervals in a problem is dependent on the dynamics and control used in the specific scenario. GPOPS uses a method called *hp*-adaptive mesh. To understand *hp*-adaptive mesh, a short background of other meshes is important. The original method used was called *h*-method, which is a Euler or Runge-Kutta method. The key feature of the *h*-method is the same fixed-degree polynomial is used in each mesh interval. This means that convergence has to come from increasing the number of mesh points or adjusting their placement [39].

The next method that came along is Gaussian quadrature orthogonal collocation. This method is described in detail in Section 2.10.4. Gaussian quadrature orthogonal collocation originally utilized the *p*-method using a single interval. This meant convergence occurred by increasing the degree of the polynomial approximation [39].

GPOPS utilizes a combination of these methods called *hp*-adaptive Gaussian quadrature collocation. In this method, both the number of mesh intervals and the degree of the approximating polynomials within each mesh interval are adapted. The major benefit of combining the two methods is it takes advantage of exponential convergence in smooth regions and increases mesh points in areas with discontinuities [39].

2.12 Literature Review of Relevant Works

The main body of work relevant to the current research deal with using direct collocation methods to find optimal trajectories in the CR3BP. Through a review of current research, it was found that the vast majority of the topics also happen to deal with low-thrust transfers, as opposed to impulsive. There are many bodies of work on general optimization of low-thrust trajectories in the CR3BP. Zhang and Zao provide a minimum-fuel low thrust trajectory design in the Earth-Moon CR3BP using an

indirect approach and solving the Two-Point Boundary Value Problem (TPBVP) with a shooting method [21]. Ozimek and Howell use hybrid optimization to determine low-thrust transfers in the Earth-Moon CR3BP. Optimal control theory is used to produce a parametrized control law, which is then solved using both a single and multiple-shooting direct method [40]. Stuart et al. also use a hybrid approach to finding optimal minimum propellant trajectories for a tour of Jupiter's Trojan asteroids [41].

There is also a vast body of work dealing with direct methods of solving for minimum-fuel optimal trajectories in the CR3BP. A multiple shooting differential corrections method is used to construct complex orbit chains and periodic orbits in [5]. However, collocation has been an area of prominent recent research in the CR3BP, including pseudospectral methods [42, 43]. Grebow et al. utilize a seventh-degree Gauss-Lobatto collocation scheme to approximate a low-thrust trajectory to perform a lunar pole-sitter mission in [44, 45]. Mingotti et al. search for optimal transfers to Earth-moon halo orbits by parametrizing the manifolds of the Halo Orbits and targeting the parametrized manifolds in a direct collocation scheme [46]. Minimum-fuel optimal low-thrust trajectories in the CR3BP are found in [47] with two optimization schemes: Legendre psuedospectral and Hermite-Simpson. The robustness of the collocation scheme to poor initial guesses was noted in [47], but did not utilize automatic mesh-refinement schemes.

Despite the robustness of the collocation schemes, there is a large amount of current work evaluating different initial guesses. Pritchett et al. used a method of "trajectory stacking" where both the initial and starting orbit are propagated for several revolutions and then the states of both orbits are concatenated, even if the states do not align [43]. The "trajectory stacking" method was designed to require little intuition for the problem, but [43] noted that some scenarios may require more accurate initial guesses. Pritchett et al. also found that after a single optimal trajectory

has been found, the optimal trajectory can be used as the initial guess to develop a family of orbit, where only one parameter, such as thrust level, is changed [43]. Anderson and Lo noted that the invariant manifold may be useful in developing an initial guess, since optimal trajectories across energy levels appear to follow the same types of paths as the invariant manifold [48]. Furthermore, Parker et al. demonstrate that a spacecraft requires substantially less Δv to perform single-maneuver transfers if the transfers are near heteroclinic connections in the corresponding phase space [5].

Research on low-thrust optimal trajectories in the CR3BP has delved into collocation schemes, but has not fully evaluated the usefulness of a direct pseudospectral method with an adaptive mesh-refinement scheme, specifically as it applies to objective functions that do not deal with minimum-fuel. Evaluating the efficacy of a direct pseudospectral method with an adaptive mesh-refinement scheme in the CR3BP would be beneficial to allow a more variety of circumstances to be evaluated using a collocation scheme. Additionally, initial guesses are of particular interest in the relevant research and an inherent problem for optimal control. While collocation schemes appear to be robust in some scenarios, many relevant works noted that a better initial guess would be useful in ensuring the collocation scheme can be applied to other scenarios. The literature suggests invariant manifolds and heteroclinic connections seem to be useful initial guesses that could potentially expand to a large range of scenarios. Determining the usefulness of the heteroclinic connection as an initial guess in comparison with other initial guess could allow the usefulness of collocation schemes to be expanded to further situations. To the extent of this author's knowledge, an evaluation of the efficacy of a direct pseudospectral method with an adaptive mesh-refinement scheme in a scenario dealing with a compound objective function with both minimum-fuel and minimum-time has not been investigated. The current research will also analyze the effectiveness of several initial guesses, including

the heteroclinic connection, on the resulting optimal trajectories.

III. Research Methodology

3.1 Chapter Overview

The test plan is developed in this chapter to evaluate the efficacy of the pseudospectral method. The specific configuration used for analysis is defined in this chapter by discussing the selected parameters used in the research and introducing the equations of motion (EOMs) with the control form and scaling included. The initial guesses used in the pseudospectral method are developed, specifically the generation of a heteroclinic trajectory. Finally, the Two-Body Problem (2BP) benchmark is introduced to allow the optimal trajectories in the Circular-Restricted Three-Body Problem (CR3BP) to be compared to a nominal scenario.

The current investigation utilizes MATLAB® version 9.0.0.341360 (R2016a). The optimization software is General Purpose Optimal Control Software (GPOPS) Version 2.3. All simulations were performed on a Windows 10 computer operating with an Intel® Xeon® Central Processing Unit E3-1245 v3 @ 3.50 GHZ with 32 GB of RAM.

3.2 Test Plan Overview

The test plan is intended to test multiple facets of the employed direct pseudospectral method. The first step in this process is verifying the pseudospectral method works by comparing the output from the pseudospectral method to a known result. The heteroclinic trajectory is the known minimum-fuel solution in the CR3BP. By providing the heteroclinic trajectory as the initial guess, the pseudospectral method should converge on the heteroclinic trajectory, since the heteroclinic trajectory is the minimum-fuel solution.

The next steps in the process will be evaluating the pseudospectral method in a variety of scenarios in the CR3BP. This evaluation will require testing different initial

guesses, testing "nearby" trajectories, applying a compound objective functional, and comparing the results to a 2BP benchmark.

The initial guesses used in the current research are found in Section 3.5. The two initial guesses utilized will be called the minimum-fuel initial guess, which is a heteroclinic trajectory, and a minimum-time-initial guess. For details on the initial guesses, see Section 3.5.

"Nearby" trajectories will be evaluated to determine the robustness of the pseudospectral method. Initial guesses from the original orbits will be used to determine if the pseudospectral method is robust enough to find an optimal result. "Nearby" trajectories are discussed in detail in Section 3.6.

A compound objective functional is an objective functional that takes the form

$$J = \alpha f_1 + (1 - \alpha) f_2 \tag{3.1}$$

where f_1 and f_2 are different functions and $\alpha \in [0, 1]$ is the balance of each function. In the current research, f_1 and f_2 will be functions that represent the minimum-fuel and minimum-time, respectively. The compound objective function is discussed in detail in Section 3.7. The compound objective functional will also test the robustness of the pseudospectral method as the balance of the functions within the cost functional change.

Finally, the results of the pseudospectral method in the CR3BP will be compared to a 2BP benchmark to determine the benefits of analysis in the CR3BP and to ensure the results from the pseudospectral method are reasonable. The 2BP benchmark is discussed in Section 3.8. A summary of the specific tests to be run are given in Table 3.1.

Table 3.1. Test Plan Overview

Test	Initial Guess	α in Objective Function	Orbit
1	Min-Fuel (Heteroclinic Trajectory)	$\alpha = 0$	Original
2	Min-Fuel (Heteroclinic Trajectory)	$\alpha = 0$	Nearby
3	Min-Fuel (Heteroclinic Trajectory)	varies	Original
4	Min-Time	varies	Original
5	Min-Fuel (Heteroclinic Trajectory)	varies	Nearby
6	Min-Time	varies	Nearby

3.3 Selection of Parameters

An important decision in the research was the parameters used to characterize the satellite's thruster performance. Low-thrust capabilities are suited for the CR3BP environment because low-thrust requires a long time period to accelerate and is therefore better suited for missions of longer length [49], The heteroclinic trajectories in the CR3BP are low-fuel trajectories at the expense of longer time of flights, which is similar to the advantages for low-thrust capabilities. This makes the low-thrust capabilities and the CR3BP well-suited problems to combine. The specific impulse, I_{sp}, was selected to provide an exhaust velocity that is typical of electrostatic and electromagnetic vehicles [49]. The exhaust velocity, c, is equal to the thruster specific impulse multiplied by the gravitational acceleration at sea level, g_0 [21] as follows

$$c = I_{sp} g_0. \tag{3.2}$$

To characterize the satellite's thruster, a thrust level must be selected. In electric propulsion, the value of the thrust is usually low, between 0.005 N to 1 N [49]. The current research allows the thrust to change with a throttle, but a maximum thrust

must be defined as seen in Table 3.2.

Finally, the mass of the spacecraft must be defined. The mass of the spacecraft will affect how fast the spacecraft can accelerate. The mass was selected to be 1500 kg. The value of the specific impulse, thrust, and mass in the current research are summarized in Table 3.2.

Table 3.2. Selected Parameter Values

Parameter	Symbol	Selected Value
Specific Impulse	I_{sp}	3000 s
Gravitational Acceleration at Sea Level	g_0	9.80665 m/s^2
Max Thrust	T_{max}	1 N
Satellite Mass	m	1500 kg

3.4 Equations of Motion

While the EOMs are defined for the CR3BP in Section 2.3.2, the specific form of the equations must be formulated to incorporate outside factors, such as problem scaling, control input, and mass loss. The way outside factors are incorporated into the optimal control problem can affect the rate of convergence and the ability for the optimization problem to be solved.

To begin this process, the method to incorporate the control into the equation of motion was evaluated. Section 2.2.1, shows how outside accelerations can be incorporated into the motion as a perturbation as long as they satisfy the fundamental assumption of perturbation theory. Continuous low-thrust can be accounted for as a perturbation in the CR3BP, since it is assumed to be a small effect, which obeys the fundamental assumption of perturbation theory.

The form that the control input takes is also an important aspect of setting up the EOMs. In the current research, the problem has been simplified to a planar

problem, so the control only has two-dimensions, x and y. The first form of the control attempted to write the control as an angle α as seen in Fig. 3.1.

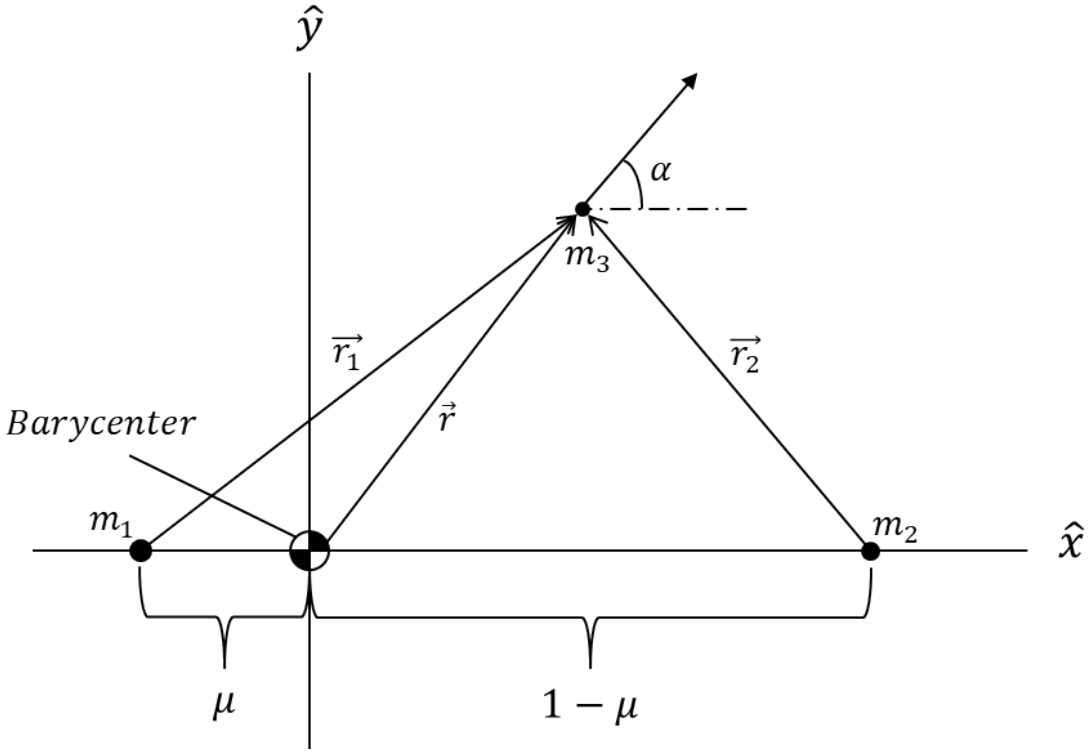


Figure 3.1. Form of the Control as an Angle α

This meant the control only had one parameter

$$u = \alpha, \tag{3.3}$$

where $\alpha \in [0, 2\pi]$.

The EOMs took the form

$$\ddot{x} = 2\dot{y} + x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} + \frac{T_{max}}{m}\cos(\alpha) \tag{3.4}$$

$$\ddot{y} = -2\dot{x} + y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} + \frac{T_{max}}{m}\sin(\alpha), \tag{3.5}$$

where T_{max} is the maximum thrust magnitude and m is the mass of the spacecraft.

The form of the control in Eq. (3.3) seemed the easiest form to utilize, since there is only one control parameter, which reduced the required number of parameters to solve the problem. In practice, this control form's main issue was the thrust magnitude could not be reduced, so the control always had to have the thrusters continually burning. This made it difficult to arrive at the selected final state because the control would constantly flip directions as the spacecraft approached the end of the trajectory to try and slow the motion of the spacecraft.

To counteract this issue, the control was changed to incorporate a control component in the x- and y- direction as follows

$$\vec{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \tag{3.6}$$

where u_x and $u_y \in [0, 1]$. The control can be seen in Fig. 3.2.

