

LAGRANGIAN COHERENT STRUCTURES
IN THE CIRCULAR RESTRICTED THREE-BODY PROBLEM

A Thesis

Submitted to the Faculty

of

Purdue University

by

Cody R. Short

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2010

Purdue University

West Lafayette, Indiana

“But why dwell upon these things as though human wit were satisfied with earthly regions and had not dared advance beyond?”

—Galileo Galilei, *Sidereus Nuncius*, 1610

“...for all the difficulty of philosophy seems to consist in this—from the phænomena of motions to investigate the forces of nature, and then from these forces to demonstrate the other phænomena;”

—Isaac Newton, *Principia* (author’s preface), 1686

On a more personal note, thesis writing always falls in proximity to child birth. We were expecting our first son while I was writing my undergraduate thesis at BYU, and while that work was never officially dedicated to him, it really should have been, at least in part. Now our second boy has arrived during the *process* of authoring this document (remember, correlation doesn’t *necessarily* prove causation), and so I’d like to dedicate this to him. However, I cannot let myself do so without recognition of the other two most important people in my life, my sweet wife and our first son:

—for Grayson, Jackson and Lori

ACKNOWLEDGMENTS

I’ve always been a student. I enjoy being a student, and hope to always be a student (although not always in such a formal setting). It all started with my mom. Her love, encouragement and tireless efforts to ensure that I would be safe and provided for cultivated in me the sense that I could do well in my personal studies. Later, my grandparents and uncle reinforced this personal belief, and by the time my high school friends and teachers got a hold of me, it was a lost cause. I knew I would go to college. My professors and colleagues at BYU, including (in no small way) my undergraduate advisor Michael Joner and my favorite instructor Lawrence Rees, kept up the trend. While my undergraduate work *stretched* me, in the end, I was still ready to come back for more. I was manifestly fortunate to continue my education under the auspices of Armstrong Hall at Purdue.

Since arriving at Purdue, I’ve always felt like this is where I should be. While I wasn’t sure if I was personally of the requisite caliber to join Kathleen Howell’s research group, I knew that I wanted to work for her from the moment Lindsay Millard and Chris Patterson introduced me to the work. I still feel, at times, that I’m not quite “there” yet, but Professor Howell’s extensive patience, countless hours of personal advising, and uncanny ability to teach the most complicated ideas to practically anyone, have convinced me that she can *make* a high-caliber graduate student out of any compliant raw material, even me. I have immense respect for Professor Howell and continually strive to emulate her careful methods and understanding nature. There are few people in my life whose opinion of me personally matters, Professor Howell is one of these people. I’m especially grateful to her for all of her efforts in my behalf, including the enormous time commitment represented by her careful review of this thesis.

At this point, it seems appropriate to thank the other members of my committee, namely, James Longuski and Anastasios Lyrintzis. Professor Longuski helped me navigate my initial semesters at Purdue and was extensively involved as an instructor in what I affectionately call my first year: “dynamics boot-camp”. Professor Lyrintzis was, perhaps, most directly involved in me coming to Purdue. I still recall the excitement that my wife and I shared

after learning from Professor Lyrintzis that he had, as the Graduate Chair for Aeronautics and Astronautics, secured funding for me. This was the clarion call for me to come to Purdue. I appreciate both of these professors serving on my committee.

I am especially grateful to Xavier Tricoche who is largely responsible for the idea behind this work. He introduced me to many underlying concepts as well as computational and visualization methods. His advice and instruction have been tremendously helpful throughout my Masters degree. I look forward to continued work with Professor Tricoche.

I hold our research group in high esteem because of its great members and excellent advisor. While some group members have had a larger impact on me, each of them has provided support and assistance to me on various occasions. Therefore, I would like to acknowledge them all, past and present: Todd Brown, Diane Craig Davis, Loïc Chappaz, Dawn Gordon, Daniel Grebow, Amanda Haapala, Aurelie Heritier, Jayme Howsman, Lucia Irrgang, Masaki Kakoi, Amanda Knutson, Lindsay Millard, Zubin Olikara, Martin Ozimek, Chris Patterson, Tom Pavlak, Raoul Rausch, Wayne Schlei, Rohan Sood, Jeff Stuart, Mar Vaquero, Matt Vavrina and Geoff Wawrzyniak. Apart from our group, I would also like to acknowledge a few additional colleagues that made certain day-to-day aspects of graduate life more enjoyable, namely, Kara Cunzeman and Hugo Pelaez.

The School of Aeronautics and Astronautics supported me with a Teaching Assistantship and a departmental Estus H. and Vashti L. Magoon award for my first year of study at Purdue. First Year Engineering Honors provided support for the subsequent two years. Partial support was provided by a summer Research Assistantship in the Rune and Barbara Eliassen Visualization Laboratory and a Frederic Miller Graduate Scholarship from the Purdue Division of Financial Aid. Currently, support for my efforts is provided by First Year Engineering. I am grateful for each of these opportunities and the associated support.

Not unlike the more renowned actors who assume an honored last billing in the credits of a production, I've reserved my most important acknowledgments for the end. I do everything I do for my wife and children. Although I enjoy graduate school immensely, I would not have the personal motivation to pursue such an endeavor without the three people that I go home to each night – even if they are frequently asleep by the time I get there. At the risk of becoming repetitive, I would like to extend the greatest acknowledgment to the same individuals for whom this effort is dedicated, Gray, Jack and Lori.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
ABSTRACT	x
1 INTRODUCTION	1
1.1 Problem Scope and Definition.....	1
1.2 This Work in Context	1
1.2.1 History of the Three-Body Problem.....	2
1.2.2 The Advent of Lagrangian Coherent Structures.....	4
1.3 Overview of the Present Work	5
2 GENERAL FORMULATIONS AND THE RESTRICTED PROBLEM	8
2.1 Critical Assumptions.....	8
2.2 Model Formulation.....	10
2.3 Equations of Motion	11
2.4 The Jacobi Integral.....	16
2.5 Equilibrium Points.....	17
2.6 Curves and Surfaces of Zero Relative Velocity.....	22
2.7 Linear Variational Equations Relative to the Collinear Points	25
2.8 The State Transition Matrix.....	31
2.9 Periodic Orbits, a Shooting Example.....	34
3 LAGRANGIAN COHERENT STRUCTURES	41
3.1 Lagrangian Analysis.....	41
3.2 Lyapunov Exponents.....	43
3.3 Poincaré Maps.....	46
3.4 Finite-Time Lyapunov Exponents.....	48
3.5 Invariant Manifolds.....	54
3.5.1 Stable and Unstable Manifolds Associated with L_1 and L_2	55
3.5.2 Manifolds Associated with a Periodic Orbit (Maps)	58
3.6 Lagrangian Coherent Structures	62
3.6.1 A Formal Definition of Ridges.....	62
3.6.2 The Nature of Lagrangian Coherent Structures	63

	Page
4 COMPUTATIONAL METHODS.....	66
4.1 Numerical Considerations	66
4.1.1 Integration.....	66
4.1.2 Interpolation.....	67
4.1.3 Regularization	68
4.2 Ridge Calculation.....	69
4.3 Parallel Processing	73
4.3.1 CPU Implementation.....	73
4.3.2 GPU Implementation	75
4.4 Visualization	76
5 ANALYSIS AND RESULTS.....	80
5.1 Lagrangian Coherent Structures in the Earth-Moon System.....	80
5.1.1 Structures Associated with a Lyapunov Orbit	81
5.1.2 Structures In the L_4 Region with $C_{L_4} < C < C_{L_2}$	101
5.2 Selected Lagrangian Coherent Structures in the Sun-Saturn System	105
5.2.1 Potential Sun-Saturn Transit Trajectories.....	105
5.2.2 Long-Term Trajectory Analysis in the Sun-Saturn System.....	114
6 CONCLUSION	124
6.1 Summary.....	124
6.2 Future Work.....	125
6.3 Concluding Remarks	125
LIST OF REFERENCES	126

LIST OF TABLES

Table	Page
2.1 Earth-Moon Libration Points	21
2.2 Jacobi Constant Values for Earth-Moon Libration Points	21
3.1 The Lyapunov Spectrum and Steady State Behavior.....	45
3.2 LCS and Manifold Correspondence	64
4.1 Speed Improvements with OpenMP Multi-Core Parallelization.....	74
5.1 Data Points Selected from Figure 5.15 for Closer Inspection Via Plotting	95
5.2 Point Data Selected Manually from Figure 5.25 for Inspection.....	109

LIST OF FIGURES

Figure	Page
2.1 Model for the circular restricted three-body problem	10
2.2 Equilibrium points in the CR3BP	20
2.3 Earth-Moon ZVC for various C values	23
2.4 More Earth-Moon ZVC(S) for various C values	24
2.5 A linear ellipse about L_1	30
2.6 Perturbed and reference solutions	31
2.7 Correcting a linear solution with the STM	33
2.8 Transition from a linear solution to nonlinear solutions	37
2.9 Families of planar orbits about each of the libration points separately	39
2.10 Families of planar orbits about each of the libration points	40
3.1 Stretching of an ellipsoid represented by the Lyapunov exponents	45
3.2 An illustration of a one-sided Poincaré map	47
3.3 Backward FTLE for $C \approx C_{L_2}$ in the Earth-Moon system	51
3.4 Forward FTLE for $C \approx C_{L_2}$ in the Earth-Moon system	51
3.5 Forward (red) and backward (blue) FTLE for $C \approx C_{L_2}$	52
3.6 Poincaré map of manifold crossings overlain on a forward FTLE field	55
3.7 Local manifolds associated with L_1 and L_2 in the Earth-Moon CR3BP	57
3.8 Position-space projections of eigenvectors along an L_1 Lyapunov orbit	59
3.9 Poincaré map of manifolds associated with fixed point along a periodic orbit	60
3.10 Lyapunov orbit stable/unstable manifolds represented by trajectories	61
3.11 “Ridges” of FTLE values represented as relative heights	64
4.1 An example of Gaussian blurring applied to a noisy FTLE field	70
4.2 Five point stencil for central differencing	71
4.3 Annotated user interface screenshot	78
4.4 User interface screenshot with examples	79
4.5 An example of results from an extended interactive session	79
5.1 Lyapunov orbit and zero velocity curves corresponding to $C \approx C_{L_2}$	81
5.2 L_1 Lyapunov orbit corresponding to $C = 3.17216$	82
5.3 Configuration-space projections of stable (blue) and unstable (red) manifolds	84
5.4 An illustration of the first two crossings of the map Σ	85
5.5 Stable manifold crossings associated with a Lyapunov orbit	86
5.6 Unstable crossings; counterpart of Figure 5.5	86
5.7 Both stable (blue) and unstable (red) manifold crossings	87
5.8 FTLE for three crossings of the Poincaré map, $\Sigma = y = 0$	88

Figure	Page
5.9 FTLE for five crossings of the Poincaré map, $\Sigma = y = 0$	89
5.10 FTLE for ten crossings of the Poincaré map, $\Sigma = y = 0$	89
5.11 FTLE ($5\times$ of $\Sigma = y = 0$) with manifold crossings.....	90
5.12 Ridge points on the three-crossing FTLE map	91
5.13 Ridge points on the five-crossing FTLE map	92
5.14 Ridge points on the ten-crossing FTLE map	92
5.15 Zoom of the five-crossing FTLE map	94
5.16 Trajectory arcs associated with Point 1 from Table 5.1	97
5.17 Trajectory arcs associated with Point 2 from Table 5.1	98
5.18 Trajectory arcs associated with Point 3 from Table 5.1	99
5.19 Zoom of Point 3 trajectory arcs focused in the region near the Moon.....	100
5.20 The L_4 region of interest explored in Section 5.1.2	102
5.21 FTLE in the EM L_4 region for various C values (I)	103
5.22 FTLE in the EM L_4 region for various C values (II)	104
5.23 Periapse Poincaré map of manifolds (with permission: A. Haapala [69]).....	106
5.24 FTLE map corresponding to Figure 5.23	107
5.25 Zoomed insets focused on regions of interest from Figure 5.24	108
5.26 Trajectories from green point in top inset of Figure 5.25.....	110
5.27 Zoom focused on Saturn region from Figure 5.26	111
5.28 Trajectories from green point in bottom inset of Figure 5.25	112
5.29 Zoom focused on Saturn region from Figure 5.28	113
5.30 FTLE for $3\times$ of the periapse map in the vicinity of Saturn.....	114
5.31 FTLE for $10\times$ of the periapse map in the vicinity of Saturn.....	116
5.32 Periapse Poincaré map (with permission: D. C. Davis and K. C. Howell [70])...	117
5.33 Zoom focused on distinctive structures of the 3-crossing FTLE map	118
5.34 Zoom focused on distinctive structures of the 10-crossing FTLE map.....	119
5.35 Wide numerical divergence induced by close proximity Saturn passages	121
5.36 Trajectory from 9-crossing stable region indicated in Figure 5.34.....	122
5.37 Saturn focused view of Figure 5.36	123

ABSTRACT

Short, Cody R. M.S., Purdue University, December 2010. Lagrangian Coherent Structures in the Circular Restricted Three-Body Problem. Major Professor: Kathleen C. Howell.

The mathematical formulation that represents the motion of a particle under the simultaneous influence of two gravitational fields is identified as the Circular Restricted Three-Body Problem (CR3BP). This model of an autonomous dynamical system displays both ordered and chaotic behavior. For some behaviors, simple linear analysis relative to a numerically determined point solution in the problem is sufficient to reveal interesting aspects of the motion, while other scenarios require more extensive procedures to capture the unique features comprising the dynamical behavior. This balance between predictability and complexity that is exhibited in the CR3BP, along with the typical approaches for its analysis, supplies excellent justification for the application of relatively straightforward techniques typically applied in more complicated problems. Among these, a number of versatile analysis tools are based on the concepts of the Finite-Time Lyapunov Exponent (FTLE) and Lagrangian Coherent Structures (LCS).

Lagrangian coherent structures appear as height ridges, or curves of constrained maxima, in a field of FTLE values. Application of interactive visualization, numerical methods, and parallel computation is employed to obtain FTLE data and the associated LCS. These results are compared with known structures to further establish LCS as a useful tool for application in the CR3BP and to demonstrate LCS as a seed for a variety of additional research questions. Results associated with potential applications to mission design are supplied, and comparisons between LCS methods and concurrent research efforts involving other, more familiar, approaches are generated with particular focus on the advantages of FTLE and LCS methods. Ultimately, this analysis serves to validate the concepts of FTLE and LCS as an effective means to further understand the complex behavior in the CR3BP.

1. INTRODUCTION

Even the earliest of recorded history reflects humanity's desire and effort to understand its place and purpose. A major avenue for the progression of this understanding has been critical thought and the resulting practical applications. Throughout the centuries, critical thinkers of all statures have contributed piece-by-piece to the present body of knowledge. This effort reflects an attempt to assimilate many facets of that knowledge and contribute. In this introduction, the current problem is established, placed into context, and its analysis briefly outlined.

1.1 Problem Scope and Definition

This analysis serves a dual purpose. First, the relatively recent emergence of Lagrangian Coherent Structures (LCS) for the exploration of dynamical systems is applied to the Circular Restricted Three-Body Problem (CR3BP). This type of application is not an entirely novel effort as previous researchers have applied LCS to similar problems. However, the ultimate results represent new applications that highlight and establish the applicability and advantages of the concepts. No less important, the secondary purpose is the compilation and organization of the extensive idealogical framework that is necessary to expand this application to more detailed models of multi-body motion. While LCS continue to draw significant attention as a new avenue of research in general, their application in the astrodynamics community has been relatively rare. Therefore, a comprehensive development of the associated theory with respect to orbital mechanics and trajectory design is timely, and the applications in this analysis serve as a foundation to expand their use.

1.2 This Work in Context

A consequence of melding concepts from different areas is a natural split in the ancestries of those ideas. Such a separation is exhibited here and, as such, the major historical

contributions to the framework for both concepts are provided. First, the astrodynamical contributions that are key in the formulation of the circular restricted three-body problem and in the investigation of the resulting behavior are summarized. The contributions that ultimately lead to Lagrangian coherent structures are then elaborated.

1.2.1 History of the Three-Body Problem

The context for the three-body problem is expansive. Efforts to describe and explain the motion of celestial bodies are evident in the earliest recorded history. However, despite the combined efforts of generations of astronomers and philosophers, not until Johannes Kepler empirically established his laws of planetary motion in 1609 [3], did a reasonably correct description of planetary motion emerge. Kepler postulated that the motion of a planet is elliptical, but he did not explain the cause. Galileo’s investigations occurred concurrently and resulted in the methodological paradigm [4] that would explain the general basis of Kepler’s kinematical description in terms of forces and motion. Employing the scientific process of Galileo, Newton deduced the law of gravitation [2], one of the chief results included in his *Principia* published in 1687. However, to fully establish the theoretical foundations of gravitation, the model should correctly solve for the combined motion of multiple bodies and, thus, examination of the n -body problem was initiated. Given the force model for gravity, Newton solved the two-body problem using geometry and calculus. Subsequently, in 1710, Johann Bernoulli justified Kepler’s empirical ellipses by proving that conics solve the two-body problem in general [5]. With the solution of the two-body problem, the next logical effort toward solving the n -body problem is to consider three bodies. The addition of a single mass to the problem seemingly represents only a minor increase in complexity but, unfortunately, this is not the case. No closed-form solution to the three-body problem is currently known.

Due to the inherent difficulties in analysis of the three-body problem and the lack of a solution, alternative methods for approaching the problem were considered. Euler [6] reformulated the problem into the “restricted” problem (as later termed by Poincaré) in 1772. A significant component in this formulation was the introduction of a rotating reference frame [5]. Additionally, Euler identified three particular solutions of the problem, the

collinear libration points; Lagrange, concurrently, identified both the collinear and the two triangular libration points for a total of five equilibrium solutions.

In 1843, Jacobi reduced the order of the problem from eighteen (three positions and three velocities for three bodies) to six (positions and velocities for the body of most interest). At the same time, he introduced an integral of the motion, later named the Jacobi integral [7]. With this constant, boundaries on the motion were distilled by Hill in 1878. These boundaries appear as curves or surfaces in two- or three-dimensional configuration space, respectively. An infinite number of periodic orbits in the restricted problem were predicted by Poincaré, who originally introduced surfaces of section (later termed Poincaré sections) to gain insight into the behavior in the CR3BP. Poincaré also demonstrated the existence of deterministic chaos in the restricted problem. These significant contributions by Poincaré are found in his major treatise, *New Methods of Celestial Mechanics* (1893) [8]. The close of the 19th century saw further development in periodic orbits with the efforts of Darwin, who built on ideas of Lagrange and Hill [5], beginning in 1897.

In 1912, Sundman [9] first offered a solution in the general three-body problem by means of a convergent power series. Several decades later, in 1991, Quidong Wang offered a similar solution for the general n -body problem with $n > 3$ [10, 11]. Unfortunately, both series solutions are characterized by slow convergence and are generally of little practical use [10]. In 1920, Moulton published *Periodic Orbits* [12], offering the status of work on periodic orbits to that time; the calculation of periodic orbits was included, thereby confirming their existence. Moulton, unable to predict future computational tools, indicated that the numerically intensive nature of identifying such orbits would prohibit periodic orbits from serious future study [13]. In 1966, Hénon employed Poincaré sections in his examination of the three-body problem, but it was a surprisingly rare application [14]. In 1967, Szebehely's *Theory of Orbits: The Restricted Problem of Three Bodies* [7] proved to be *the* comprehensive treatise on the three-body problem up to that point. Szebehely offered an extensive and cohesive overview of the aspects of the restricted problem, as well as numerous analytical and numerical results. His *Theory of Orbits* continues to be an authoritative and widely cited text.

With one of the most famous speeches in recent history, President John F. Kennedy effectively transitioned the three-body problem from a curiosity to a very practical concern,

when, on September 12, 1962, he stated, “We choose to go to the Moon.” [15, 16] While the early investigators of the three-body problem, including Newton, had considered the Earth, Moon, and a third body as a suitable configuration for investigation, it wasn’t until space exploration expanded into a multi-body regime, where a spacecraft would be under the continuous influence of two significant gravitational fields, that the three-body problem gained practical interest, typified by the Earth-Moon-spacecraft configuration. In the early 1970’s, toward the end of the Apollo program, and in support of options for the last few planned Apollo flights, Breakwell et al. [17] and Farquhar [18] explored trajectories that emerge within the context of the three-body problem. For a manned mission to the far side of the Moon, communications support was to be supplied via a relay satellite in a “halo” orbit about the trans-lunar collinear libration point. These analyses were largely theoretical in nature and, while the Apollo program ended before such a relay option was realized, these studies renewed interest in applications of the three-body problem to mission design. Halo orbits were extensively explored in a number of investigations including a numerical study accomplished by Howell and Breakwell (1984) [19, 20]. Since these early years of applied three-body analyses, many successful missions have incorporated libration point orbits beginning with ISEE-3 [21] and continuing through Genesis [22] and others [23]. Currently, the ARTEMIS mission is underway which extensively exploits three-body analysis to repurpose two spacecraft with limited remaining fuel from a successfully completed previous mission [24, 25].

1.2.2 The Advent of Lagrangian Coherent Structures

While Lagrangian Coherent Structures (LCS) are relatively nascent when compared with the three-body problem, introduced just over a decade ago, the underlying concept of coherent structures is timeless. Peacock and Dabin [26] indicate that da Vinci was the first to capture structures in a moving fluid by sketching different flow patterns in water flowing over obstacles [27] some 500 years ago. Much more recently, in 2000, George Haller [28, 29] as well as Haller and Yuan [30] formalized the mathematics of LCS and gave the concept its name. While there are different metrics that can be employed to identify LCS, the Finite-Time Lyapunov Exponent (FTLE) is generally the most common, where relatively

high values of the FTLE indicate LCS. The FTLE measures the stretching between adjacent trajectories over a prescribed time interval. Building on his own work, Haller later identified and provided additional criteria to distinguish between actual stretching and shear [31]. Shadden et al. [32] have rigorously established the fundamental idea that LCS act as transport barriers in the flow by proving the flux across LCS is negligible. Mathur et al. [33] have improved on the criteria for extracting LCS, and Lekien et al. [34] established the relevance of LCS methodology in n -dimensional motion.

Given the well-established theoretical foundation, the popularity of LCS has blossomed recently such that active research is occurring in multiple disciplines. Simultaneous activities in computer science and visualization seek to effectively compute and extract LCS. Specifically, work by Garth et al. [35], featuring adaptive mesh refinement for the calculation of FTLE near structures, has provided one means for improving the time efficiency of methods for obtaining LCS. Additional applications of LCS have been presented regarding weather data, transport in the oceans, aeronautical computational fluid dynamics, and even human musculoskeletal biomechanics, circulation, and airway transport [26].

The application of LCS in an astrodynamical context is not entirely novel. Villac [36] along with Villac and Broschart [37] apply fast Lyapunov indicators and chaoticity indicators, both metrics similar in form to FTLE, to preliminary spacecraft trajectory design in multi-body environments and for stability analysis near small bodies, respectively. Even more directly, Gawlik et al. [38] offer an analysis of LCS in the planar elliptic restricted three-body problem. Each of these contributions serves to indicate the timeliness and relevance of this work.

1.3 Overview of the Present Work

With the extensive history of analysis in the three-body problem, a significant amount of insight is already available. Fortunately, the nature of the problem is such that many of its aspects can be examined in ever-finer levels of detail. At the same time, an already well-established direction of investigation in the three-body problem is frequently amenable to extensions or alternative perspectives. While Lagrangian concepts, in general, exhibit an equally considerable pedigree, the concept of Lagrangian coherent structures is new.

The ultimate goal of this analysis is an overarching description of how Lagrangian coherent structures are applied to deliver new insight. Further, this analysis serves as an organized compilation of theory to lay the groundwork for other discoveries.

A simple outline of the thesis follows:

- Chapter 2: General Formulations and the Restricted Problem

The circular restricted three-body problem is the primary working model for this investigation. As such, the derivation is detailed beginning with the general n -body problem and progressing through the subsequent focus, restriction, and simplification. Linear analysis and a means for predicting changes introduced by small variations are also discussed. Finally, an exposition of dynamical systems concepts, which are effectively exhibited in the CR3BP, including equilibrium points and periodic orbits, is presented.

- Chapter 3: Lagrangian Coherent Structures

The ideas of Lagrangian analysis are presented in this chapter. Additional dynamical systems concepts, including an introduction to invariant manifolds, Lyapunov exponents, and Poincaré sections are introduced as they relate to Lagrangian coherent structures. A critical modification in the definition of Lyapunov exponents applicable to LCS, that is, the definition of the finite-time Lyapunov exponent, is elaborated. This chapter concludes with a discussion of the basic concept of LCS.

- Chapter 4: Computational Methods

To some degree, the expansion of computational capabilities has revitalized research in the three-body problem. Poincaré maps and Lagrangian analysis, for example, were considered infeasible only little more than a decade ago due to their numerically intensive nature. It is not surprising, then, that the methodology associated with this analysis is largely computational. As such, an overview of the applicable computational tools is introduced including numerical methods, parallel approaches, and visualization.

- Chapter 5: Analysis and Results

Results generated for several sample cases are presented. Initially, a comprehensive example of the ideas presented in previous chapters is discussed. These ideas are presented as they apply to a specific orbit in the CR3BP, one that is characteristic of an Earth-Moon-spacecraft configuration. Additional analysis in this same system illustrates the potential of LCS to support the design of trajectory transfers. Comparison between FTLE/LCS results and other concurrent research efforts is also offered.

- Chapter 6: Conclusion

Finally, a summary of the entire investigation is presented. Additionally, comments are included regarding potential future investigation associated with this analysis, and concluding remarks are given.

2. GENERAL FORMULATIONS AND THE RESTRICTED PROBLEM

The Restricted Three-Body Problem (R3BP) is an excellent example of a commonly occurring natural configuration that is modeled well as a dynamical system. The formulation of the R3BP as a dynamical system is explored in this chapter, as well as some of the common aspects of such systems as applied to the R3BP. Initially, the R3BP is defined and the equations of motion are derived. Then, some of the significant mathematical and dynamical concepts illustrated by the problem are introduced.

2.1 Critical Assumptions

Newton posed the n -body problem as a natural consequence of the observed motions of bodies in the solar system. Simply put, the n -body problem models the motion of n bodies under their mutual gravitational influences. Given that each body in an n -body system affects the behavior of each other body, it is quickly apparent that a complete model for such a system is very complex and only becomes more so as n increases. No closed-form solution to the n -body problem is currently known; an infinite series solution exists, but it is of limited practical value [10]. Given the lack of a useful analytical solution, only a few viable options remain to investigate the n -body problem. Additional assumptions and simplifications, as well as particular solutions and numerical simulations, are all reasonable starting points. The introduction of additional assumptions to simplify the model is the first step in this analysis.

The first simplification in exploring the n -body problem, is a narrowing of the scope. Rather than investigating n mutually gravitating bodies, consideration of a smaller subset offers effective insight. For example, a two-body formulation has been studied extensively and applied to many important practical applications. However, the two-body problem yields an analytical solution which does not capture some of the features of interest in this analysis. On the other hand, formulating the problem in terms of three bodies produces a model sufficiently complex to reveal many important characteristics while remaining

tractable. Three-body configurations are often representative of significant physical systems. For example, a planet-moon-spacecraft or Sun-planet-moon system is frequently a focus of interest in astronomy and astrodynamics.

In the restricted three-body problem, the next key assumption concerns the relative masses. Of the three bodies in the general three-body problem, no restriction is imposed on the value of their masses. However, as noted above, a common three-body system could involve a Sun-planet-moon or a planet-moon-spacecraft configuration. In either case, an obvious disparity in terms of mass is evident. Assuming the mass of the third body (m_3) to be much smaller than the mass of either larger body (m_1 or m_2) yields an advantage in terms of the analysis, yet represents a number of systems very well. Szebehely [7] illustrates that this assumption ($m_3 \ll m_1, m_2$) is reasonable in some situations through an example for a 6000 kg spacecraft in the Earth-Moon system; neglecting the spacecraft mass introduces a relative error on the order of 10^{-16} in terms of the motion of the system. Under this assumption, that is, neglecting the mass of the third body, the model is denoted the “restricted” three-body problem. The significant effect of this assumption on the motion of the two massive bodies is the reduction of their relative orbit to the well-known conic solutions of the two-body problem, while maintaining the framework of the three-body problem.

One final simplification serves to bound the scope of the R3BP. Since the previous assumption reduced the motion of the massive primaries to conics, and considering that parabolic and hyperbolic orbits are not closed, the motion of the two massive bodies is modeled as elliptical. However, elliptical orbits add complexities that are frequently unnecessary and a further assumption that the bodies move on circular orbits is common. Modeling the motion of the massive bodies as circles or ellipses leads to two common forms of the restricted three-body problem that are more generally termed the Circular Restricted Three-Body Problem (CR3BP) and the Elliptic Restricted Three-Body Problem (ER3BP). This analysis will focus on the CR3BP, but its applications are not limited to this regime. The context for this work is thus established, given these assumptions and simplifications.

2.2 Model Formulation

The circular restricted three-body problem has been briefly introduced but, to derive an accurate mathematical model, a more careful definition is necessary. The model in the CR3BP represents the motion of an essentially massless particle under the gravitational influence of two bodies of relatively larger mass. The bodies of finite mass orbit their common barycenter on circular paths, while no restriction on the motion of the third body, other than that of the natural dynamics, is imposed. The “massless” third body can move in all three spatial dimensions. To gain additional insight and exploit some useful dynamical aspects, the system is observed in a rotating reference frame, a consideration first introduced by Euler [5, 6]. A general depiction of the CR3BP, highlighting the rotating and inertial reference frames, appears in Figure 2.1. In Figure 2.1, the three primaries are indicated as

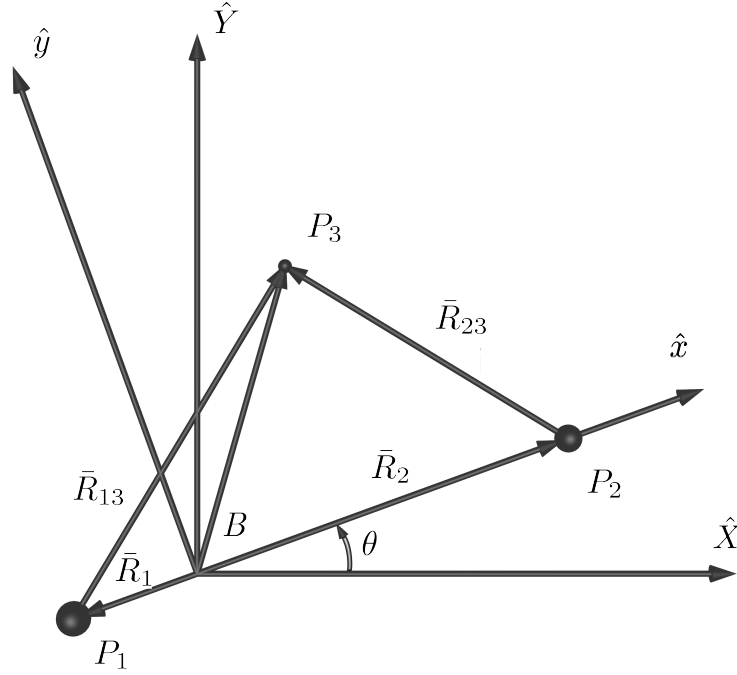


Figure 2.1. Model for the circular restricted three-body problem

solid spheres. Specifically, P_1 refers to the first primary (largest mass, m_1) in the system, P_2 , the second primary, is the smaller mass (m_2), and P_3 is the essentially massless particle ($m_3 \approx 0$) of interest whose motion is described. The inertial frame I is indicated by the

X and Y axes which are oriented consistent with the unit vectors \hat{X} and \hat{Y} (where carets indicate vectors of unit magnitude). The inertial Z axis (unit vector \hat{Z}) is directed out of the page, completing the right-hand triad. The barycenter, B , serves as the common origin of both the inertial and rotating reference frames since it is conveniently fixed in both. The rotating frame R is based on the instantaneous orientation of the larger bodies in their orbits. Thus, the line formed by the two larger masses, through the barycenter, serves as the x axis of the rotating system as indicated by the \hat{x} unit vector. The y axis is defined orthogonal to the x axis and positive in the general direction of velocity. Thus, \hat{y} is oriented 90° from the positive x axis. Both \hat{x} and \hat{y} are specifically defined to be in the plane of motion of the massive primaries. The positive z axis completes the right-handed triad and is coincident with the Z axis of the inertial frame. The rotating frame is oriented with respect to the inertial frame through the angle θ . Given the circular orbits, the angular velocity is constant and the rotating frame moves with the angular rate $\dot{\theta} = N$. This rotational velocity is defined such that ${}^I\bar{\omega}^R = \dot{\theta}\hat{z}$ and is positive when ${}^I\bar{\omega}^R$ is directed consistent with \hat{z} . Since the angular velocity of the rotating frame is selected to be coincident with the rotational velocities of the primaries, the positions of P_1 and P_2 are fixed relative to the rotating frame.

Several important position vectors also appear in the figure. The vectors \bar{R}_i represent the positions of the primaries with respect to the barycenter, while \bar{R}_{ij} are relative vectors denoting the position of the third primary with respect to each of the other primaries. All of these position vectors are expressed in terms of components parallel to each of the \hat{x} , \hat{y} , and \hat{z} directions (i.e., rotating coordinates). The reference frames and vectors discussed here create the framework for the development of the equations of motion.

2.3 Equations of Motion

Given the definition of the system, a natural next step is to mathematically formulate the equations that govern the motion of the particle of interest, P_3 . For convenience and to ultimately generalize the results, the significant parameters and variables are nondimensionalized using a set of characteristic quantities. For a specific three-body configuration, the distances involved and the relative masses of the primaries are generally disparate in

magnitude. Nondimensionalization generalizes the problem and ultimately yields a more convenient mathematical model that can be adapted to a wide range of physical systems. Characteristic quantities, indicated by a superscript “*”, are defined for the fundamental quantities of mass, length, and time, as follows:

$$m^* = m_1 + m_2, \quad (2.1)$$

$$l^* = R_1 + R_2, \quad (2.2)$$

$$t^* = \sqrt{\frac{(R_1 + R_2)^3}{\tilde{G}(m_1 + m_2)}} = \sqrt{\frac{l^{*3}}{\tilde{G}m^*}}. \quad (2.3)$$

The constant \tilde{G} is the Newtonian gravitational constant, where the tilde indicates a dimensional quantity. The time parameter, t^* , is selected such that the resulting nondimensional value of G is unity. The characteristic quantities, m^* , l^* , and t^* represent yet another key component in the development of the equations of motion.

Nondimensional parameters are defined using the characteristic quantities. The nondimensional mass parameter, μ , represents the mass fraction of the system attributable to the second primary,

$$\mu = \frac{m_2}{m^*}. \quad (2.4)$$

The remaining system mass is attributable to the first primary and is also defined in terms of μ , i.e.,

$$1 - \mu = \frac{m_1}{m^*}. \quad (2.5)$$

The vectors defined in Figure 2.1 are nondimensionalized with respect to characteristic length,

$$\bar{r}_i = \frac{\bar{R}_i}{l^*} \text{ and } \bar{r}_{ij} = \frac{\bar{R}_{ij}}{l^*}. \quad (2.6)$$

Of particular note is the nondimensional position vector of the particle P_3 written explicitly in terms of rotating frame coordinates,

$$\bar{r}_3 = \frac{\bar{R}_3}{l^*} = x\hat{x} + y\hat{y} + z\hat{z}, \quad (2.7)$$

where the symbol \bar{r} with any subscripts indicates a dimensionless distance vector. The distance between the primaries, \tilde{a} , is equal to the dimensional radius of the circular orbit

of primary motion. This distance is nondimensionalized and, consequently, normalized by the characteristic length,

$$a = \frac{\tilde{a}}{l^*} = \frac{R_1 + R_2}{l^*} = 1. \quad (2.8)$$

The nondimensional gravitational constant, G , and constant mean motion, n , of the primaries also reduce to unity upon nondimensionalization, that is,

$$G = \frac{\tilde{G}m^*t^{*2}}{l^{*3}} = 1, \quad (2.9)$$

$$\begin{aligned} N &= \sqrt{\frac{\tilde{G}(m_1 + m_2)}{\tilde{a}^3}} = \sqrt{\frac{\tilde{G}m^*}{al^{*3}}} = \sqrt{\frac{\tilde{G}m^*}{l^{*3}}} \\ \Rightarrow n &= Nt^* = \sqrt{\frac{\tilde{G}m^*}{l^{*3}}} \sqrt{\frac{l^{*3}}{\tilde{G}m^*}} = 1. \end{aligned} \quad (2.10)$$

The final characteristic quantity is time – the nondimensional time is defined such that,

$$\tau = \frac{t}{t^*}. \quad (2.11)$$

With these nondimensional parameters, derivation of the equations of motion is straightforward.

From Figure 2.1 and Newton's Second Law, the equations of motion for P_3 can be written by inspection,

$$m_3 \bar{R}_3'' = -\frac{\tilde{G}m_3m_1}{R_{13}^3} \bar{R}_{13} - \frac{\tilde{G}m_3m_2}{R_{23}^3} \bar{R}_{23}. \quad (2.12)$$

The left side of the equation is comprised of the acceleration of the particle of interest with respect to an inertial observer. Derivatives with respect to dimensional time are denoted by primes. The right side is the sum of the gravitational forces acting on the particle. Fortunately, the mass of P_3 easily cancels. The derivatives can also be written explicitly with superscript I or R indicating the frame of differentiation (inertial or rotating, respectively),

$$\frac{{}^I d^2 \bar{R}_3}{dt^2} = -\frac{\tilde{G}m_1}{R_{13}^3} \bar{R}_{13} - \frac{\tilde{G}m_2}{R_{23}^3} \bar{R}_{23}. \quad (2.13)$$

Multiplying both sides by $\frac{t^{*2}}{l^*}$ and recalling $t^* = \sqrt{\frac{l^{*3}}{\tilde{G}m^*}}$,

$$\frac{{}^I d^2 \frac{\bar{R}_3}{l^*}}{d\frac{t}{t^*}^2} = -\frac{\tilde{G}m_1}{R_{13}^3} \frac{\bar{R}_{13}}{l^*} \frac{l^{*3}}{\tilde{G}m^*} - \frac{\tilde{G}m_2}{R_{23}^3} \frac{\bar{R}_{23}}{l^*} \frac{l^{*3}}{\tilde{G}m^*}. \quad (2.14)$$

Next, nondimensional quantities are substituted yielding,

$$\frac{{}^I d^2 \bar{r}_3}{d\tau^2} = -\frac{(1-\mu)\bar{r}_{13}}{r_{13}^3} - \frac{\mu\bar{r}_{23}}{r_{23}^3}. \quad (2.15)$$

The familiar “dot” notation is employed to represent differentiation with respect to nondimensional time as indicated by,

$$\ddot{\bar{r}}_3 = \frac{{}^I d^2 \bar{r}_3}{d\tau^2} = -\frac{(1-\mu)\bar{r}_{13}}{r_{13}^3} - \frac{\mu\bar{r}_{23}}{r_{23}^3}. \quad (2.16)$$

Recall, from Equation (2.7), that r_3 is the nondimensional form of the position vector with basepoint at the barycenter (fixed in both the inertial and rotating frames) and terminal point at P_3 . This nondimensional position vector is expressed in terms of rotating frame coordinates,

$$\bar{r}_3 = x\hat{x} + y\hat{y} + z\hat{z}.$$

The well known kinematic expansion (e.g., [39] p. 46) relating a vector derivative as observed in different reference frames is,

$$\frac{{}^I d\bar{r}_3}{d\tau} = \frac{{}^R d\bar{r}_3}{d\tau} + {}^I \bar{\omega}^R \times \bar{r}_3, \quad (2.17)$$

where ${}^I \bar{\omega}^R = n\hat{z}$ is the nondimensional angular velocity between the frames. A first and second application of Equation (2.17) yields the following expansions,

$$\dot{\bar{r}}_3 = (\dot{x} - ny)\hat{x} + (\dot{y} + nx)\hat{y} + \dot{z}\hat{z}, \quad (2.18)$$

$$\ddot{\bar{r}}_3 = (\ddot{x} - 2n\dot{y} - n^2x)\hat{x} + (\ddot{y} + 2n\dot{x} - n^2y)\hat{y} + \ddot{z}\hat{z}. \quad (2.19)$$

Further inspection of Figure 2.1 allows the relative position vectors \bar{r}_{13} and \bar{r}_{23} , as well as their respective norms, to be written in rotating frame coordinates as applicable, i.e.,

$$\bar{r}_{13} = (x + \mu)\hat{x} + y\hat{y} + z\hat{z}, \quad (2.20)$$

$$\bar{r}_{23} = (x - 1 + \mu)\hat{x} + y\hat{y} + z\hat{z}, \quad (2.21)$$

$$r_{13} = \sqrt{(x + \mu)^2 + y^2 + z^2}, \quad (2.22)$$

$$r_{23} = \sqrt{(x - 1 + \mu)^2 + y^2 + z^2}. \quad (2.23)$$

Combining the respective coordinates from Equations (2.16), (2.20), and (2.21) yields three second-order, coupled nonlinear scalar differential equations of motion:

$$\ddot{x} - 2n\dot{y} - n^2x = -\frac{(1-\mu)(x+\mu)}{r_{13}^3} - \frac{\mu(x-1+\mu)}{r_{23}^3}, \quad (2.24a)$$

$$\ddot{y} + 2n\dot{x} - n^2y = -\frac{(1-\mu)y}{r_{13}^3} - \frac{\mu y}{r_{23}^3}, \quad (2.24b)$$

$$\ddot{z} = -\frac{(1-\mu)z}{r_{13}^3} - \frac{\mu z}{r_{23}^3}. \quad (2.24c)$$

Equations (2.24a)–(2.24c) fully describe the evolving path of P_3 given any initial state.

The equations of motion can also be derived by incorporating the gradient of a potential function. As with any set of differential equations, a potential function can be defined for the CR3BP, that is,

$$U^* = \frac{1-\mu}{r_{13}} + \frac{\mu}{r_{23}} + \frac{1}{2}n^2(x^2 + y^2). \quad (2.25)$$

Differentiation solely of the gravitational potential function does not fully recover the EOM. Given the problem formulation relative to a rotating frame, additional terms are introduced in Equation (2.25) to accommodate the Coriolis acceleration producing a potential function with no direct physical significance. Thus, Equation (2.25) is more commonly designated a “pseudo”-potential. The resulting differential equations are then also represented in terms of the pseudo-potential, i.e.,

$$\ddot{x} = \frac{\partial U^*}{\partial x} + 2\dot{y}, \quad (2.26a)$$

$$\ddot{y} = \frac{\partial U^*}{\partial y} - 2\dot{x}, \quad (2.26b)$$

$$\ddot{z} = \frac{\partial U^*}{\partial z}. \quad (2.26c)$$

These two forms of the equations of motion (Equations (2.24) and (2.26)) are equivalent and offer a mathematical formulation for the CR3BP. The equations of motion represent a mathematical model of significant importance, and serve as a basis for a numerical investigation in a problem of sufficient complexity to display both ordered motion and chaotic behavior.

2.4 The Jacobi Integral

The differential equations in the CR3BP possess an integral of the motion, commonly termed the Jacobi integral. The Jacobi integral is derived from Equations (2.26a)–(2.26c) as the scalar product of the vector EOM with the relative velocity vector associated with P_3 . The relative velocity vector is defined and expressed,

$$\bar{v} = {}^R\dot{\bar{r}}_3 = \dot{x}\hat{x} + \dot{y}\hat{y} + \dot{z}\hat{z}. \quad (2.27)$$

The result from multiplication for the scalar product of Equation (2.27) with Equations (2.26a)–(2.26c) is three scalar terms,

$$x\ddot{x} - 2\dot{x}\dot{y} = \frac{\partial U^*}{\partial x}\dot{x}, \quad (2.28)$$

$$y\ddot{y} + 2\dot{y}\dot{x} = \frac{\partial U^*}{\partial y}\dot{y}, \quad (2.29)$$

$$z\ddot{z} = \frac{\partial U^*}{\partial z}\dot{z}, \quad (2.30)$$

which, added together, produce,

$$\dot{x}\ddot{x} + \dot{y}\ddot{y} + \dot{z}\ddot{z} = \frac{\partial U^*}{\partial x}\frac{dx}{d\tau} + \frac{\partial U^*}{\partial y}\frac{dy}{d\tau} + \frac{\partial U^*}{\partial z}\frac{dz}{d\tau}. \quad (2.31)$$

Integrating Equation (2.31) and designating the integration constant as \tilde{C} yields,

$$\frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = U^* + \tilde{C}. \quad (2.32)$$

From Equation (2.27), $v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$, which allows Equation (2.32) to be written in the recognizable form of the Jacobi integral, that is,

$$C = 2U^* - v^2, \quad (2.33)$$

where $\tilde{C} = -\frac{C}{2}$. The existence of an integral of the motion in the CR3BP is exploited to gain a better understanding of the system.

The Jacobi integral serves in a number of capacities. Fundamentally, the Jacobi integral provides a check on the accuracy of the numerical integration process. The Jacobi constant should remain constant throughout any numerical integration process. A consequence of the existence of an integral is also a potential reduction of order in the phase space. A 6th order system, like the CR3BP, can be reduced to a 5th order system, while the 4D phase

space in the planar CR3BP can be reduced to three. In this way, the Jacobi constant can also be employed to recover a full state given three of the four state variables (or five of six in the full CR3BP). Additionally, the Jacobi integral can be used as a parameter to help organize analysis in the CR3BP. These examples illustrate the value of such an integral.

2.5 Equilibrium Points

When observed in a rotating reference frame, the CR3BP is known to possess equilibrium points [7]. The equilibrium points are identified as solutions to the differential equations where the associated velocity and acceleration fields are zero. There are five such equilibrium point solutions of the CR3BP EOM, designated L_1 through L_5 . Euler first identified the points designated as L_1 through L_3 while Lagrange identified all the equilibrium points, including L_4 and L_5 [5]. The five points associated with the equilibrium solutions are commonly termed the Lagrange points (after Lagrange) or libration points after the observed behavior of bodies existing naturally near these locations.

The libration points are determined as the roots of equations that result when the gradient of the pseudo-potential (Equation (2.25)) is equal to zero; thus, the equilibrium equations become,

$$x = \frac{(1-\mu)(x+\mu)}{r_{13}^3} + \frac{\mu(x-1+\mu)}{r_{23}^3}, \quad (2.34a)$$

$$y = \frac{(1-\mu)y}{r_{13}^3} + \frac{\mu y}{r_{23}^3}, \quad (2.34b)$$

$$0 = \frac{(1-\mu)z}{r_{13}^3} + \frac{\mu z}{r_{23}^3}. \quad (2.34c)$$

Inspection of Equation (2.34c) implies that $z = 0$ for all equilibrium points, that is, all five equilibrium points exist in the plane of motion of the primaries. Further, substituting $z = 0$ into Equations (2.34a) and (2.34b) results in a coupled linear system of two equations in two unknowns. With $y \neq 0$, simplifying the system via substitution leads to $r_{13} = r_{23} = 1$. Fixing $r_{13} = r_{23} = 1$, squaring Equations (2.22) and (2.23), and equating the two equations reveals expressions for the x and y coordinates of the equilibrium points that appear off of the x axis with $y \neq 0$, that is,

$$x = \frac{1}{2} - \mu, \quad (2.35)$$

$$y = \pm \frac{\sqrt{3}}{2}. \quad (2.36)$$

These x and y values (with $y \neq 0$) correspond to points L_4 and L_5 and represent locations at the third vertex of equilateral triangles formed by P_1 , P_2 , and the respective point. The two equilateral points are denoted by convention such that L_4 corresponds to the equilibrium point with a positive y coordinate and L_5 corresponds to the equilibrium point with a negative y value. Three more equilibrium points exist when $y = 0$, thus these additional points lie along the rotating x axis and are termed the collinear points. When $y = z = 0$, the equilibrium equations simplify to a single 3rd-order equation,

$$x = \frac{(1 - \mu)(x + \mu)}{|x + \mu|^3} + \frac{\mu(x - 1 + \mu)}{|x - 1 + \mu|^3}. \quad (2.37)$$

It is evident from Equation (2.37) that singularities exist for $x = -\mu$ and $x = 1 - \mu$. These singularities bound three regions that each contain a root of the equation. These three roots correspond to the remaining equilibrium points with the common convention designating L_1 as the point interior to both primaries, L_2 is the point exterior to P_2 in the positive \hat{x} direction, and L_3 is the point exterior to P_1 in the negative \hat{x} direction. To simplify a numerical root-finding scheme, it is common to substitute the distance between the primaries and the respective equilibrium points as the variable of interest into Equation (2.37). Specifically, the following substitutions are employed,

$$x_{L_1} = 1 - \mu - \gamma_1, \quad (2.38a)$$

$$x_{L_2} = 1 - \mu + \gamma_2, \quad (2.38b)$$

$$x_{L_3} = -\mu - \gamma_3. \quad (2.38c)$$

Equation (2.37) can be further simplified by considering the three distinct regions near each collinear point independently, that is,

$L_1 : -\mu < x < 1 - \mu :$

$$x - \frac{1 - \mu}{(x + \mu)^2} + \frac{\mu}{(x - 1 + \mu)^2} = 0, \quad (2.39a)$$

$L_2 : x > 1 - \mu :$

$$x - \frac{1 - \mu}{(x + \mu)^2} - \frac{\mu}{(x - 1 + \mu)^2} = 0, \quad (2.39b)$$

$L_3 : x < -\mu :$

$$x + \frac{1 - \mu}{(x + \mu)^2} + \frac{\mu}{(x - 1 + \mu)^2} = 0. \quad (2.39c)$$

Substituting from Equation (2.38a) into Equation (2.39a) and multiplying to obtain a common denominator results in the following equation,

$$(1 - \mu - \gamma_1)(1 - \gamma_1)^2(-\gamma_1)^2 - (1 - \mu)(-\gamma_1)^2 + \mu(1 - \gamma_1)^2 = 0. \quad (2.40)$$

Expanding and collecting like terms yields a quintic polynomial that can be employed to obtain the x coordinate of L_1 to arbitrary precision. Similar steps produce relationships for the x coordinates of L_2 and L_3 . In summary,

L_1 :

$$\gamma_1^5 - (3 - \mu)\gamma_1^4 + (3 - 2\mu)\gamma_1^3 - \mu\gamma_1^2 + 2\mu\gamma_1 - \mu = 0, \quad (2.41a)$$

L_2 :

$$\gamma_2^5 + (3 - \mu)\gamma_2^4 + (3 - 2\mu)\gamma_2^3 - \mu\gamma_2^2 - 2\mu\gamma_2 - \mu = 0, \quad (2.41b)$$

L_3 :

$$\gamma_3^5 + (2 + \mu)\gamma_3^4 + (1 + 2\mu)\gamma_3^3 - (1 - \mu)\gamma_3^2 - 2(1 - \mu)\gamma_3 - (1 - \mu) = 0. \quad (2.41c)$$

Solving these equations for γ_i and then back-substituting yields the physical x coordinate corresponding to the respective libration point. The exact location of each point depends upon the relative masses of the two primaries, but a general representation of the locations appears in Figure 2.2 where the equilibrium points are depicted along with the two massive primaries (P_1 and P_2) in the rotating frame. Equilateral triangles associated with L_4 and L_5 are included to further illustrate the geometry.

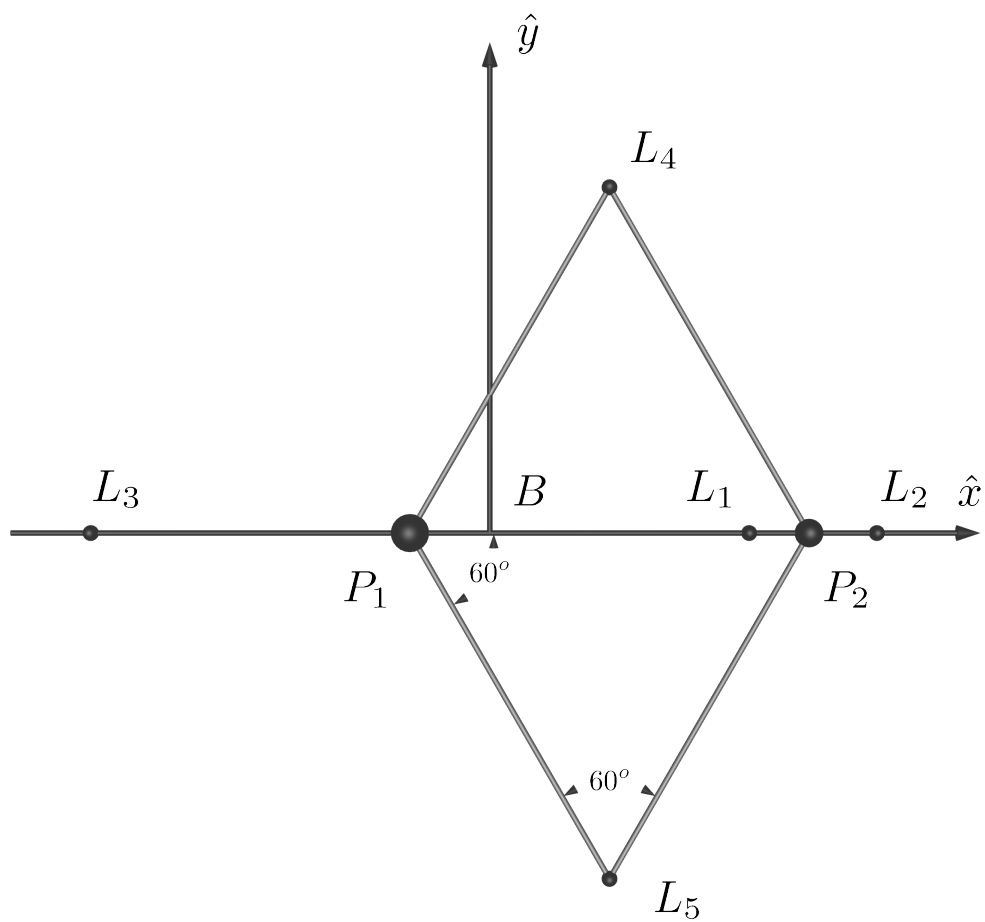


Figure 2.2. Equilibrium points in the CR3BP

Accurate knowledge of the locations of the equilibrium points in the CR3BP is critical to most numerical schemes in the study of this problem. These points become the foundation for both linear and topological analysis. This analysis, in turn, leads to a much fuller understanding of the dynamical characteristics in the problem. The equilibrium point locations in the Earth-Moon CR3BP, ($\mu \approx 0.012151$), are listed in Table 2.1 while the Jacobi constant values associated with these points are tabulated in Table 2.2.

Table 2.1
Earth-Moon Libration Points

Libration Point	x_{L_i} (non-dim)	x_{L_i} (km)	y_{L_i} (non-dim)	y_{L_i} (km)
L_1	0.836915	321700	0.0	0.0
L_2	1.155682	444230	0.0	0.0
L_3	-1.005062	-386334	0.0	0.0
L_4	0.487849	187524	0.866025	332890
L_5	0.487849	187524	-0.866025	-332890

Table 2.2
Jacobi Constant Values for Earth-Moon Libration Points

Libration Point	C (dimensionless)
L_1	3.188340986998163
L_2	3.172160349057863
L_3	3.012147136509916
L_4	2.987997064955494
L_5	2.987997064955494

2.6 Curves and Surfaces of Zero Relative Velocity

The Jacobi constant is an “energy-like” quantity and can be employed to bound the motion of P_3 given a particular C value. Recall the five equilibrium point solutions of the equations of motion for the CR3BP where both velocity and acceleration are equal to zero. If only velocity is fixed to be zero and acceleration is allowed to vary, additional observations can be made by inspection of the integral of motion. The Jacobi constant is a function of position and velocity as indicated by Equation (2.33), $C = 2U^* - v^2$, where the pseudo-potential, U^* , is a function of position only. If Equation (2.33) is rearranged to express the velocity term as a function of the pseudo-potential and the Jacobi constant, i.e.,

$$v^2 = 2U^* - C, \quad (2.42)$$

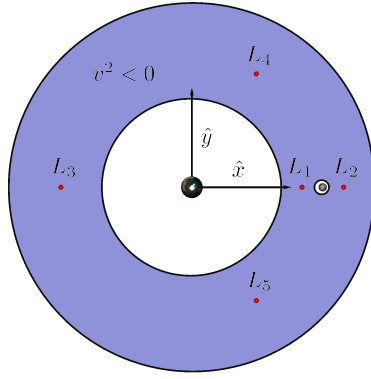
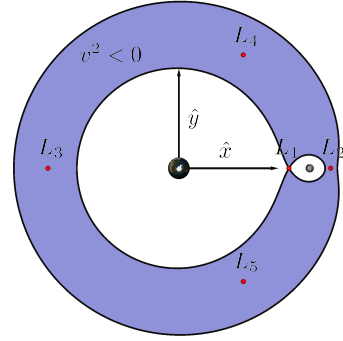
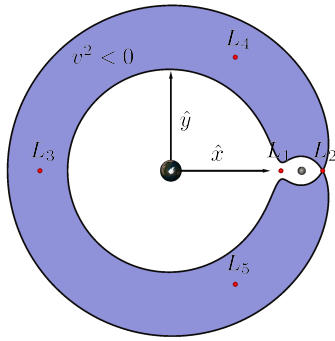
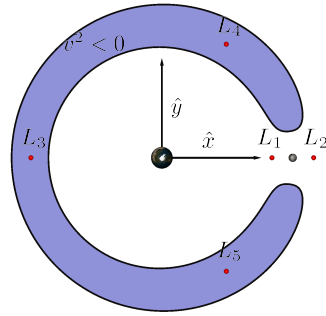
it is immediately apparent that for some value of $C > 2U^*$, the square of the velocity magnitude is negative, that is $v^2 < 0$, implying an imaginary velocity. Since P_3 exists in a physical regime, an imaginary velocity indicates a spatial region that is barred to P_3 , a *forbidden* region. If $C < 2U^*$, $v^2 > 0$, and the resulting velocity is real allowing P_3 to move in the corresponding region of configuration space. The boundary that distinguishes the region of permissible motion from the forbidden region corresponds to a zero value of velocity. Fixing velocity at zero and expanding the pseudo-potential produces an equation for a curve along which velocity is everywhere zero,

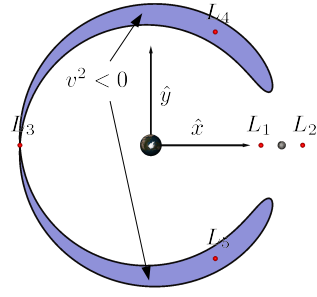
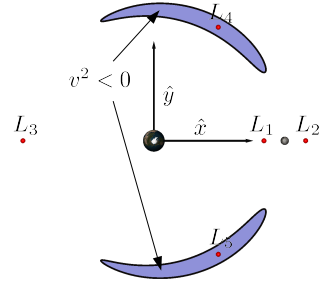
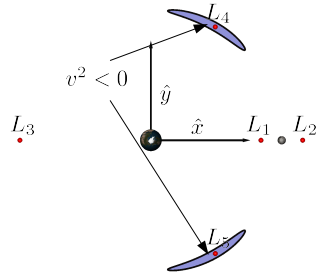
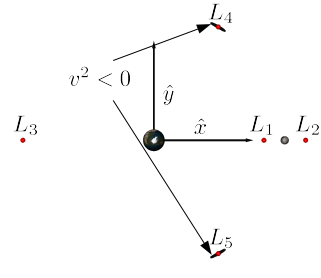
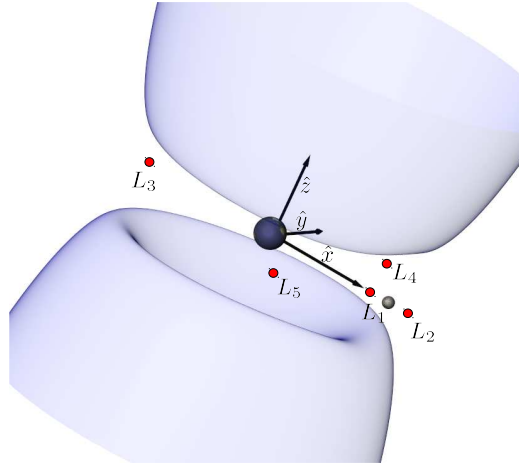
$$C = x^2 + y^2 + \frac{2(1-\mu)}{r_{13}} + \frac{2\mu}{r_{23}}. \quad (2.43)$$

These Zero Velocity Curves (ZVC) are defined by selecting a fixed value of C and solving for the positions that satisfy Equation (2.43) for a given value of C . For five distinct values of C , the acceleration in the equations of motion is also zero; these values of C correspond to the five equilibrium point positions. Finally, recall that r_{13} and r_{23} are also functions of z and Equation (2.43) can be solved with $z \neq 0$ producing a three-dimensional surface for a given value of C .

The existence of boundaries governing the motion of P_3 provides mathematical insight that helps define where P_3 can move under the influence of the natural dynamics in the system. The equilibrium points, the Jacobi constant, and zero velocity surfaces sculpt the dynamical environment providing a structure for the solution space.

A series of ZVC(S) for a range of values for Jacobi constant in the Earth-Moon system appear in Figures 2.3 and 2.4. In Figure 2.3 (a–d) and Figure 2.4 (a–d), the shaded areas represent the forbidden region for the given value of C . In Figure 2.4 (e), a 3D view demonstrates the appearance of the ZVS for $C = 2.9000$. (Also note, the Earth and Moon in these figures are not to scale.) As indicated in the figures, each of the libration points lies on a ZVC for a specific value of the Jacobi constant.