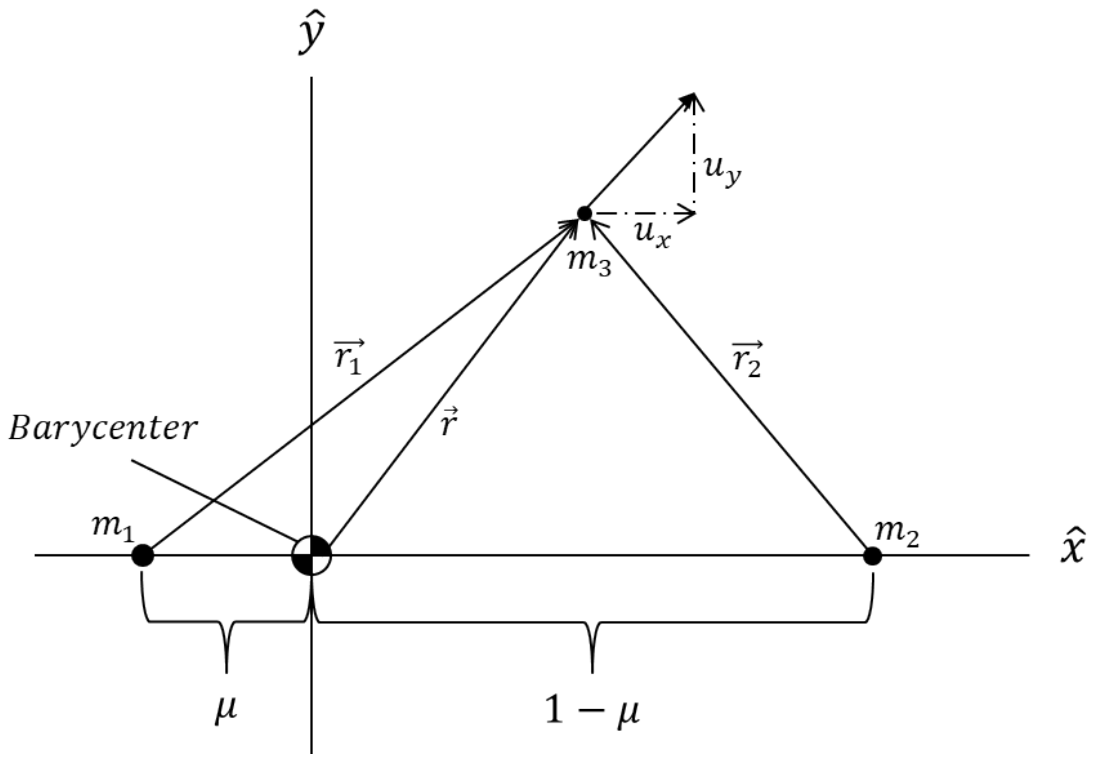


Figure 3.2. Form of the Control in the x- and y-direction

To ensure the magnitude of the control stayed between $[0, 1]$, a path constraint was introduced to force the magnitude of the control to always be less than or equal to 1 as follows

$$\sqrt{u_x^2 + u_y^2} \leq 1. \tag{3.7}$$

The EOMs for the control in Eq. (3.6) are written as

$$\ddot{x} = 2\dot{y} + x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} + \frac{T_{max}}{m}u_x \tag{3.8}$$

$$\ddot{y} = -2\dot{x} + y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} + \frac{T_{max}}{m}u_y. \tag{3.9}$$

The new control in Eq. (3.6) was intended to allow the magnitude of the thrusters

on the spacecraft to be reduced. Thus, the thrusters would not have to constantly be going full throttle. However, the pseudospectral method did not converge in some scenarios. This is potentially due to the path constraint on the control. A final form of the control was found by Zhang [21]. Rather than incorporate the throttle as a path constraint, the problem was solved by adding the throttle directly into the control as follows

$$\vec{u} = \begin{bmatrix} u_x \\ u_y \\ \gamma \end{bmatrix}, \tag{3.10}$$

where $\gamma \in [0, 1]$ is the throttle. This form of the control is the same as in Fig. 3.2 with γ equal to the length of the thrust vector.

This made the EOMs

$$\ddot{x} = 2\dot{y} + x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} + \gamma\frac{T_{max}}{m}u_x \tag{3.11}$$

$$\ddot{y} = -2\dot{x} + y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} + \gamma\frac{T_{max}}{m}u_y. \tag{3.12}$$

The advantage to the control taking the form in Eq. (3.10), is the control form removes the path constraint and also allows the throttle to be used as a metric for calculating mass loss in an objective function for minimum-fuel optimal control problems. A final equation is included to account for mass loss. Some authors, such as Caillau, decided against incorporating mass loss into the EOMs assuming mass loss to be negligible [50]. Others, such as Zhang, determined the mass loss will affect the acceleration of the spacecraft and produce a different trajectory [21]. The EOMs for mass loss is taken from [21] and is given as

$$\dot{m} = -\gamma \frac{T_{max}}{c}, \tag{3.13}$$

where c is the exhaust velocity. The exhaust velocity can be calculated as $c = I_{sp} g_0$.

Equations (3.11) - (3.13) make up the EOMs used in the current research with the control being of the form given in Eq. (3.10). However, the EOMs still need to be scaled to be effectively utilized.

The final issue dealt with in developing the EOMs is the scaling of the problem. In direct optimal control it is important to keep values scaled to relatively the same order of magnitude and the closer to the order of $\mathcal{O}(1)$ is beneficial. By scaling in to the $\mathcal{O}(1)$, higher order terms do not make lower order terms insignificant. Most of the terms in Eq. (3.11) and (3.12) are already nondimensionalized as seen in Section 2.3.1 and remain relatively close to being of $\mathcal{O}(1)$. However, $\frac{T_{max}}{m}$ is not nondimensionalized, meaning the value could affect the scaling of the problem.

The unscaled version of Eq. (3.11)-(3.13) were evaluated and were only capable of converging on solutions where there was a high T_{max}, or when no control was needed, such as in the case of finding a heteroclinic connection. To resolve this issue, $\frac{T_{max}}{m}$ was nondimensionalized to match the scaling of the CR3BP. Since T_{max} has units of Newtons and m has units of kg, the entire term $\frac{T_{max}}{m}$ has units of m/s^2. To nondimensionalize $\frac{T_{max}}{m}$, the characteristic quantities l^* and t^* were used as follows

$$\ddot{x} = 2\dot{y} + x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} + \gamma \frac{T_{max}}{m} \frac{{t^*}^2}{c_1} u_x \tag{3.14}$$

$$\ddot{y} = -2\dot{x} + y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu y}{r_2^3} + \gamma \frac{T_{max}}{m} \frac{{t^*}^2}{c_1} u_y. \tag{3.15}$$

where $c_1 = 1000 l^*$. The characteristic length l^* is multiplied by 1000 to convert it from units of km to m. By performing this nondimensionalization, all terms in Eq.

(3.14) and (3.15) have units of the same order of magnitude.

Equation (3.13) still needs to be scaled to the nondimensionalized units of the CR3BP. Even though $\dot{m}$ has units of kg/s, getting rid of the kg is not necessary and would actually affect Eq. (3.14) and (3.15) due to the fact that mass was assumed to have units of kg. Therefore, Eq. (3.13) only needs to have the time nondimensionalized by using the characteristic time t^* as follows

$$\dot{m} = -\gamma \frac{T_{max}}{c} t^*. \tag{3.16}$$

Equations (3.14) - (3.16) represent the dynamics used in the current research. They are nondimensionalized properly so that all the necessary terms remain near $\mathcal{O}(1)$.

3.5 Initial Guess in GPOPS

Direct pseudospectral methods require an initial guess. The accuracy required of the initial guess differs between optimization methods as some methods are more robust to initial guesses than others. However, a more accurate initial guess can help converge on an accurate solution. Insight into the CR3BP helps attain an accurate initial guess that can be beneficial in evaluating the pseudospectral method. A heteroclinic connection was described in Section 2.8 as a "free" transfer between two periodic orbits. Since the heteroclinic connection is a "free" transfer, it is a known minimum-fuel solution and makes a good initial guess to use for a minimum-fuel trajectory.

The details behind finding a heteroclinic connection are explained in Sections 2.6-2.8, and are summarized here for completion. To begin the process of finding a heteroclinic connection, the initial and final periodic orbit must be selected. These orbits can be between any locations in the CR3BP, but are required to be the same

Jacobi constant. The orbits chosen for the current research were taken to match the orbits from [5]. The two orbits selected are an orbit about the L1 libration point in the Earth-Moon system and a Distant Prograde Orbit (DPO) about the Moon. The initial conditions used for for the two orbits is given in Table 3.3. Both of these orbits have a Jacobi constant of 3.1331.

Table 3.3. Initial Conditions for L1 Periodic Orbit and DPO

	L1 Periodic Orbit	DPO
x (DU)	0.812255	1.061692
y (DU)	0	0
$\dot{x}$ (DU/TU)	0	0
$\dot{y}$ (DU/TU)	0.248312	0.403877

Using differential corrections described in Section 2.6.2 and the initial conditions from Table 3.3, periodic orbits are generated. A zoomed in version of the L1 periodic orbit and the DPO can be seen in Fig. 3.3. Since the L1 periodic orbit and the DPO orbit will be used throughout the current research, a view of the orbits in the entire CR3BP synodic reference frame is shown in Fig. 3.4. The orbits are also shown in the inertial are in Fig. 3.5. The L1 Periodic Orbit is the starting orbit, while the DPO is the ending orbit.

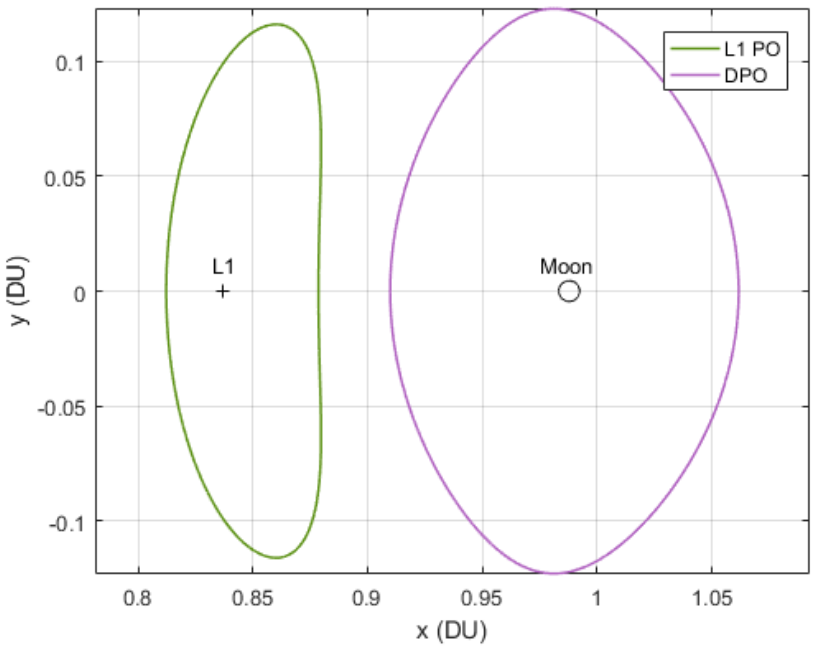


Figure 3.3. Periodic Orbits in the CR3BP

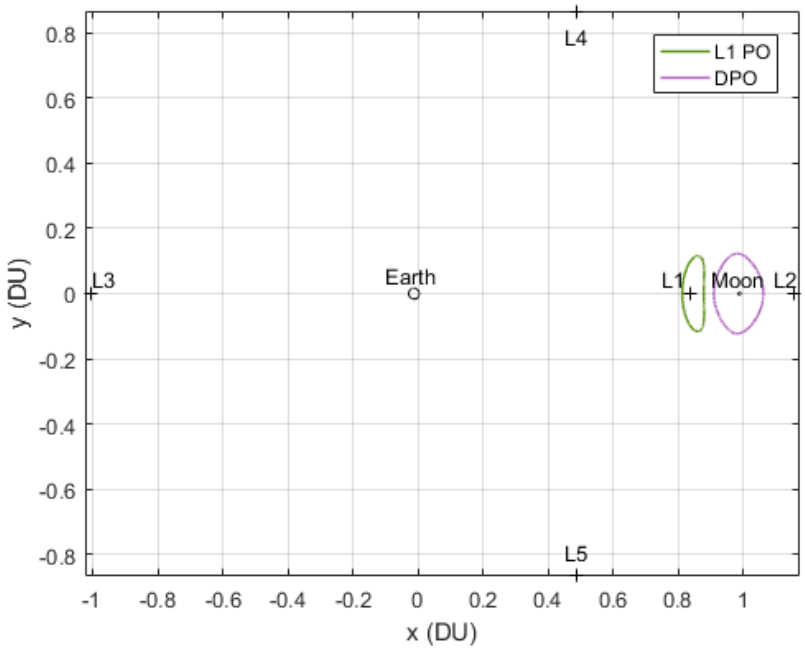


Figure 3.4. Periodic Orbits in the Entire CR3BP Synodic Reference Frame

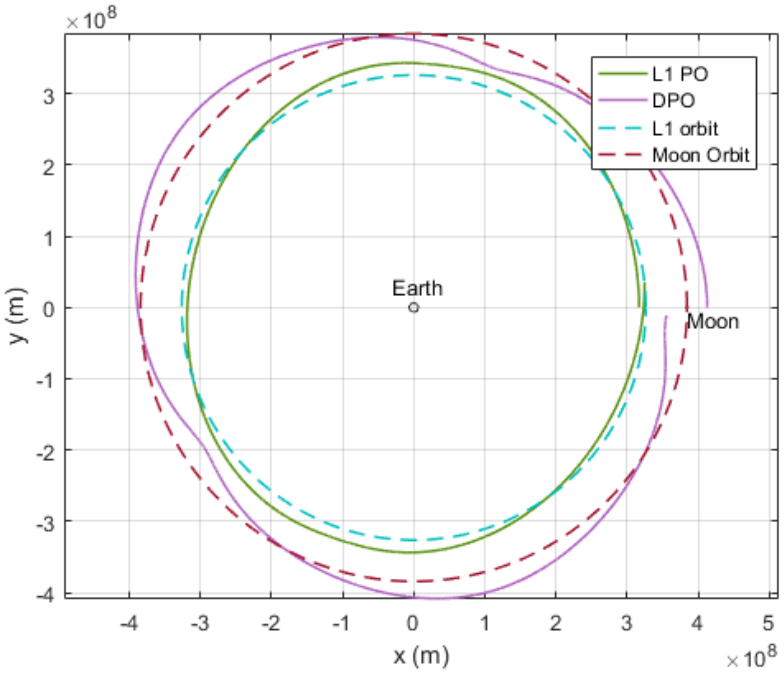


Figure 3.5. Periodic Orbits in the Inertial Reference Frame

With the periodic orbits selected, the invariant manifolds associated with the orbits must be propagated. As discussed in Section 2.7, the orbit is perturbed in the direction of the stable or unstable eigenvector at selected points around the orbit. This creates a manifold tube as seen in Fig. 3.6. The manifold shown in Fig. 3.6 is the unstable manifold of the L1 periodic orbit perturbed in the direction of the unstable eigenvector in both the positive and negative direction. The negative direction is only propagated to the surface of section and is thus only shown on Fig. 3.6 at values of $x > 0.8$.