(a) $C = 3.4000$ (b) $C = 3.1883 \approx C_{L_1}$ (c) $C = 3.1722 \approx C_{L_2}$ (d) $C = 3.0722$ Figure 2.3. Earth-Moon ZVC for various C values

(a) $C = 3.0121 \approx C_{L_3}$ (b) $C = 3.0000$ (c) $C = 2.9910$ (d) $C = 2.9881 \approx C_{L_4}/C_{L_5}$ (e) $C = 2.9000$, 3D ViewFigure 2.4. More Earth-Moon ZVC(S) for various C values

2.7 Linear Variational Equations Relative to the Collinear Points

Given that the equilibrium points are already identified, characterizing motion in their local neighborhood is useful. The behavior near L_i is initially explored by linearizing relative to a collinear equilibrium point. The linear system is frequently a good first approximation to the actual nonlinear behavior in a sufficiently small vicinity near the reference solution. This first approximation can then be used to generate an initial state for numerical integration using the nonlinear EOM. To linearize, first introduce small perturbations relative to the libration points. For example, if an equilibrium point is defined in position space by a specific set of cartesian coordinates, i.e., $(x, y, z) = (a, b, c)$, perturbations are introduced such that $(x, y, z) = (a + \xi, b + \eta, c + \zeta)$. The new variables of interest are now differentiated with respect to nondimensional time, τ ,

$$\dot{x} = \dot{\xi}, \quad (2.44a) \quad \ddot{x} = \ddot{\xi}, \quad (2.44d)$$

$$\dot{y} = \dot{\eta}, \quad (2.44b) \quad \ddot{y} = \ddot{\eta}, \quad (2.44e)$$

$$\dot{z} = \dot{\zeta}, \quad (2.44c) \quad \ddot{z} = \ddot{\zeta}. \quad (2.44f)$$

Linearization is accomplished via a Taylor series expansion evaluated at one of the libration points. The Taylor series expansion of a function $f(x)$ relative to a constant reference solution a , has the following general form,

$$f(x) = f(a) + \left. \frac{df}{dx} \right|_a (x - a) + \frac{1}{2!} \left. \frac{d^2f}{dx^2} \right|_a (x - a)^2 + \frac{1}{3!} \left. \frac{d^3f}{dx^3} \right|_a (x - a)^3 + \dots \quad (2.45)$$

Note that the partials are evaluated at the reference, a . For multiple variables, the Taylor series expansion conveniently reduces to,

$$\begin{aligned} f(a + \xi, b + \eta, c + \zeta) = f(a, b, c) + \left(\xi \frac{\partial f}{\partial x} + \eta \frac{\partial f}{\partial y} + \zeta \frac{\partial f}{\partial z} \right) \Big|_{(a,b,c)} + \\ \frac{1}{2!} \left(\xi^2 \frac{\partial^2 f}{\partial x^2} + \eta^2 \frac{\partial^2 f}{\partial x \partial y} + \zeta^2 \frac{\partial^2 f}{\partial x \partial z} \right) \Big|_{(a,b,c)} + \dots \end{aligned} \quad (2.46)$$

Now, the first partial of the pseudo potential (U^*) with respect to x , to first order, is written,

$$\frac{\partial U^*}{\partial x} = \left. \frac{\partial U^*}{\partial x} \right|_{L_i} + \left. \frac{\partial^2 U^*}{\partial x^2} \right|_{L_i} \xi + \left. \frac{\partial^2 U^*}{\partial x \partial y} \right|_{L_i} \eta + \left. \frac{\partial^2 U^*}{\partial x \partial z} \right|_{L_i} \zeta + \dots \quad (2.47)$$

The value of $\frac{\partial U^*}{\partial x}$ evaluated at any libration point is zero by definition, so the 2nd–4th terms of Equation (2.47) represent the first-order expression for $\frac{\partial U^*}{\partial x}$. Similarly, expansions for $\frac{\partial U^*}{\partial y}$ and $\frac{\partial U^*}{\partial z}$ lead to the following linear equations of motion, i.e.,

$$\ddot{\xi} - 2\dot{\eta} = \left. \frac{\partial^2 U^*}{\partial x^2} \right|_{L_i} \xi + \left. \frac{\partial^2 U^*}{\partial x \partial y} \right|_{L_i} \eta + \left. \frac{\partial^2 U^*}{\partial x \partial z} \right|_{L_i} \zeta, \quad (2.48a)$$

$$\ddot{\eta} + 2\dot{\xi} = \left. \frac{\partial^2 U^*}{\partial y \partial x} \right|_{L_i} \xi + \left. \frac{\partial^2 U^*}{\partial y^2} \right|_{L_i} \eta + \left. \frac{\partial^2 U^*}{\partial y \partial z} \right|_{L_i} \zeta, \quad (2.48b)$$

$$\ddot{\zeta} = \left. \frac{\partial^2 U^*}{\partial z \partial x} \right|_{L_i} \xi + \left. \frac{\partial^2 U^*}{\partial z \partial y} \right|_{L_i} \eta + \left. \frac{\partial^2 U^*}{\partial z^2} \right|_{L_i} \zeta. \quad (2.48c)$$

Or, more succinctly,

$$\ddot{\xi} - 2\dot{\eta} = U_{xx}^*|_{L_i} \xi + U_{xy}^*|_{L_i} \eta + U_{xz}^*|_{L_i} \zeta, \quad (2.49a)$$

$$\ddot{\eta} + 2\dot{\xi} = U_{yx}^*|_{L_i} \xi + U_{yy}^*|_{L_i} \eta + U_{yz}^*|_{L_i} \zeta, \quad (2.49b)$$

$$\ddot{\zeta} = U_{zx}^*|_{L_i} \xi + U_{zy}^*|_{L_i} \eta + U_{zz}^*|_{L_i} \zeta, \quad (2.49c)$$

where $U_{ij}^* = \frac{\partial^2 U^*}{\partial i \partial j}$. Equations (2.49a)–(2.49c) are first-order linear variational equations with constant coefficients. The required partials in Equations (2.49a)–(2.49c) are evaluated directly from the pseudo-potential. Recall Equation 2.25,

$$U^* = \frac{1-\mu}{r_{13}} + \frac{\mu}{r_{23}} + \frac{1}{2}n^2(x^2 + y^2),$$

where,

$$r_{13} = \sqrt{(x+\mu)^2 + y^2 + z^2},$$

$$r_{23} = \sqrt{(x-1+\mu)^2 + y^2 + z^2}.$$

The associated first and second partials are

$$U_x^* = x - \frac{(1-\mu)(x+\mu)}{r_{13}^3} - \frac{\mu(x-1+\mu)}{r_{23}^3}, \quad (2.50a)$$

$$U_y^* = y - \frac{(1-\mu)y}{r_{13}^3} - \frac{\mu y}{r_{23}^3}, \quad (2.50b)$$

$$U_z^* = -\frac{(1-\mu)z}{r_{13}^3} - \frac{\mu z}{r_{23}^3}, \quad (2.50c)$$

$$U_{xx}^* = 1 - \frac{1-\mu}{r_{13}^3} - \frac{\mu}{r_{23}^3} + \frac{3(1-\mu)(x+\mu)^2}{r_{13}^5} + \frac{3\mu(x-1+\mu)^2}{r_{23}^5}, \quad (2.50d)$$

$$U_{xy}^* = U_{yx}^* = \frac{3(1-\mu)(x+\mu)y}{r_{13}^5} + \frac{3\mu(x-1+\mu)y}{r_{23}^5}, \quad (2.50e)$$

$$U_{xz}^* = U_{zx}^* = \frac{3(1-\mu)(x+\mu)z}{r_{13}^5} + \frac{3\mu(x-1+\mu)z}{r_{23}^5}, \quad (2.50f)$$

$$U_{yy}^* = 1 - \frac{1-\mu}{r_{13}^3} - \frac{\mu}{r_{23}^3} + \frac{3(1-\mu)y^2}{r_{13}^5} + \frac{3\mu y^2}{r_{23}^5}, \quad (2.50g)$$

$$U_{yz}^* = U_{zy}^* = \frac{3(1-\mu)yz}{r_{13}^5} + \frac{3\mu yz}{r_{23}^5}, \quad (2.50h)$$

$$U_{zz}^* = -\frac{1-\mu}{r_{13}^3} - \frac{\mu}{r_{23}^3} + \frac{3(1-\mu)z^2}{r_{13}^5} + \frac{3\mu z^2}{r_{23}^5}. \quad (2.50i)$$

The linear variational equations model behavior near the constant equilibrium points, L_i , and the associated partials are evaluated at the libration points. In evaluating the partials, any term multiplied by z in the numerator vanishes, since $z_{L_i} = 0$, so the linear variational equations are simplified, i.e.,

$$\ddot{\xi} - 2\dot{\eta} = U_{xx}^* \xi + U_{xy}^* \eta, \quad (2.51a)$$

$$\ddot{\eta} + 2\dot{\xi} = U_{yx}^* \xi + U_{yy}^* \eta, \quad (2.51b)$$

$$\ddot{\zeta} = U_{zz}^* \zeta. \quad (2.51c)$$

By inspection, it is apparent that Equation (2.51c) is uncoupled from Equations (2.51a) and (2.51b). Furthermore, the form of Equation (2.51c) reflects a simple harmonic oscillator with purely imaginary characteristic roots. Thus, the linear out-of-plane motion near the libration points is a stable harmonic oscillation with a period equal to $P = \frac{2\pi}{\sqrt{|U_{zz}^*|}}$. The two second-order linear differential equations, Equations (2.51a) and (2.51b), are easily represented as a system of four first-order linear equations and written in matrix form,

$$\dot{\bar{x}} = A_{4 \times 4} \bar{x}, \quad (2.52)$$

where,

$$\bar{x} = \begin{bmatrix} \xi \\ \eta \\ \dot{\xi} \\ \dot{\eta} \end{bmatrix}, \quad (2.53)$$

and,

$$A_{4 \times 4} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ U_{xx}^* & U_{xy}^* & 0 & 2 \\ U_{yx}^* & U_{yy}^* & -2 & 0 \end{bmatrix}. \quad (2.54)$$

Or, A can be expressed more compactly in the form,

$$A = \left[\begin{array}{c|c} 0_{2 \times 2} & I_{2 \times 2} \\ \hline U_{2 \times 2}^* & \Omega_{2 \times 2} \end{array} \right], \quad (2.55)$$

where the 2×2 sub-matrices are directly defined in the matrix in Equation (2.54). General solutions for a first-order system of linear equations are of the form,

$$\xi = \sum_{i=1}^4 A_i e^{\lambda_i t}, \quad (2.56a)$$

$$\eta = \sum_{i=1}^4 B_i e^{\lambda_i t}, \quad (2.56b)$$

where A_i and B_i are coupled constants of integration and λ_i are the roots of the characteristic equation of $A_{4 \times 4}$. Two of the characteristic roots (λ_1 and λ_2) are real, with one root greater than zero and the other negative. These real roots represent attracting motion associated with the negative value and repelling motion corresponding to the positive root. The other two roots (λ_3 and λ_4) are imaginary, indicating corresponding oscillatory modes. Thus, the collinear points possess both oscillatory and hyperbolic modes. The relationship between A_i and B_i is developed by differentiating and substituting the solutions (Equations (2.56a) and (2.56b)) into the linear variational equations (Equations (2.51a) and (2.51b)), ultimately producing the relationship,

$$B_i = \alpha_i A_i, \quad (2.57)$$

with,

$$\alpha_i = \frac{\lambda_i^2 - U_{xx}^*}{2\lambda_i}. \quad (2.58)$$

Notwithstanding the existence of hyperbolic behavior near the collinear points, initial conditions can be selected to suppress the exponential decay and divergence associated with λ_1 and λ_2 . For $A_1 = A_2 = 0$ the linear equations are further reduced, noting that $\lambda_4 = -\lambda_3$ and, by Equation (2.58), $\alpha_4 = -\alpha_3$,

$$\xi = A_3 e^{\lambda_3 t} + A_4 e^{-\lambda_3 t}, \quad (2.59a)$$

$$\eta = \alpha_3 A_3 e^{\lambda_3 t} + \alpha_4 A_4 e^{-\lambda_3 t}. \quad (2.59b)$$

Given the following definitions,

$$\beta_1 = 2 - \frac{U_{xx}^* + U_{yy}^*}{2}, \quad (2.60)$$

$$\beta_2 = \sqrt{-U_{xx}^* U_{yy}^*}, \quad (2.61)$$

$$s = \sqrt{\beta_1 + \sqrt{\beta_1^2 + \beta_2^2}}, \quad (2.62)$$

$$\beta_3 = \frac{s^2 + U_{xx}^*}{2s}. \quad (2.63)$$

Equations (2.59a) and (2.59b) are conveniently expressed in terms of the initial parameters, ξ_0 and η_0 ,

$$\xi = \xi_0 \cos(s(t - t_0)) + \frac{\eta_0}{\beta_3} \sin(s(t - t_0)), \quad (2.64a)$$

$$\eta = \eta_0 \cos(s(t - t_0)) - \beta_3 \xi_0 \sin(s(t - t_0)). \quad (2.64b)$$

Equations (2.64a) and (2.64b) describe an ellipse about any one of the collinear points with the associated point at its center. The elliptical motion possesses a period of $P = \frac{2\pi}{s}$ and an eccentricity such that $e = \sqrt{1 - \beta_3^{-2}}$, and represents the analytical solution to the linear variational equations with parameters A_1 and A_2 selected to suppress divergent and damping behavior. Two of the four initial conditions are constrained by this selection, and if ξ_0 and η_0 reflect the free parameters, then $\dot{\xi}_0$ and $\dot{\eta}_0$ are constrained by Equations (2.64a) and (2.64b),

$$\dot{\xi}_0 = \frac{\eta_0 s}{\beta_3}, \quad (2.65a)$$

$$\dot{\eta}_0 = -\beta_3 \xi_0 s. \quad (2.65b)$$

Such a linear elliptical solution near the L_1 libration point in the Earth-Moon ($\mu \approx 0.01215$) system is represented in Figure 2.5. The ellipse is generated from an initial state such that $\xi_0 = 0.01$ nondimensional units and $\eta_0 = 0$ ($x \approx x_{L_1} + 3843.88\text{km}$); $\dot{\xi}_0$ and $\dot{\eta}_0$ result from Equations (2.65a) and (2.65b). The period of the elliptical orbit is approximately 5.5 days. The direction of motion along the ellipse is clockwise on the figure as indicated by the color scale with time increasing from blue to red. This process of judiciously selecting initial conditions extends to the unstable modes near the collinear points to elicit solutions that diverge from the vicinity of the collinear points. Additionally, this type of analysis is applied near the equilateral points where similar elliptical results are observed.

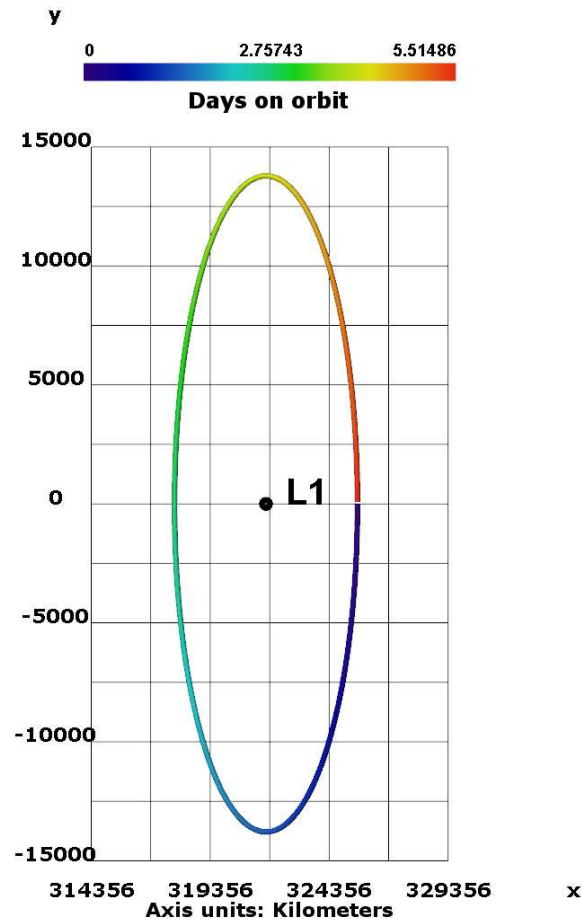


Figure 2.5. A linear ellipse about L_1

In the vicinity of the collinear points, the linear analysis supplies absolutely no guarantee for the behavior in the nonlinear system. However, the analysis does produce a useful first approximation for the nonlinear behavior given a sufficiently small region of interest. In this case, using the linear solution to extract initial conditions close to an equilibrium point yields an initial state suitable for numerical integration in the nonlinear system. Once this state is propagated it can be corrected to yield a solution in the nonlinear system. In turn, the initial nonlinear solutions can be employed to generate other nonlinear solutions. It is evident that linear analysis is a critical step in the progression for obtaining nonlinear solutions beyond the equilibrium points.

2.8 The State Transition Matrix

Linear variational equations are not strictly confined to the neighborhood near a libration point. In fact, variational equations can predict behavior relative to any reference arc in the CR3BP, and the reference need not be constant. This concept is illustrated in Figure 2.6 where an initial perturbation at time t_0 , $\delta\bar{x}_0$, is added to the reference resulting in an eventual variation, $\delta\bar{x}_f$, at t_f . The relationship between an initial variation and the

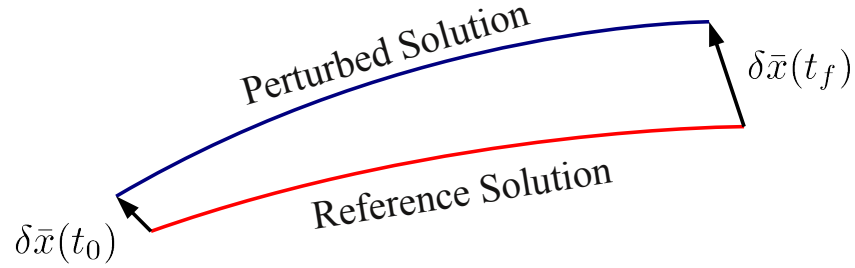


Figure 2.6. Perturbed and reference solutions

variation along a neighboring arc is determined by perturbing the reference solution, $\bar{x}_0(t)$,

$$\bar{x}(t) = \bar{x}_0(t) + \delta\bar{x}(t). \quad (2.66)$$

The equations of motion are of the form, $\dot{\bar{x}} = \bar{f}(\bar{x}, t)$, and the first derivative of Equation (2.66) is written,

$$\dot{\bar{x}} = \dot{\bar{x}}_0 + \delta\dot{\bar{x}} = \bar{f}(\bar{x}, t). \quad (2.67)$$

It is convenient to expand Equation (2.67) as a Taylor series relative to the reference solution. Truncating to first-order yields,

$$\dot{\bar{x}}_0 + \delta\dot{\bar{x}} \approx \bar{f}(\bar{x}_0, t) + \left. \frac{\partial \bar{f}}{\partial \bar{x}} \right|_{\bar{x}_0(t)} \delta\bar{x}. \quad (2.68)$$

Since $\dot{\bar{x}}_0 = \bar{f}(\bar{x}_0, t)$, Equation 2.68 simplifies,

$$\delta\dot{\bar{x}} = \left. \frac{\partial \bar{f}}{\partial \bar{x}} \right|_{\bar{x}_0(t)} \delta\bar{x}, \quad (2.69)$$

where $\frac{\partial \bar{f}}{\partial \bar{x}} \big|_{\bar{x}_0(t)} = A_{6 \times 6}(t)$, consistent with Equation (2.52), but now all six dimensions are represented and time is expressed as an explicit parameter. Thus, $A_{6 \times 6}(t)$ is of the form,

$$A_{6 \times 6}(t) = \left[\begin{array}{c|c} 0_{3 \times 3} & I_{3 \times 3} \\ \hline U_{3 \times 3}^* & \Omega_{3 \times 3} \end{array} \right], \quad (2.70)$$

where $U_{3 \times 3}^*$ includes the partials that are evaluated on the reference arc. Thus, $U_{3 \times 3}^*$ and, as a consequence $A_{6 \times 6}$, vary as a function of time. The matrix $\Omega_{3 \times 3}$ is 3×3 such that,

$$\Omega_{3 \times 3} = \begin{bmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (2.71)$$

The submatrices, $I_{3 \times 3}$ and $0_{3 \times 3}$ are the three-dimensional identity and zero matrices, respectively. The general solution to Equation (2.69) is,

$$\delta \bar{x}(t) = \Phi(t, t_0) \delta \bar{x}(t_0), \quad (2.72)$$

where $\Phi(t, t_0)$ is the matrix formed from the basis of the fundamental set of solutions. Within the context of the CR3BP, this is the State Transition Matrix (STM). The six-dimensional variational state vectors, $\delta \bar{x}(t)$ and $\delta \bar{x}(t_0)$, are defined,

$$\delta \bar{x}(t) = \begin{bmatrix} \delta x_f & \delta y_f & \delta z_f & \delta \dot{x}_f & \delta \dot{y}_f & \delta \dot{z}_f \end{bmatrix}^T, \quad (2.73)$$

$$\delta \bar{x}(t_0) = \begin{bmatrix} \delta x_0 & \delta y_0 & \delta z_0 & \delta \dot{x}_0 & \delta \dot{y}_0 & \delta \dot{z}_0 \end{bmatrix}^T, \quad (2.74)$$

where the notation \square^T indicates the matrix transpose. The matrix $\Phi(t, t_0)$ consists of the first partials, $\phi_{ij} = \frac{\partial x_i}{\partial x_{j,0}}$,

$$\Phi(t, t_0) = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} & \phi_{15} & \phi_{16} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} & \phi_{25} & \phi_{26} \\ \phi_{31} & \phi_{32} & \phi_{33} & \phi_{34} & \phi_{35} & \phi_{36} \\ \phi_{41} & \phi_{42} & \phi_{43} & \phi_{44} & \phi_{45} & \phi_{46} \\ \phi_{51} & \phi_{52} & \phi_{53} & \phi_{54} & \phi_{55} & \phi_{56} \\ \phi_{61} & \phi_{62} & \phi_{63} & \phi_{64} & \phi_{65} & \phi_{66} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial y_0} & \frac{\partial x}{\partial z_0} & \frac{\partial x}{\partial \dot{x}_0} & \frac{\partial x}{\partial \dot{y}_0} & \frac{\partial x}{\partial \dot{z}_0} \\ \frac{\partial y}{\partial x_0} & \frac{\partial y}{\partial y_0} & \frac{\partial y}{\partial z_0} & \frac{\partial y}{\partial \dot{x}_0} & \frac{\partial y}{\partial \dot{y}_0} & \frac{\partial y}{\partial \dot{z}_0} \\ \frac{\partial z}{\partial x_0} & \frac{\partial z}{\partial y_0} & \frac{\partial z}{\partial z_0} & \frac{\partial z}{\partial \dot{x}_0} & \frac{\partial z}{\partial \dot{y}_0} & \frac{\partial z}{\partial \dot{z}_0} \\ \frac{\partial \dot{x}}{\partial x_0} & \frac{\partial \dot{x}}{\partial y_0} & \frac{\partial \dot{x}}{\partial z_0} & \frac{\partial \dot{x}}{\partial \dot{x}_0} & \frac{\partial \dot{x}}{\partial \dot{y}_0} & \frac{\partial \dot{x}}{\partial \dot{z}_0} \\ \frac{\partial \dot{y}}{\partial x_0} & \frac{\partial \dot{y}}{\partial y_0} & \frac{\partial \dot{y}}{\partial z_0} & \frac{\partial \dot{y}}{\partial \dot{x}_0} & \frac{\partial \dot{y}}{\partial \dot{y}_0} & \frac{\partial \dot{y}}{\partial \dot{z}_0} \\ \frac{\partial \dot{z}}{\partial x_0} & \frac{\partial \dot{z}}{\partial y_0} & \frac{\partial \dot{z}}{\partial z_0} & \frac{\partial \dot{z}}{\partial \dot{x}_0} & \frac{\partial \dot{z}}{\partial \dot{y}_0} & \frac{\partial \dot{z}}{\partial \dot{z}_0} \end{bmatrix}, \quad (2.75)$$

or, more compactly,

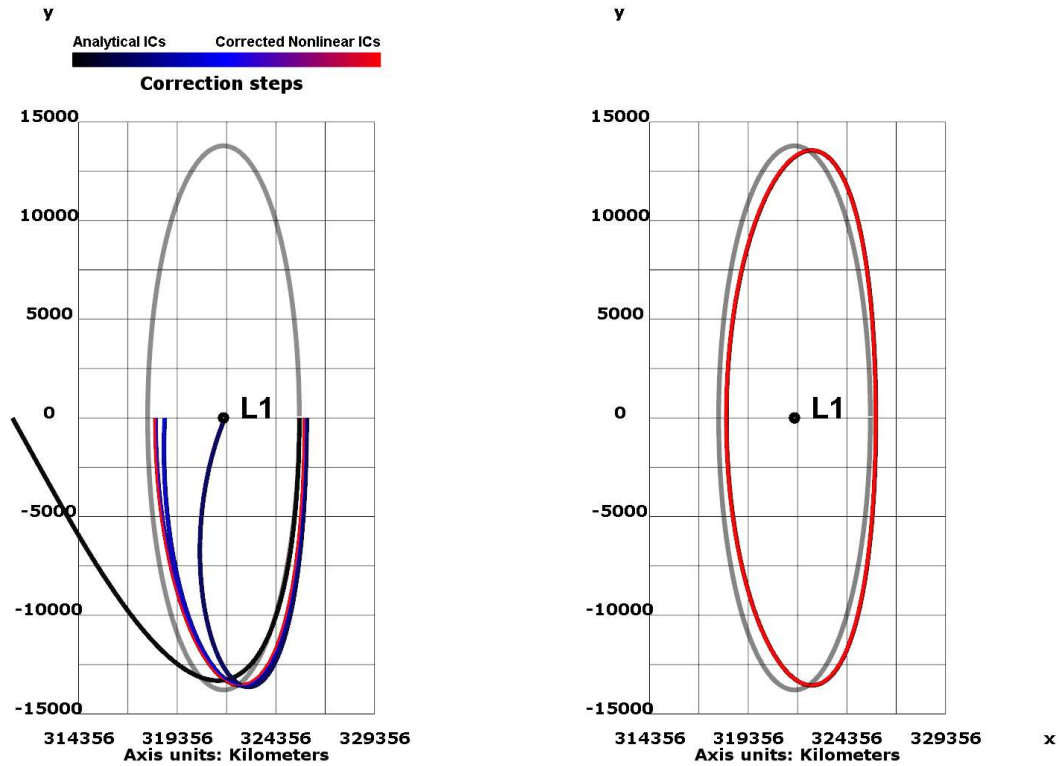
$$\Phi(t, t_0) = \left[\begin{array}{c|c} \bar{\phi}_{rr} & \bar{\phi}_{rv} \\ \hline \bar{\phi}_{vr} & \bar{\phi}_{vv} \end{array} \right], \quad (2.76)$$

where the 3×3 submatrices indicate partials representing the relationships between position and velocity vectors at different times. Since $\Phi(t, t_0)$ is also a solution of the differential equations, its evolution is governed by the same relationship,

$$\dot{\Phi}(t, t_0) = A_{6 \times 6}(t)\Phi(t, t_0). \quad (2.77)$$

This matrix equation is easily resolved into 36 scalar differential equations. These equations are numerically integrated along with the state to produce first-order variational information indicating the effect of a change in the initial state on the final state at a future time. To integrate the variational equations, initial conditions are defined such that $\Phi(t_0, t_0) = I_{6 \times 6}$, the six-dimensional identity matrix.

With the information supplied by the STM, a differential corrections scheme is formulated. Figure 2.7 illustrates one such corrections strategy. In Figure 2.7 (a), the linear



(a) Correcting the linear solution to produce a periodic orbit in the nonlinear model
 (b) Linear approximation (gray) and final nonlinear periodic solution (red)

Figure 2.7. Correcting a linear solution with the STM

initial conditions are propagated to the x axis in the nonlinear model and then numerically corrected. The linear approximation is gray and subsequent corrections steps are colored in order from black to blue to red. In Figure 2.7 (b), the corrected nonlinear orbit (red) is overlain on the linear ellipse (gray). In both cases, the direction of motion along the respective arcs is clockwise starting on the x axis to the right of L_1 . Because the STM is a linear approximation, its accuracy decreases for large initial variations or over lengthy time integration, and is, therefore, most effectively employed over appropriate integration times or in an iterative algorithm to correct initial conditions. In the previous illustration, Figure 2.7, the initial conditions obtained from the linear solution are updated based on the correction required to obtain a perpendicular crossing of the x axis, as estimated by the STM.

2.9 Periodic Orbits, a Shooting Example

Dynamical systems often possess solutions in the form of periodic orbits. Since the state variables in the differential equations that model the CR3BP are position and velocity vectors, the term *periodic orbit* does, in fact, represent a repeating trajectory in configuration space. The complexity in the CR3BP yields an infinite number of periodic orbits as well as families of such orbits, a consideration first examined by Poincaré [8] and further established later by Moulton and others [12]. In the general topological sense, the term *manifold* indicates a structured subspace of a larger space, and periodic orbits in the CR3BP can be envisioned as one-dimensional manifolds embedded in a three-dimensional configuration space. For comparison, the equilibrium points are zero-dimensional manifolds embedded in the same 3D space. Considering periodic orbits as manifolds conveys one aspect of the value such orbits supply by further scoping the structure of the problem.

One family of periodic orbits that is of particular interest in the CR3BP is the planar family of periodic Lyapunov orbits associated with the collinear libration points. This family of periodic orbits are named in honor of Aleksandr Lyapunov whose work on stability analysis in dynamical systems is well known [40]. Lyapunov families exist about each of the collinear points and, therefore, are categorized by their respective point. For example, the

L_1 Lyapunov family is a family of planar periodic orbits that exist about L_1 . Additional, related planar families of orbits exist in the vicinity of the equilateral libration points.

Given an initial orbit, for example, a linear approximation or nonlinear solution such as those plotted in Figures 2.5 and 2.7 respectively, additional solutions are generated by applying a targeting scheme. An entire family of solutions is produced through a continuation process. The targeting strategy is an example of a single shooting approach where an initial state is integrated for some prescribed duration $(t_f - t_0)$, and then the error between the final state and the targeted final state is employed to update the initial state iteratively until some convergence criteria is met. Specifically, in the previous example, the STM appears in an update equation to modify the initial state based on the difference between an integrated final state on the x axis and a perpendicular crossing indicated by a state with $\dot{x} = \dot{z} = 0$. In this particular case, only certain parameters of the initial state are allowed to vary. Specifically, \dot{y}_0 is fixed while x_0 is allowed to shift based on the update information supplied by the STM. The full details of deriving such a targeter from the STM relationship follow. Variations in time can be incorporated as appropriate, that is,

$$\delta\bar{x} = \left. \frac{\partial\bar{x}}{\partial\bar{x}_0} \right|_{\bar{x}_0(t)} \delta\bar{x}_0 + \left. \frac{\partial\bar{x}}{\partial t} \right|_{\bar{x}_0(t)} \delta t, \quad (2.78)$$

or,

$$\delta\bar{x} = \Phi(t, t_0)\delta\bar{x}_0 + \dot{\bar{x}}|_t \delta t. \quad (2.79)$$

A matrix expression is often more convenient,

$$\delta\bar{x} = \begin{bmatrix} \Phi & \dot{\bar{x}}(t) \end{bmatrix} \begin{bmatrix} \delta\bar{x}_0 \\ \delta t \end{bmatrix}. \quad (2.80)$$

Many targeting schemes are available to correct a given parameter by varying one or several initial parameters, but only a subset of the relationships represented by Equation (2.80) apply to the current example. Specifically, the relationship that yields the correct \dot{x}_f by varying x_0 is required. Using $\delta\dot{x}_f$ from $\delta\bar{x}$ and the fourth row of the augmented matrix in Equation (2.80), produces the necessary expression,

$$\delta\dot{x}_f = \phi_{(4,1)}\delta x_0 + \phi_{(4,2)}\delta y_0 + \phi_{(4,3)}\delta z_0 + \phi_{(4,4)}\delta\dot{x}_0 + \phi_{(4,5)}\delta\dot{y}_0 + \phi_{(4,6)}\delta\dot{z}_0 + \ddot{x}\delta t. \quad (2.81)$$

For the planar Lyapunov families, the z and \dot{z} terms are always zero. Choosing to constrain \dot{y}_0 implies $\delta\dot{y}_0 = 0$, and the requirement, $\dot{x}_0 = 0$, for a perpendicular departure indicates

that $\delta\dot{x}_0 = 0$. Finally, considering that the departure is always from the x axis, $\delta y_0 = 0$. Incorporating these considerations simplifies Equation (2.81), that is,

$$\delta\dot{x}_f = \phi_{(4,1)}\delta x_0 + \ddot{x}\delta t. \quad (2.82)$$

Next, enforcing $y_f = 0$ allows for the elimination of δt from Equation (2.82). Following steps similar to those above, with $\delta y_f = 0$, and solving for δt ,

$$\delta y_f = \phi_{(2,1)}\delta x_0 + \dot{y}\delta t, \quad (2.83)$$

$$\delta t = -\phi_{(2,1)} \frac{\delta x_0}{\dot{y}}. \quad (2.84)$$

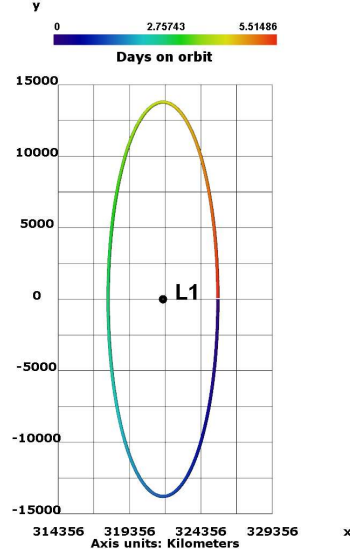
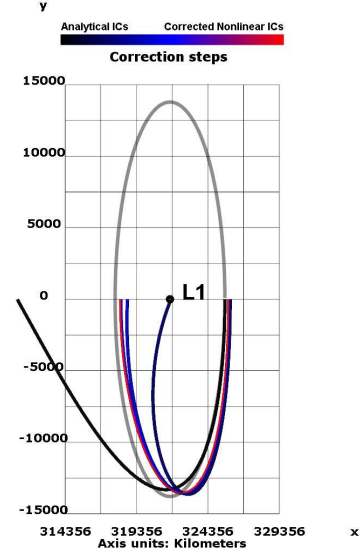
Finally, substitution back into Equation (2.82) produces the targeting equation for this example, that is,

$$\delta\dot{x}_f = \left(\phi_{(4,1)} - \frac{\ddot{x}}{\dot{y}} \phi_{(2,1)} \right) \delta x_0. \quad (2.85)$$

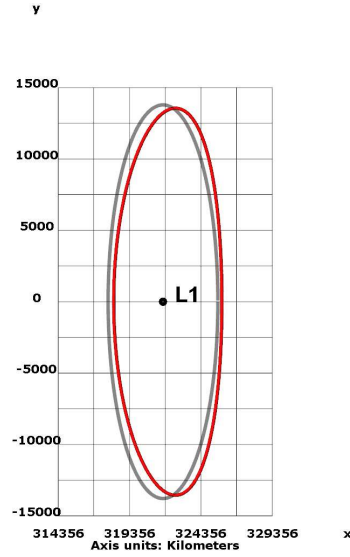
This targeting equation is employed in an iterative algorithm, and, in general, such an algorithm is described by the following steps.