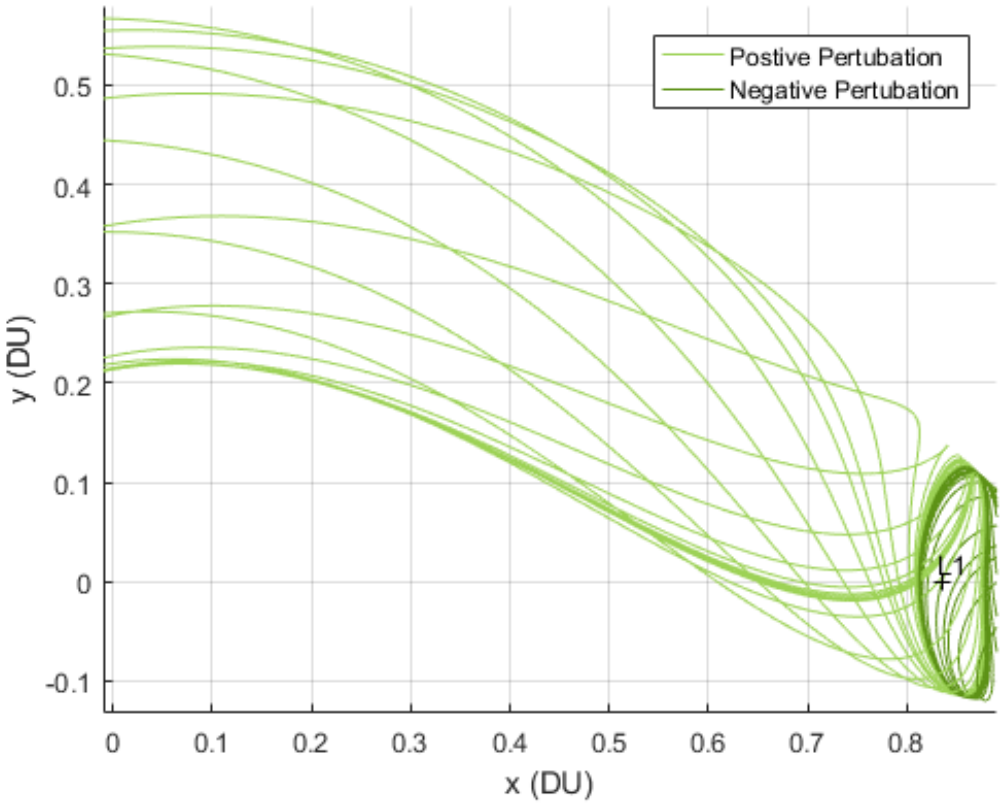


Figure 3.6. L1 Periodic Orbit Unstable Manifold from Perturbations in both the Positive and Negative Direction

The manifold tube will need to be created for both the L1 periodic orbit and the DPO. However, since the L1 periodic orbit is the starting orbit, only the unstable manifold needs to be found as unstable manifolds depart from the orbit. Similarly, the DPO is the final orbit and only needs to have the stable manifold propagated. When a connection is found between the unstable manifold of the L1 periodic orbit and the stable manifold of the DPO, a spacecraft traveling from the L1 periodic orbit along the unstable manifold will be able to transfer to the stable manifold and arrive at the DPO.

To find a connection between manifolds, a surface of section is chosen at a selected x-value. In the current research, the surface of section is chosen at $x = 0.890940$ to

match the value chosen in [5]. A surface of section is discussed in Section 2.8. The manifolds are propagated until they intersect this surface. The L1 periodic orbit manifolds in Fig. 3.6 only need to be propagated in the direction of the negative perturbation to intersect the surface of section. The propagation of the stable and unstable manifolds to the surface of section can be seen in Fig. 3.7.

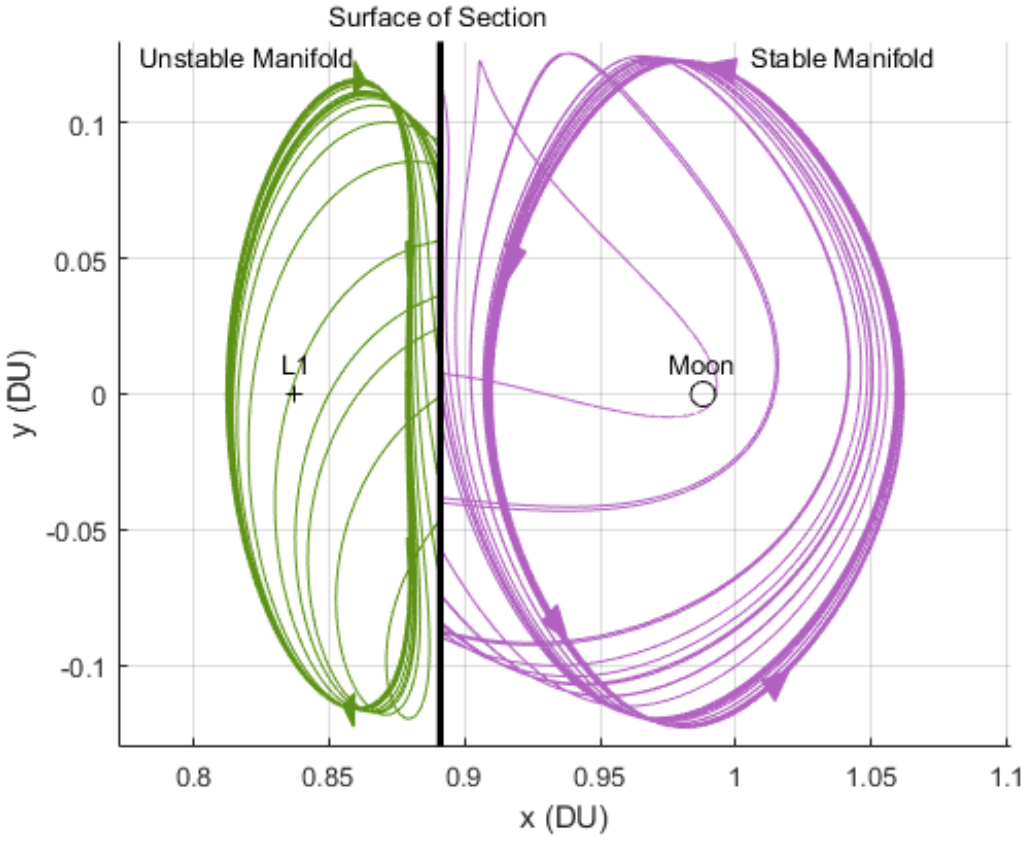


Figure 3.7. Unstable and Stable Manifolds Propagated until Intersection with the Surface of Section

At the intersection of the surface of section, the manifolds are guaranteed to have the same x-value. Two of the three remaining states must be chosen to be plotted on a Poincaré map as described in Section 2.8. A two-sided Poincaré map is utilized to show the intersection of the manifolds from either direction. The two states selected to be plotted on the Poincaré map are y and $\dot{y}$. The Poincaré map from the periodic

orbits can be seen in Fig. 3.8. Each point on the map represents the y- and $\dot{y}$-values of a single manifold as it intersects the surface of section.

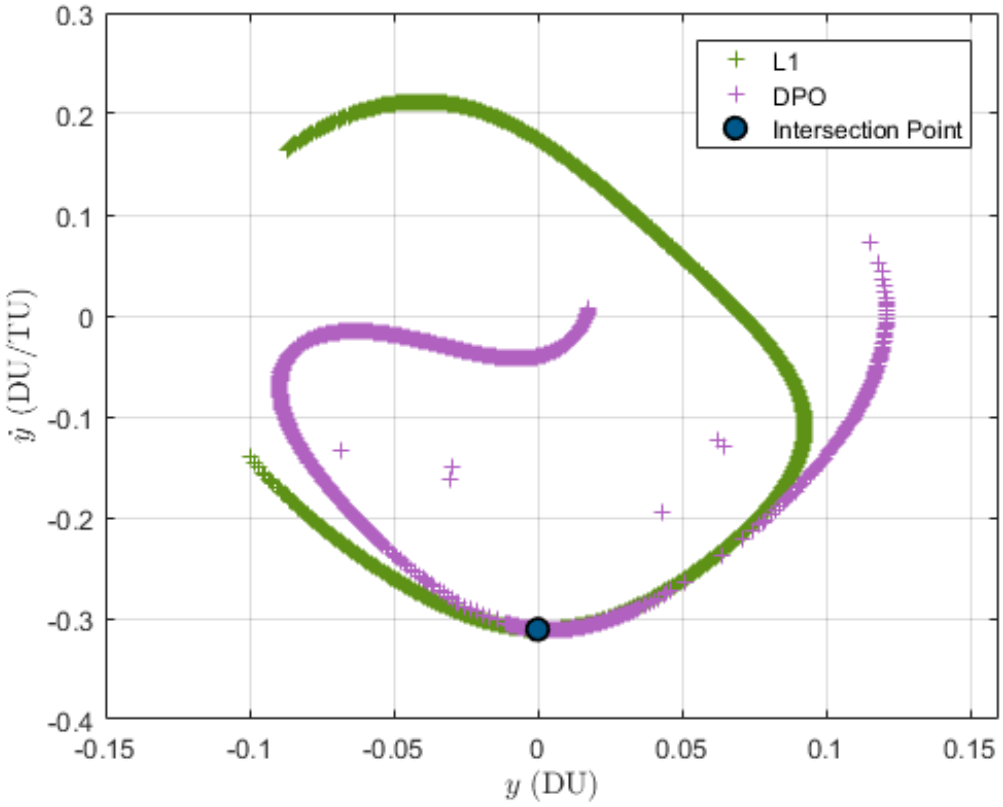


Figure 3.8. Poincaré Map Showing the Intersection Point between the stable and Unstable Manifold

Any points on the Poincaré map that overlap guarantee that the stable and unstable manifold intersecting the surface of section at that point have the same state. The x-values are the same due to being on the surface of section; the y- and $\dot{y}$-values are the same due to intersecting on the Poincaré map; and the $\dot{x}$-values are guaranteed to be aligned due to the orbits having a shared Jacobi constant. Since all the values of the state are the same, a heteroclinic connection has been found. The closest point from the stable and unstable manifold on the Poincaré map in Fig. 3.8 is labeled as the intersection point. There are many points that are close to overlapping, and any

of these points could be labeled as a heteroclinic connection depending on the defined tolerance. The point selected as the heteroclinic connection in the current research was the point where two manifolds were the closest out of all the intersection with the surface of section. After, finding the heteroclinic connection, the unstable and stable manifold that intersected on the Poincaré map are the manifolds that comprise the heteroclinic trajectory. The heteroclinic trajectory is seen in Fig. 3.9.

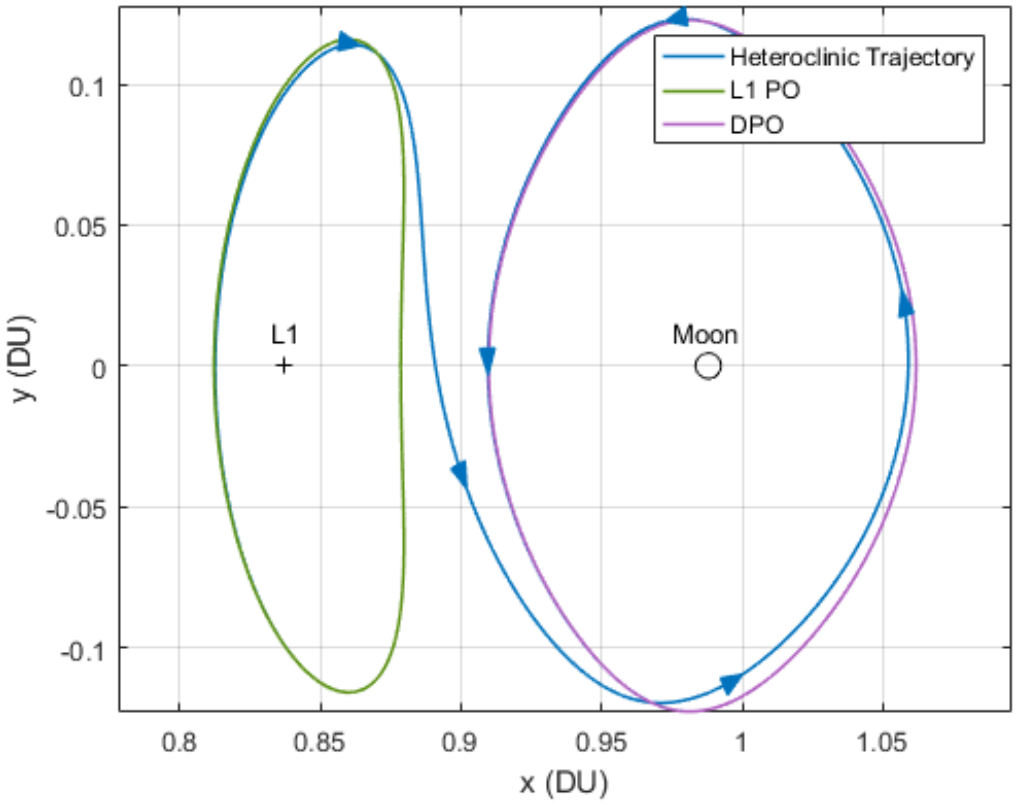


Figure 3.9. The Heteroclinic Trajectory Found between the L1 Periodic Orbit and the DPO

The heteroclinic trajectory is a "free" transfer from the L1 periodic orbit to the DPO. Since the heteroclinic trajectory is the known minimum-fuel trajectory, it is used as the initial guess for the optimal control problem. The heteroclinic connection only found the required states needed for the optimal control problem, so the control

initial guess still needs to be defined. Since this is the min-fuel solution, there should be minimal throttle from the control, γ. The direction of the control, u_x and u_y, in the initial guess does not matter, since there is minimal throttle. Therefore, the throttle in the control, γ, is assigned to be zero for the entire trajectory and the direction of the control, u_x and u_y, are arbitrarily assigned, since u_x and u_y will have negligible impact on the resulting trajectory.

There are two initial guesses used in the current research. The first initial guess is the heteroclinic trajectory, which will be called the minimum-fuel initial guess. Another initial guess is also used, but is not found from insight into the CR3BP. The other initial guess is a resulting trajectory found in the process of applying a compound objective functional to balance the weights of minimum-time versus minimum-fuel. The best way to describe this initial guess is the minimum-time solution for transfer between the periodic orbits in Fig. 3.3. Since this initial guess cannot be derived from insight into the CR3BP, the initial guess will be introduced in the results section. The second initial guess will be called the minimum-time initial guess. There are now two initial guesses: the minimum-fuel initial guess, which is a heteroclinic trajectory, and a minimum-time initial guess. Other initial guesses were attempted, such as using an iterative process, where the previous result is used as the next initial guess. However, the iterative initial guess was problematic due to issues with a single orbit arriving at a local minimum and causing all successive orbits to also arrive at a local minimum or failing to converge altogether. Thus, the iterative initial guess was no longer utilized as an initial guess, leaving only the minimum-fuel and the minimum-time initial guesses.

3.6 Continuation Method

In Section 3.5, the initial conditions for the L1 Periodic Orbit and the DPO are given. However, these are not the only two orbit evaluated in the current research. Orbits "nearby" to the original L1 Periodic orbit are also analyzed. All "nearby" orbits are in the same family of orbits as the original L1 periodic orbit and each "nearby" orbit has a different Jacobi constant. In the current research, "nearby" orbits are defined as within 0.0307 DU or 11,784 km of the original orbit. This distance is slightly smaller than the diameter of the Earth. The distance was selected because the L1 periodic orbits in Fig. 3.10 begin to significantly overlap the DPO beyond this distance.

To find the nearby orbits, a continuation method was used. A continuation method calculates a family of orbits by varying the initial state of a previous periodic orbit slightly in a certain direction and performing the differential corrections on the new initial conditions [4]. This process can be iterated for multiple orbits to develop a family of orbits. In the current research, the original orbit is varied by Δx, such that the new state is $\vec{X} = [x + \Delta x, y, z, \dot{x}, \dot{y}, \dot{z}]^T$. This process is iterated until changing the Δx moves the orbit outside 0.0307 DU from the original orbit. The continuation method and the "nearby" orbits used in this thesis are shown in Fig. 3.10.

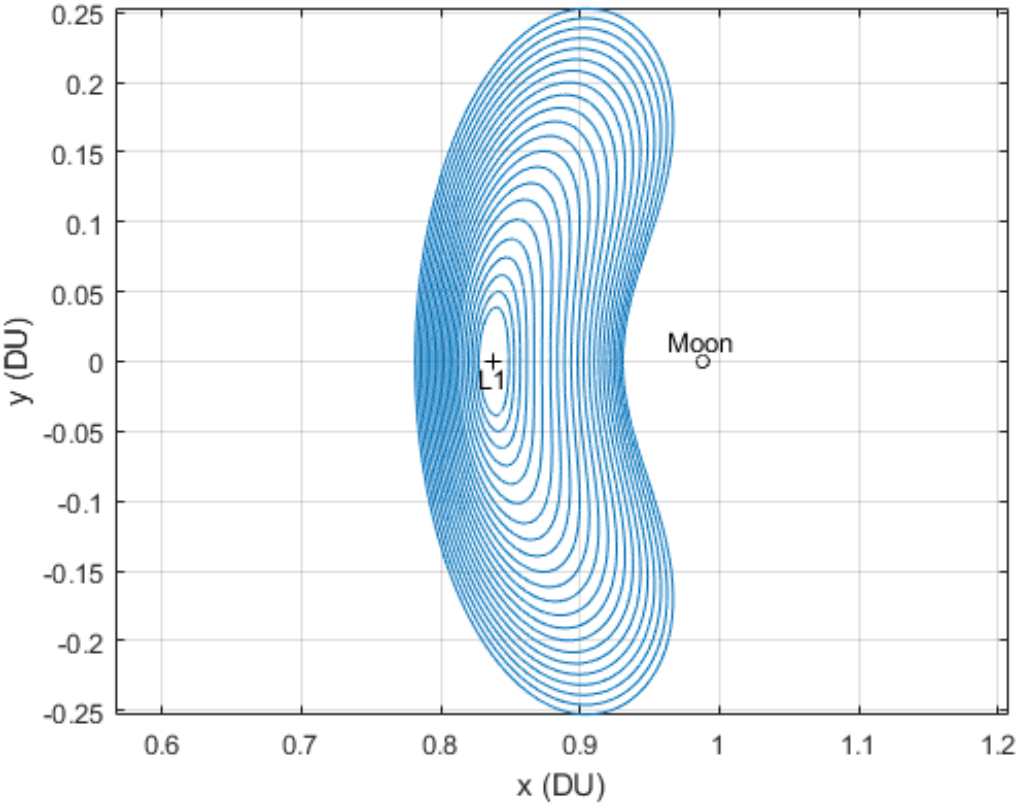


Figure 3.10. The Family of Orbits Produced from the Continuation Method

3.7 Objective Functions

The objective function in the current research is actually a compound objective functional. The form of the compound objective functional is given in Eq. (3.1). The compound objective functional allows different weights to balance the cost functional between two functions. In the current research, the two functions being balanced are fuel and time.

The minimum-fuel cost function is defined as

$$J = \frac{T_{max}}{c} \int_{t_0}^{t_f} \gamma dt \tag{3.17}$$

and the minimum-time cost function is defined as

$$J = \int_{t_0}^{t_f} dt. \tag{3.18}$$

To combine Eq. (3.17) and (3.18) into one compound objective functional, a weighting factor, α, must be introduced. α will determine if the compound objective functional is balanced toward minimum-fuel or minimum-time. Since the $\frac{T_{max}}{c}$ in Eq. (3.17) will affect the weighting in the problem, $\frac{T_{max}}{c}$ will be removed since it is simply multiplying the minimum-fuel cost function by a constant. The compound objective functional then becomes

$$J = \int_{t_0}^{t_f} \left[\alpha + (1-\alpha)\gamma\right] dt. \tag{3.19}$$

The compound objective functional can be varied to balance between minimum-fuel and minimum-time by changing the value of α. If $\alpha = 0$, then the compound objective functional is the minimum-fuel solution. If $\alpha = 1$, then the compound objective functional is the minimum-time solution. This compound objective functional can be used to evaluate the trade-off between minimum-fuel trajectories and minimum-time trajectories.

3.8 2BP Benchmark

In Section 3.2, a 2BP comparison was listed as part of the test plan. The most basic Three-Body Problem (3BP) comparison is the optimal maneuver between two circular orbits using an impulsive maneuver, called a Hohmann Transfer [2]. A Hohmann transfer is performed by making two impulsive burns in the trajectory. The first burn occurs to take the spacecraft out of the starting circular orbit and place the spacecraft on an elliptical transfer orbit. The second burn is used to re-circularize

the orbit upon arriving at the second orbit. Since the orbits used in the current research are not circular when converted to the 2BP, an exact comparison between trajectories in the 2BP and the CR3BP is difficult. To simplify the comparison, two circular orbits will be used in the analysis. This basic approximation will at least be able to demonstrate if the CR3BP trajectories are generally performing better than the 2BP trajectories.

The circular orbits used in the approximation will be the location of the L1 point and the location of Moon when these locations are converted to an Earth-Centered Inertial (ECI) reference frame. The process for converting from the synodic reference frame to the ECI frame is described in Section 2.3.3.As seen in Fig. 3.5, the L1 orbit and the moon's orbit in the inertial frame are approximately equivalent to the L1 periodic orbit and the DPO used in the current research.

To calculate the Δv required to perform the Hohmann transfer, the following equation is used [2]

$$\Delta v = \Delta v_1 + \Delta v_2 \tag{3.20}$$

where

$$\Delta v_1 = \sqrt{\frac{2\mu}{r_1} - \frac{2\mu}{r_1 + r_2}} - \sqrt{\frac{\mu}{r_1}} \tag{3.21}$$

and

$$\Delta v_2 = \sqrt{\frac{\mu}{r_2}} - \sqrt{\frac{2\mu}{r_2} - \frac{2\mu}{r_1 + r_2}}. \tag{3.22}$$

Equation (3.21) shows that Δv_1 is the first burn to put the spacecraft on the transfer trajectory and Eq. (3.22) shows that Δv_2 is the second burn to re-circularize the orbit at the final orbit. The distances r_1 and r_2 are the radius of the first and

second circular orbits, respectively, and μ is the gravitational parameter for the two-body problem, which is different from the μ used in the CR3BP.

Transfer time is also an important comparison metric used in the current research. For a Hohmann transfer, the transfer time is [2]

$$\Delta t = \pi \sqrt{\frac{(r_1 + r_2)^3}{8\mu}} \tag{3.23}$$

3.9 Summary

Chapter 3 discussed the research methodology used in the current research. A test plan was introduced to describe the scenarios that will be used to analyze the pseudospectral method. The parameters, EOMs, and controls specific to the current research were established. An initial guess was generated by finding a heteroclinic trajectory. Finally, the "nearby" periodic orbits, the compound objective functional, and a 2BP benchmark were presented.

IV. Results

4.1 Chapter Overview

Chapter 4 performs the test plan developed in Section 3.2. First, the heteroclinic connection will be given as an initial guess to validate that the pseudospectral method can find a known minimum-fuel solution in the Circular-Restricted Three-Body Problem (CR3BP). The research will then examine the robustness of the pseudospectral method for "nearby" periodic orbits using the heteroclinic initial guess. Finally, a compound objective functional will examine the balance of min-time and min-fuel for the original periodic orbit and "nearby" periodic orbits. An evaluation of the initial guess and the pseudospectral method will be provided based on the results from the tests.

4.2 Heteroclinic Trajectory in General Purpose Optimal Control Software (GPOPS)

A heteroclinic trajectory is the known minimum-fuel result for the transfer between two periodic orbits of the same Jacobi Constant. As an initial test, the pseudospectral method was given the heteroclinic connection as an initial guess and the objective function was set to find the minimum-fuel solution. The result is shown in Fig. 4.1.

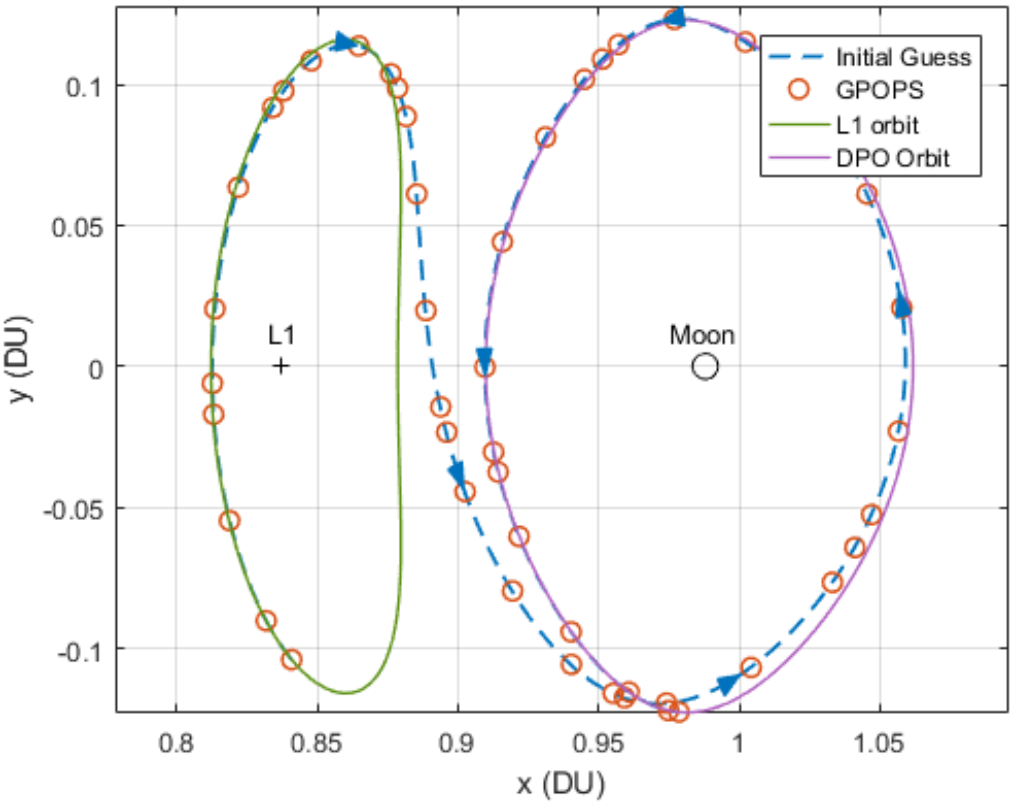


Figure 4.1. Heteroclinic Connection Found by Pseudospectral Method in GPOPS

As expected, the pseudospectral method successfully converges on the heteroclinic trajectory when solving for the minimum-fuel. While the convergence on the heteroclinic trajectory helps to validate the initial guess in GPOPS, the resulting control and Hamiltonian are also important to examine. The throttle in the control should be essentially zero since almost no Δv should be expended for a heteroclinic trajectory. There should also be negligible mass loss since the heterolcinic trajectory is a "free" transfer. Therefore, the Hamiltonian should be constant because the cost function and equations of motion (EOMs) do not explicitly contain time when mass loss is negligible. The control and Hamiltonian are shown in Fig. 4.2 for the minimum-fuel trajectory found by GPOPS in Fig. 4.1.

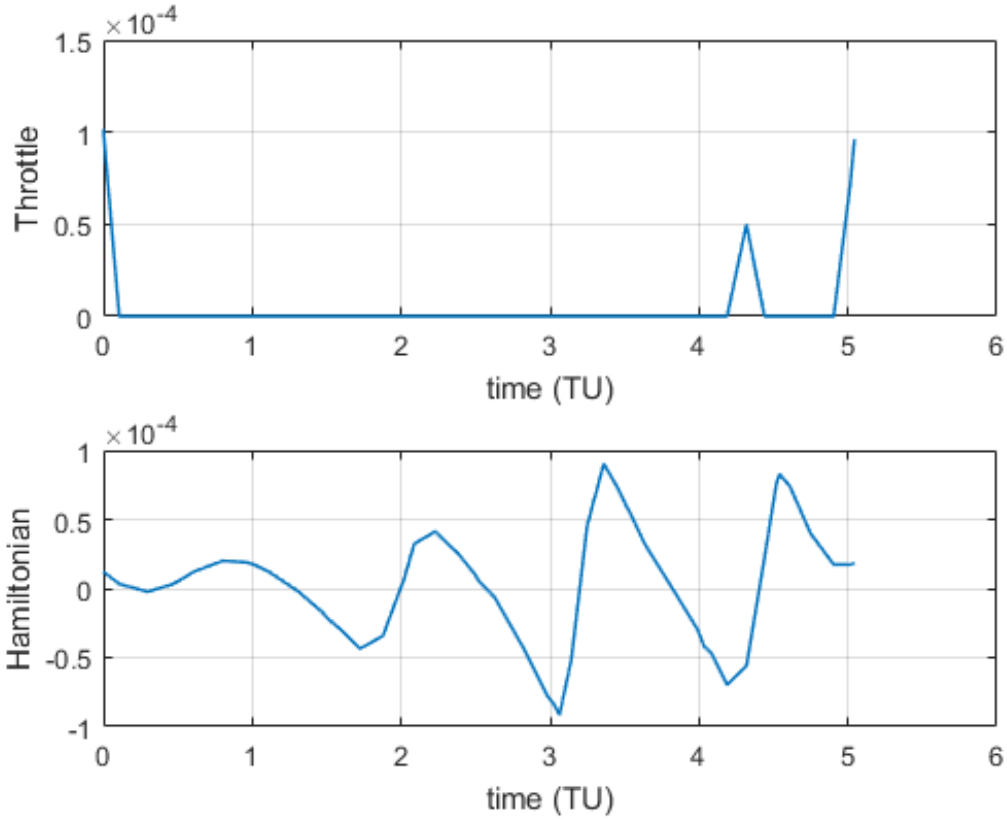


Figure 4.2. Control and Hamiltonian for the Heteroclinic Connection Found by GPOPS

In Fig. 4.2, the control and Hamiltonian are all of the order 10^{-4}. Thus, the control is practically zero and the Hamiltonian is very close to constant. The Δv expended in this maneuver is 0.8677 m/s, which is also very low. This demonstrates that the pseudospectral method used in GPOPS effectively determined the min-fuel solution when provided with an accurate initial guess. Further analysis will be done in the current research to determine the robustness of the pseudospectral method in the CR3BP.

4.3 Nearby Periodic Orbits

Determining the ability for the pseudospectral method to converge on an accurate optimal solution in different scenarios is important to evaluate the pseudospectral

method's robustness. The ability of the pseudospectral method to converge on an optimal solution can also be improved by providing an accurate guess. Thus, determining if the heteroclinic trajectory can be used as an initial guess for other scenarios, such as "nearby" orbits, will provide a valuable insight into effective initial guesses to ensure convergence of the pseudospectral method. The "nearby" orbits used in the current research are described in Section 3.6. Several of the minimum-fuel trajectories resulting from an initial guess of the heteroclinic connection are shown in Fig. 4.3.