1. Obtain an initial guess, either from linear analysis or continuation.
2. Propagate the initial guess and the STM via numerical integration until reaching some state near the desired final condition (in the current example, propagate to the subsequent x -axis crossing).
3. Evaluate the difference between the actual final state and the desired final state. Use this difference to generate a correction via the update equation.
4. Update the initial guess and iterate until a desired criteria for accuracy is met (e.g., a final error on the order of 10^{-12} or similar).

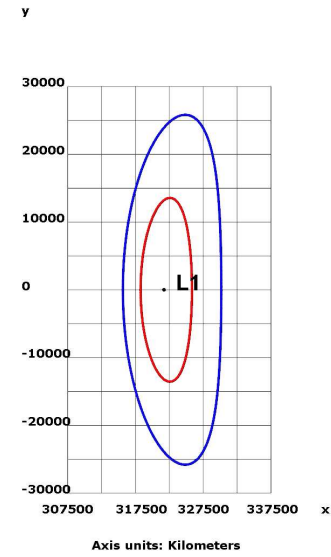
In Figure 2.8, the process is illustrated. The sequence of frames in Figures 2.8 (a–c) illustrates the computation of a periodic orbit near the L_1 libration point in the Earth-Moon ($\mu \approx 0.01215$) system. From a linear approximation, the corrections process produces a numerical solution. The nonlinear result is then employed as an initial guess to generate other nonlinear solutions. Recall that in Figure 2.8 (a), an analytical linear ellipse about L_1 is plotted ($\xi_0 = 0.01$, $\eta_0 = 0$). In Figure 2.8 (b), the linear initial conditions are propagated to the x axis in the nonlinear model and then numerically corrected. The linear ellipse is

(a) A linear ellipse about L_1 

(b) Correcting the nonlinear solution



(c) Linear (gray) and nonlinear (red) solutions



(d) Numerically determined periodic orbits

Figure 2.8. Transition from a linear solution to nonlinear solutions

gray and subsequent iterations are colored in order from black to blue to red. In Figure 2.8 (c), the corrected nonlinear Lyapunov orbit (red) is overlain on the linear approximation (gray). Finally, in Figure 2.8 (d), the first nonlinear solution is plotted in red with a second, related nonlinear orbit in blue.

The success of the previously described algorithm, that is, generating periodic orbits that are symmetric across the x axis, is due to the “mirror theorem”. The mirror theorem (Roy and Ovenden [41]) is based in the fact that the symmetric form of the equations of motion allow for mirroring across the x axis. After a single “half-orbit” is generated in the nonlinear problem via single-shooting, it is reflected across the x axis to form a full periodic orbit. The family of orbits is extended using continuation; the natural continuation parameter is selected as appropriate (for example, \dot{y}_0 or C are two of many available choices in the Earth-Moon CR3BP). Families of planar orbits about each of the five libration points appear in Figures (2.9) and (2.10). The collinear Lyapunov families of planar periodic orbits are plotted near each collinear libration point. Planar families of periodic orbits in the vicinity of the equilateral points also appear in Figure 2.9 (d). (Note that the corrections algorithm is modified to produce periodic orbits near L_4/L_5 since these orbits are asymmetric.) The direction of motion along all the orbits in Figure 2.9 is indicated by black arrows in each frame. These families are comprised of an infinite set of orbits but are represented by a finite number of orbits for clarity. The orbits in the figures are colored consistent with the associated value of Jacobi constant. Other families of periodic orbits can be determined by locating bifurcating orbits in the planar families and then growing those orbits out-of-plane by continuation.

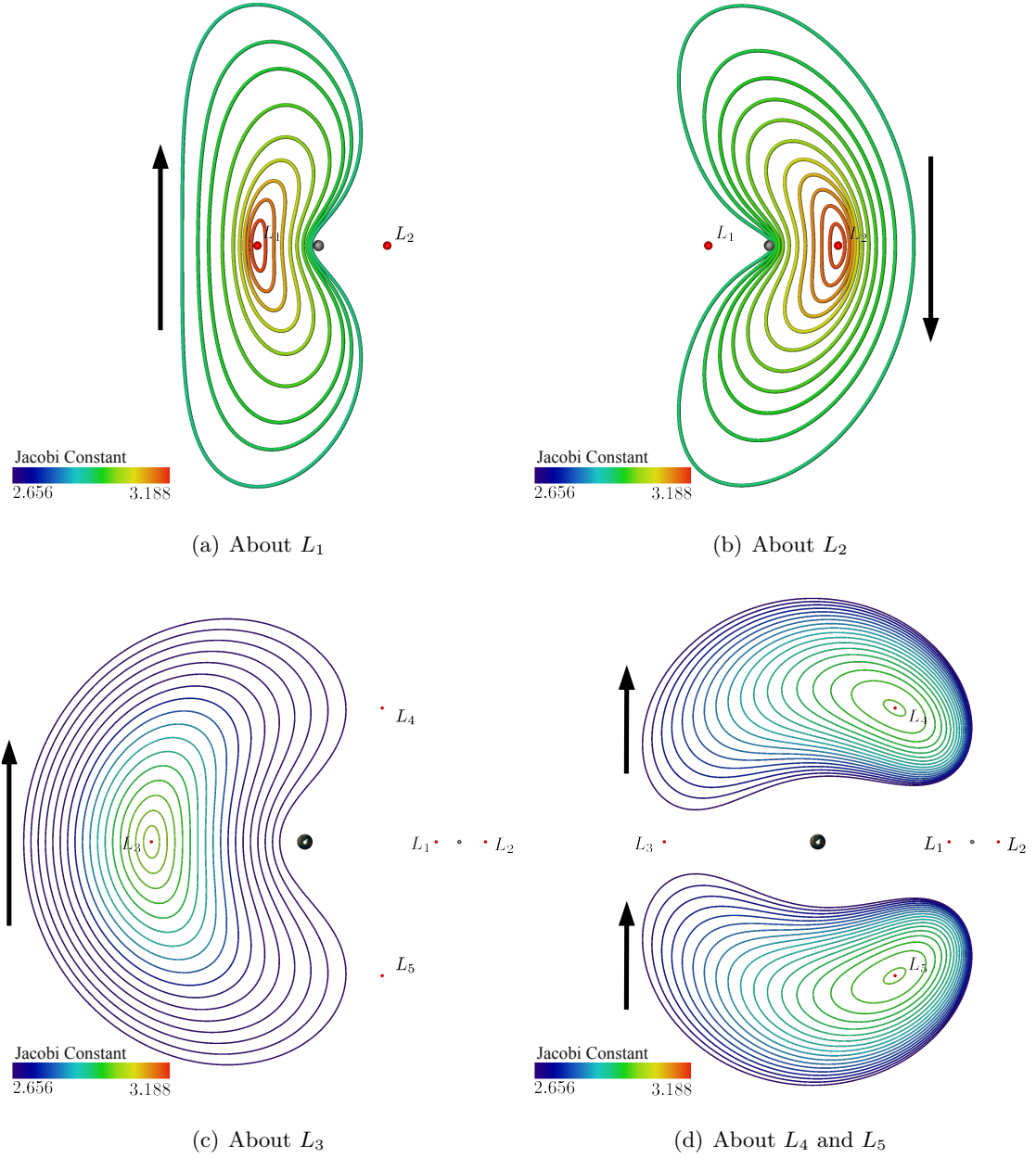


Figure 2.9. Families of planar orbits about each of the libration points separately

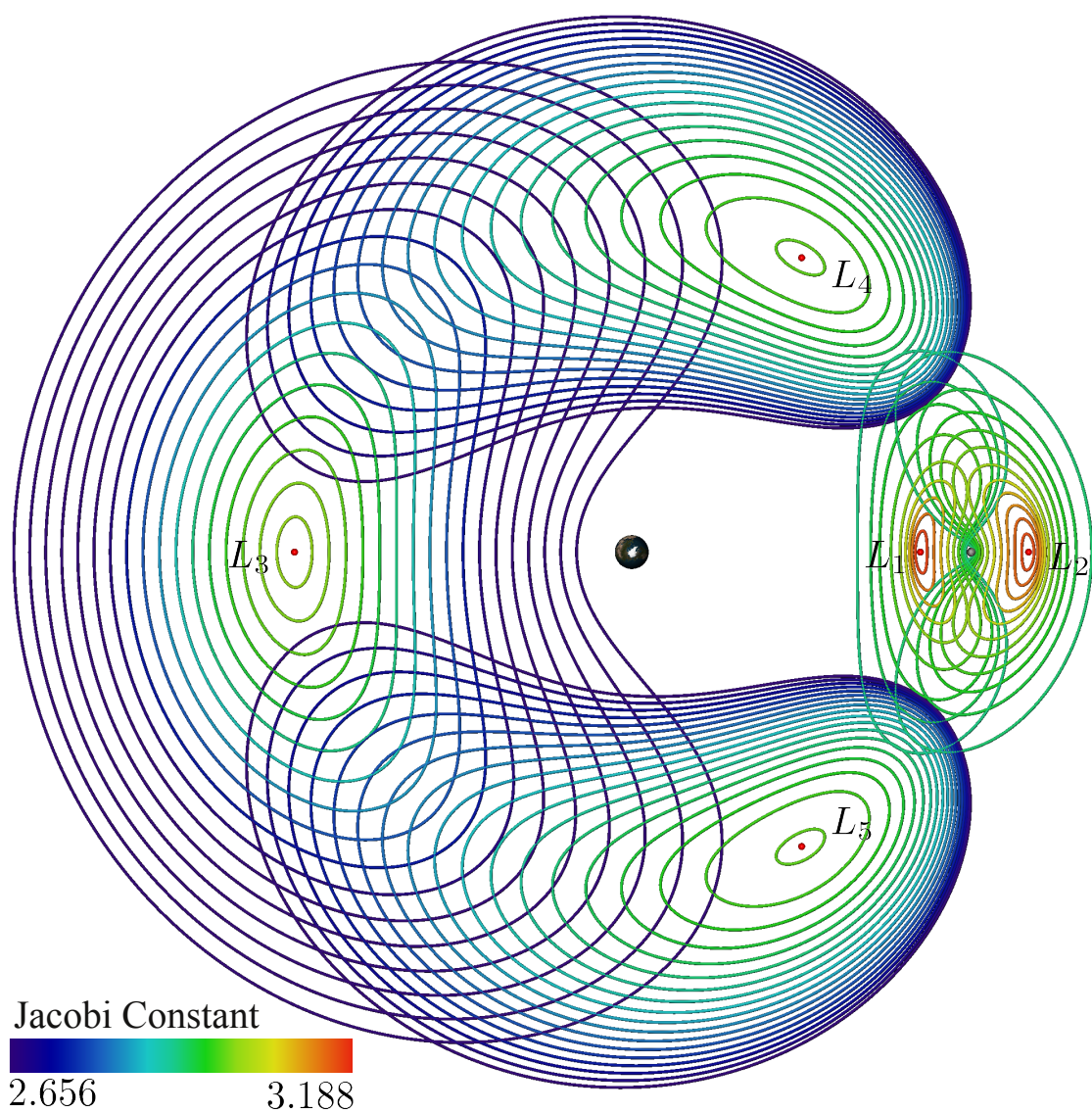


Figure 2.10. Families of planar orbits about each of the libration points

3. LAGRANGIAN COHERENT STRUCTURES

The concept of Lagrangian Coherent Structures (LCS) represents a relatively new approach for analyzing behavior in a nonlinear system. The foundations meld dynamical systems theory with ideas commonly employed in the study of fluid flows. In this chapter, the Lagrangian description of motion is introduced as well as the fundamentals of LCS. Additional notions are discussed to highlight the application of LCS to dynamical systems, particularly the CR3BP.

3.1 Lagrangian Analysis

The main focus of dynamical analysis is describing behavior and predicting motion. It is, therefore, not surprising that more than one approach can be successful in accomplishing this dynamical imperative. Of particular interest to this investigation is the Lagrangian description of motion as it relates to an Eulerian perspective. Bennett [42] indicates that the Lagrangian formulation is embodied by the “concept of conservation of particle identity, which is perhaps *the* intrinsically Lagrangian concept.”¹ Essentially, in a Lagrangian perspective, the focus is a description of the properties of an infinitesimal volume of space, that is, a *parcel* or *material point*, as these properties evolve with time. For example, the density in a fluid parcel may change with time as the volume moves in a flow or as the surroundings otherwise change. In an Eulerian viewpoint, the same intrinsic properties are described in terms of a snapshot of the entire domain. For example, an Eulerian view reflects the density at a specific point for a given time, regardless of the “material” currently present in the immediate vicinity of that point. For the specific example of fluid density, different mathematical representations exist for both perspectives. In the Lagrangian description, the density ρ for parcel p is represented at time t as $\rho_p(t)$, while an Eulerian description models the density ρ for a particular position \bar{r} at time t and employs the notation $\rho(\bar{r}, t)$.

¹A. Bennett, *Lagrangian Fluid Dynamics*, p. xiv

Lagrangian analysis offers additional insight that is not readily available from an Eulerian perspective.

A fundamental idea associated with the Lagrangian viewpoint of the behavior is a capability to incorporate the intermediate changes associated with a particle as it evolves between two points, say points a and b , into a measurement. The Eulerian description defines the velocity relative to a given frame at point a and the velocity with respect to the same frame at point b independently. If the velocity at b is changing with respect to the velocity at a , the associated difference is described as the acceleration of point b with respect to point a as observed from the common frame. In contrast, the corresponding Lagrangian view considers the parcel that originates at point a and moves toward point b , continually tracking the difference in velocity of one point relative to the other. Thus, the Lagrangian acceleration is a function of the instantaneous Eulerian acceleration as well as the acceleration occurring in traversal. Burr [43] develops this idea with the *substantial* or *material* derivative. If a vector field is represented as $\bar{f}(\bar{r}, t)$, the Lagrangian time derivative for parcel p is evaluated as,

$$\frac{D\bar{f}}{Dt} = \left. \frac{d\bar{f}}{dt} \right|_p = \lim_{\delta t \rightarrow 0} \frac{\bar{f}(\bar{r} + \bar{v}\delta t, t + \delta t) - \bar{f}(\bar{r}, t)}{\delta t}, \quad (3.1)$$

while p instantaneously moves with Eulerian velocity $\bar{v} = \frac{d\bar{r}}{dt}$. Expanding to form the Taylor series for $\bar{f}(\bar{r} + \bar{v}\delta t, t + \delta t)$, evaluated at (\bar{r}, t) , yields,

$$\bar{f}(\bar{r} + \bar{v}\delta t, t + \delta t) = \bar{f}(\bar{r} + \delta\bar{r}, t + \delta t) \approx \bar{f}(\bar{r}, t) + (\bar{r} + \delta\bar{r} - \bar{r}) \frac{\partial \bar{f}}{\partial \bar{r}}(\bar{r}, t) + (t + \delta t - t) \frac{\partial \bar{f}}{\partial t}(\bar{r}, t) + \dots, \quad (3.2)$$

or,

$$\bar{f}(\bar{r} + \bar{v}\delta t, t + \delta t) \approx \bar{f}(\bar{r}, t) + \delta\bar{r} \nabla \bar{f}(\bar{r}, t) + \delta t \frac{\partial \bar{f}}{\partial t}(\bar{r}, t). \quad (3.3)$$

Rearranging, dividing by δt , and evaluating in the limit yields the substantial derivative,

$$\frac{D\bar{f}}{Dt} = \frac{\partial \bar{f}}{\partial t} + \bar{v} \cdot \nabla \bar{f}. \quad (3.4)$$

Equation (3.4) also serves as an expression relating derivatives of the same quantity with respect to different observers, i.e., Lagrangian versus Eulerian. In this equation, the Lagrangian derivative, $\frac{D\bar{f}}{Dt}$, of a function $\bar{f}(\bar{r}, t)$ with respect to time is evaluated in terms of the Eulerian derivative, $\frac{\partial \bar{f}}{\partial t}$, plus the inner product of the velocity at \bar{r} and the gradient of

\bar{f} at \bar{r} . For example, recall parcel p . The Lagrangian acceleration for position \bar{r} can be written,

$$\frac{D\bar{v}}{Dt} = \frac{\partial\bar{v}}{\partial t} + \bar{v} \cdot \nabla\bar{v}. \quad (3.5)$$

Alternatively, given the Lagrangian derivative of some quantity, the associated Eulerian derivative is recovered from Equation (3.4), assuming that the velocity and the gradient of the desired quantity are available at the respective position. While the mathematics in this section are not directly employed in this analysis, the basic components that comprise Lagrangian coherent structures represent a quantity embodying the ideas of Lagrangian analysis – in this case, a quantity of particular interest to the investigation within the context of the CR3BP.

3.2 Lyapunov Exponents

The Lyapunov exponents arise as the chief result of Aleksandr Lyapunov’s study of stability in dynamical systems [40]. The most common approach for computation of the Lyapunov exponents traces back to Olseledec [44], and the necessary mathematical background of the Lyapunov exponents, as presented here, is based on both Olseledec and Anderson [45]. Olseledec defines a norm, $\|M\|$, as the largest eigenvalue ($\tilde{\lambda}_{max}$) of $\sqrt{M^T M}$ where M is an arbitrary matrix and M^T is its transpose. When M is a square, nonsingular matrix with elements that are a function of time, t , or $M = M(t)$, the following relationship is valid,

$$\tilde{\lambda}_{max} \left(\sqrt{M(t)^T M(t)} \right) = \sqrt{\tilde{\lambda}_{max} (M(t)^T M(t))}. \quad (3.6)$$

Thus, given these conditions, Olseledec’s norm is equivalent to the matrix spectral norm. Above, the non-standard functional notation indicated by $\tilde{\lambda}_{max}()$ represents the operation of calculating the largest eigenvalue associated with the given operand. Since $M(t)$ is nonsingular, $M(t)^T M(t)$ is a symmetric positive definite matrix with real eigenvalues.

Lyapunov exponents represent a metric that, in the limit, describes the relative balance of order and chaos in a system. In a given dynamical system, there are as many Lyapunov exponents as dimensions in the phase space, and collectively they comprise the Lyapunov spectrum. However, one exponent typically dominates and is often the only one of interest, that is, the characteristic or maximal Lyapunov exponent. On a local scale,

the Lyapunov exponents supply a measure of the contraction or expansion relative to a neighboring trajectory given an initial variation, $\delta\bar{x}$. The effect of such initial variations is typically approximated via the linear variational equations as discussed in Sections 2.7 and 2.8. Specifically, an initial variation, $\delta\bar{x}(t_0)$, results in a final variation $\delta\bar{x}(t)$ consistent with Equation (2.72), that is,

$$\delta\bar{x}(t) = \Phi(t, t_0)\delta\bar{x}(t_0),$$

where $\Phi(t, t_0)$ is the STM. Assuming that the variations grow exponentially, the overall expansion or contraction indicated by the STM is dominated by the rate Λ_{max} , and the norms of the initial and final variations can be compared, i.e.,

$$|\delta\bar{x}(t)| = |\delta\bar{x}(t_0)| e^{\Lambda_{max}t}, \quad (3.7)$$

where Λ_{max} represents the characteristic Lyapunov exponent (note, $\tilde{\lambda}$ generally indicates an eigenvalue). Solving for Λ_{max} , normalizing by time, and evolving the system such that time approaches infinity, yields the characteristic Lyapunov exponent, that is,

$$\Lambda_{max} = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{|\delta\bar{x}(t)|}{|\delta\bar{x}(t_0)|}. \quad (3.8)$$

The argument of the natural logarithm in Equation (3.8) is, effectively, “the norm of the STM”. Thus, the spectral norm also produces the characteristic Lyapunov exponent,

$$\Lambda_{max} = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \|\Phi(t, t_0)\|. \quad (3.9)$$

Parker and Chua [46] define the entire Lyapunov spectrum similarly, that is,

$$\Lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \left| \tilde{\lambda}_i(t) \right| \quad i = 1, \dots, n, \quad (3.10)$$

where $\left| \tilde{\lambda}_i(t) \right|$ are the n normalized eigenvalues of the STM, computed from the n eigenvalues of $\sqrt{\Phi(t, t_0)^\top \Phi(t, t_0)}$. In total, the Lyapunov exponents essentially represent a hyperellipsoid, reflecting the stretch and contraction of the system as time increases. Assuming a phase space of dimension four, including two positions (say, x and y) as well as two velocities, such a hyperellipsoid viewed in position space appears as an evolving ellipse. This concept is illustrated generally in Figure 3.1, where stretching in one dimension is colored blue while the stretching in another dimension appears in red. In the limit, the Lyapunov exponents define the chaoticity of the local neighborhood in the vicinity of \bar{x}_0 .

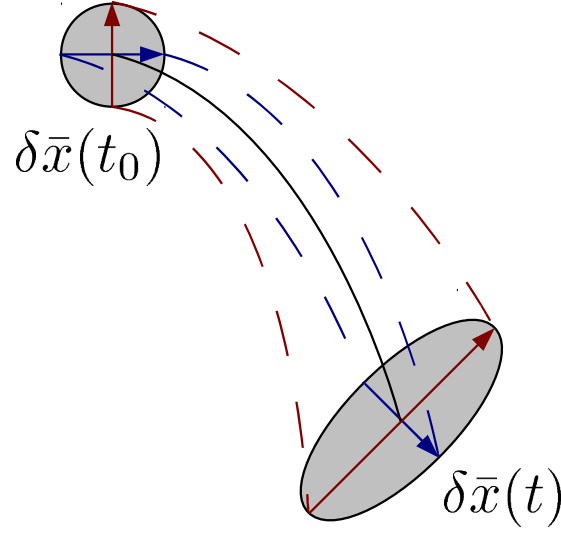


Figure 3.1. Stretching of an ellipsoid represented by the Lyapunov exponents

In Table 3.1, similar to the table appearing in Parker and Chua [46], the relative values of the Lyapunov exponents are summarized in terms of the corresponding steady state behavior in a local region. In Table 3.1, $\Lambda_1 = \Lambda_{max}$, is the characteristic Lyapunov exponent.

Table 3.1
The Lyapunov Spectrum and Steady State Behavior

Lyapunov Exponents	Associated Steady State
$0 > \Lambda_1 \geq \dots \geq \Lambda_n$	Equilibrium Point
$\Lambda_1 = 0,$ $0 > \Lambda_2 \geq \dots \geq \Lambda_n$	Periodic Orbit
$\Lambda_1 = \Lambda_2 = 0,$ $0 > \Lambda_3 \geq \dots \geq \Lambda_n$	Two-periodic Torus
$\Lambda_1 = \dots = \Lambda_K = 0,$ $0 > \Lambda_{K+1} \geq \dots \geq \Lambda_n$	K-periodic Torus
$\Lambda_1 > 0, \quad \sum_{i=2}^n \Lambda_i < 0$	Chaotic

The Lyapunov exponents characterize the future behavior in a particular region that is part of a system or, in some cases, the local region may, in fact, comprise the entire system. This characterization is inherently associated with some initial position and time, a common notion associated with the Lagrangian description of motion (e.g., labeling a parcel with its initial position and time and then following that parcel as it evolves). However, the definition of the Lyapunov exponents implies knowledge of the system evolution as time approaches infinity. In some systems, such a limit is not readily available, and computing a metric such as a Lyapunov exponent requires modification of the definition for practical application.

3.3 Poincaré Maps

Often, the complexity of a system like the three gravitational bodies whose evolution is modeled in the CR3BP, can hinder analysis by obscuring its salient aspects. Poincaré sections, also denoted Poincaré maps or surfaces of section, represent the evolution of a system by reducing the complexity thereby revealing useful information. These maps were first conceived by Henri Poincaré in his work on the three-body problem [5,8], and have since been widely employed for various applications that exhibit repeating behavior. Poincaré sections also provide a convenient format for communicating other information via mapping techniques, for example, a map of a metric such as the characteristic Lyapunov exponent.

Poincaré sections reduce a n -dimensional phase space to a phase space of dimension $n - 1$. This reduction is accomplished by introducing a hyperplane, Σ of dimension $n - 1$, such that the flow possesses a component that is generally transverse to the plane. This hyperplane serves as both the starting point for initiating motion, and the stage upon which the behavior is observed. A hyperplane is defined in terms of any set of quantities that are insightful to the problem. Certainly, a subset of the state variables are one option that is frequently convenient, but other dynamical quantities are also easily employed to define a hyperplane. For a 3D system of first-order equations, such a hyperplane is effectively 2D, because one of the three initial states is constrained by selection of the plane and the remaining states are fully represented in two dimensions. Thus, an initial state originally possessing three degrees of freedom is now restricted to two. Any constants, or integrals of

the motion, can also be employed to further reduce the phase space for ease in representation on a Poincaré map. Initial states originating from the plane, Σ , are then observed at each subsequent crossing of the hyperplane. Poincaré sections may include crossings of the plane in one direction only (a one-sided map), in both directions (a two-sided map), or by using a fixed-time sampling (generally termed a stroboscopic map), depending on the application. Frequently, periodicity is a significant factor in a system and an associated Poincaré section is therefore concerned with crossings in only one direction. In any case, crossings of a Poincaré section reflect the iterations of the Poincaré map. Thus, a Poincaré map is a discrete mapping since a state is mapped to its subsequent iterations. Letting \bar{x} represent an initial state on the hyperplane, the N^{th} iteration of the Poincaré map is defined,

$$P^N(\bar{x}) = \phi(\bar{x})_{t_0}^{\Delta t}, \quad (3.11)$$

where Δt represents the time required for the flow map, ϕ , to advance to the N^{th} crossing of the hyperplane. Given a single initial state, a one-sided Poincaré map, as it evolves over two iterations, is illustrated in Figure 3.2. In practice, many initial states are simultaneously allowed to populate the map, and inspection of the resulting patterns may reveal various regions of the map with notable aspects of the system behavior. For example, a fixed point on the map, $P^N(\bar{x}^*) = \bar{x}^* \forall N$, identifies a periodic orbit in the system (green in Figure 3.2).

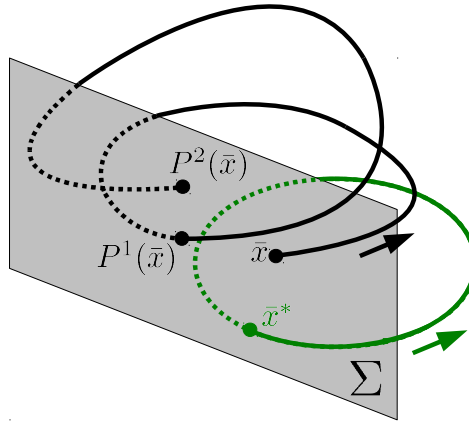


Figure 3.2. An illustration of a one-sided Poincaré map

Poincaré sections are very useful in the four-dimensional phase space associated with planar motion in the CR3BP. A hyperplane is defined such that Σ corresponds to a fixed value of one of the state variables; in every subsequent crossing of the hyperplane, the number of free state variables is reduced to three. Two of the remaining three state variables are employed to produce a grid of initial states. The third state variable is typically constrained by fixing the Jacobi constant to a specific value. Constraining the state as described is a convenient approach to define a hyperplane; as a consequence the system evolves to an acceptable stopping condition for a large number of trajectories. The effect of the flow map, as applied to a number of initial states over a given time, may produce widely varying path lengths. But, a Poincaré map can be employed to require each initial state to continue evolving until N crossings of the maps are achieved, regardless of the time interval required by each individual trajectory to traverse an equal number of crossings. Of particular interest to this investigation is the correlation between the features of a Poincaré map and LCS.

3.4 Finite-Time Lyapunov Exponents

The Finite-Time Lyapunov Exponents (FTLE) differ only slightly from the more formally defined Lyapunov exponents. The entire FTLE spectrum is defined,

$$\lambda_i = \frac{1}{|T|} \ln \left| \tilde{\lambda}_i(t) \right| \quad i = 1, \dots, n, \quad (3.12)$$

where the unadorned symbol λ is used to represent the FTLE. The major difference between Equation (3.12) and Equation (3.10) is the absence of the limit. Clearly, the calculation of the FTLE is truncated at some prescribed final time, T . Further, the absolute value of T is incorporated simply to reflect the fact that the calculation of the FTLE is accomplished in forward or backward time. Consistent with the Lyapunov exponents computed in infinite time, there is a single FTLE in the entire spectrum that generally dominates, and this dominant value is typically the FTLE value of most interest. The characteristic or maximal FTLE is often denoted as “the” FTLE for a given evaluation. It is common to discuss forward FTLE and backward FTLE, referring to the characteristic forward time

and backward time FTLE, rather than the FTLE spectrum. The characteristic FTLE is calculated in a process analogous to the characteristic Lyapunov exponent, that is,

$$\lambda_{max} = \frac{1}{|T|} \ln ||\Phi(t, t_0)||, \quad (3.13)$$

where the STM is normalized consistent with Equation (3.6). Again, the characteristic FTLE differs from the characteristic Lyapunov exponent in terms of its evaluation at a finite time.

The conclusions that result from an examination of the FTLE values are less precise than those from values of the Lyapunov exponents, as summarized in Table 3.1, because FTLE values are computed over a finite-time duration. Despite the truncated nature of the FTLE, the associated value actually represents the relative behavior of neighboring trajectories across practical time domains, which is more useful than the ultimate steady-state information supplied by the limits required in the evaluation of the Lyapunov exponents. Scheeres et al. [47] and Anderson et al. [48] both exploit time-truncated Lyapunov exponents for stability analysis and maneuver design, respectively. Another important FTLE application is the identification of material boundaries, or flow-separating structures, in a system. The identification of material boundaries is accomplished by computing the FTLE over a specified region and isolating the locally constrained maxima in this field; the constrained maxima are a direct representation of these structures (the material boundaries). Abarbanel [49] introduces the concept of the local Lyapunov exponent (synonymous with FTLE) and discusses the application to strange attractors. Anderson [45] even suggests the use of the FTLE to determine a stationkeeping or trajectory corrections process. All of these direct applications of the FTLE support the practical value of truncating Lyapunov exponents to form FTLE values.

In preparation for discussion of the computation of the FTLE, one important term requires a more complete definition. The concept of FTLE frequently focuses on the *flow map* that is commonly designated $\phi_{t_0}^t(\bar{x})$. The flow map represents the state of the system that has evolved to a final time t from an initial state \bar{x} at time t_0 . This terminology and compact notation is a convenient expression for any particular solution in a system.

Given the flow map, the calculation of the FTLE is accomplished in one of two ways. The first approach requires numerical integration of an initial state along with the STM.

After a finite duration, the resulting STM is normalized and the value of the FTLE is obtained from Equation (3.13). The more common alternative method for computing the FTLE arises as a consequence of the lack of a set of variational equations. In such a case, it is common practice to produce a numerical approximation to the STM via finite differencing directly from the flow map, evaluated after a specified time duration. The flow map is obtained through numerical integration or from empirical data. Obtaining the FTLE from finite differencing yields the advantages of increased numerical stability and fewer differential equations to integrate (thus avoiding any numerical issues associated with integrating the STM equations). However, the finite-difference method can also be inherently noisy depending on the specific application and the parameters involved in the simulation. Shadden [50] effectively describes the process for calculating the FTLE via finite differencing. Shadden uses the notation, $\frac{d\phi_{t_0}^t(\bar{x})}{d\bar{x}_0}$, to represent the derivative of the flow map with respect to its initial state. The resulting matrix is an approximation to the STM, $\Phi(t, t_0)$. Rather than generating this STM from numerical integration, Shadden demonstrates the calculation of the approximate STM from an existing flow map. This finite-difference approach requires that the flow map be defined for several initial states separated by known initial variations. It is convenient, but not required, that these initial states are defined on a regular grid. If a 2D grid of (x, y) positions is defined with spacing specified by some integer mapping, then the finite-difference STM, in terms of its elements, is defined as,

$$\left. \frac{d\phi_{t_0}^t(\bar{x})}{d\bar{x}_0} \right|_{(i,j)} = \begin{bmatrix} \frac{x_{i+1,j}(t) - x_{i-1,j}(t)}{x_{i+1,j}(t_0) - x_{i-1,j}(t_0)} & \frac{x_{i,j+1}(t) - x_{i,j-1}(t)}{y_{i,j+1}(t_0) - y_{i,j-1}(t_0)} \\ \frac{y_{i+1,j}(t) - y_{i-1,j}(t)}{x_{i+1,j}(t_0) - x_{i-1,j}(t_0)} & \frac{y_{i,j+1}(t) - y_{i,j-1}(t)}{y_{i,j+1}(t_0) - y_{i,j-1}(t_0)} \end{bmatrix}. \quad (3.14)$$

The result is then used to calculate the FTLE via Equation (3.13). A drawback of the finite-difference approach is the requirement for five evaluations of the flow map to generate one FTLE value. However, for some applications, such as the identification of material structures within the system, evaluation of the entire grid is usually required and this overhead is quickly absorbed. While finite-differencing produces a less accurate result, this concern is offset by the fact that this strategy directly employs the actual stretching induced by the flow map rather than an indirect estimate of the system evolution from a propagated linear system of equations.