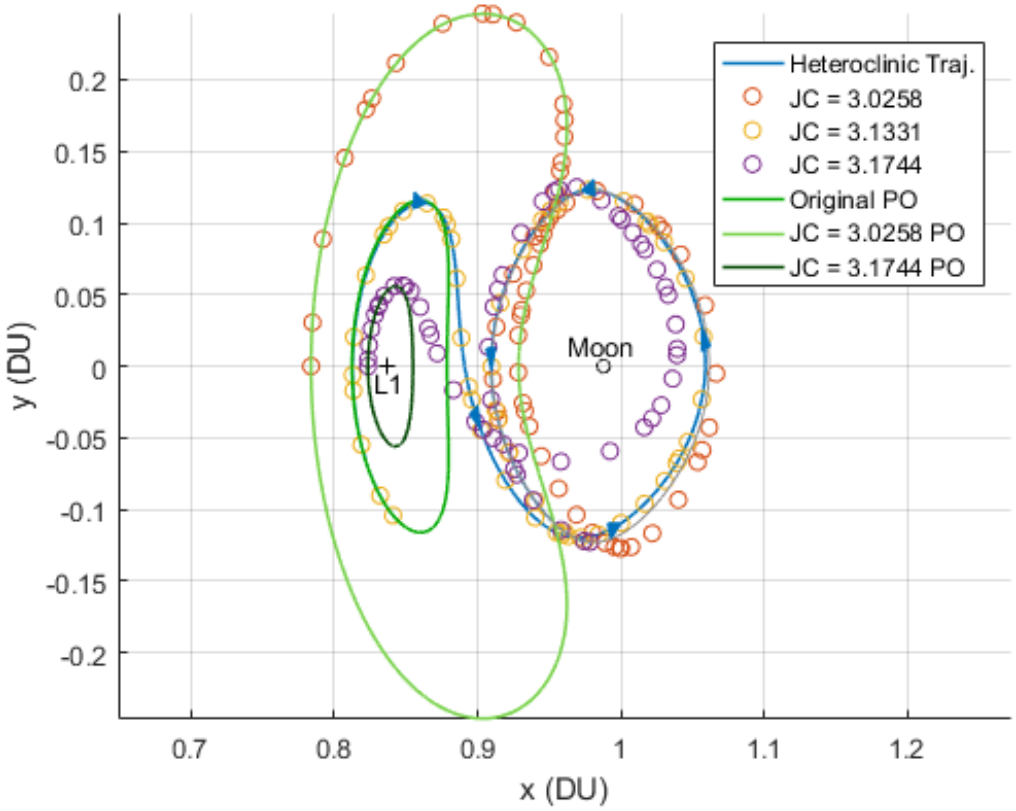


Figure 4.3. Min-Fuel Optimal Trajectories Found by GPOPS for "Nearby" Periodic Orbits of different Jacobi Constants

Each of the periodic orbits in Fig. 4.3 have different Jacobi constants. If the required Δv is compared to to the Jacobi Constant, a close to linear relationship appears as seen in Fig. 4.4.

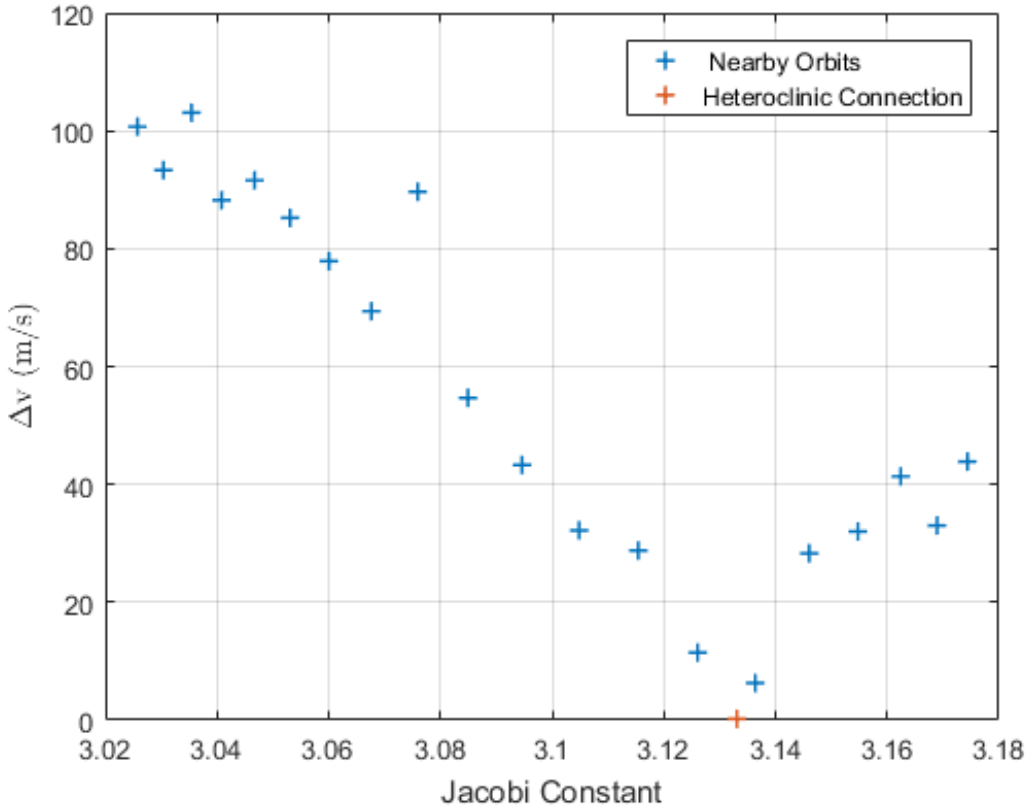


Figure 4.4. Optimal Δv Found by GPOPS for "Nearby" Periodic Orbits of different Jacobi Constants

Note that the Δv increases in either direction when moving away from the Jacobi constant of the heterolcinic trajectory. This makes intuitive sense because the further away from the "free" transfer an orbit is, the more expensive a transfer should be. However, an important aspect of Fig. 4.4 is that not all of the Δv's are on the line that would create a linear relationship. These points likely indicate that the optimal trajectory for that periodic orbit converged to a local minimum. A global minimum, or at least a lower local minimum, could likely be found that reduces the Δv required to make the transfer. This indicates that while the pseudospectral method being tested might be robust at converging on a solution, the solution may not be most optimal solution that could be found. Thus, it is possible that a better initial guess

may need to be provided to the pseudospectral method.

4.4 Compound Objective Function Initial Guesses

In this section, the compound objective functional

$$J = \int_{t_0}^{t_f} [\alpha + (1-\alpha)\gamma]\, dt, \tag{4.1}$$

introduced in Section 3.7 will be evaluated as the balance between min-fuel and min-time is tested by varying α between 0 and 1. The balance of min-fuel and min-time provides a unique scenario for the pseudospectral method to find an optimal solution. Comparing the balance of min-fuel and min-time will demonstrate the robustness of the pseudopsectral method. The robustness of the provided initial guesses will also be evaluated by comparing the outputs in different scenarios.

4.4.1 Compound Objective Functional: Original Periodic Orbit

Varying α from 0 to 1 in the compound objective functional, the optimal control problem is solved using the heteroclinic trajectory, the minimum-fuel solution, as the initial guess. The optimal trajectories generated by GPOPS are shown in Fig. 4.5.

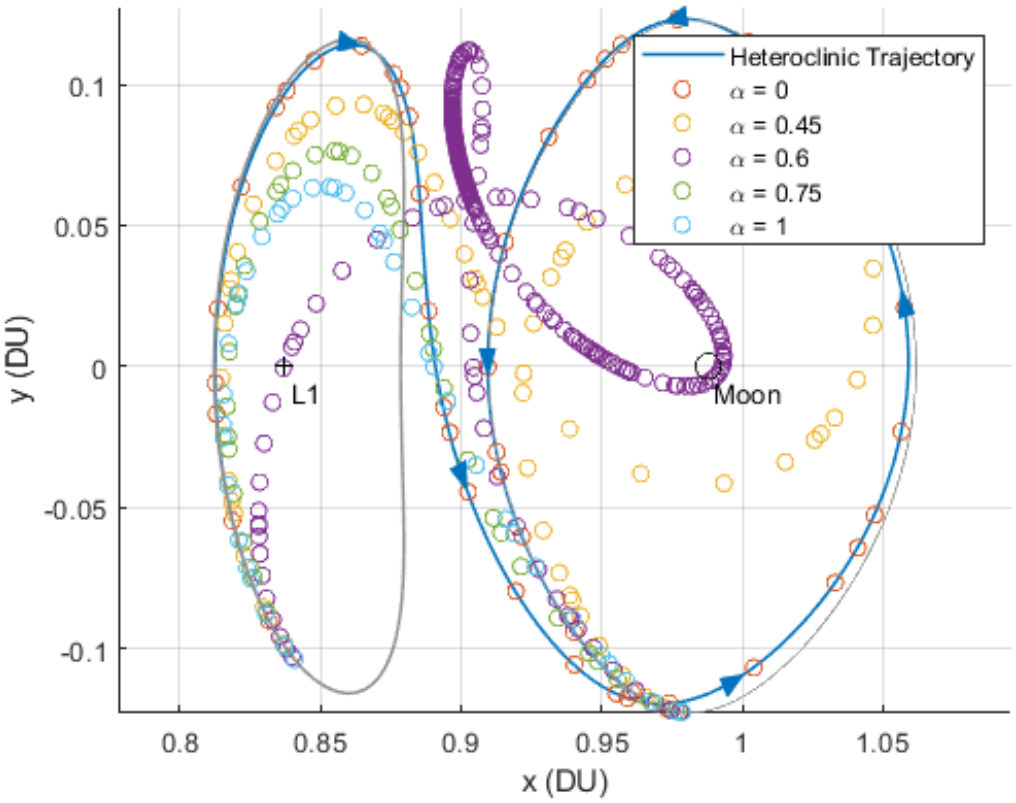


Figure 4.5. Optimal Trajectories Found in GPOPS by varying the balance of α between Min-Fuel and Min-Time with an initial guess of the heteroclinic trajectory

Figure 4.5 only shows selected trajectories from the varying of α between $[0, 1]$. Although it cannot be seen in Fig. 4.5, the pseudospectral method converged on the heteroclinic connection for $\alpha = [0, 0.4]$. After $\alpha = 0.4$, the trajectory follows a similar shape as the heteroclinic trajectory by traversing around the moon before reaching the final state. However, as α increases, the trajectory passes closer and closer to the Moon. This can be seen in the trajectories $\alpha = 0.45$ and $\alpha = 0.6$ in Fig. 4.5.

An interesting aspect of the pseudospectral method with the adaptive mesh is the adaptive mesh's ability to account for rapidly changing dynamics. When the trajectory at $\alpha = 0.6$ passes very close by the moon, which is an area of rapidly changing dynamics. In order for the dynamics to be approximated correctly near the

moon, many collocation points are used. A large number of collocation points allows a trajectory to still be found even though the dynamics are changing very rapidly.

Despite the pseudospectral method's ability to converge on solutions with rapidly changing dynamics, it is unable to converge at $\alpha = 0.65$. The balance between min-fuel and min-time is difficult for the pseudospectral method at this point. At $\alpha = 0.6$, the trajectory passes very close to the moon, but still passes around the moon before ending at the final state, whereas at $\alpha = 0.75$ a new shape for the trajectory is found, where the trajectory no longer passes around the moon and, instead, takes a more direct route to the final state. A bifurcation point appears to exist at $\alpha = 0.65$, where the shape of the trajectories change and create a new family of orbits. At this supposed bifurcation point, not only do the shape of the trajectories change, but also the required Δv and time of flight change. At $\alpha = 0.6$, $\Delta v = 359.05$ m/s and $t = 12.03$ days. At $\alpha = 0.75$, $\Delta v = 275.16$ m/s and $t = 8.96$ days. The new family of orbits after $\alpha = 0.75$ have lower Δv and time. Since the new family of orbits seems to have better characteristics, the new family of orbits would be a beneficial initial guess to reevaluate the balance between min-fuel and min-time. This new family of orbits could potentially have lower Δv and transfer time for all α if the pseudospctral method happened to converge on a local minimum when the heteroclinic connection was provided as the initial guess. Thus, the min-time solution, when $\alpha = 1$, was chosen as the initial guess to rerun the analysis.

With the minimum-time solution as the initial guess for the pseudospectral method, the optimal trajectories were found again. The optimal trajectories found from the min-time initial guess are shown in Fig. 4.6.

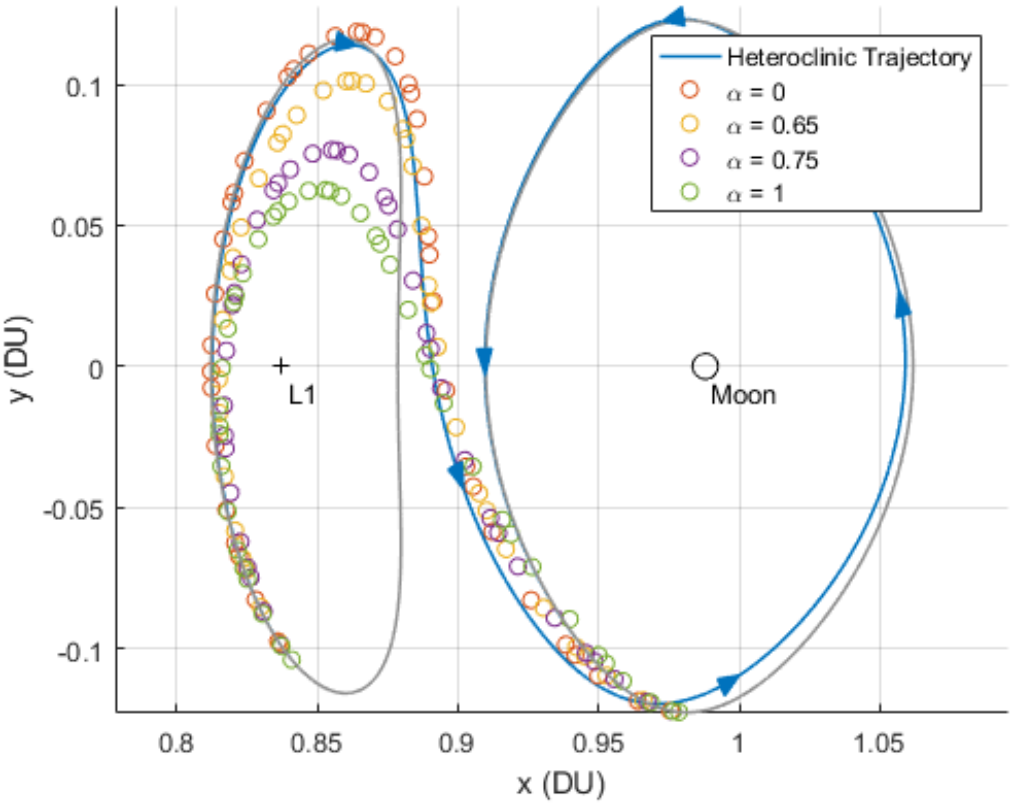


Figure 4.6. Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of Minimum-Time

The trajectories found from the min-time initial guess shown in Fig. 4.6 appear to be all of the same family of orbits. All of the orbits from the min-time initial guess appear to be of the same shape, where the trajectory goes directly to the final state without passing around the Moon. There also seems to be a consistent trade-off between required Δv and transfer time. Both initial guesses results are plotted in Fig. 4.7 to demonstrate the trade-off between Δv and transfer time. Values for α are shown in Fig. 4.7, so the associated trajectory can be seen in Fig. 4.5 and 4.6.

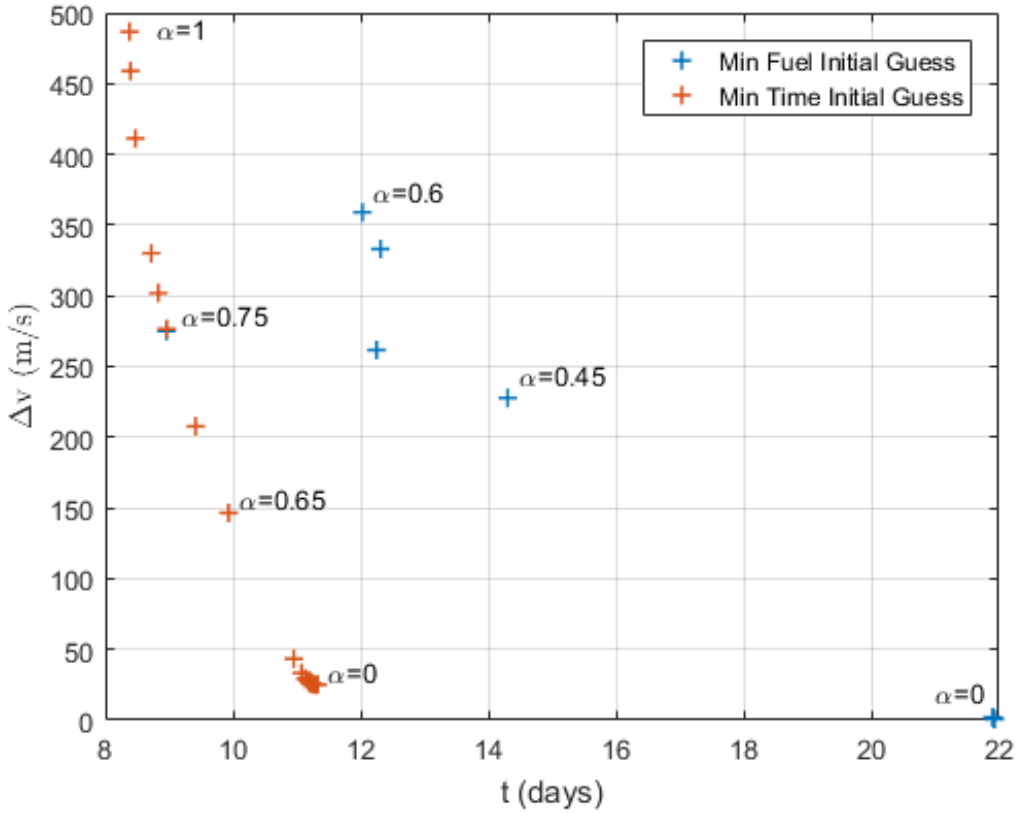


Figure 4.7. Trade-off between Δv and Transfer Time for Optimal Trajectories Found in GPOPS

The most beneficial solutions can be determined as the location where either Δv or time must be given up to reach another optimal solution, but not both. These solutions are called a Pareto Front. As can be seen in Fig. 4.7, the min-time initial guess has a consistent trade-off between Δv and transfer time, whereas the min-fuel initial guess lacks this structure. The min-time initial guess is more beneficial at each respective α. Thus, the min-time optimal trajectories lie on the Pareto Front. The only min-fuel trajectories that lies on the Pareto Front are at $\alpha = 0$ and for α's that are greater than 0.7. The trajectories found when $\alpha > 0.7$ for the min-fuel initial guess happen to be the same trajectories found from the min-time initial guess.

The min-time initial guess effectively finding an optimal solution, while the min-

fuel initial guess does not, provides insights into what makes a good initial guess for the pseudospectral method. While the pseudospectral method struggled to converge at certain values when given the minimum-fuel initial guess, it did not have issues when given the minimum-time initial guess. The min-time initial guess also resulted in more globally optimal trajectories for the majority of the values of α. The min-time solution appears to be the better initial guess for the compound objective function at all values of α, except for $\alpha = 0$.

When examining the Pareto Front in Fig. 4.7, the location of a comparable Two-Body Problem (2BP) transfer needs to be examined. The 2BP results are those of a Hohmann transfer as described in Section 3.8. The Hohmann transfer was found to have a $\Delta v = 79.14$ m/s and a transfer time of $t = 12.29$ m/s. From examining Fig. 4.7, the Hohmann transfer would not fall on the Pareto Front and would be a more expensive transfer in both Δv and transfer time as compared to an optimal solution on the Pareto Front in the CR3BP. This demonstrates the applicability and cost savings from exploiting the dynamics of the CR3BP and utilizing continuous low-thrust. The optimal solutions found from the pseudospectral method are also shown to be a more optimal solution, then the results from the 2BP.

4.4.2 Compound Objective Functional: Nearby Periodic Orbit

Due to the minimum-time initial guess appearing more robust in the pseudospectral method for the original orbit, it is beneficial to check these results in another scenario. Thus, the initial guesses are evaluated for a "nearby" orbit. Beginning with the min-fuel (heteroclinic trajectory) initial guess, the balance of min-fuel and min-time in the compound objective function is evaluated. The resulting trajectories can be seen in Fig. 4.8 and 4.9. The optimal trajectories are shown on two images because many of the trajectories had unique characteristics that could not be seen

when all the trajectories were on one plot.

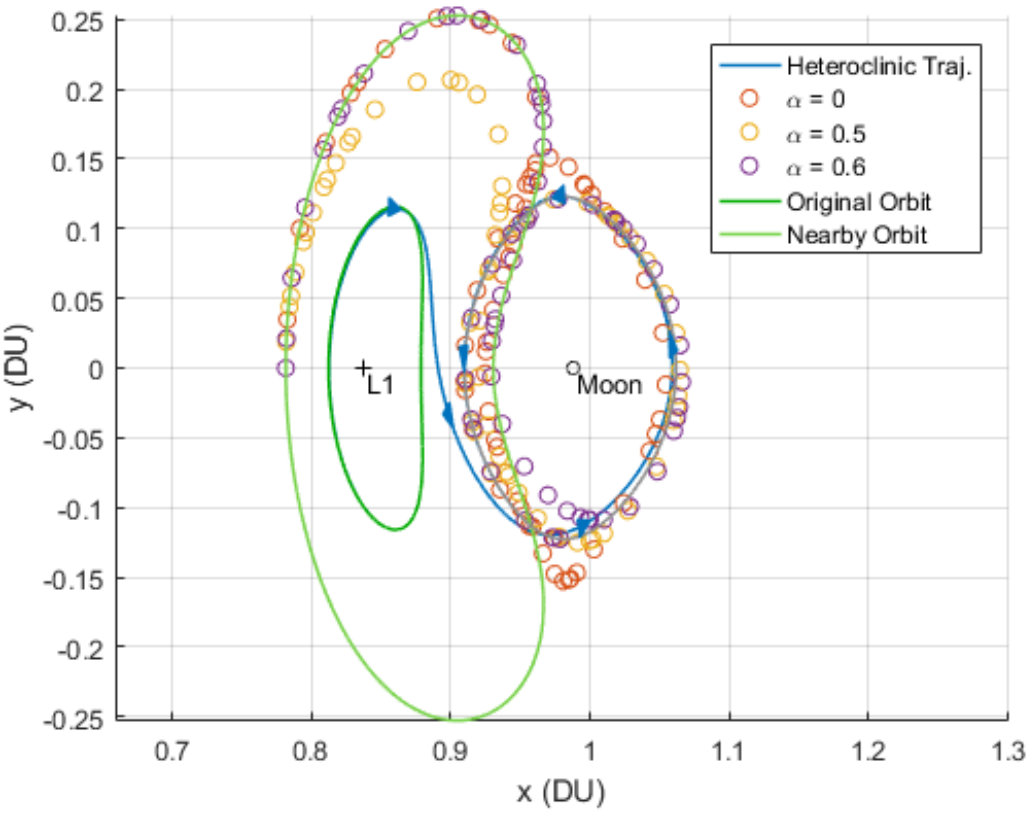


Figure 4.8. (1 of 2) Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Heteroclinic Trajectory for a "Nearby" Orbit

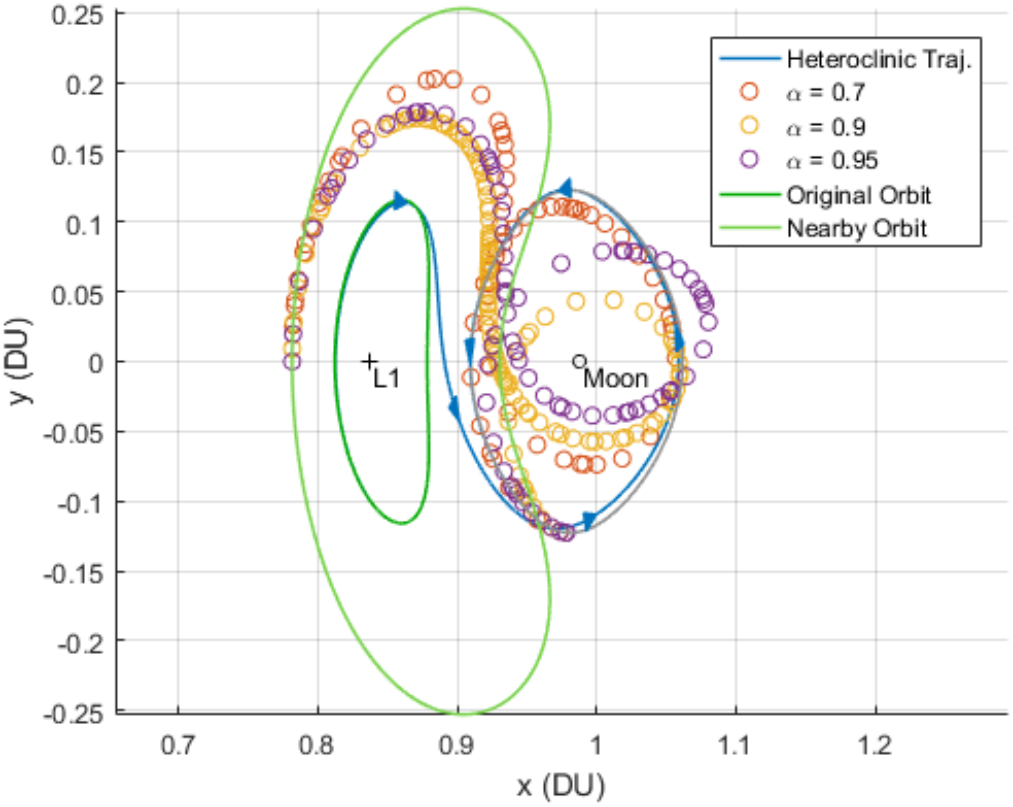


Figure 4.9. (2 of 2) Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Heteroclinic Trajectory for a "Nearby" Orbit

The min-fuel initial guess appears to converge on optimal trajectories that appear to be in the same family, where the trajectories make one orbit around the moon before arriving at the final state. Even though these orbits appear to be of the same family, optimization failed at several values, $\alpha = 0.4$, $\alpha = 0.75$, and $\alpha = 1$. The reason for the failed convergence is difficult to determine, but it should be noted that each trajectory in Fig. 4.8 and 4.9 seemingly differ in form. Thus, the trajectories found at these particular values of α could have required a maneuver that was too different from the heteroclinic trajectory to allow the pseudospctral method to converge on a solution.

Regardless of the failed trajectories, it appears that the optimal trajectories with

lower values of α attempted to remain on the L1 Nearby Periodic orbit and the Distant Prograde Orbit (DPO) for as long as possible. The solution at $\alpha = 0$ seems to have two noticeable burns: one as the trajectory departs the L1 Nearby Periodic orbit and one as the trajectory attempts to continue following the DPO at the DPO's highest point. Alternatively, the near min-time solution at $\alpha = 0.9$ departs the L1 Nearby Periodic orbit immediately, passes nearby by the Moon, and only enters the DPO at the final state.

The min-time initial guess was also used in the pseudospectral method to determine if the alternative initial guess improved the robustness of the pseudospetral method in a different scenario. The optimal solution found for the min-time initial guess is shown in Fig. 4.10.

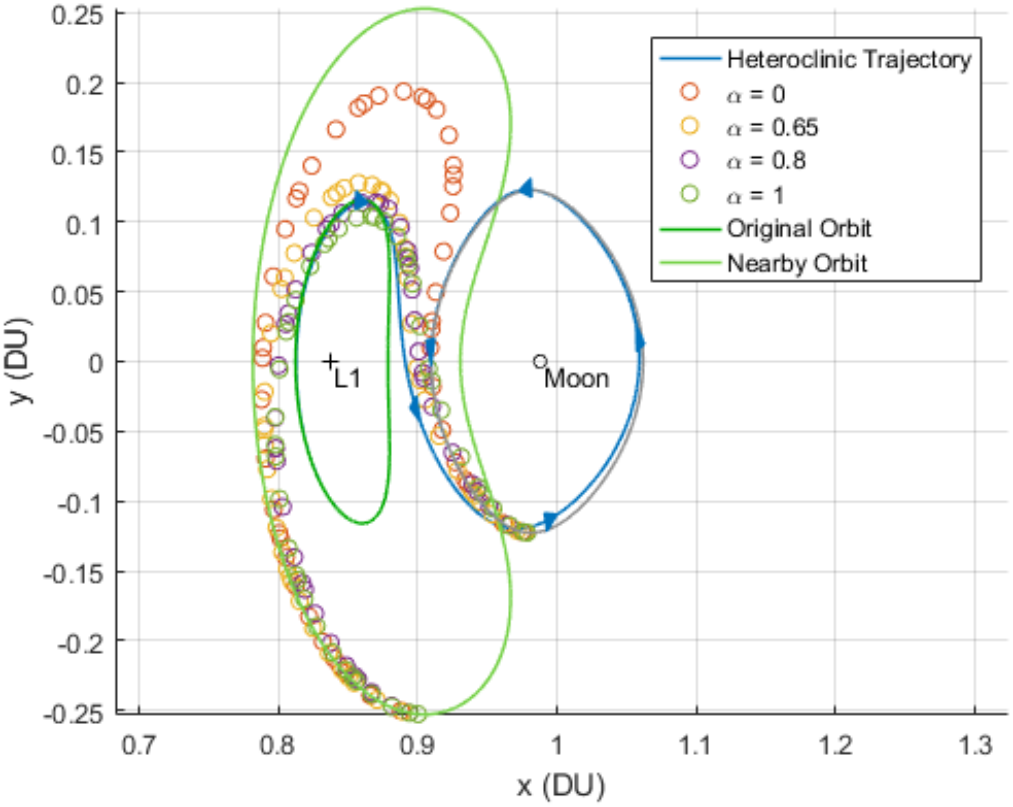


Figure 4.10. Optimal Trajectories Found in GPOPS by Varying the balance of α between Min-Fuel and Min-Time with an Initial Guess of the Min-Time Solution for a "Nearby" Orbit

Similar to the min-time initial guess in the original orbit, all the trajectories in the nearby orbit take the same shape by going directly to the final state. The trajectories for lower values of α follow the L1 "Nearby" Periodic Orbit shape for longer than higher values of α, but all of the trajectories take the same form. The trade-off between Δv and time for both initial guesses is shown in Fig. 4.11.