Examples that demonstrate the direct comparison of both the finite-difference FTLE method and an integrated STM algorithm appear in Figures 3.3–3.5. A 512×512 grid of initial (x, \dot{x}) values, defined in a specific CR3BP environment, is employed to generate the resulting FTLE maps. In Figure 3.3, only the forward FTLE is depicted, with high values indicated by white saturation. The result from a backward FTLE computation is plotted in Figure 3.4 and, again, high FTLE values are indicated by a higher saturation of white. In Figure 3.5, both forward and backward FTLE are plotted, and to distinguish

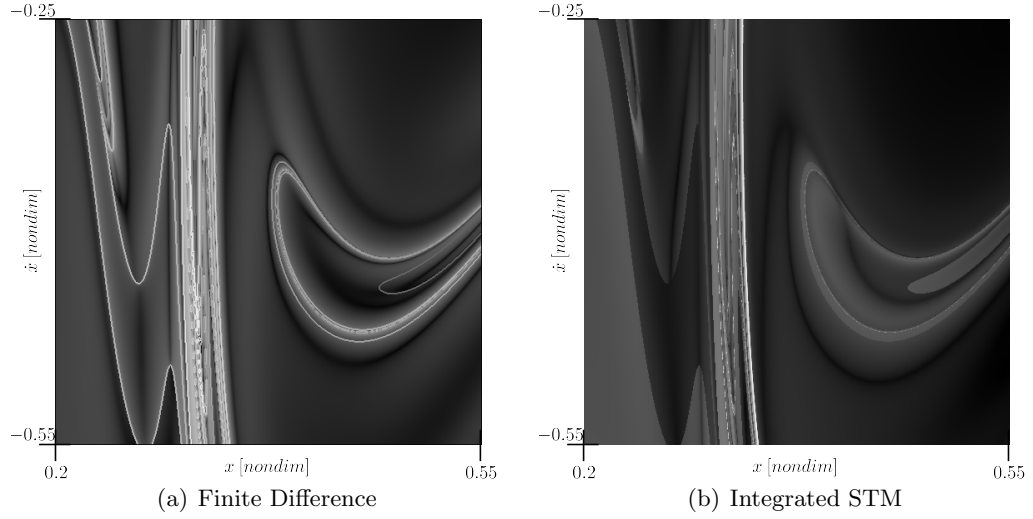


Figure 3.3. Backward FTLE for $C \approx C_{L_2}$ in the Earth-Moon system

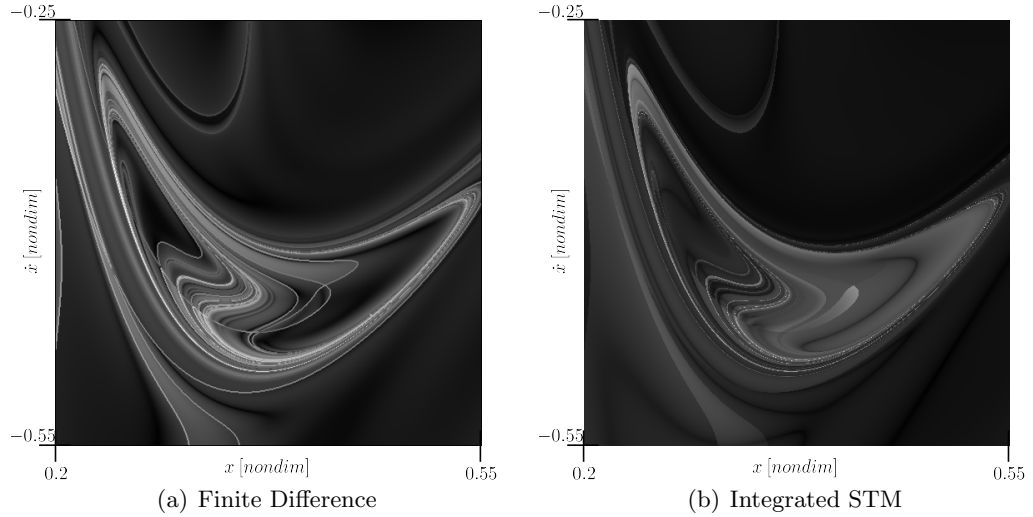
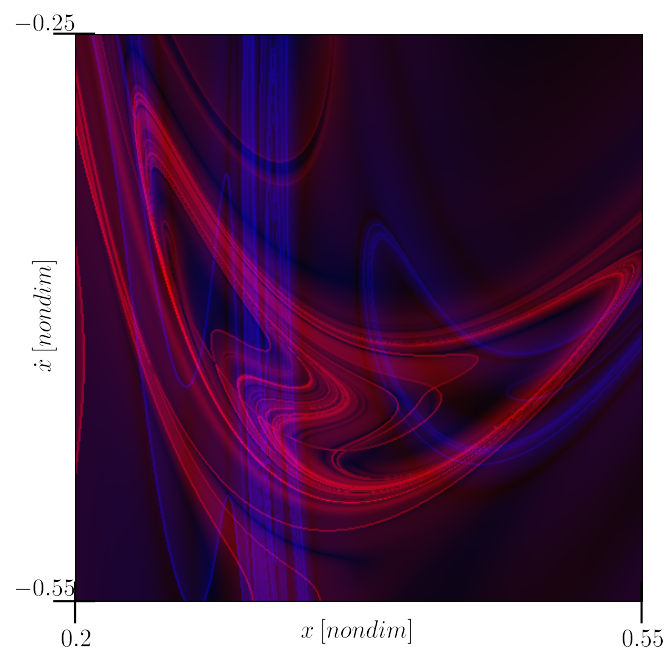
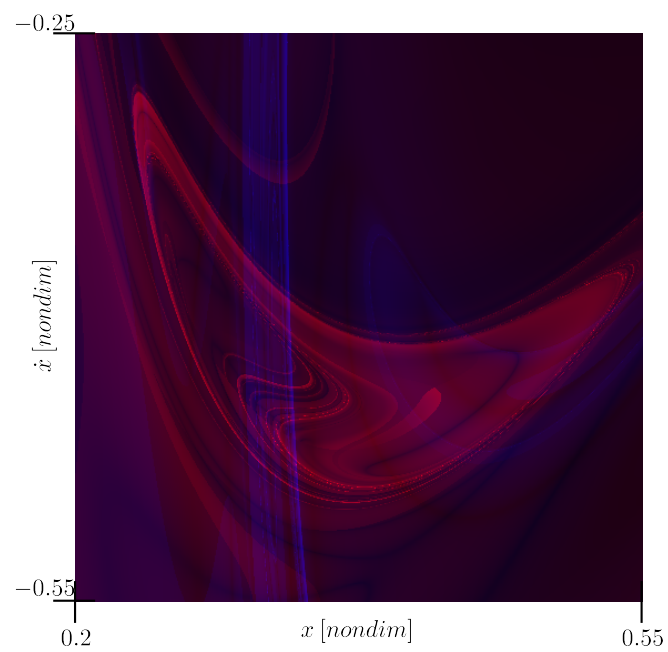


Figure 3.4. Forward FTLE for $C \approx C_{L_2}$ in the Earth-Moon system



(a) Finite Difference



(b) Integrated STM

Figure 3.5. Forward (red) and backward (blue) FTLE for $C \approx C_{L_2}$

the two, forward FTLE are colored with red saturation corresponding to high relative values. Backward FTLE are colored by a similar measure with high relative values represented with a greater saturation of blue¹. In all cases, dark regions of the field represent regions of relatively small values of the FTLE and bright regions correspond to relatively high FTLE values. The images use 8-bit quantization for visualization by mapping the smallest FTLE values to black and the largest FTLE values to full saturation (white, red, or blue, as appropriate) through 256 equally spaced gradations in saturation. These maps are defined in the Earth-Moon system ($\mu \approx 0.01215$) with a value of Jacobi constant, $C \approx 3.17216$, where x and \dot{x} ranges are indicated in nondimensional units. The motion associated with these maps is restricted to the x - y plane since $z_0 = \dot{z}_0 = 0$. Additionally, the initial state is constrained to the x axis since all $y_0 = 0$, and \dot{y} is determined from the Jacobi constant. Thus, the figures represent a Poincaré map reflecting FTLE values on the hyperplane.

Some discussion of the computation involved in producing the maps in Figures 3.3–3.5 is necessary to further compare the two methods. In both cases, the 512×512 (262,144) grid of initial states are integrated forward and backward in time until their resulting flow reaches the fifth crossing of the map. The numerical integration process uses a Runge-Kutta-Fehlberg (RKF) 4th-5th-order integration scheme with an integration tolerance of 10^{-12} . In the finite-difference approach, integration involves the evolution of a system of four differential equations (the planar CR3BP equations). The total computation time that is required to numerically integrate and calculate the FTLE values using compiled C++ routines with a serial implementation is approximately 37 minutes (~ 19 minutes for the forward map and ~ 18 minutes for the backward map). To produce the FTLE map from an integrated STM strategy, a system of twenty equations (the planar CR3BP equations and the 16 associated scalar STM differential equations) requires numerical integration. The computational time using the same serial C++ implementation with a slight modification to accommodate the STM requires about 29 hours (~ 19 hours to produce the forward map and ~ 10 hours to produce the backward map). Aside from the obvious time advantage associated with the finite-difference scheme, another positive aspect is immediately apparent in the resulting FTLE maps. From an examination of Figures 3.3–3.5, it is observed that the finite-difference method produces much more definitive outlines of various regions; this

¹Red and blue colors, used here arbitrarily, in no way imply instability/stability as in other sections/chapters.

clarity is an important component in the identification of Lagrangian coherent structures. Unless otherwise required for a specific situation, this analysis focuses on the computation of FTLE values via the finite-differencing technique.

In any Lyapunov exponent calculation, truncation occurs and an important issue is the appropriate simulation length to produce meaningful FTLE results. Since the FTLE is based on the characteristic Lyapunov exponent, a limit-based metric, a longer simulation implies a more representative FTLE value. However, in practical applications, calculation of the FTLE for a length of time corresponding to a time frame of interest is a compelling argument. But, the normalization, that is, division by the simulation time, represents a numerical difficulty. For systems with periodic behavior, such as the CR3BP, an FTLE map is sometimes defined to correspond to a specified number of map crossings. In any case, the simulation length must yield an FTLE value that is adequate to reveal pertinent information for the current application. The current investigation focuses exclusively on the CR3BP.

3.5 Invariant Manifolds

In the broadest sense, a manifold is a structured subset of a space with a dimension less than the space itself. Moreover, such a manifold locally resembles the Euclidean space corresponding to its dimension [46]. For example, the equilibrium points that correspond to the differential equations in the rotating frame formulation of the CR3BP are zero-dimensional manifolds existing in a six-dimensional space. Similarly, periodic orbits are one-dimensional manifolds in the same six-dimensional space. However, within the context of CR3BP analysis, the term manifold often refers to higher-dimensional manifolds, which are frequently represented by plotting their projection onto physical space. These projections typically appear as invariant surfaces taking the form of tubes that, in the planar CR3BP, act as separatrices of the flow. The contour depicted in Figure 3.6 is an example of one such manifold tube, represented discretely by 1024 trajectories crossing the hyperplane on a Poincaré map. In general, invariant manifolds offer significant insight into the behavior in the CR3BP. Moreover, the separatrix characteristics of the manifolds closely relate to the ideas of material boundaries associated with the FTLE [32].

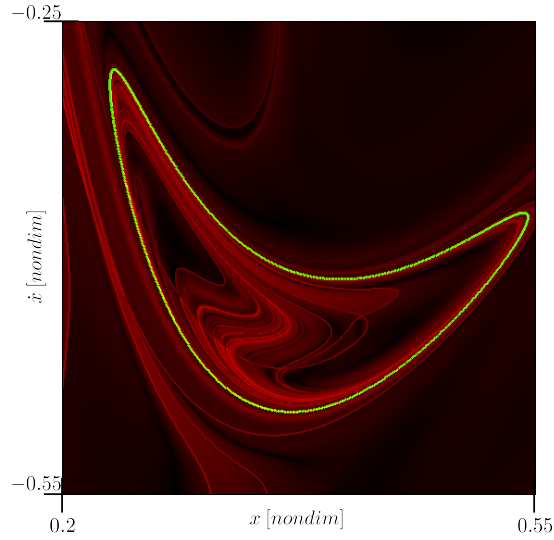


Figure 3.6. Poincaré map of manifold crossings overlain on a forward FTLE field

3.5.1 Stable and Unstable Manifolds Associated with L_1 and L_2

Manifolds associated with a periodic orbit, and the process to generate these manifolds, are of specific interest. However, the approach to produce manifolds of periodic orbits results from a generalization of methods for computing the manifolds of critical points. Recall that the stability of the equilibrium points is characterized by the Jacobian matrix evaluated at the equilibrium points. Further, the collinear equilibrium points are characterized as hyperbolic, possessing both a stable and an unstable mode indicated by one positive and one negative eigenvalue. The process for identifying the manifolds of the critical points originates with a perturbation of the 4D initial state along the position-normalized eigenvectors of the Jacobian. Propagating the state resulting from this perturbation along the normalized eigenvector corresponding to the positive eigenvalue results in an approximation to the unstable manifold. If, on the contrary, an initial state is perturbed along the normalized eigenvector associated with the negative eigenvalue and propagated *backward* in time, the result is an approximation of the stable manifold. In both cases, the state must be perturbed in two directions along the eigenvector, and subsequently evolved in time, as appropriate, to obtain the entire globalized manifold.

Proper normalization of the eigenvectors corresponding to the Jacobian is critical to ensure an effective perturbation of the state. Thus, normalization is accomplished to retain the general velocity characteristics of the eigenvector but normalizing with the position components allowing for a displacement that is representative in physical space. Given such a position-normalized eigenvector, the required perturbation is accomplished by the addition of a relatively small scalar multiple of the normalized eigenvector to the initial state (in this case \bar{x}_{L_i}). Specifically, the normalized “stable” or “unstable” eigenvector, \bar{v}^S or \bar{v}^U , respectively, is produced by dividing the original eigenvector, $\bar{Y}^{S/U}$, by the magnitude of the position, that is,

$$\bar{v}^{S/U} = \frac{\bar{Y}^{S/U}}{r}, \quad (3.15)$$

where $r = \sqrt{x^2 + y^2 + z^2}$. Some judicious displacement step, d , yields the perturbed initial state vector,

$$\bar{x}_0 = \bar{x}_{L_i} \pm d \cdot \bar{v}^{S/U}. \quad (3.16)$$

The perturbation displacement, d , is usually determined as some appropriate physical distance relative to the equilibrium point. In the Earth-Moon system, such a distance is typically a few dozen kilometers, while in the Sun-Earth system a more reasonable displacement might be a few hundred kilometers. The accuracy of the approximation to the manifold is inversely related to the magnitude of the displacement, d . However, since the manifolds arrive and depart from the point of interest asymptotically, a very small displacement requires extensive integration to move away from the point and increases the numerical error. The selection of the displacement is, therefore, somewhat subjective, but must be sufficiently small to produce a reasonably accurate approximation to the manifold but large enough to avoid excessive integration time. Given an appropriate initial perturbation, propagation of the new initial state generates an approximation to the actual manifold.

In Figure 3.7, the stable and unstable manifolds associated with the equilibrium points are generated in the neighborhood of the Earth-Moon L_1 and L_2 points and projected onto configuration space. The figure also includes other appropriate vectors and subspaces. Specifically, the position-space projections of the stable and unstable eigenvectors, \bar{v}^S and \bar{v}^U , respectively, are depicted with arrow heads indicating the direction of the flow associated near these vectors. The stable and unstable eigenspaces, E^S and E^U , respectively, are indicated by the dashed lines extending from the points. The local stable and unstable

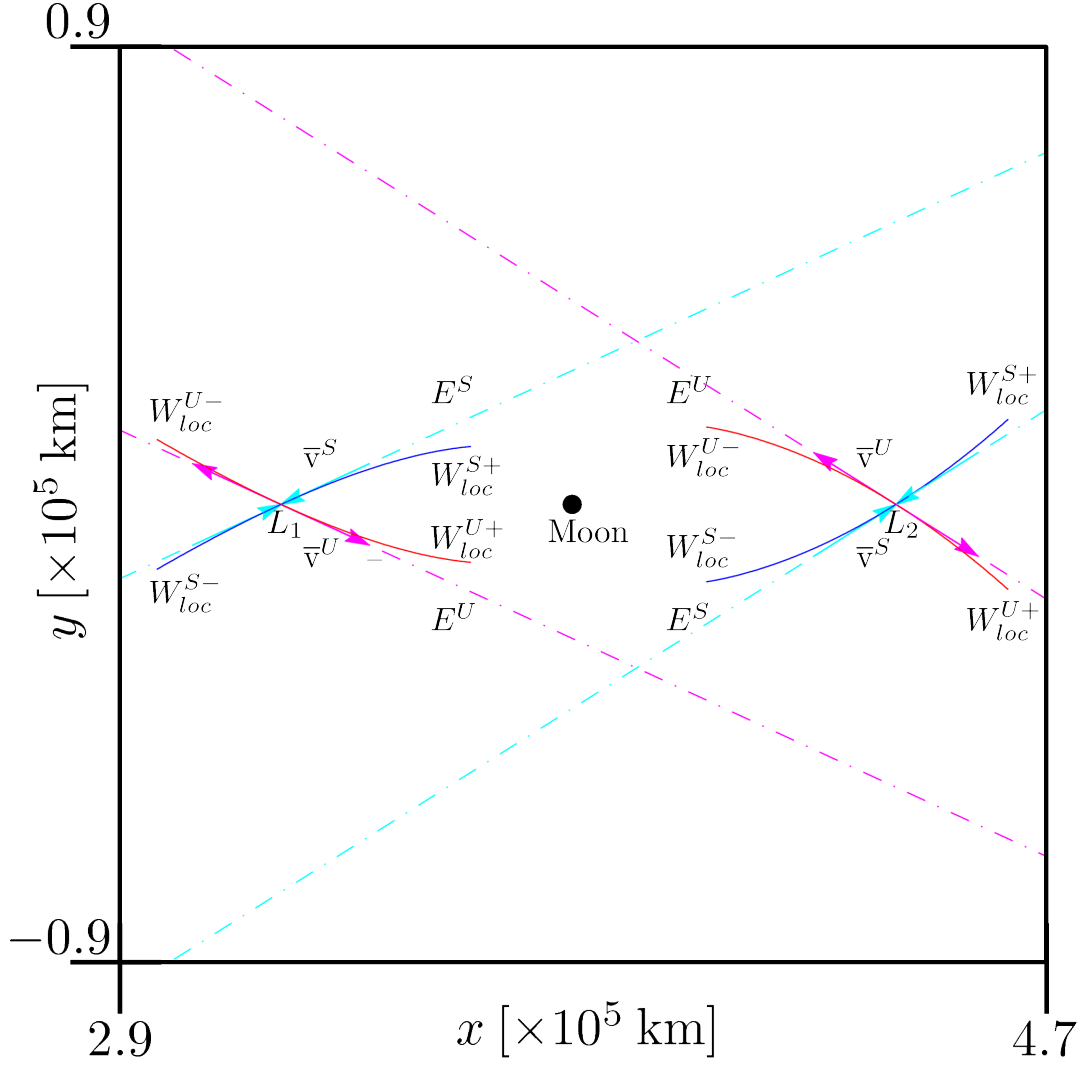


Figure 3.7. Local manifolds associated with L_1 and L_2 in the Earth-Moon CR3BP

manifolds are tangent to their respective eigenspaces and are indicated in red (W_{loc}^S) and blue (W_{loc}^U), respectively. Further, + and - identify the half-manifolds and the notation simply reflects the direction of the flow, consistent with $+\bar{v}^{S/U}$ or $-\bar{v}^{S/U}$, respectively, along the position-normalized eigenvector the perturbations are made. The localized manifolds in Figure 3.7 are propagated for an arbitrary duration to globalize the manifolds with associated symbols W^U and W^S . All of the quantities and symbols in the figure are distinct for

each libration point. Moreover, the approximations to the manifolds associated with each libration point L_1 and L_2 , possess Jacobi constant values that are close but not precisely equal to C_{L_1}/C_{L_2} given the initial offset relative to the points that is required for computation. The theoretical manifolds possess Jacobi constant values equal to their respective point and act as separatrices for motion, consistent with C_{L_1} and C_{L_2} .

3.5.2 Manifolds Associated with a Periodic Orbit (Maps)

The process for identifying manifolds that correspond to equilibrium points is extended to periodic libration point orbits by introducing maps. Stability information for a periodic orbit is available through the STM, in particular, the monodromy matrix that is, in fact, a first-return Poincaré map [40]. This matrix is the analog of the Jacobian matrix for an equilibrium point. The eigenvalues of the monodromy matrix reveal stable and unstable modes, if they exist, and the associated eigenvectors supply the necessary perturbation directions to generate individual manifold trajectories associated with an orbit. The eigenvalues of the monodromy matrix are independent of the starting point or fixed point. If a given orbit is discretized into a number of representative fixed points along the orbit, and the eigenvalues and eigenvectors of the monodromy matrix are calculated at each, then the resulting manifold trajectories produce a “wireframe” for the manifold surface. Properties of the STM as described by Katsiaris and Goudas [51] are used to transform the eigenvectors of the monodromy matrix for the first fixed point to each successive fixed point as described by Marchand [52]. Given the eigenvectors for the first fixed point, the eigenvectors for the remaining $N - 1$ points are computed from,

$$\bar{\mathbf{v}}^S(n\Delta t) = \Phi(n\Delta t, 0)\bar{\mathbf{v}}^S(0), \quad (3.17)$$

$$\bar{\mathbf{v}}^U(n\Delta t) = \Phi(n\Delta t, 0)\bar{\mathbf{v}}^U(0), \quad (3.18)$$

with $n = 1 \cdots N - 1$. The position-space projections of the eigenvectors for 20 fixed points along an L_1 Lyapunov orbit are illustrated in Figure 3.8, with stable vectors colored blue and unstable vectors colored red. The Lyapunov orbit is sized in terms of its amplitudes. The amplitude in x is $A_x \approx 7000$ km and the corresponding amplitude in y is $A_y \approx 26,000$ km. After perturbing the state at each fixed point by 50 km in the appropriate directions and

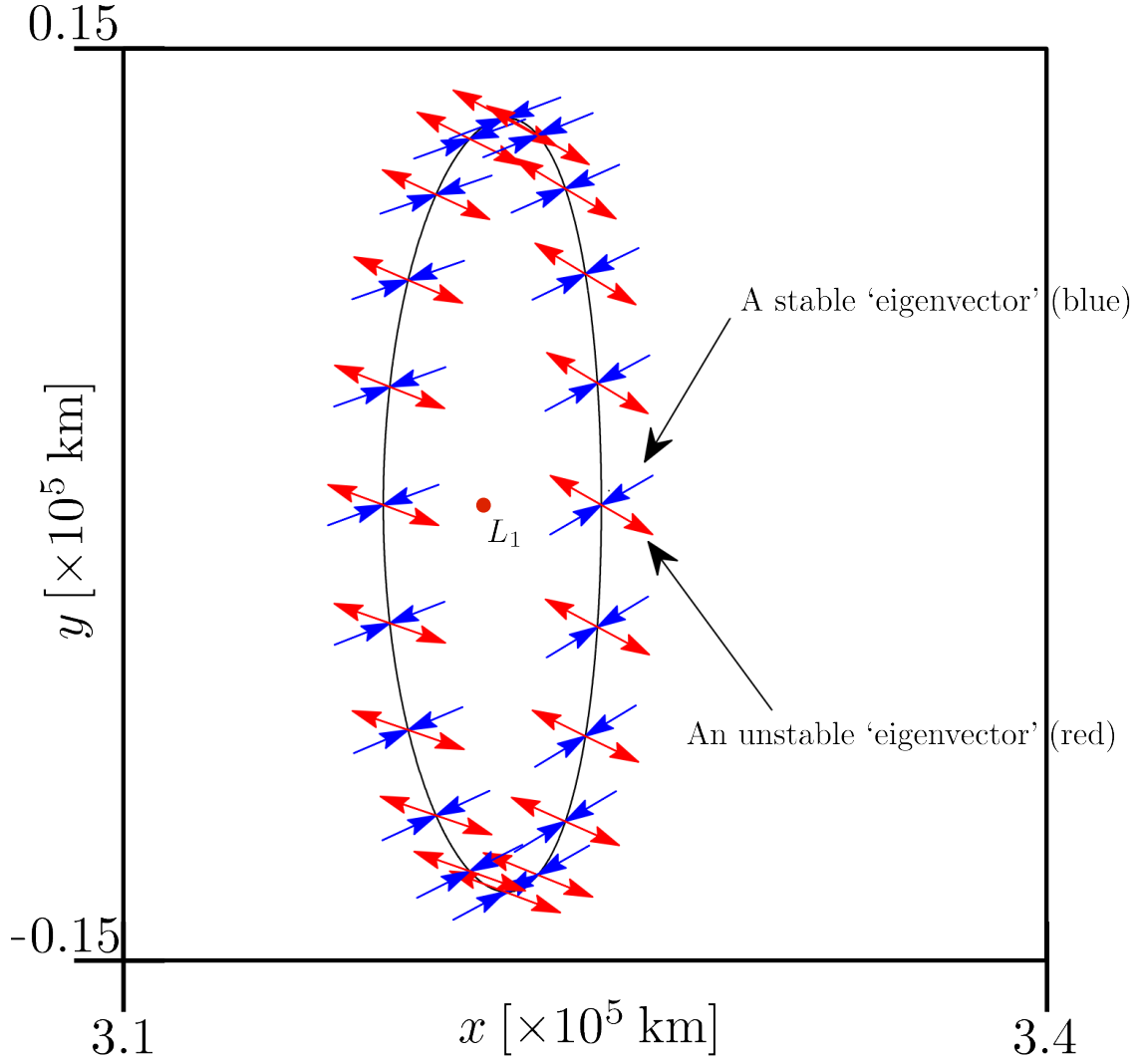


Figure 3.8. Position-space projections of eigenvectors along an L_1 Lyapunov orbit

integrating forward and backward in time, the first crossing of the manifolds associated with the periodic orbit are plotted on a Poincaré map in Figure 3.9. In this case, the discrete map fully represents the manifolds (recall each point on the map is a projection of a full 4D state). Additionally, the trajectories representing the manifolds are plotted in configuration space in Figure 3.10. Again, red is used in both Figures 3.9 and 3.10 to indicate the unstable manifold and blue is used to indicate the stable manifold. Arrows reflecting the general

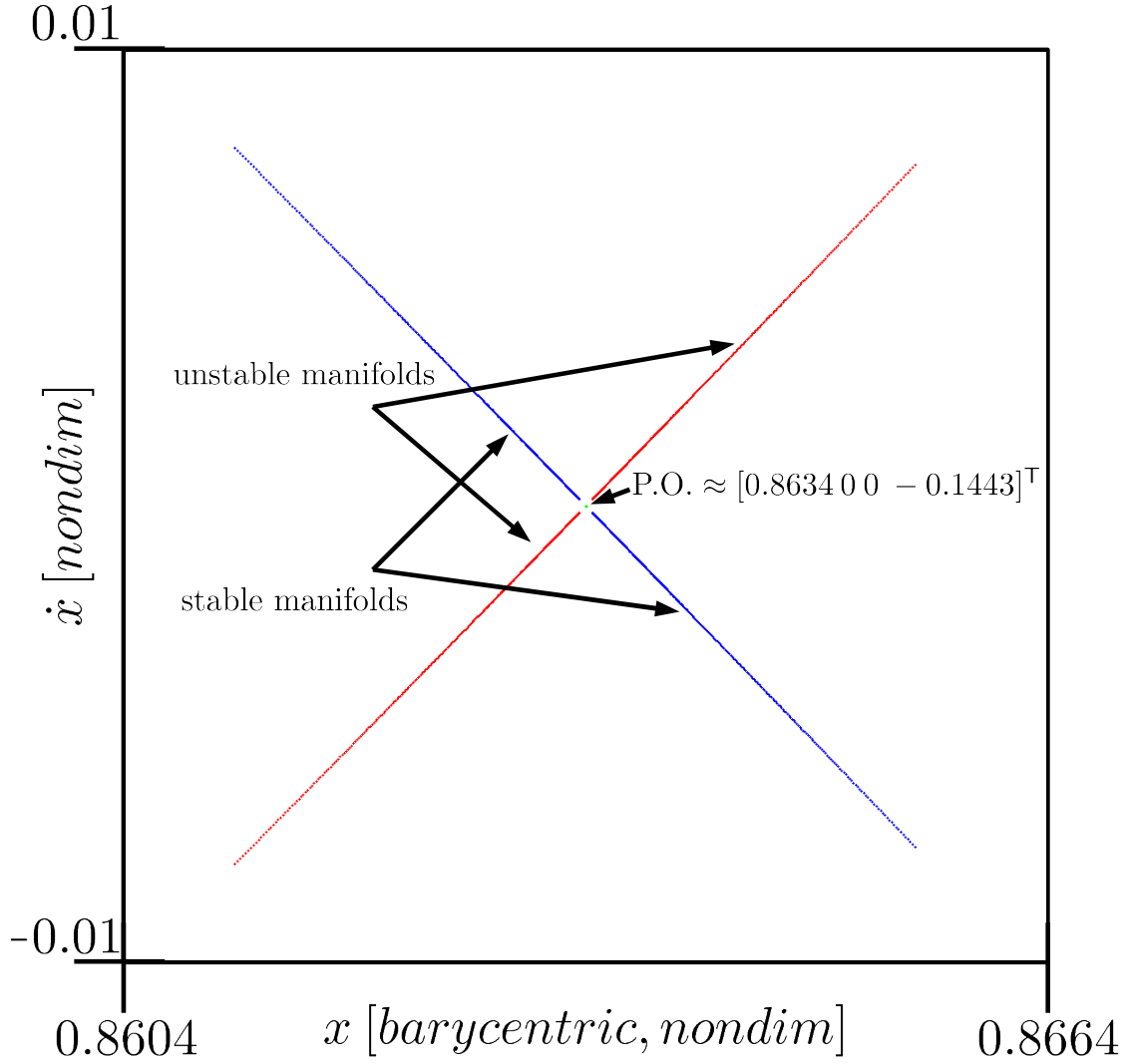


Figure 3.9. Poincaré map of manifolds associated with fixed point along a periodic orbit

direction of the flow are included in Figure 3.10 along with representations of the Earth and Moon for perspective. The manifolds are terminated at the x values corresponding to the Earth and Moon, however, this is only for purposes of illustration and does not represent a physical termination of the manifold trajectories. Ultimately, the concept of invariant manifolds represents a characteristic of the natural dynamics that is exploited for effective mission design. The usefulness of manifolds, particularly stable and unstable manifolds,

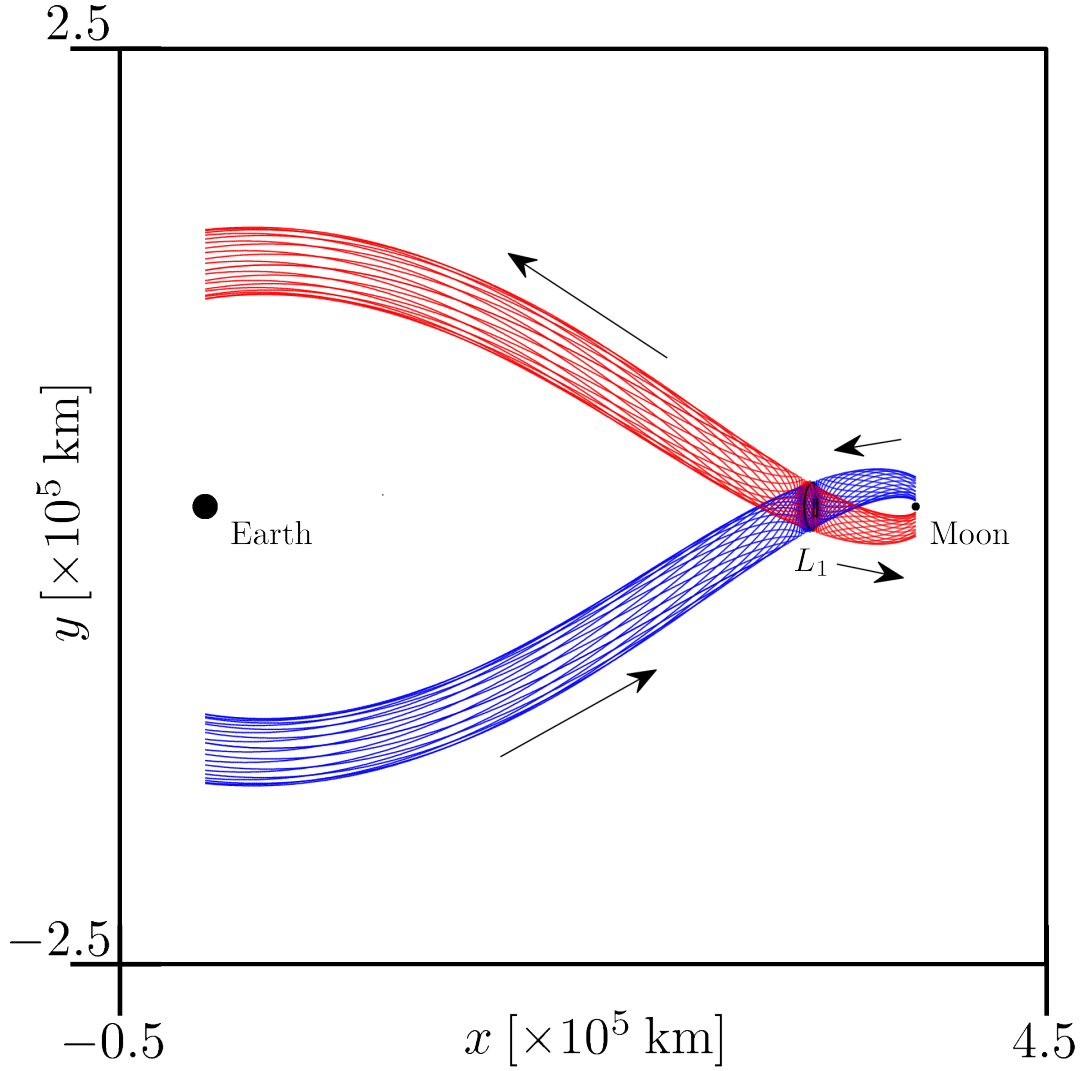


Figure 3.10. Lyapunov orbit stable/unstable manifolds represented by trajectories

supplies an impetus for identifying such structures. The methods described in this section represent only one way of identifying manifolds, and require a prescribed process and *a priori* knowledge of the existence of the manifold. The requirement for knowledge that the stable and unstable manifolds exist, represents part of the motivation for this analysis. Effectively identifying manifolds without prior knowledge of their existence is an overarching goal, and Lagrangian coherent structures provide a means to accomplish this objective.

3.6 Lagrangian Coherent Structures

Lagrangian Coherent Structures (LCS) appear as ridges, or curves of constrained maxima, on an FTLE map. The concept of LCS originates with Haller [28] and is well-illustrated by Shadden et al. [32] and Shadden [50]. More important to this analysis is the general interpretation that LCS are the Lagrangian analog of Poincaré map crossings as the invariant manifolds evolve. This correlation is introduced in Figure 3.6. Poincaré map crossings are Eulerian in nature, that is, a snapshot of the evolution of the trajectory at discrete points in time. The observation that LCS and Poincaré map crossings expose the same structure and yield the same fundamental information from different perspectives, supplies a potential opportunity to exploit the advantages of both. Specifically, FTLE maps provide an immediate view of manifold crossings in a given region and invariant manifold theory allows the immediate interpretation of these FTLE structures.

3.6.1 A Formal Definition of Ridges

The direct identification of Lagrangian coherent structures in a data field depends on the underlying concept of ridges. In terms of visual analysis, various definitions for ridges exist. The description that is typically employed to identify LCS is consistent with the formulation of height ridges as discussed by Eberly et al. [53]. This discussion establishes ridges, in general, as locally defined and preferably invariant, that is, “...ridges should be *invariant* under the following transformations:

- translations in the spatial variables,
- rotations in the spatial variables,
- uniform magnification in the spatial variables, and
- monotonic transformations of the intensity function.”¹

Eberly et al. indicate that height ridges only satisfy the first three invariances, but such ridges remain suitable when obtained from data that does not experience monotonic transformations of the intensity function. The first two invariance transformations indicate

¹Eberly et al., “Ridges for Image Analysis”, p. 3

that the ridges corresponding to a rotated or translated function are equivalent to rotated/translated ridges from the original function. The third invariance is particularly important to this work, because it indicates that a uniform scaling (such as that imposed by converting from nondimensional units to dimensional units) preserves ridge structure. The formal definition of a height ridge further indicates that the identification of points along the ridge is accomplished by evaluating constrained local maxima, that is, local maxima of principal curvature where the associated differentiation occurs along principal directions [53]. These maxima are not true local maxima because the gradient does not completely vanish at the corresponding points; the gradient vanishes in only a subset of the principal directions. Given a curve of ridge points, the points can only be considered as maxima with respect to the directions transverse to the ridge. Any of the points along the ridge may possess a height value that is greater or less than one of its neighbors on the ridge.

3.6.2 The Nature of Lagrangian Coherent Structures

Lagrangian coherent structures and invariant manifolds are correlated, and to clarify that relationship it is important to elaborate on the nature of LCS. Lagrangian coherent structures are aptly named considering their appearance as tight ridges, or coherent structures, in the scalar field of a Lagrangian metric, the FTLE. The concept of a “ridge” in a given scalar field, as discussed in Section 3.6.1, indicates a contour in the field that possesses a relatively higher value than its transverse neighbors. This concept equates directly to the topographical definition of ridges from which it draws its name. Relative heights, or ridges, in a flow field are illustrated in Figure 3.11; FTLE data from a region in a flow passing behind several obstacles is rendered. (The data from which this image is produced and the associated FTLE calculations are from Christoph Garth at the University of California at Davis [54].) The ridges in this image represent LCS.

The methodology for producing invariant manifolds corresponds inversely to the techniques for generating LCS in a key aspect. Forward FTLE fields yield LCS that correspond to stable manifolds while LCS from backward FTLE are associated with unstable manifolds. Recall that stable manifolds are produced by integrating backwards in time. The simulation originates from a state that is perturbed relative to a reference state on a particular

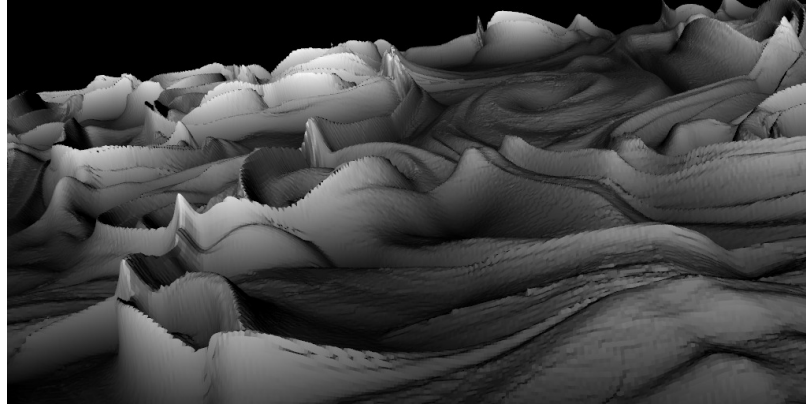


Figure 3.11. “Ridges” of FTLE values represented as relative heights

solution. Forward FTLE are produced by integrating forward in time. This apparently opposing behavior is reconciled by considering the fact that the associated manifold represents, to some degree, a boundary separating qualitatively different regions of motion. Specifically, two neighboring trajectories on either side of a manifold will ultimately evolve to different regions of the solution space and, as a result, experience a relatively large amount of stretching leading to a high FTLE value. A trajectory directly on the stable manifold approaches its associated particular solution. Of course, the system is numerically sensitive so the manifold is an infinitely thin surface. A slight perturbation to one side or the other ultimately results in a shift towards vastly different regions. Detail of the correspondence between manifolds and LCS is summarized in Table 3.2.

Table 3.2
LCS and Manifold Correspondence

Structure	Generated via	Analog
Forward LCS	Forward time integration	Stable manifold
Backward LCS	Reverse time integration	Unstable manifold
Stable manifold	Reverse time integration	Forward LCS
Unstable manifold	Forward time integration	Backward LCS

The potential for applications of Lagrangian coherent structures within the astrodynamics community is far-reaching as both an analysis tool and as a means to identify important system structures, such as manifolds, without *a priori* knowledge. However, calculation of the FTLE is still resource intensive and the numerical identification of LCS in an FTLE field can be challenging. Despite these difficulties, the advantage of the predictive abilities of LCS warrants careful investigation and consideration. This analysis is based upon all of the previous concepts with a number of numerical and computational tools for application in the CR3BP.

4. COMPUTATIONAL METHODS

Considering the physical regime associated with astrodynamics applications, it is not surprising that nearly all of the design and analysis occurs in a virtual setting. Thus, computational modeling and simulation are common in trajectory and mission design. Some of the computational aspects of the analysis are examined, including numerical integration schemes, methods to expedite computation, and various tools employed for visualization of the resulting data.

4.1 Numerical Considerations

Many aspects of analysis and design in astrodynamics require the application of numerical methods. Prominent throughout this investigation is numerical integration. Additionally, interpolation also plays a key role in identifying crossings on a Poincaré map and the ridge points in a scalar field. Finally, some consideration of the numerical difficulties posed by the CR3BP vector field, due to the presence of the two massive primaries, is assessed using regularization. Each of these numerical tools is employed in the analysis as appropriate.

4.1.1 Integration

Several numerical integration schemes are employed in this analysis. Initially simple, lower-accuracy integrators are used for exploratory efforts. A simple trapezoid method is employed for implementation in settings that exploit simpler mathematical operations. Lower-order Runge-Kutta integration algorithms are also employed. After the regions of interest are identified and designated for a more careful inspection, an adaptive step-size Runge-Kutta-Fehlberg (RKF) 4th/5th-order integrator, one included in the GNU Scientific Library (GSL) [55], is generally applied. Additionally, a Runge-Kutta Prince-Dormand 8th/9th-order scheme is available, also from the GSL, and is employed to generate output

for specific values of the integration parameter (e.g., time or a state variable). Alternatively, when output is required for crossings of a Poincaré section, the RKF integrator, coupled with linear interpolation at plane crossings, is sufficient if the required step size for a specified accuracy is incorporated. Despite the varied number of numerical integrators employed for various steps in this analysis, all of the ultimate results are generated using the same integrator with consistent parameters, namely the RKF algorithm, and the accuracy required of the integrator is always on the order of 10^{-12} .

4.1.2 Interpolation

Often, it is necessary to determine a value between discrete data points defined on a line or a grid. To accomplish this determination, an interpolation process is employed. Interpolation methods are applied specifically for locating crossings on a Poincaré map as well as for ridge-finding algorithms in a scalar field. Two standard interpolation schemes [56] are incorporated: (1) linear interpolation in a 1D parameter space; and (2) bilinear interpolation in a 2D parameter space. While the idea of linear interpolation is well-known, its role in a more complicated bilinear interpolation scheme necessitates a specific notation associated with the linear interpolator, that is,

$$f(x) = x(f_1 - f_0) + f_0, \quad (4.1)$$

where the value of the function, f , at some point, x , interior to the interval $[x_0, x_1]$ is a function of the value of $f(x_0) = f_0$ and $f(x_1) = f_1$. Bilinear interpolation arises as a natural extension of linear interpolation, and essentially represents linear interpolation repeated twice over two parameters. The key formula for bilinear interpolation produces the value of a 2D function, $f(u, v)$, where u and v are the coordinates of a point interior to the four points on the corners of a grid,

$$f(u, v) = (1 - v) ((1 - u)f_{00} + uf_{10}) + v ((1 - u)f_{01} + uf_{11}). \quad (4.2)$$

Consistent with linear interpolation, the values at the points on the corners of the grid are represented by $f(u_0, v_0) = f_{00}$, $f(u_0, v_1) = f_{01}$, $f(u_1, v_0) = f_{10}$ and $f(u_1, v_1) = f_{11}$. Together, these two interpolators are sufficient for this analysis.

4.1.3 Regularization

Another numerical consideration is represented by the process of regularization of the equations of motion. Regularization is employed, as necessary, in the present analysis when the numerical integrator progresses slowly due to numerical difficulties. The EOM associated with the CR3BP include singularities since the distance to each of the two massive primaries appears in the denominator. Specifically, as a state evolves into close proximity of a primary, smaller time steps are required; reducing the step-size imposes an increasing effect on the evolution steps of the integrator. The smaller time steps lead to inaccuracies and, eventually, even an adaptive integration algorithm stalls as it continually attempts to reduce the step size to maintain accuracy. In a process of regularization available from Sundman [9], the system EOM are transformed into a different, but related, system of equations with decreased effects from the singularity. Olikara [57] offers a concise description of the generalized Sundman transformation as used to regularize the equations of motion in the CR3BP. This regularization is accomplished by the introduction of an alternate time parameter, s , that is related to the usual time variable through the following,

$$\frac{ds}{dt} = g(\bar{x}), \quad (4.3)$$

where g is a scalar function of the state vector, \bar{x} . The new time variable, s , is calculated by integrating, i.e.,

$$s(t) = \int_0^t g(\bar{x}(u)) du. \quad (4.4)$$

The inverse, for recovering time t is,

$$t(s) = \int_0^s \frac{1}{g(\bar{x}(u))} du. \quad (4.5)$$

The function, $g(\bar{x})$, depends on the system EOM and, in the CR3BP, is selected as,

$$g(\bar{x}) = \frac{1 - \mu}{r_{13}^{3/2}} + \frac{\mu}{r_{23}^{3/2}}, \quad (4.6)$$

with the distances from the first and second primaries indicated by r_{13} and r_{23} in the denominators.

Each of these numerical tools is significant in producing meaningful FTLE and LCS results and allows greater flexibility in the overall implementation of the required numerical operations. Moreover, the availability of various numerical approaches increases efficiency and produces interactive feedback to guide the process.

4.2 Ridge Calculation

The general concept of ridges in a scalar field is defined in Section 3.6.1, but a more careful mathematical definition is required to generate ridge points. Consider a height map that represents a 2D scalar field; the scalar value is used as an appropriately scaled z component. This height mapping concept is illustrated in Figure 3.11. Recall that ridges are essentially lines along the “tops” of the features in the height map. Bilinear interpolation is incorporated in the steps to calculate the ridge points. Generation of the relative values for the points along the ridge is accomplished by comparing the eigenvectors of the Hessian associated with the scalar field to the gradient of the same field. In this case, the scalar field is a FTLE field and the associated ridges are LCS.

Frequently, some pre-processing is necessary to smooth FTLE data before calculating the Hessian and gradient fields. Recall that calculation of FTLE values is accomplished via central difference derivatives to produce an STM, either as a numerical approximation or simply an empirical value. The process of differentiation inherently amplifies any noise in the data, a situation potentially exacerbated by numerical derivatives. However, this same issue actually contributes to sharper features in the scalar field. Unfortunately, the adverse effects of this noisy process are perpetuated by ridge-finding methodologies involving the Hessian and the gradient of the FTLE scalar field since the calculation of these values also requires further differentiation. To offset this noise and extract various ridge features, it is common to apply a low-pass filter to the data either before or after computing the gradient of the field. An optimal low-pass filter is embodied in Gaussian blurring [58]. For a 2D scalar field, which can be interpreted as an image, the Gaussian kernel is defined,

$$G(\bar{u}, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{m^2+n^2}{2\sigma^2}}, \quad (4.7)$$

where $\bar{u} = [m \ n]^T$ represents the relative horizontal and vertical coordinates of a given point in the FTLE field and σ is the standard deviation of the Gaussian distribution. This kernel is convolved with the image to produce a smoothed image, i.e.,

$$\tilde{i}(\bar{u}) * G(\bar{u}, \sigma) = \int_{-\infty}^{\infty} \tilde{i}(\bar{u} - \bar{t}) G(\bar{t}, \sigma) d\bar{t}, \quad (4.8)$$

where \tilde{i} is the value of the scalar field, in this case $\tilde{i} = \lambda$, the FTLE value, at the relative point, \bar{u} . A discrete convolution sum is available as a sample approximation to the continuous convolution [59],

$$i(\bar{u}) = \sum_{\bar{t}} \tilde{i}(\bar{u} - \bar{t}) G(\bar{t}, \sigma), \quad (4.9)$$

where the domain of the summation in \bar{t} is the extent of the Gaussian distribution, and i is the smoothed data. The effect of this low-pass filtering is a preservation of the strong features in the image while blurring out the small features associated with the noise introduced from differentiation. An example of a noisy FTLE field, along with its corresponding smoothed image, appears in Figure 4.1. The smoothing employs an exaggerated Gaussian blur for illustration. Given a smoothed image, the gradient field (a field of 2-vectors) is calculated via central differencing. From the gradient field, subsequent central differencing

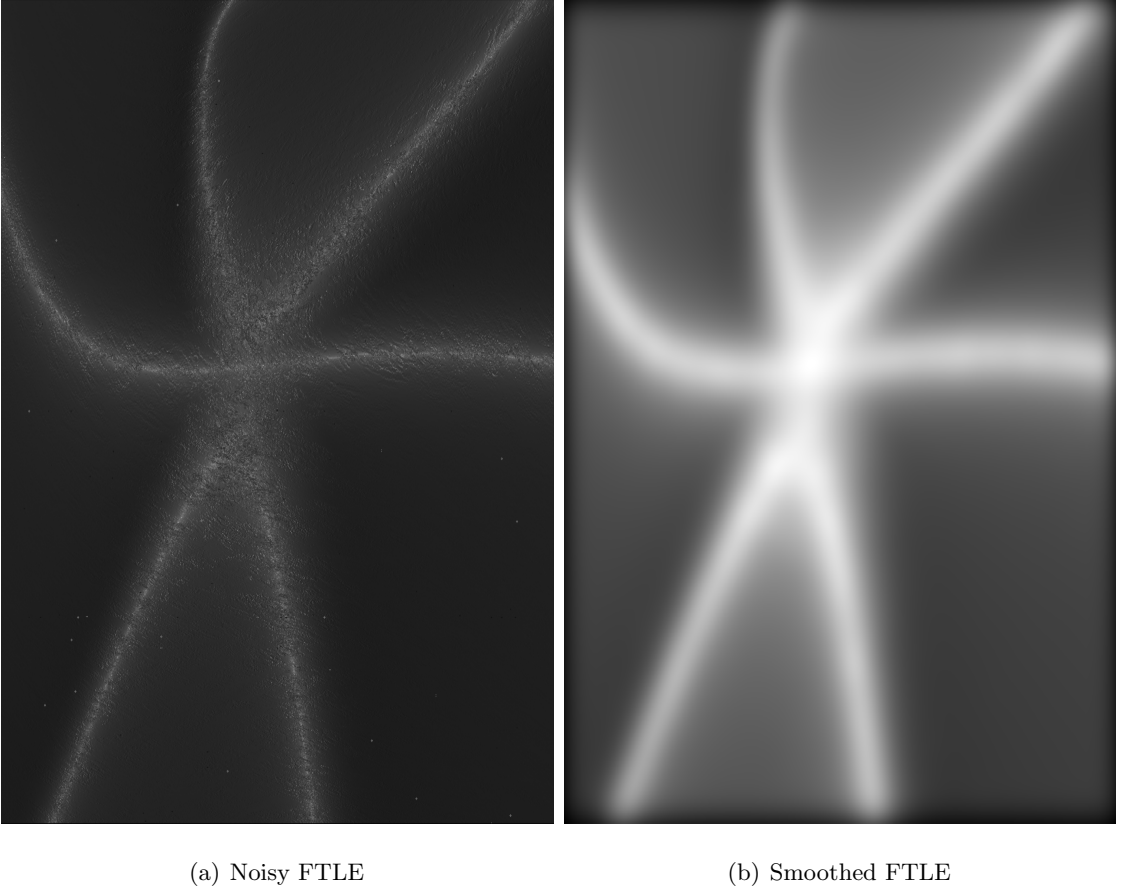


Figure 4.1. An example of Gaussian blurring applied to a noisy FTLE field

produces the Hessian field (a field of 2×2 matrices). The values from the smoothed field that are associated with the difference derivatives employed to calculate the gradient and Hessian are reflected in Figure 4.2. Specifically, the four values employed to perform the

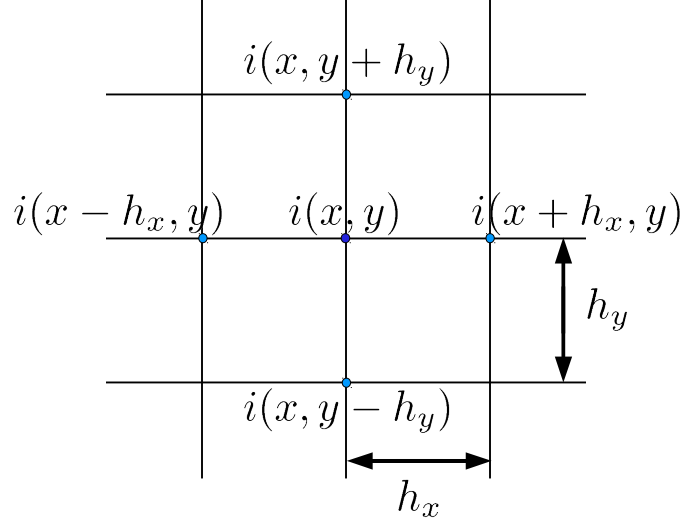


Figure 4.2. Five point stencil for central differencing

central differencing associated with $i(x, y)$ appear as points on the grid in the figure. Note that the spacings h_x and h_y may differ. The gradient is calculated for a particular point in the image such that,

$$\bar{g}(x, y) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial i(x, y)}{\partial x} \\ \frac{\partial i(x, y)}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{i(x+h_x, y) - i(x-h_x, y)}{2h_x} \\ \frac{i(x, y+h_y) - i(x, y-h_y)}{2h_y} \end{bmatrix}. \quad (4.10)$$

The Hessian is then evaluated as,

$$H(x, y) = \begin{bmatrix} \frac{\partial g_x}{\partial x} & \frac{\partial g_x}{\partial y} \\ \frac{\partial g_y}{\partial x} & \frac{\partial g_y}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{g_x(x+h_x, y) - g_x(x-h_x, y)}{2h_x} & \frac{g_x(x, y+h_y) - g_x(x, y-h_y)}{2h_y} \\ \frac{g_y(x+h_x, y) - g_y(x-h_x, y)}{2h_x} & \frac{g_y(x, y+h_y) - g_y(x, y-h_y)}{2h_y} \end{bmatrix}. \quad (4.11)$$

With the Hessian and the gradient available for every point in the FTLE map, ridge points are calculated as described by Eberly et al. [53] and Mathur et al. [33]. The eigenvectors associated with the Hessian align with the extremal directions of the second directional

derivative [60]. Further, these eigenvectors from the symmetric Hessian are orthogonal, thus, the eigenvector (\bar{e}_1) associated with eigenvalue of largest magnitude ($\tilde{\lambda}_1$) aligns with the direction of maximum curvature while the other eigenvector (\bar{e}_2), the eigenvector corresponding to the eigenvalue of smallest magnitude ($\tilde{\lambda}_2$), aligns with the direction of minimum curvature [61]. Recall that the gradient is aligned with the direction of steepest descent. Therefore, for a point to be a ridge point its gradient should be orthogonal to \bar{e}_2 , that is,

$$\bar{g} \perp \bar{e}_2. \quad (4.12)$$

Unfortunately, the lack of a directional sense associated with eigenvectors gives rise to implementation issues with the requisite interpolation. These implementation issues arise from the fact that a numerical algorithm must arbitrarily select a direction for the eigenvector. Then, the possibility exists that eigenvectors on adjacent grid corners may, in reality, point in nearly opposite directions. Therefore, noting that the previous condition in Equation (4.12) is consistent with $\bar{g} \parallel \bar{e}_1$, it follows that \bar{g} satisfies the eigenvalue-eigenvector relationship,

$$H\bar{g} = \tilde{\lambda}_1\bar{g}. \quad (4.13)$$

Ridge points in a 2D scalar field are then deduced via a search for points where $H\bar{g} \times \bar{g} = 0$; in this 2D case, the cross product produces a scalar. The condition that the cross product equals zero indicates an extremum but does not distinguish between ridge and valley points. To ensure that $\bar{g} \perp \bar{e}_2$, evaluate Equation 4.13 such that,

$$||H\bar{g}|| - ||\tilde{\lambda}_1\bar{g}|| < \epsilon, \quad (4.14)$$

where ϵ is some tolerance indicating that the difference is essentially zero.

Recall that the FTLE image is dense up to the sampling size, thus, values are defined for each pixel in the image. But it is unlikely that the conditions describing a ridge point occur precisely on a specific pixel defined on the associated grid. The original FTLE field, its smoothed counterpart, the gradient field, and the Hessian field are all defined on the discrete grid consistent with the selected sampling. While this sampling can always be increased, the associated spacing must be less than ϵ to expose a ridge point. However, bilinear interpolation succeeds in producing an estimate for a ridge point from the $H\bar{g}$ field. Note, however, that $H\bar{g}$ is a nonlinear quantity and bilinear interpolation is a linear process. Nevertheless, given a sufficiently small grid spacing, bilinear interpolation produces

adequate results. Finally, recall that the interpolation process necessitated an alternate criterion for identifying ridge points, due to the lack of direction associated with eigenvectors. Fortunately, such an alternate criterion is available in a 2D field. In 3D fields, the eigenvector directions must be tracked from corner to corner on a given 3D cell.

4.3 Parallel Processing

The computations associated with an FTLE map often prove expensive in terms of time and function evaluations. Various options exist to improve this computational cost and deliver results quickly. Parallel computational approaches offer such advantages. Two types of implementation for parallel computing are discussed in support of this investigation.

4.3.1 CPU Implementation

Computational architectures have undergone a drastic paradigm shift in the last several years. The limitations that constrain the speed of the Central Processing Units (CPU) in most off-the-shelf computers have served as the impetus for multi-core architectures previously unavailable to end users. These multi-core processing units allow access to parallel implementations, originally serial in nature, through the use of widely available programming Application Program Interfaces (API).

One specific API that offers a straightforward parallelization implementation on a multi-core architecture is the OpenMP library that is available for some high-level programming languages including C/C++ and Fortran [62]. This programming library utilizes a shared-memory parallelization approach where each processing core uses the memory available in a single system. Further, harnessing the power of the API is accomplished easily by specifying processor directives in the software source code, linking to various associated OpenMP libraries, and employing a supported compiler with OpenMP compilation flags enabled. If the software is compiled on a system with only a single processing core, the program simply executes in serial.

The algorithmic methodology associated with generating FTLE maps is well-suited to a parallel implementation. Recall that the computation of a single FTLE value involves

the evolution of five trajectories. Generally, large sets of initial conditions, defined on a grid, are employed to produce an FTLE map. This en masse approach serves to alleviate the overhead associated with integrating $(5n)^2$ trajectories across an $n \times n$ grid, reducing the numerical requirements to only n^2 trajectories. However, the integration cost associated with even n^2 simulations can be extensive. Fortunately, each individual initial state evolves independently, allowing concurrent integration. In a system with multiple cores, the OpenMP API is utilized to wrap the applicable looping structures involving integration and an immediate speed advantage is realized. Each additional core is harnessed to improve the processing speed. While the speed upgrade typically does not scale proportionately to the number of cores, the improvement is always appreciable. For example, numerical simula-

Table 4.1
Speed Improvements with OpenMP Multi-Core Parallelization

Run	No. of Cores	Core Speed	CPU Time	Real Time	Speedup
(itrs of $n \times n$ map)		(GHz)	(hr)	(hr)	
3× of 4096×4096	6	2.00	10.38	2.46	4.21×
3× of 4096×4096	10	2.67	7.45	1.18	6.33×
10× of 4096×4096	6	2.00	34.80	7.63	4.56×
10× of 4096×4096	10	2.67	23.85	3.45	6.91×
5× of 2048×2048	4	2.40	5.65	1.82	3.10×
5× of 2048×2048	4	2.40	6.20	1.99	3.12×
10× of 2048×2048	8	2.00	9.33	1.38	6.76×

tion to produce a 4096×4096 grid of FTLE values associated with 10 crossings of a specific periapse Poincaré map is accomplished in ~ 3.45 hours using 10 cores (2.67 GHz per core) on a Linux workstation. The total associated CPU time for this example is ~ 23.85 hours, indicating a speed increase of $\sim 6.91\times$, not the $10\times$ that might be expected. However, producing such a map with double precision accuracy in a few hours, versus an entire day, is indeed significant. In contrast, while CPU parallelization delivers a highly accurate and

direct mode for increasing efficiency, it is typically limited to a relatively small number of CPU cores. Additional examples of speed improvements due to multi-core parallelization with OpenMP appear in Table 4.1.

4.3.2 GPU Implementation

Before investing a significant amount of real time in a computational effort, a targeted purpose and range for the calculations is desirable. A quick, rough picture is invaluable for guiding a computational approach. This quick, approximate view of some region in the map is realized via an alternate parallel computational approach, that is, leveraging the Graphics Processing Unit or GPU. The GPU allows for thousands of concurrent executions to effectively generate an immediate view of a user-defined FTLE bounding area. Then, based on the view in this particular area, the bounding box is resized or re-focused by zooming in or out or even shifting the focus to an entirely different region. Finally, once a suitable area is identified, a more extensive and accurate CPU map is generated. This process supplies nearly immediate feedback and a much more effective approach.

The GPU is generally available under two different hardware configurations. The GPU is (1) the primary component of a computer's internal graphics card; or, (2) an entirely separate internal computing module existing solely to support extensive parallel implementations. Originally intended for concurrently computing millions of graphical primitives to render visual scenes, over the past decade GPU technology has been adapted for scientific work. Multiple aspects of this adaptation are a focus for research in the computer science and visualization fields. However, recently, graphics card manufacturers have begun to recognize that these technical advances are a viable application of their hardware and are now releasing programming tools to support these efforts. One such programming interface is the NVIDIA[®] Compute Unified Device Architecture (CUDA[™]) [63]. A single CUDA-capable computing card may possess several hundred processing cores, each capable of multi-threaded computation allowing for thousands of concurrent operations. The essence of a CUDA GPU implementation is the decomposition of the algorithm into as many parallel operations as possible. The calculation of the flow map associated with an FTLE map is well-suited for parallelization with the GPU. Each of the trajectories on a

grid are effectively integrated simultaneously yielding the entire result in approximately the same time interval that is required to integrate a single trajectory on a single CUDA core (with typical clock speeds around 500 MHz). While this single integration is slower than a single CPU integration, the parallel advantage of concurrent integration is immediately apparent.

The CUDA GPU is best-suited to algorithms involving small mathematical steps with little or no associated branching logic. So, a fourth-order Runge-Kutta integrator with a fixed step-size is more suitable to a CUDA implementation than any adaptive step-size algorithm. Additionally, the CUDA hardware is typically restricted to single precision operations, consistent with the fact that single precision operations are sufficient to deliver visual information. This apparent trade of speed for accuracy is becoming less distinct as NVIDIA continues to improve the general purpose computational abilities of their graphics hardware. Recently, CUDA-capable cards to support double precision operations with limited effectiveness, relative to single precision performance, have been introduced. A GPU approach is employed in this investigation, along with an interactive visualization application, to guide the generation of various FTLE maps and the identification of the associated LCS.

4.4 Visualization

Simplification of the analysis is accomplished by judicious application of various methods to produce FTLE maps and investigate the associated LCS. To this end, an interactive visualization process ensures that the regions of interest can be quickly identified. The visualization interface integrates a wide array of numerical and computational tools, including:

1. Teem [64], “a coordinated group of libraries for representing, processing, and visualizing scientific raster data.”¹: The Teem library principally interfaces directly with the generated FTLE data. Tools from Teem’s Nearly Raw Raster Data (nrrd)² library subset, along with Teem’s Utah Nrrd Utilities (unu)² command line tool, are applied

¹“Teem: Tools to process and visualize scientific data and images”, p. 1

²Acronym capitalization taken directly from the author of the Teem library

to map the data onto an appropriate scale, combine the data into red-green-blue (RGB) three-vectors to color points, and to represent the data as an image.