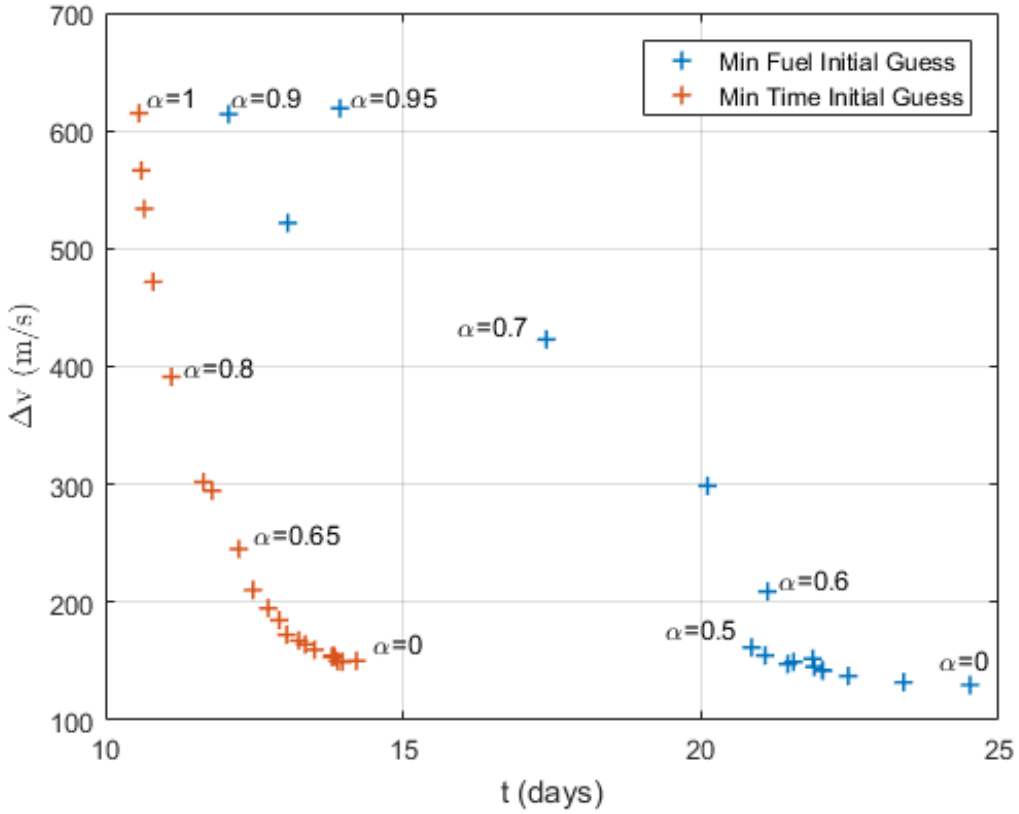


Figure 4.11. Trade-off between Δv and Transfer Time for Optimal Trajectories Found in GPOPS for a "Nearby" Periodic Orbit

Figure 4.11 shows comparable results to the original orbit. The optimal trajectories found from the min-time initial guess all lie on the Pareto Front. However, the optimal trajectories from the min-fuel initial guess also lie on the Pareto front for $\alpha = 0$ to around $\alpha = 0.4$. The min-fuel initial guess trajectories produce a longer time of flight in exchange for a lower Δv. Since each of these points cannot improve in both transfer time and Δv, the points lie on the Pareto Front. Unlike in the original orbit, higher values of α for the min-fuel initial guess do not lie on the Pareto Front. Thus, for higher values for α the min-time initial guess is clearly favored, whereas at lower values of α the initial guess would depend on whether lower transfer times or lower Δv are desired.

Several other unique trajectories should be noted, such as $\alpha = 0.6$ and $\alpha = 0.95$ for the min-fuel initial guess. The values $\alpha = 0.6$ and $\alpha = 0.95$ produce slightly different trajectories than neighboring α values. When examining Fig. 4.11, both of the values of α have a more optimal solution, where there would be less Δv expended for a lower transfer time. These trajectories obviously converged on local minimum. This demonstrates the importance of a good initial guess on the robustness of the pseudospectral method.

4.5 Evaluation of Methods used in Current Research

4.5.1 Evaluation of Inclusion of Mass Loss

Several assumptions were made in Chapter 3 that can be evaluated. The first choice was to incorporate mass lass into the EOMs. This decision was made because the amount of mass being lost in a trajectory was unknown, so incorporating mass loss ensured that this factor would be accounted for if the mass loss was not negligible. Mass loss was also incorporated to check if the results were comparable to reality and if the maneuvers could actually be performed with the allocated amount of mass. After completing the results for a compound objective function with the heteroclinic trajectory as the initial guess in Section 4.4.1, the same analysis was performed, but with mass assumed to be constant. Table 4.1 gives the values for the min-fuel ($\alpha = 0$) and min-time ($\alpha = 1$) when mass is assumed constant and when mass loss is incorporated.

Table 4.1. Resulting Time and Δv for Constant Mass and Mass Loss

Case	α	Time (days)	Δv (m/s)
Constant Mass	$\alpha = 0$ (Min-Fuel)	21.934	1.2777
Mass Loss	$\alpha = 0$ (Min-Fuel)	21.934	0.8677
Constant Mass	$\alpha = 1$ (Min-Time)	8.3783	482.59
Mass Loss	$\alpha = 1$ (Min-Time)	8.3600	485.52

From the results from Table 4.1, the inclusion of mass loss does not seem necessary for the scenarios examined in the current research. The difference in Δv is negligible even in in the minimum-time case, which requires much higher Δv. By including mass loss, the pseudospectral method had to account for an additional variable. Reducing the number of variables is beneficial because it reduces the chance of converging on a local minimum. Mass loss may need to be incorporated in other scenarios where mass loss is much higher, but for the case presented in the current research, assuming constant mass would have been recommended.

4.5.2 Evaluation of Utilizing Initial Guess

Throughout the results in Chapter 4, the importance of an initial guess has been emphasized. The min-fuel initial guess frequently converged on local minimum solutions, whereas the min-time initial guess did not have this issue. The min-time initial guess was valuable in guaranteeing the pseudospectral method converged on a at least a more beneficial local minimum than the min-fuel solution. However, there were certain situations in which the min-fuel initial guess produced an optimal solution that was on the Pareto Front. Thus, another analysis could benefit from using a hybrid initial guess scheme, such as using the min-fuel initial guess to develop an original optimal solution set and then using the minimum-time solution as an initial

guess to complete the Pareto Front. In the scenario presented in the current research, the min-fuel initial guess only produced optimal results on the Pareto Front through $\alpha = 0.4$ and all of these results required trading higher transfer times for Δv. If transfer time is a major concern, then the min-fuel initial guess may not be relevant and instead the min-time initial guess should be utilized.

Another conclusion concerning initial guesses is that a single initial guess is generally acceptable to be used as an initial guess in all "nearby" scenarios. By "nearby" scenarios, it is meant that a single initial guess can be used for "nearby" orbits and also for a varying compound objective functional. The heteroclinic connection initial guess successfully found optimal trajectories for "nearby" orbits in Section 4.3. Even though several optimal trajectories converged on local minimum, the majority of the trajectories converged on advantageous results that allowed a general linear trend to be noticed. On top of this, the min-time initial guess worked superbly in the compound objective functional for all values of α. The applicability of a single initial guess to a variety of scenarios demonstrates the robustness of the pseudospectral method. The pseudospectral method is capable of converging on an optimal solution in "nearby" scenarios even with an initial guess that may not be very accurate.

4.5.3 Evaluation of Pseudospectral Method

The selected pseudospectral method utilized an *hp*-adaptive mesh refinement scheme. This mesh refinement scheme was important in ensuring the convergence of the pseudospectral method. As mentioned in Sections 4.4.1 and 4.4.2 the mesh was refined to ensure that there were enough collocation points to accurately approximate the EOMs in areas of rapidly changing dynamics. Thus, the pseuodpsectral method was able to converge on solutions even with non-accurate initial guesses in a variety of scenarios.

In some situations, the mesh-refinement method did not successfully divide the mesh frequently enough to add enough collocation points. One of these situations can be seen in Fig. 4.9 for $\alpha = 0.95$. The fact that $\alpha = 0.95$ converged to a local minimum is difficult to determine from Fig. 4.9 alone. However, Fig. 4.10 demonstrates that when $\alpha = 0.95$ there is a more optimal solution at $\alpha = 0.9$. Since the trajectories for $\alpha = 0.9$ and $\alpha = 0.95$ are becoming increasingly close to the Moon, the dynamics in the region are changing rapidly and the EOMs are approaching a singularity as $r_2 \to 0$. Thus, the inclusion of more collocation points would ensure the trajectory for $\alpha = 0.95$ does not converge on a local minimum. By increasing the required number of collocation points in each mesh in GPOPS, the pseudospectral method successfully converged on a more optimal solution (potentially a globally optimal solution). Figure 4.12 shows the the old trajectory for $\alpha = 0.95$, the new trajectory for $\alpha = 0.95$, and the trajectory for $\alpha = 0.9$ for comparison. The old trajectory for $\alpha = 0.95$ had a $\Delta v = 619.05$ m/s and a transfer time of $t = 13.94$ days. The new trajectory for $\alpha = 0.95$ had a $\Delta v = 673.76$ m/s and a transfer time of $t = 11.84$ days. The new trajectory would be an optimal solution that is on the Pareto Front, which demonstrates the success of adding more collocation points when the dynamics are rapidly changing.

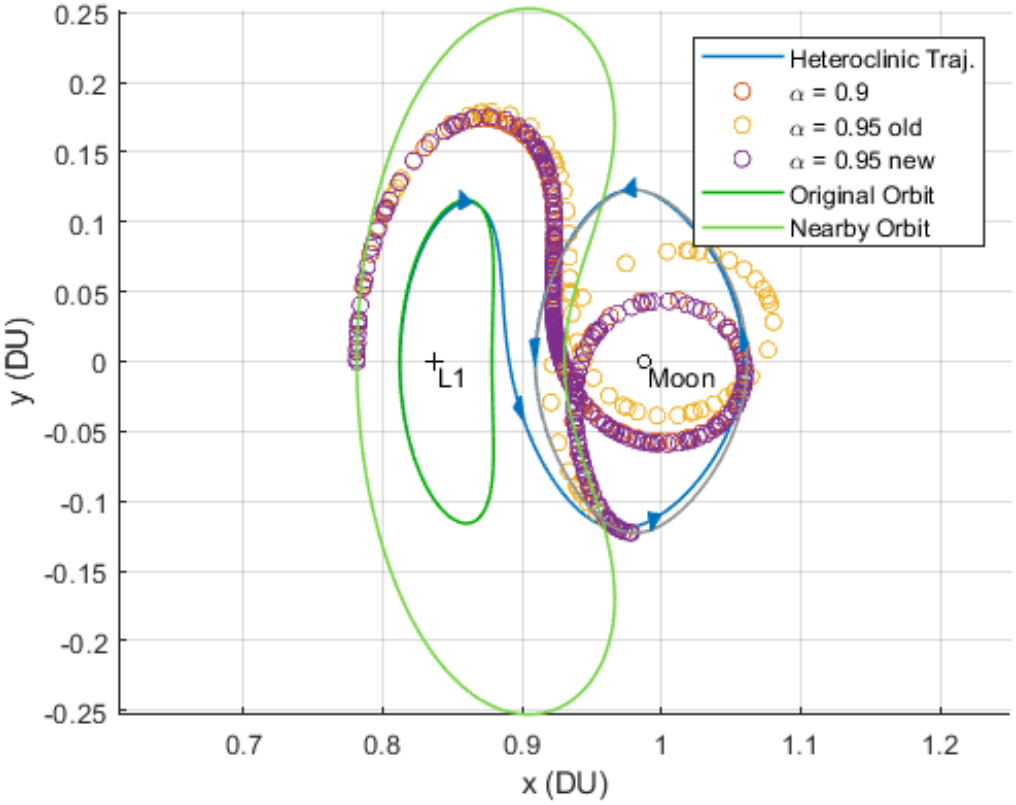


Figure 4.12. Increasing Collocation Points Results in a New Trajectory for $\alpha = 0.95$

Figure 4.12 demonstrates an effective method of correcting a trajectory that converged on a local minimum. By increasing the number of collocation points in each mesh interval, the pseudospectral method can successfully solve for an optimal trajectory, even in a region of rapidly changing dynamics and approaching a singularity.

V. Conclusions and Recommendations

5.1 Summary of Work

The current research evaluated the efficacy of the pseudospectral method utilized in General Purpose Optimal Control Software (GPOPS) in determining optimal trajectories in the Circular-Restricted Three-Body Problem (CR3BP). The pseudospectral method was evaluated by applying the method to a variety of scenarios.

The first scenario used the pseudospectral method to solve for the heteroclinic trajectory, which is a known minimum-fuel solution. Converging on the known minimum-fuel solution verified that the pseudospectral method is viable in the CR3BP system.

Next, the research examined if the pseudospectral method could converge without an accurate initial guess. The minimum-fuel solution was evaluated for "nearby" periodic orbits using the heteroclinic connection as an initial guess. In the majority of the cases, the pseuospectral method converged on a beneficial local minimum. Several trajectories did converge on less beneficial local minima, but these minima could be clearly picked out by examining the trend (discussed in the research) between the Jacobi constant of the periodic orbit and the Δv required for the optimal trajectory.

Finally, this research evaluated a compound objective function by investigating the balance between the minimum-fuel and minimum-time trajectories. A minimum-fuel initial guess versus a minimum-time initial guess was evaluated, along with an evaluation of the original periodic orbit and a "nearby" periodic orbit. A Pareto Front was generated to compare the trade-off between Δv and transfer time. The minimum-time initial guess allowed the pseudospectral method to converge on optimal solutions that were all on the Pareto Front, whereas the minimum-time initial guess only had optimal solutions on the Pareto Front for low values of α. This demonstrates the importance of an initial guess on ensuring the robustness of the pseudospectral

method. The recommendation is to use the minimum-time initial guess to generate the first part of the Pareto Front and then use the minimum-fuel initial guess up to approximately $\alpha = 0.4$ to generate the remainder of the Pareto Front.

The pseudospectral method successfully converged on optimal solutions when provided a proper initial guess, such as a minimum-time solution. The pseudospectral method was aided by the *hp*-adaptive mesh refinement scheme used in GPOPS. The mesh refinement allowed for an increased number of collocation points to be placed in locations of rapidly changing dynamics. With insight into the CR3BP, the pseuodspectral method can be utilized to determine optimal trajectories in the chaotic CR3BP environment, but accurate initial guesses are required and an effective mesh-refinement scheme can help the robustness of the pseudospectral method..

5.2 Future Work

Future work for the current research should involve evaluating the pseudospectral method in further scenarios and comparing it to other methods. Recommended future work include:

- Provide a more accurate initial guess to the pseudospectral method, such as through the use of a heuristic method. The pseudospectral method was highly dependent on the initial guess provided to it, so it may provide better solutions if accurate initial guesses are provided specific to each scenario.