2. The Visualization Toolkit (VTK) [65]: As the name suggests, VTK exists as a library of visualization algorithms and tools. Each aspect of the toolkit is written in C++ and is extensively integrated into various other programming frameworks. The VTK library is used to create an interactive “widget”, or **w**indow **g**adget element, as part of a Graphical User Interface (GUI), with the images generated by Teem. This VTK widget is embedded in a Qt GUI.
3. Qt [66], “a cross-platform application and [User Interface] UI framework.”¹: This toolset consists of C++ libraries, a graphical integrated development environment, and an intermediate meta-object compiler (MOC). Qt serves as the overarching framework for the visualization environment used in this analysis.

Various computational tools are embedded in the visualization interface. Including GSL, OpenMP, and CUDA. Other user-defined functions are employed as required. One advantage of this collection of visualization tools is that the algorithms are largely open-source or otherwise “free” software.

Two views of the user interface, developed as part of this investigation, appear in Figures 4.3 and 4.4. In Figure 4.3, five regions of the interface are labeled. Region A includes controls to set and modify CR3BP system parameters such as μ , C , and the bounding box for the FTLE field. In region B, control options are available to manipulate the integration parameters, including the choice to use the CPU or the GPU for integration. The FTLE image is displayed in region C where ridge data is subsequently overlaid. Region D offers options for Gaussian blurring and ridge detection tolerances. Finally, the blurred image and ridge data, after each operation, appears in region E. In Figure 4.4, an example of the GUI in operation appears. The top frame displays the FTLE data associated with a region similar to that in Section 3.4. Overlain on the FTLE are ridge points, represented as red dots, added after blurring and ridge identification. In the bottom frame, the blurred image and the ridge is apparent with ridge points overlaid on the image. The various control boxes display the parameters employed to produce the featured results. The imperfect correlation

¹“Qt: A cross-platform application and UI framework”, document title

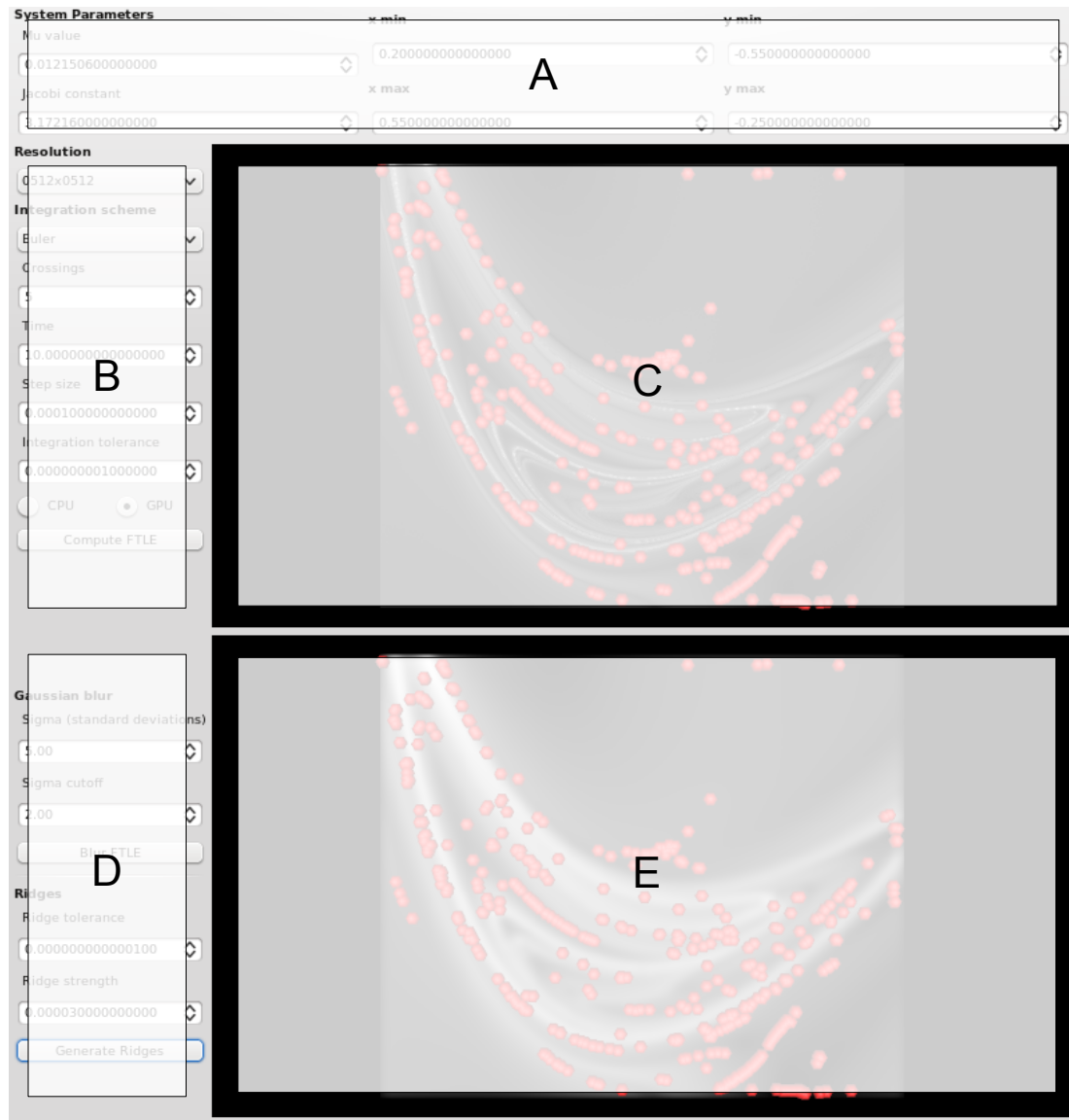


Figure 4.3. Annotated user interface screenshot

of the calculated points with structures in the image serves to reinforce the utility of such an interactive environment that allows for adjustment of parameters “on-the-fly” to achieve better results. Figure 4.5 illustrates a better correspondence associated with the FTLE map from Section 4.2, and represents sample results from a less complex FTLE field, produced with a more extended interactive visualization effort.

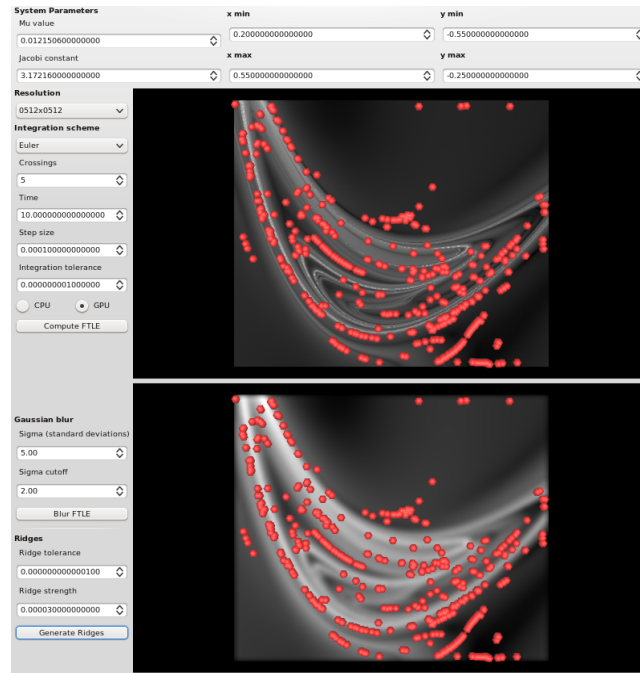
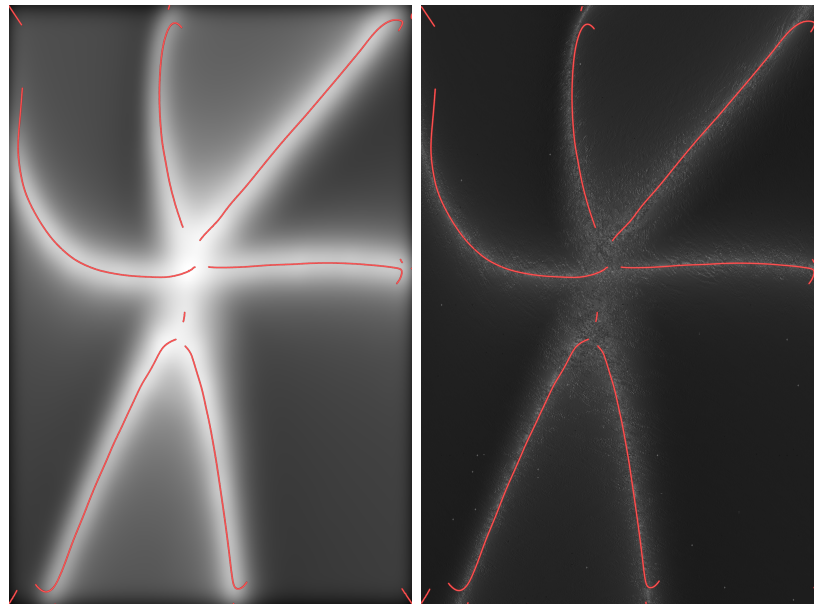


Figure 4.4. User interface screenshot with examples



(a) Ridges on smoothed FTLE

(b) On the original noisy FTLE

Figure 4.5. An example of results from an extended interactive session

5. ANALYSIS AND RESULTS

The applicability of Lagrangian coherent structures to the CR3BP is demonstrated through a series of examples. Examination of LCS that are associated with known invariant manifolds exemplifies the direct correlation of the two concepts. Specifically, a simple comparison between LCS and the manifolds corresponding to a Lyapunov orbit in the Earth-Moon system suggests additional features. Immediately apparent in the FTLE field, structure that is not observed from the associated Poincaré map emerges. Additional applications of LCS in the Earth-Moon system are also presented. Several examples of LCS in the Sun-Saturn system correlate with parallel research efforts focused on the identification of transit trajectories and the characterization of long-term trajectory behavior. Both FTLE maps and LCS are useful in many systems to explore regions of qualitatively different behavior.

5.1 Lagrangian Coherent Structures in the Earth-Moon System

A system of continuing interest involves the Earth and the Moon and their gravitational influence on the motion of a spacecraft. The Earth-Moon (EM) CR3BP is characterized, for this analysis, by a system mass parameter of $\mu \approx 0.01215$, a relatively large value compared with other naturally occurring planet-moon three-body configurations in the solar system. Additionally, Sun-planet three-body mass parameters are generally much smaller as well. Characteristics of dynamical behavior in the EM CR3BP, coupled with the general appeal of this system, create a convenient test bed for various analysis strategies. Two particular examples in the Earth-Moon system follow: (1) investigation of LCS and related concepts corresponding to an L_1 Lyapunov orbit and (2) exploration of FTLE associated with the L_4 region.

5.1.1 Structures Associated with a Lyapunov Orbit

An initial investigation to examine FTLE and LCS associated with invariant manifolds in the EM CR3BP supplies an overview of each concept. As discussed previously, certain values of Jacobi constant yield zero velocity curves that restrict the regions of possible motion. If the value of the Jacobi constant is selected such that motion is restricted to a region encompassing the two massive primaries, for example, $C = 3.17216 \approx C_{L_2}$, a specific Lyapunov orbit is known to exist near L_1 . This configuration is illustrated in Figure 5.1, where the region forbidden by the ZVC is colored in lavender and the Lyapunov orbit near L_1 is colored consistent with its time evolution – the associated zoomed view in Figure 5.2 further quantifies this color scale.

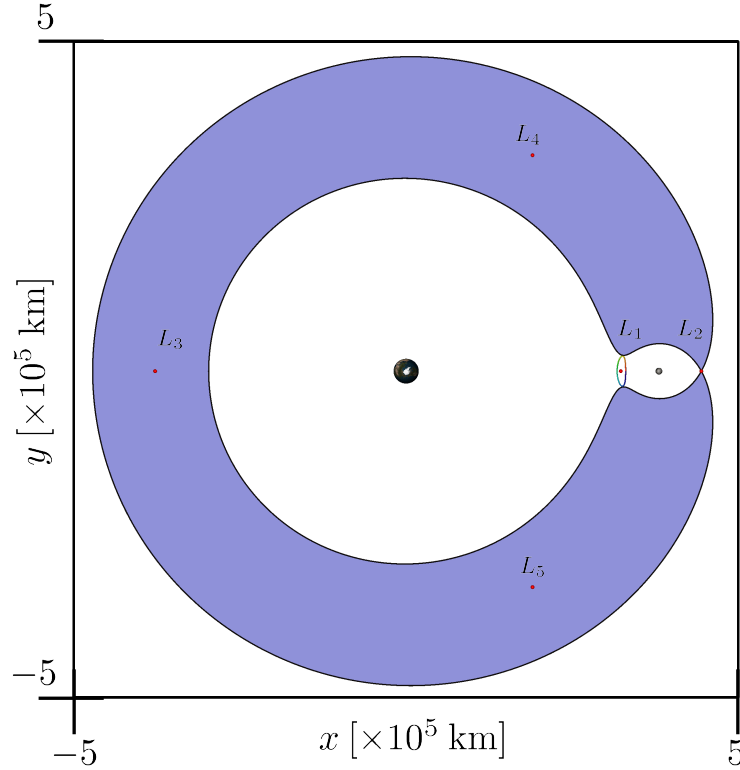


Figure 5.1. Lyapunov orbit and zero velocity curves corresponding to $C \approx C_{L_2}$

To establish the Lyapunov orbit in Figure 5.2, several characteristics are notable. The planar periodic orbit possesses a period of ~ 11.95 days. Recalling the definition of the state, \bar{x} , such that,

$$\bar{x} = \begin{bmatrix} \bar{r} & \bar{v} \end{bmatrix}^T = \begin{bmatrix} x & y & z & \dot{x} & \dot{y} & \dot{z} \end{bmatrix}^T,$$

the initial conditions that generate the Lyapunov orbit are,

$$\bar{x}_0 = \begin{bmatrix} \bar{r}_0 [\text{km}] & \bar{v}_0 [\frac{\text{km}}{\text{s}}] \end{bmatrix}^T = \begin{bmatrix} 329180.457017 & 0.0 & 0.0 & 0.0 & -0.147860687495 & 0.0 \end{bmatrix}^T.$$

Further, the stability characteristics are represented by the eigenvalues of the monodromy matrix. The eigenvalues indicate one stable ($\tilde{\lambda}_s = 0.0004$) and one unstable ($\tilde{\lambda}_u = 2314$) mode. The presence of stable and unstable modes in the linear subspace implies the existence of stable and unstable invariant manifolds associated with this orbit. The color bar in Figure 5.2 supplies a visual indication of the direction of motion along the orbit, where the evolution is marked by color with the initial and final states displayed in blue and red,

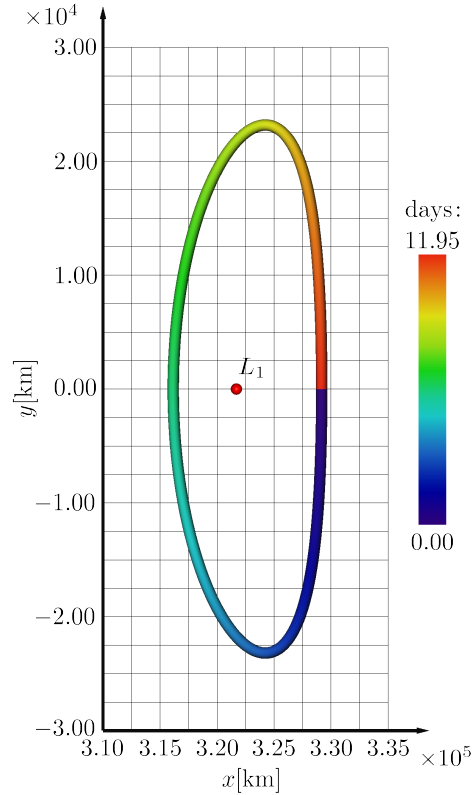


Figure 5.2. L_1 Lyapunov orbit corresponding to $C = 3.17216$

respectively. This coloring further indicates a general clockwise direction of motion around the orbit.

As a first step toward investigating the correspondence between LCS and the invariant manifolds, consider a Poincaré map. Several manifold crossings of a Poincaré map, $\Sigma := y = 0$, are generated using a value of $d = 50$ km as the orbital offset. At each fixed point, the offset is directed along the appropriate position-normalized eigenvectors consistent with Equations (3.15) and (3.16). This position-normalized perturbation is introduced at 1024 discrete fixed points about the orbit. Each propagation produces a trajectory and, combined, they represent the stable/unstable manifolds corresponding to the Lyapunov orbit in Figure 5.2. The manifolds are projected onto configuration space and appear in Figure 5.3; note that the trajectories are plotted to their “second” intersection with the x axis, such that crossing is consistently in the same direction, and occurs between the Earth and L_1 . The “first” intersection represents crossings associated with an “unwinding” of the trajectories from the orbit. Subsequently, the trajectories are further propagated for an extended time as represented through multiple iterations of a map. Crossings of the map, Σ , are defined in the $x-\dot{x}$ phase space between the Earth and L_1 when the state variable y evolves from a negative to a positive value (for the unstable manifold) or from a positive to a negative value (the stable case). The map (Σ) and the first and second crossings are illustrated in Figure 5.4. Crossings associated with both stable and unstable manifolds appear in Figures 5.5–5.7 (note that the axes associated with these figures are marked in nondimensional barycentric coordinates). Each point on the map is constrained by Jacobi constant and the planar formulation. Thus, each point on the map fully embodies a single trajectory in the system, but each point does not necessarily represent an independent simulation given that a single trajectory crosses the map multiple times. As each of the 1024 trajectories evolve at different temporal rates, the crossing numbers indicated in the figures represent the crossings for which a majority of the trajectories begin to fill out the associated contour. In Figure 5.5, the stable manifold crossings appear and the unstable manifold crossings are plotted in Figure 5.6. Note the symmetry across the x axis when comparing Figures 5.5 and 5.6 or directly apparent in Figure 5.7 where both the stable and unstable manifolds appear, colored in blue and red, respectively.

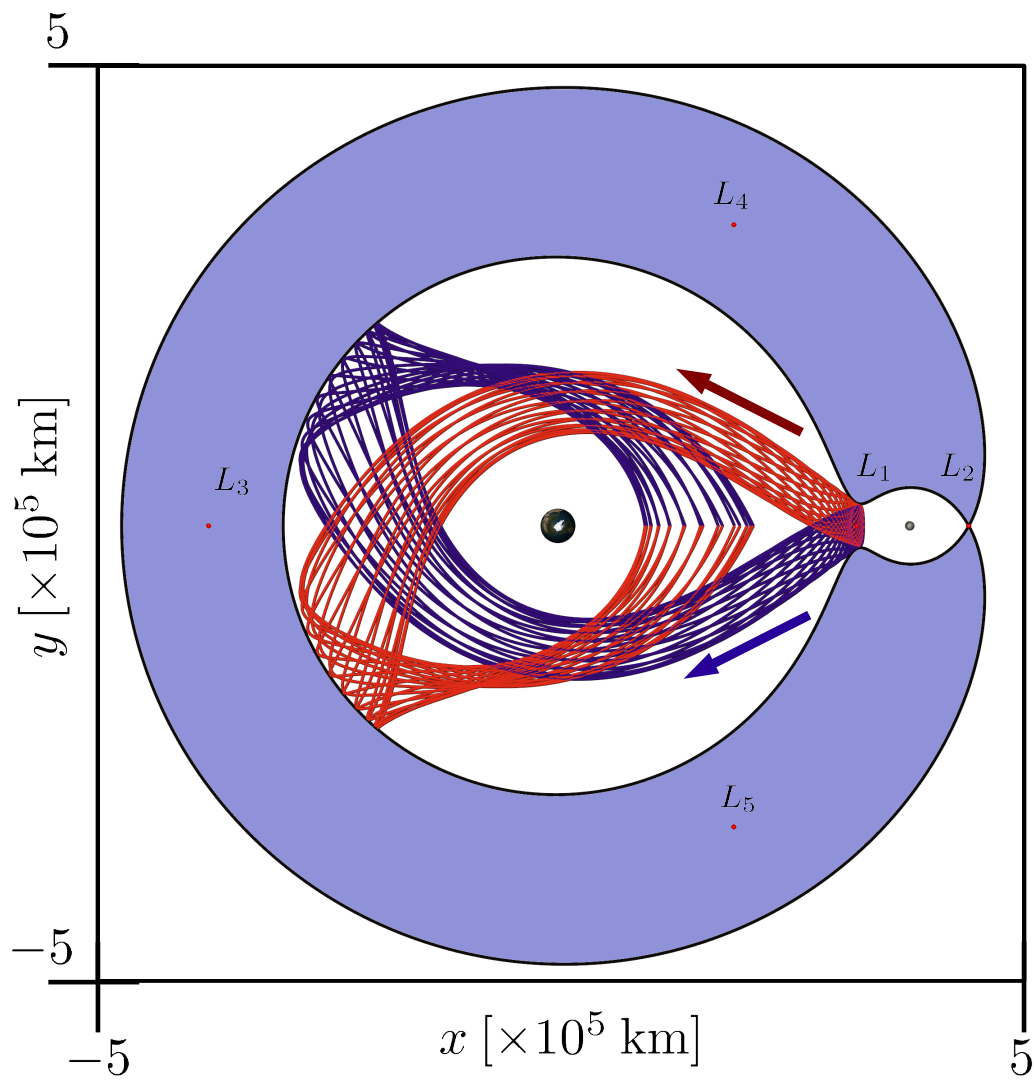


Figure 5.3. Configuration-space projections of stable (blue) and unstable (red) manifolds (arrows imply direction of integration)

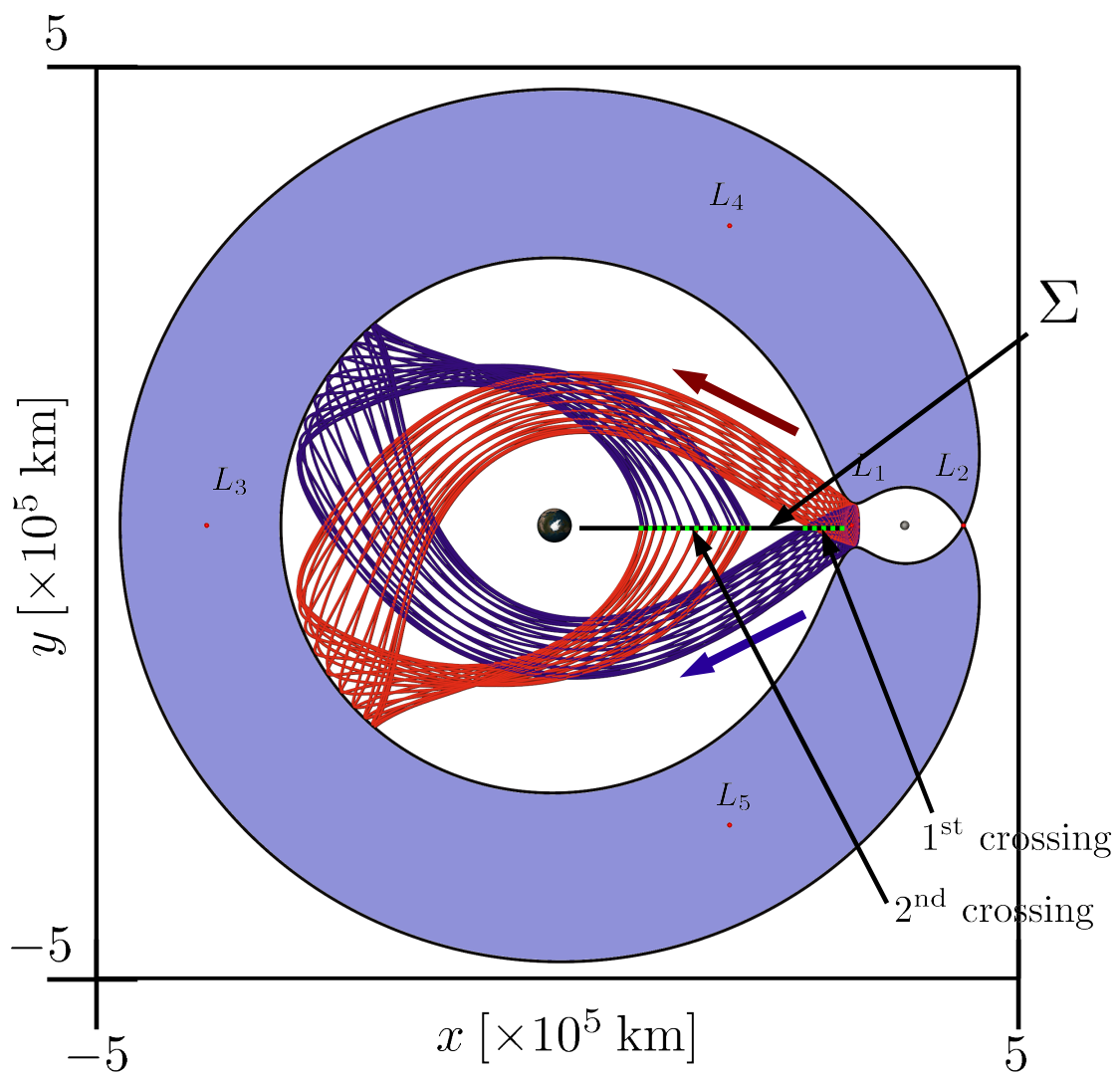


Figure 5.4. An illustration of the first two crossings of the map Σ

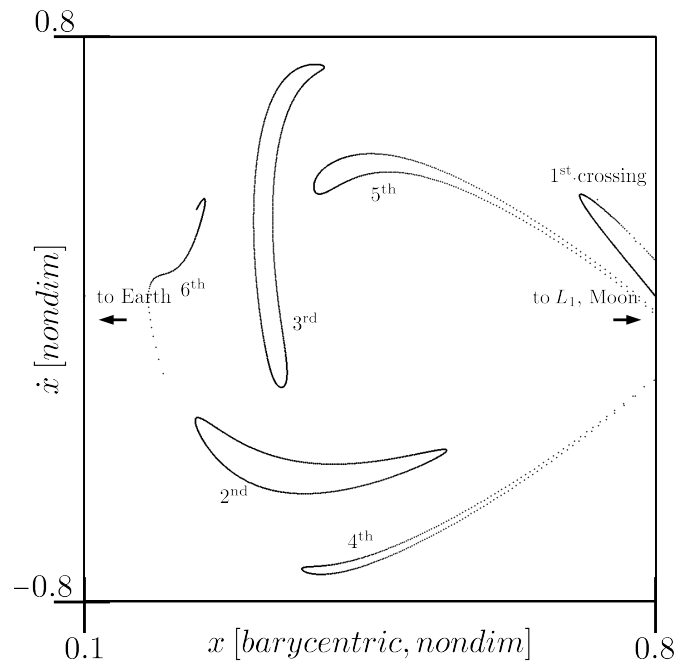


Figure 5.5. Stable manifold crossings associated with a Lyapunov orbit

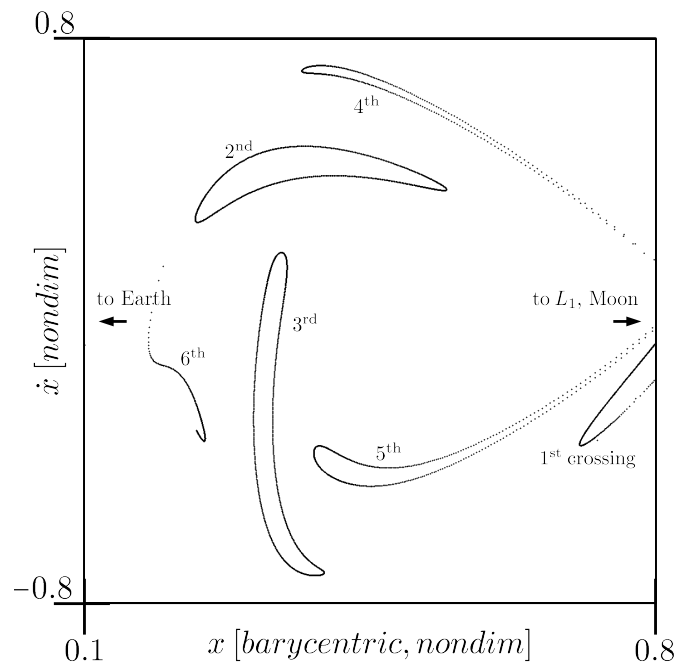


Figure 5.6. Unstable crossings; counterpart of Figure 5.5

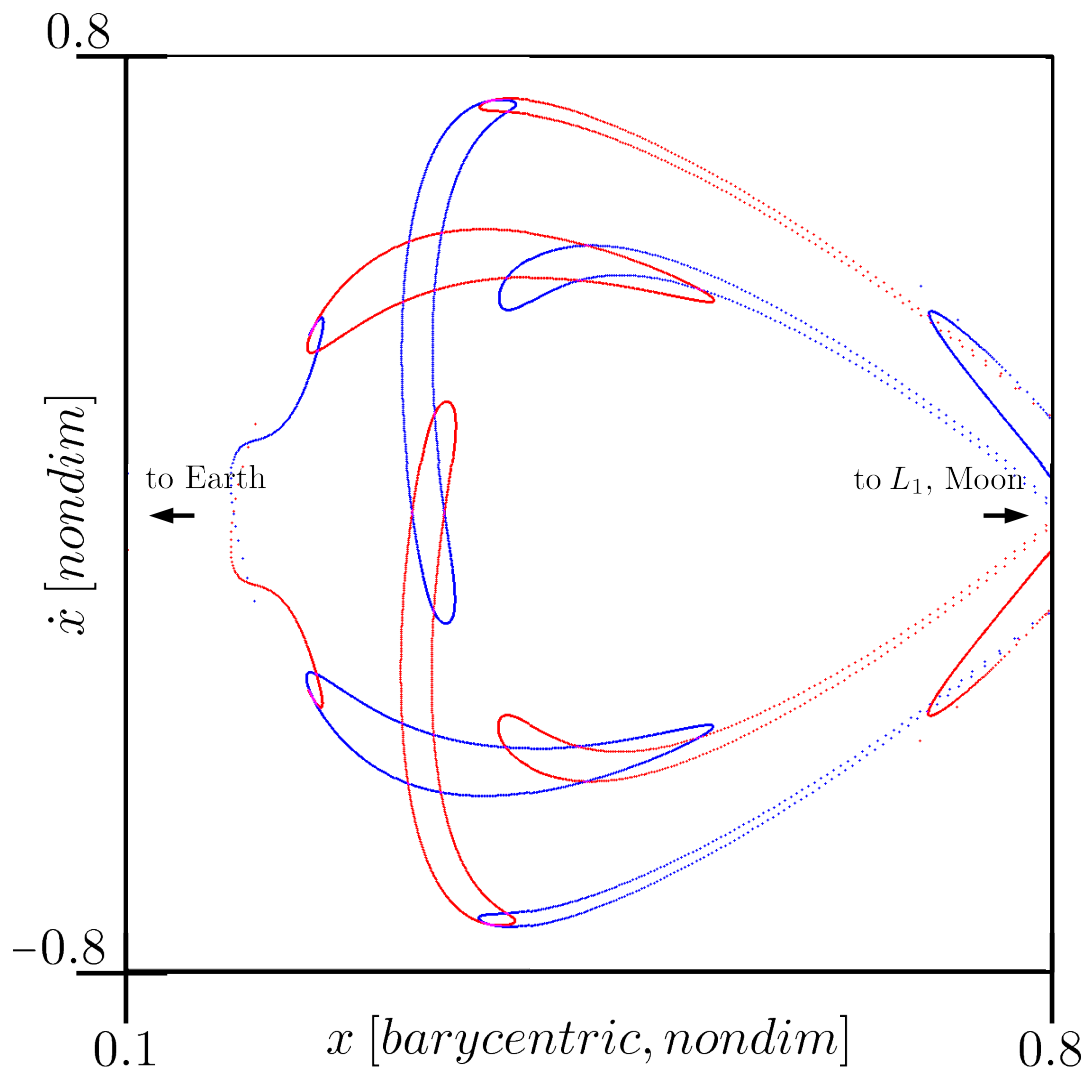


Figure 5.7. Both stable (blue) and unstable (red) manifold crossings

The FTLE maps associated with the region from Figures 5.5–5.7 reveal immediate similarities to the manifolds in the figures while also highlighting various differences. For example, in Figure 5.8, the FTLE values result from three iterations of the Poincaré map, consequently the contours corresponding to the later crossings of the Lyapunov manifolds have yet to develop. However, in Figure 5.9 (after 5 crossings of the Poincaré map), the structures associated with these later crossings are present. A comparison of the figures reflects the fundamental and important observation that FTLE values becomes more representative of the dynamics with a longer evolution. Figure 5.10 is included for further illustration.