- Perform the same analysis in a larger variety of scenarios, including different locations with the CR3BP and with different parameters, such as thrust levels and specific impulse. Even though the pseudospectral method worked for the scenario in the current research, the results may not extend to other areas in the CR3BP. Testing in other scenarios would provide further insights into the robustness of the pseudospectral method.

- Expand the problem to analyze impulsive maneuvers. Impulsive maneuvers would require the use of phases within GPOPS due to the discontinuity in states at the burn time. This would further evaluate the pseudospectral method because it would also require that constraints between phases also be met on top of the optimal control problem within each phase.

- Use another optimization method, such as a hybrid heuristic optimization method to solving indirectly with a shooting method using Particle Swarm Optimization on the costates. Another optimization method would allow a comparison to the robustness of the pseudospectral method. By seeing how another optimization method compares in a comparable scenario will allow the success and robustness of the pseudospectral method to be directly evaluated relative to other optimal control methods.

5.3 Conclusion

As the expansion of space capabilities increases and the need for more resiliency increases, the CR3BP provides an alternative environment to expand operations. It is important to understand mission design and methodologies in this environment before the CR3BP domain becomes a contested arena. The pseudospectral method used within GPOPS seems to be a beneficial tool for evaluating optimal trajectories in the CR3BP. However, further analysis needs to be completed to understand the limitations of the pseudospectral method within the CR3BP, so the pseudospectral method's efficacy can be improved within the environment. The results and analysis developed in this work will need to be expanded to determine an optimal mission design methodology in the CR3BP environment to ensure future capabilities in this region.

Bibliography

[1] US Department of Defense, "National Security Space Strategy [unclassified summary]," p. 14, 2011. [Online]. Available: https://www.hsdl.org/?view{&}did=10828

[2] W. E. Wiesel, *Spaceflight Dynamics*, 3rd ed. Beavercreek, OH: Aphelion Press, 2010.

[3] E. J. DOEDEL *et al.*, "Elemental Periodic Orbits Associated With the Libration Points in the Circular Restricted 3-Body Problem," *International Journal of Bifurcation and Chaos*, vol. 17, no. 08, pp. 2625–2677, 2007. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0218127407018671

[4] D. J. Grebow, "Generating Periodic Orbits in the Circular Restricted Three-Body Problem with Applications to Lunar South Pole Coverage," Ph.D. dissertation, Purdue University, 2006. [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Generating+Periodic+Orbits+in+the+Circular+Restricted+Three-BODY+PROBLEM+WITH+APPLICATIONS+TO+LUNAR+SOUTH+POLE+COVERAGE#0

[5] J. S. Parker *et al.*, "Chaining periodic three-body orbits in the EarthMoon system," *Acta Astronautica*, vol. 67, no. 5, pp. 623–638, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0094576510001207

[6] J. Hess, "Functional Optimization Lesson 3 : Static Optimization (1 / 5) Unconstrained Optimality," 2017.

[7] A. V. Rao, "A Primer on Pseudospectral Methods for Solving Optimal Control Problems," 2012.

[8] A. V. Rao *et al.*, "Algorithm 902: GPOPS, A MATLAB Software for Solving Multiple-Phase Optimal Control Problems Using," *ACM Transactions on Mathematical Software*, vol. 37, no. 2, pp. 1–39, 2011.

[9] M. Wilmer, "Military Applications of High-Altitude Satellite Orbits in a Multi-Body Dynamical Environment and Dynamical Systems Theory," 2016.

[10] J. N. Brick, "MILITARY SPACE MISSION DESIGN AND ANALYSIS IN A MULTI-BODY ENVIRONMENT: AN INVESTIGATION OF HIGH-ALTITUDE ORBITS AS ALTERNATIVE TRANSFER PATHS, PARKING ORBITS FOR RECONSTITUTION, AND UNCONVENTIONAL MISSION ORBITS," 2017.

[11] U. S. Joint Chiefs of Staff, "Joint Publication 3-14," *Joint Publication 3-14*, no. May, 2013. [Online]. Available: http://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3{_}14.pdf

[12] "China confirms satellite downed," January 2007, [Online; updated 23-January-2007]. [Online]. Available: http://news.bbc.co.uk/2/hi/asia-pacific/6289519.stm

[13] J. Achenbach, "Satellite collision adds to 'space junk' problem," February 2009, [Online; posted 14-February-2009]. [Online]. Available: http://www.washingtonpost.com/wp-dyn/content/article/2009/02/13/AR2009021302071.html

[14] Office of Secretary of Defense, "DOD Annual Report to Congress: Military and Security Developments Involving the People's Republic of China 2016," p. 156, 2016. [Online]. Available: https://www.defense.gov/Portals/1/Documents/pubs/2016ChinaMilitaryPowerReport.pdf

[15] G. Sousa, "Countries with the most operational satellites in orbit," April 2017, [Online; updated 25-April-2017]. [Online]. Available: https://www.worldatlas.com/articles/countries-with-the-most-operational-satellites-in-orbit.html

[16] Office of the Assistant Secretary of Defense for Homeland Defense and Global Security, "Space Domain Mission Assurance: A Resilience Taxonomy White Paper," *Federation of American Scientists*, no. September, pp. 1–10, 2015. [Online]. Available: https://fas.org/man/eprint/resilience.pdf

[17] P. D. Spudis, "The Moon: Point of Entry to Cislunar Space," *Toward a theory of spacepower: Selected essays*, pp. 241–251, 2011.

[18] W. E. Wiesel, *Modern Astrodynamics*, 2nd ed. Beavercreek, OH: Aphelion Press, 2010.

[19] ——, *Modern Orbit Determination*, 2nd ed. Beavercreek, OH: Aphelion Press, 2010.

[20] Victor Szebehely, *Theory of Orbits*. New York: Academic Press INC, 1967.

[21] C. Zhang *et al.*, "Low-Thrust Minimum-Fuel Optimization in the Circular Restricted Three-Body Problem," *Journal of Guidance, Control, and Dynamics*, vol. 38, no. 8, pp. 1501–1510, 2015. [Online]. Available: http://arc.aiaa.org/doi/10.2514/1.G001080

[22] D. L. Kunz, *Intermediate Dynamics for Aeronautics and Astronautics*. Centerville, OH: Headmaster Press, 2015.

[23] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control*. New York, NY: Springer.

[24] K. Davis, "Coordinate Frame Transformations Interplanetary Mission Design," pp. 1–3, 2010. [Online]. Available: http://ccar.colorado.edu/imd/2015/documents/Rot2Inert.pdf

[25] D. T. Greenwood, *Classical Dynamics*, dover edit ed. Mineola, NY: Dover, 1997.

[26] K. C. Howell, "Three-dimensional, periodic, 'halo' orbits," *Celestial mechanics*, vol. 32, no. 1, pp. 53–71, 1 1984. [Online]. Available: https://doi.org/10.1007/BF01358403

[27] C. Bezrouk, "Constructing Periodic Orbits," 2015. [Online]. Available: http://ccar.colorado.edu/imd/2015/Lectures/BezroukSSDC.pptx

[28] K. Nichols, "The James Webb Space Telescope & Modeling Sun-Earth L2 Halo Orbits." [Online]. Available: http://ccar.colorado.edu/asen5050/projects/projects_2015/Students/Nichols_Kristin/jwst.htm

[29] J. S. Parker and R. L. Anderson, *Low-Energy Lunar Trajectory Design.* Hoboken, NJ: John Wiley & Sons, 2014.

[30] N. Truesdale, "Using Invariant Manifolds of the Sun-Earth L2 Point for Asteroid Mining," 2012. [Online]. Available: http://ccar.colorado.edu/asen5050/projects/projects_2012/truesdale/problem.html

[31] W. S. Koon *et al.*, *Dynamical systems, the three-body problem and space mission design*, v1.2 ed., 2011, vol. 21. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.367.7054&rep=rep1&type=pdf

[32] J. Hess, "Functional Optimization Lesson 4 : Static Optimization Constrained Optimality," 2017.

[33] ——, "Functional Optimization Lesson 8 : Static Optimization Summary and Dynamic Optimization Nomenclature Dynamic Optimization - Transition," 2017.

[34] ——, "Karush-Kuhn-Tucker Conditions," 2017.

[35] A. G. Zurita Jr., "MINIMUM-FUEL TRAJECTORY DESIGN IN MULTIPLE DYNAMICAL ENVIRONMENTS UTILIZING DIRECT TRANSCRIPTION METHODS AND PARTICLE SWARM OPTIMIZATION," 2016.

[36] S. Stanton, "Finite Set Control Transcription for Optimal Control Applications," Ph.D. dissertation, 2010.

[37] J. T. Betts, "Survey of Numerical Methods for Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, vol. 21, no. 2, pp. 193–207, 1998. [Online]. Available: http://arc.aiaa.org/doi/10.2514/2.4231

[38] "GPOPS-II: Next-Generation Optimal Control Software." [Online]. Available: http://www.gpops2.com/

[39] A. V. Rao *et al.*, "GPOPS II: A MATLAB Software for Solving Multiple-Phase Optimal Control Problems Using hpAdaptive Gaussian Quadrature Collocation Methods and Sparse Nonlinear Programming," *ACM Transactions on Mathematical Software*, vol. 37, no. 2, pp. 1–39, 2010. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1731022.1731032

[40] K. C. Howell and M. T. Ozimek, "Low-thrust transfers in the earth-moon system including applications to libration point orbits," *Advances in the Astronautical Sciences*, vol. 129 PART 2, no. 2, pp. 1455–1481, 2008. [Online]. Available: https://engineering.purdue.edu/people/kathleen.howell.1/Publications/Journals/2010{_}JGCD{_}OziHow.pdf

[41] J. Stuart *et al.*, "Design of end-to-end trojan asteroid rendezvous tours incorporating potential scientific value," *Advances in the Astronautical Sciences*, vol. 152, no. January 2014, pp. 997–1016, 2014. [Online]. Available: https://www.researchgate.net/profile/Kathleen{_}Howell/publication/289393635{_}Design{_}of{_}end-to-end{_}trojan{_}asteroid{_}rendezvous{_}tours{_}incorporating{_}potential{_}scientific{_}value/links/574c5dde08ae4cada7ea8c2a.pdf

[42] J. Heiligers *et al.*, "Optimisation of solar sail interplanetary heteroclinic connections," *2nd Conference on Dynamics and Control of Space Systems*, 2014. [Online]. Available: https://strathprints.strath.ac.uk/47263/4/Heiligers{_}J{_}et{_}al{_}Pure{_}Optimisation{_}of{_}solar{_}sail{_}interplanetary{_}heteroclinic{_}connections{_}Mar{_}2014.pdf

[43] R. E. Pritchett *et al.*, "Low-Thrust Transfer Design Based on Collocation Techniques: Applications in the Restricted Three-Body Problem," *AAS/AIAA Astrodynamics Specialist Conference*, pp. 1–20, 2017. [Online]. Available: https://engineering.purdue.edu/people/kathleen.howell.1/Publications/Conferences/2017{_}AAS{_}PriHowGre.pdf

[44] D. J. Grebow *et al.*, "Design of optimal low-thrust lunar pole-sitter missions," *Journal of the Astronautical Sciences*, vol. 58, no. 1, pp. 55–79, 2011. [Online]. Available: https://www.researchgate.net/profile/Kathleen{_}Howell/publication/257288453{_}Design{_}of{_}Optimal{_}Low-Thrust{_}Lunar{_}Pole-Sitter{_}Missions/links/57703edb08ae0b3a3b7b917e.pdf

[45] ——, "Advanced modeling of optimal low-thrust lunar pole-sitter trajectories," *Acta Astronautica*, vol. 67, no. 7, pp. 991–1001, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0094576510001542

[46] G. Mingotti *et al.*, "Combined optimal low-thrust and stable-manifold trajectories to the earth-moon halo orbits," *AIP Conference Proceedings*, vol. 886, no. February, pp. 100–112, 2007. [Online]. Available: https://www.researchgate.net/profile/Francesco{_}Topputo/publication/241390635{_}Combined{_}Optimal{_}Low-Thrust{_}and{_}Stable-Manifold{_}Trajectories{_}to{_}the{_}Earth-Moon{_}Halo{_}Orbits/links/555e2db108ae6f4dcc8dd366/Combined-Optimal-Low-Thrust-and-Stable-Manifold-Trajecto

[47] N. L. Parrish *et al.*, "Low-Thrust Transfers From Distant Retrograde Orbits To L2 Halo Orbits in the Earth-Moon System," *International Conference on Astrodynamics Tools and Techniques*, no. 2, pp. 1–25, 2016. [Online]. Available: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160003314.pdf

[48] R. L. Anderson and M. W. Lo, "Role of Invariant Manifolds in Low-Thrust Trajectory Design," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 6, pp. 1921–1930, nov 2009. [Online]. Available: https://doi.org/10.2514/1.37516

[49] G. P. Sutton and O. Biblarz, *Rocket propulsion elements*. John Wiley & Sons, 2016.

[50] J. B. Caillau *et al.*, "Minimum fuel control of the planar circular restricted three-body problem," *Celestial Mechanics and Dynamical Astronomy*, vol. 114, no. 1-2, pp. 137–150, 2012.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
22-03-2018	Master's Thesis	September 2016 — March 2018

4. TITLE AND SUBTITLE

Optimal Trajectory Generation in a Dynamic Multi-Body Environment using a Pseudospectral Method

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dahlke, Jacob, A., 2d Lt

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENY-MS-18-M-248

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Intentionally Left Blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

High-altitude parking orbits could provide resiliency to the military space infrastructure by providing redundancy in key assets, allowing for rapid reconstitution of underperforming satellites. When analyzing trajectories in a high-altitude regime, two-body models of Keplerian motion become less accurate since the gravitational effects of other bodies are no longer negligible. To provide a higher fidelity model of the dynamics in a high-altitude regime, a multiple-body model can be used. In the Earth-Moon system, a spacecraft operating in the high-altitude regime can be modeled with three-body dynamics. With certain simplifying assumptions, the model is called the circular-restricted three-body problem (CR3BP). The CR3BP provides unique dynamics that could be exploited to provide beneficial trajectories unavailable and unobservable in a lower-order model. The tradeoff for using this higher-order model is there is no closed-form analytical solution and the dynamics are chaotic. Methods to search for optimal trajectories within the CR3BP are analyzed to determine viability in rapid mission development. A direct orthogonal collocation pseudospectral method is utilized to generate fuel- and time- optimal trajectories within the CR3BP. These results are compared to benchmarks from two-body dynamics, such as Hohmann transfers. Numerical approaches to finding optimal solutions are highly dependent on initial guesses to converge on candidate optimal solutions. To compound this issue, the chaotic dynamics in the CR3BP mean small variations in the initial conditions could lead to wildly varying trajectories. The results from the current research provide a methodology to establish a framework for rapid mission development in a dynamical environment, which may be essential to maintain space superiority and responsiveness.

15. SUBJECT TERMS

Circular-Restricted Three-Body Problem, Dynamical Systems, Optimal Control, Pseudospectral Methods, Heteroclinic Trajectories

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Capt Joshuah Hess, Ph.D., AFIT/ENY
U	U	U	UU	137	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636, x4713; Joshuah.Hess@afit.edu