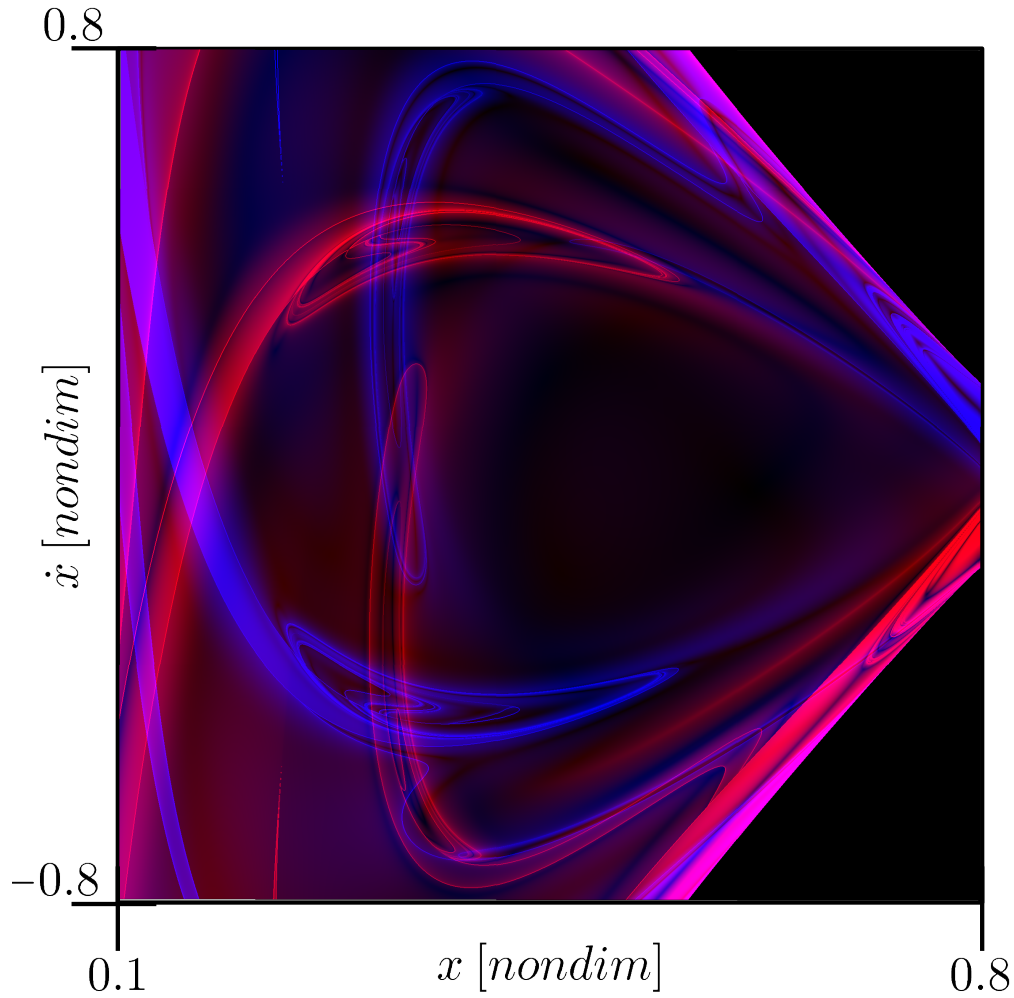


Figure 5.8. FTLE for three crossings of the Poincaré map, $\Sigma = y = 0$

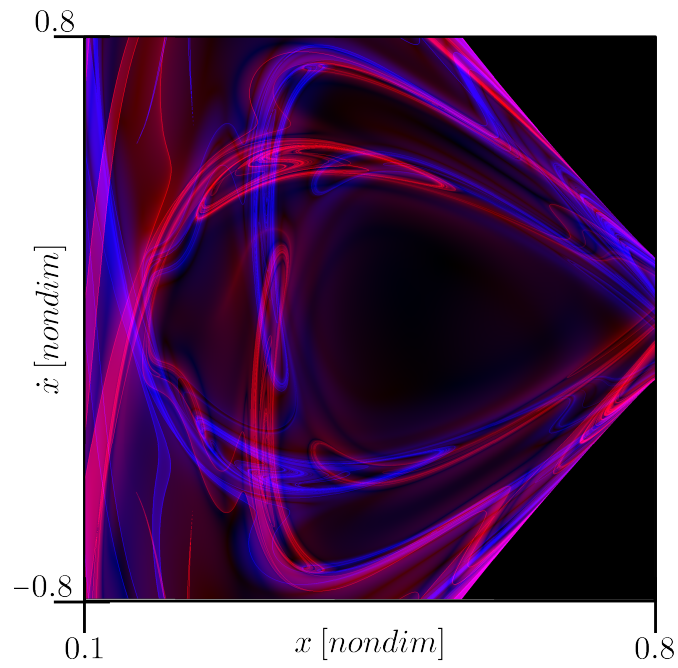


Figure 5.9. FTLE for five crossings of the Poincaré map, $\Sigma = y = 0$

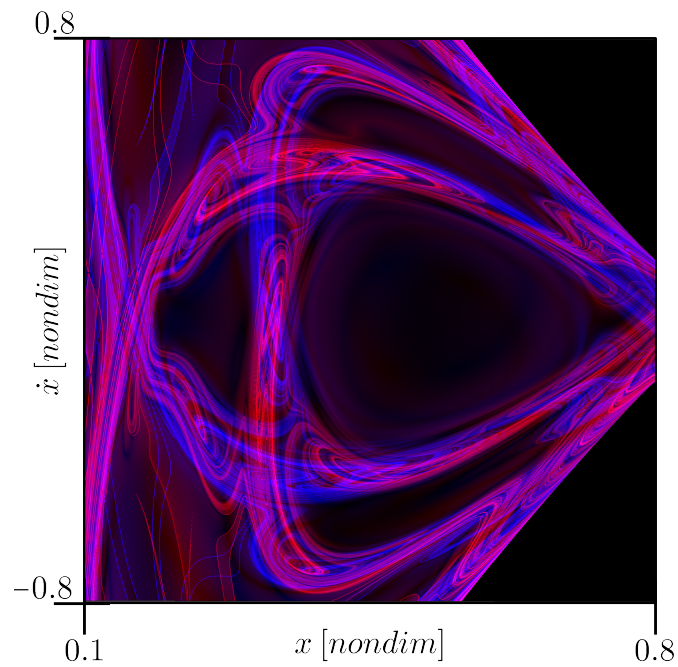


Figure 5.10. FTLE for ten crossings of the Poincaré map, $\Sigma = y = 0$

In each of Figures 5.8–5.10, the solid black triangular regions in the top and bottom right corners correspond to regions forbidden by the Jacobi limiting curves (i.e., the analog of zero velocity curves as viewed in the $x-\dot{x}$ phase space). Consistent with previous FTLE images, the relative value of the FTLE is represented by color saturation where FTLE values of zero are indicated by black regions and saturated blue and red pixels correspond to the higher values of forward and backward FTLE, respectively. Finally, the manifold crossings appear

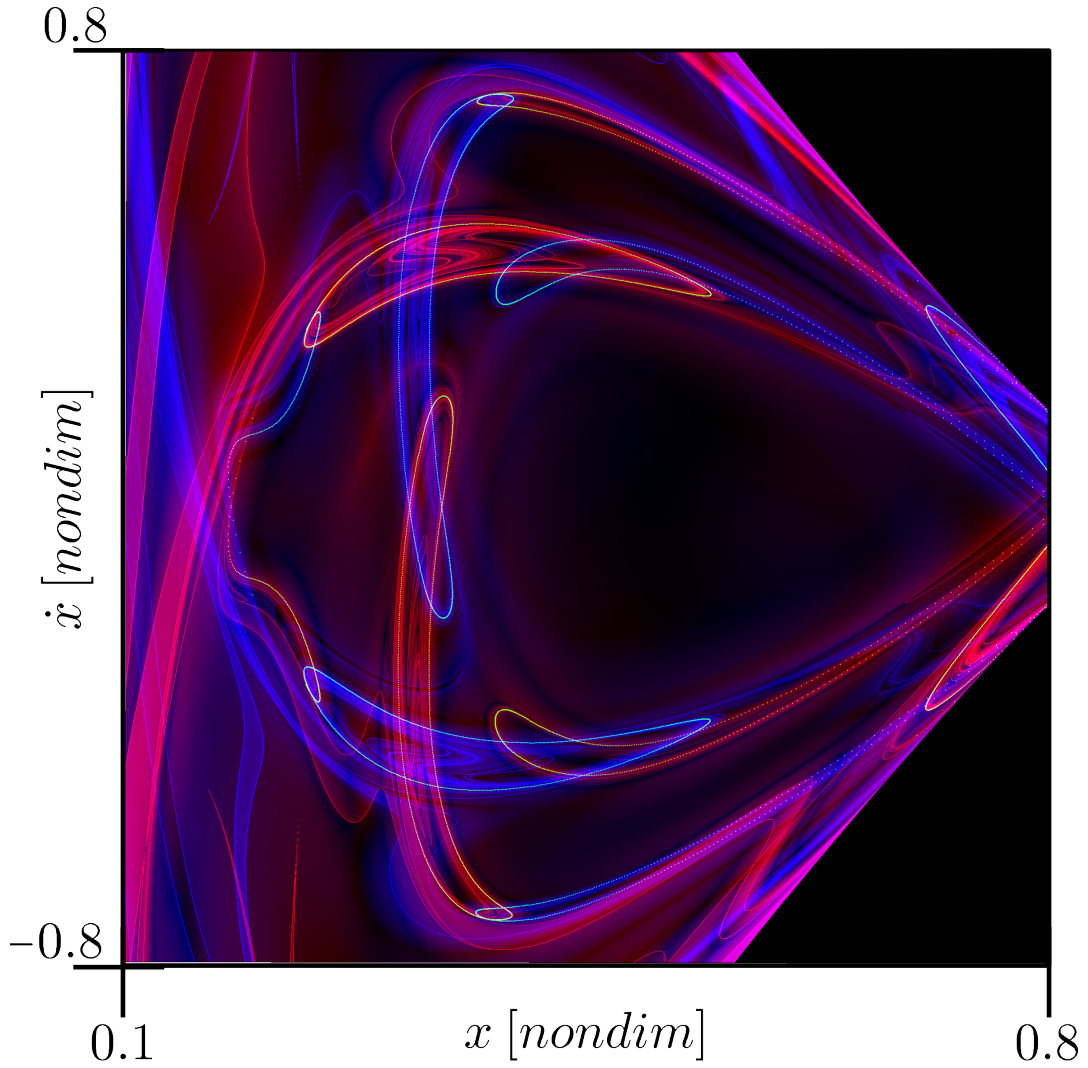


Figure 5.11. FTLE ($5\times$ of $\Sigma = y = 0$) with manifold crossings

in green on Figure 5.11 as an overlay of the contours from Figure 5.7. While the FTLE maps require integration of an initial state associated with every pixel, which is significantly many more trajectory evaluations than required to represent known manifolds, these maps reveal vastly more detail including potentially unknown structures.

Applying the ridge extraction methods described in Section 4.2, ridge points are identified in the previous FTLE maps. Ridge points are identified in green in Figures 5.12–5.14, while the red (unstable) and blue (stable) color scheme from the preceding FTLE images is maintained. The process for producing the ridge-point maps begins with a Gaussian blur where the image value at a particular pixel, along with its neighbors, is convolved with a Gaussian distribution, as described in Equation (4.9) to smooth out low-scale noise. The specific Gaussian blur that is applied is characterized by a Gaussian distribution defined over 3 standard deviations. The tolerance used to identify zeros in the modified Hessian-

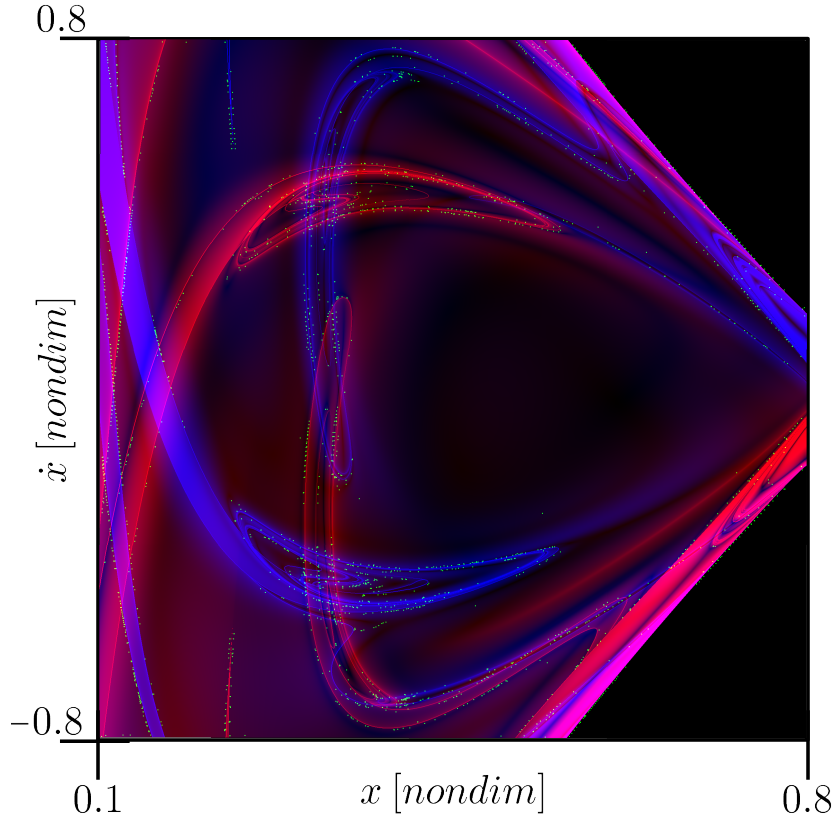


Figure 5.12. Ridge points on the three-crossing FTLE map

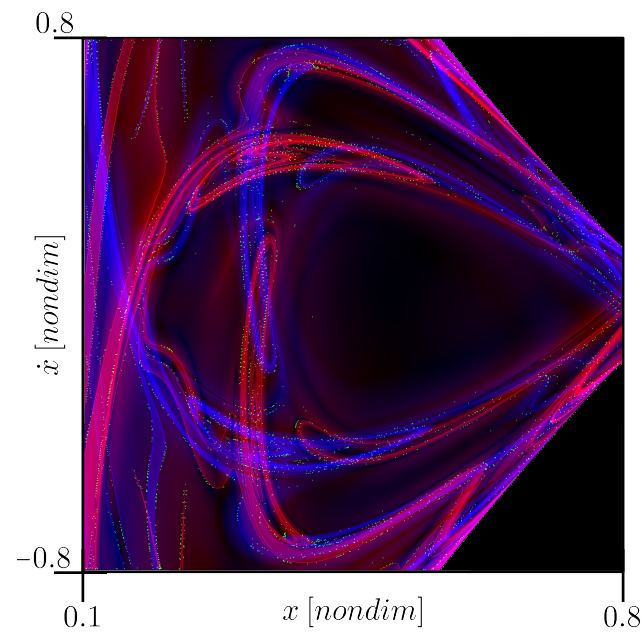


Figure 5.13. Ridge points on the five-crossing FTLE map

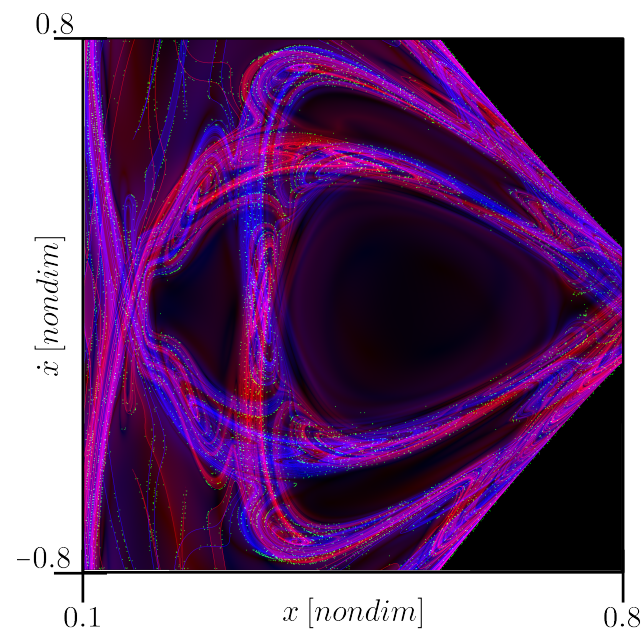


Figure 5.14. Ridge points on the ten-crossing FTLE map

gradient field, as described in Equation (4.14), is selected as 1×10^{-13} , however, this is not the only criteria available to identify ridge points in an image. While the Gaussian blur serves to dampen some low-amplitude noise, complete elimination of noisy artifacts requires excessive blurring resulting in poor correlation with the original image. Therefore, a second criteria is employed by the ridge algorithm to check against the relative strength of the ridges. Recall that larger negative values of $\tilde{\lambda}_2$, the smallest eigenvalue of the Hessian, indicate sharp downward curvature transverse to the ridge. Thus, the magnitude of the negative $\tilde{\lambda}_2$ supplies the strength of the ridge point. Setting a tolerance on this magnitude identifies only ridge points possessing at least the level of prominence associated with the cut-off value. This negative tolerance is selected such that ridge points with $\tilde{\lambda}_2 < -500$ are identified in Figures 5.12– 5.14. This magnitude is an empirically adequate value, and does not necessarily represent an optimal choice. The accuracy of the resulting ridge points highlights some of the difficulties in successfully locating ridges. Gaussian blurring dampens undesirable noise but, unfortunately, it also shifts the positions of ridge points. However, without such dampening, spurious ridge points that result from the noise, can easily saturate an image. In images with high complexity, it is difficult to identify one set of blurring parameters and ridge point tolerances that effectively capture all of the desirable features in the image. Ultimately, while a ridge-finding algorithm supplies an automated means to capture ridge points, successful implementation for a given image requires extensive user interaction, defeating a significant advantage of an automated process. Fortunately, locating ridges within an image is an active area in visualization research centers and more sophisticated ridge-finding algorithms are emerging. For this analysis, the direct investigation of the FTLE map (wherein LCS are apparent), that is, visually selecting a point along a ridge and thereby recovering its state, is often very useful.

The zoomed view of a bounded area from Figure 5.13 appears in Figure 5.15. The isolated region is indicated by the white box in the inset from Figure 5.13. Three yellow cross-hair markings indicate points on the map that are selected for illustration. These points are purposefully identified for their proximity to FTLE structures, and data associated with these points is tabulated in Table 5.1. Note that numerical data concerning the ridge points is included from visual inspection as well as ridge-finding algorithms. When the ridge point is determined from visual inspection, the six digits in the table for the initial

trajectory state corresponding to Point 1 represents the limit of accuracy attainable from the visual selection of a single pixel from the 2048×2048 image. The six digits in the table for Points 2 and 3 are actually truncations of the 12 digits available from the automated ridge-finding scheme.

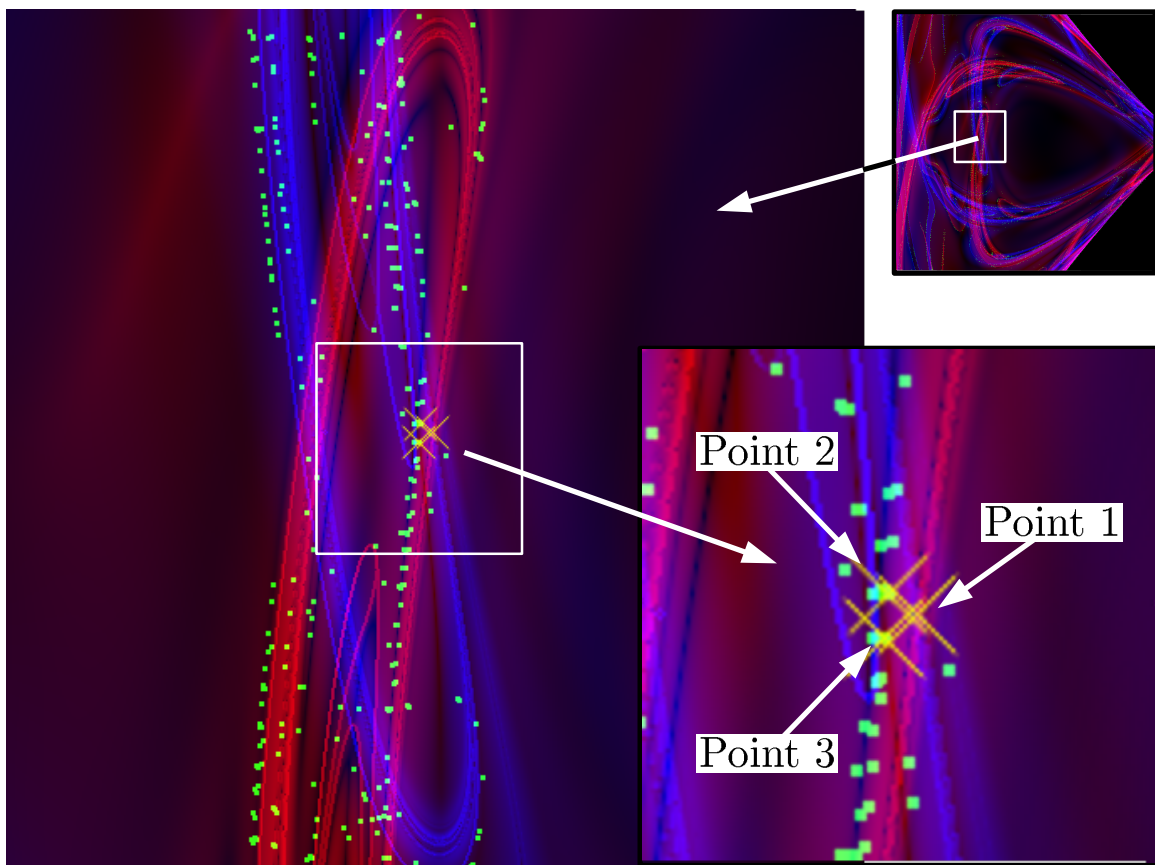


Figure 5.15. Zoom of the five-crossing FTLE map

Table 5.1
Data Points Selected from Figure 5.15 for Closer Inspection Via Plotting

Relative Location in Figure 5.15		Identified by	$[x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z}]^T$ (non-dim)	$[x \ y \ z \ \dot{y} \ \dot{z}]^T$ (km km/s)	Figure No.
1	Right-most	Inspection	$[0.340084 \ 0.0 \ 0.0 \ \dots]$	$[130724.267766 \ 0.0 \ 0.0 \ \dots]$	5.16
			$[-0.002868 \ 1.609362 \ 0.0]^T$	$[-0.002938 \ 1.648892 \ 0.0]^T$	
2	Upper-left	Algorithm	$[0.337676 \ 0.0 \ 0.0 \ \dots]$	$[129798.882374 \ 0.0 \ 0.0 \ \dots]$	5.17
			$[-0.006773 \ 1.620752 \ 0.0]^T$	$[-0.006939 \ 1.660562 \ 0.0]^T$	
3	Lower-left	Algorithm	$[0.337348 \ 0.0 \ 0.0 \ \dots]$	$[129672.583712 \ 0.0 \ 0.0 \ \dots]$	5.18 and 5.19
			$[0.002240 \ 1.622328 \ 0.0]^T$	$[0.002295 \ 1.662176 \ 0.0]^T$	

Figures 5.16–5.18 are included as trajectory examples that appear to possess an interesting nature, corresponding to the states from Table 5.1, and that are immediately identified using FTLE maps. In each of the figures, the Earth and Moon are depicted at $3\times$ their actual size, and the libration points are indicated as red spheres. The Lyapunov orbit near L_1 from Figure 5.2 is included for reference in black. Finally, the ZVC associated with the Jacobi constant employed throughout the entire subsection, that is $C = 3.17216$, is also displayed in each figure. The “forbidden” region is apparent with a solid lavender fill. In Figures 5.16–5.18, the initial state, plotted as a green point in each figure, is propagated in both forward and backward time as indicated by blue and red arrows and the resulting trajectories, respectively. Thus, the red arrow denotes the direction opposite to the natural motion, while the blue arrow is consistent with both the direction of forward integration and the natural motion. The trajectory arcs are also colored red and blue consistent with the direction of the flow. In all six cases, the trajectories are evolved until the ninth crossing of the x axis occurs. Figure 5.16 corresponds to Point 1 from the table, a point selected from the map by visual inspection. This point appears to lie at the crossing of LCS that are associated with both stable and unstable manifolds in the system. While the accuracy of the initial state is limited by the visual resolution of the image, it is apparent that forward and backward propagation of the state evolves into the vicinity of the Lyapunov orbit. A corrections algorithm could be employed to adjust the trajectory numerically to approach and remain in a Lyapunov orbit. Both Points 2 and 3 yield similar results as is apparent in Figures 5.17 and 5.18. Again, blue and red arcs correspond to forward and backward integration, respectively. Notable in the trajectories associated with these latter cases is that the stable and unstable propagations nearly mirror one another across the x axis, potentially correctable to form a closed path through the system. However, these two arcs, unfortunately, pass beneath the surface of the Moon as is visually apparent in Figure 5.19 where the Moon is displayed to scale. Despite the ultimate collision, this type of solution still represents an interesting and potentially useful result; adding a constraint for an acceptable lunar pass distance in a design process still yields a useful trajectory. All of the previous results support further investigation. Exploring nearby points, coupled with numerical corrections, offers viable options as inputs to trajectory design algorithms.

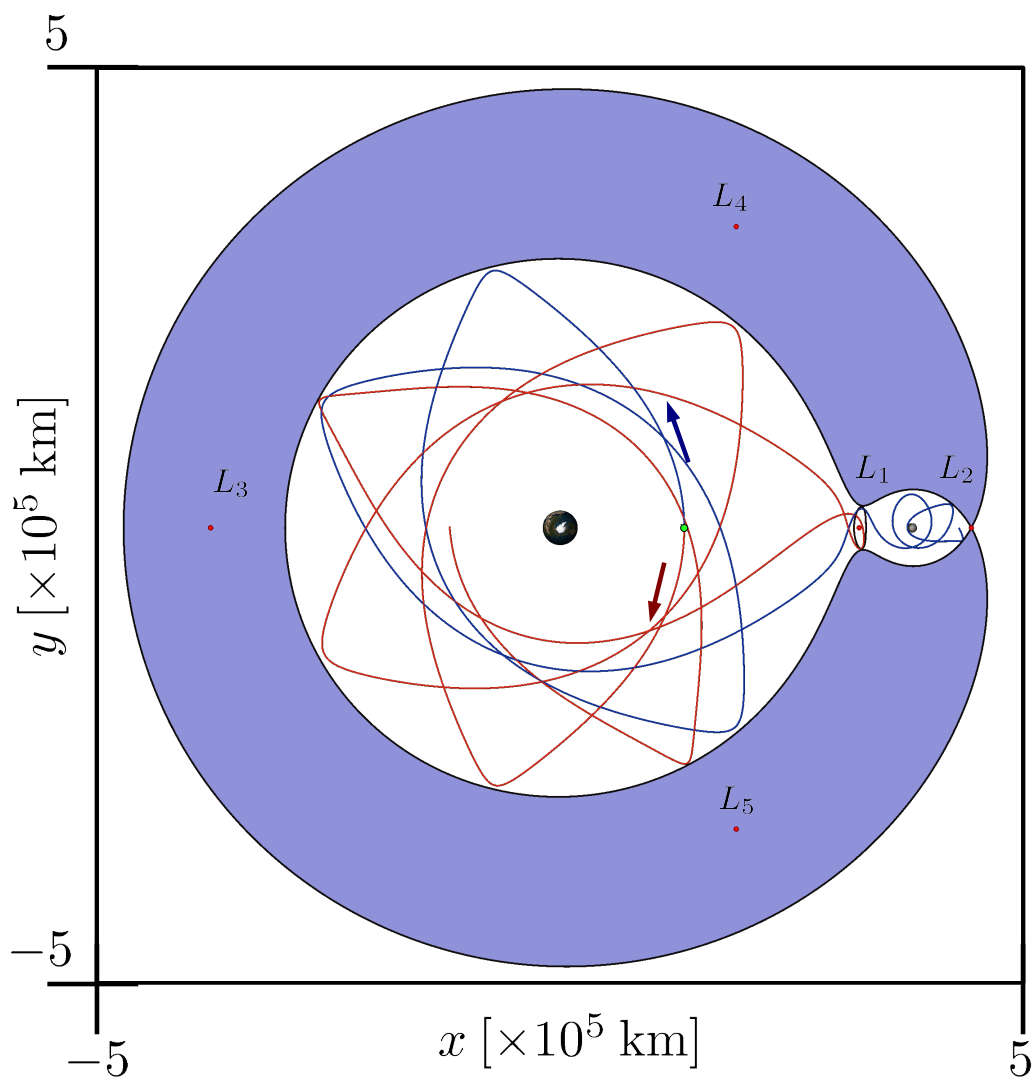


Figure 5.16. Trajectory arcs associated with Point 1 from Table 5.1

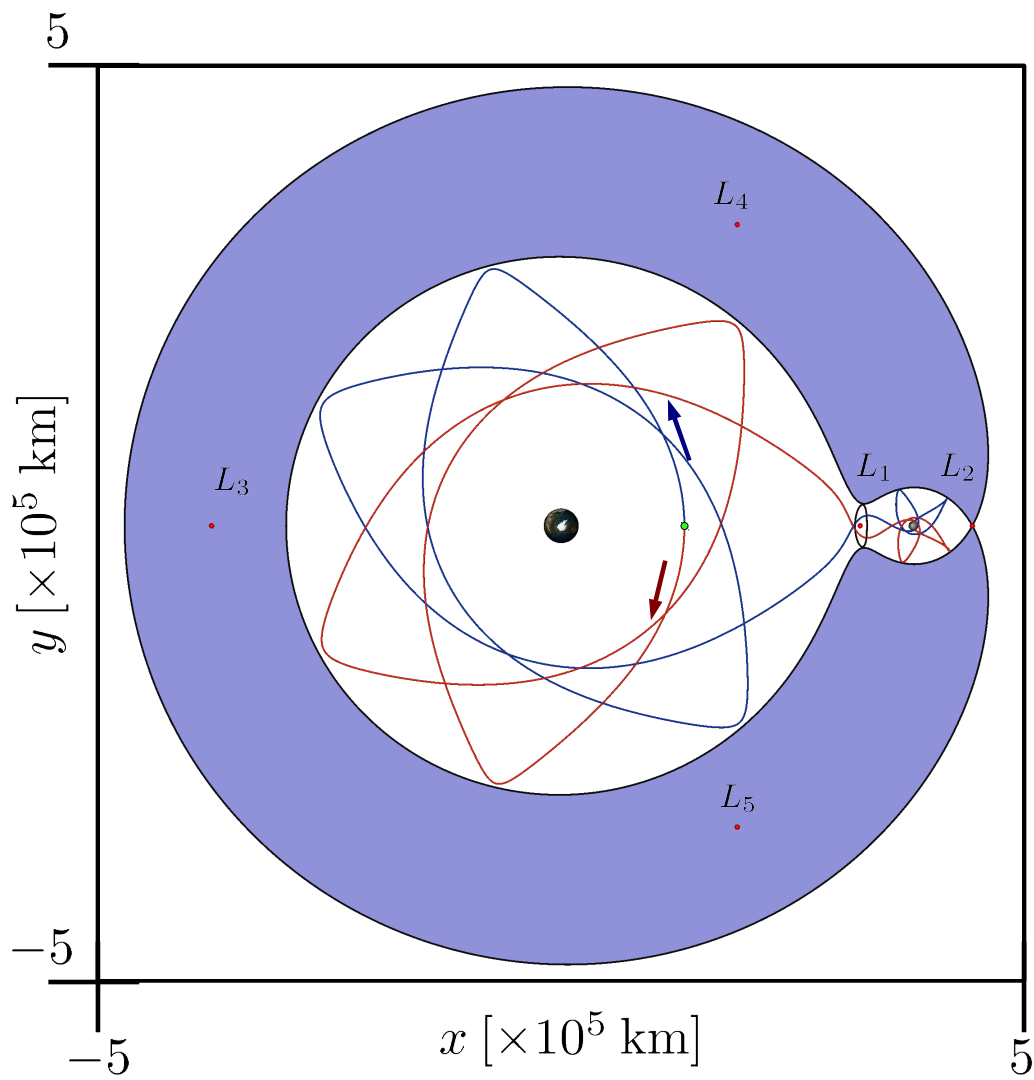


Figure 5.17. Trajectory arcs associated with Point 2 from Table 5.1

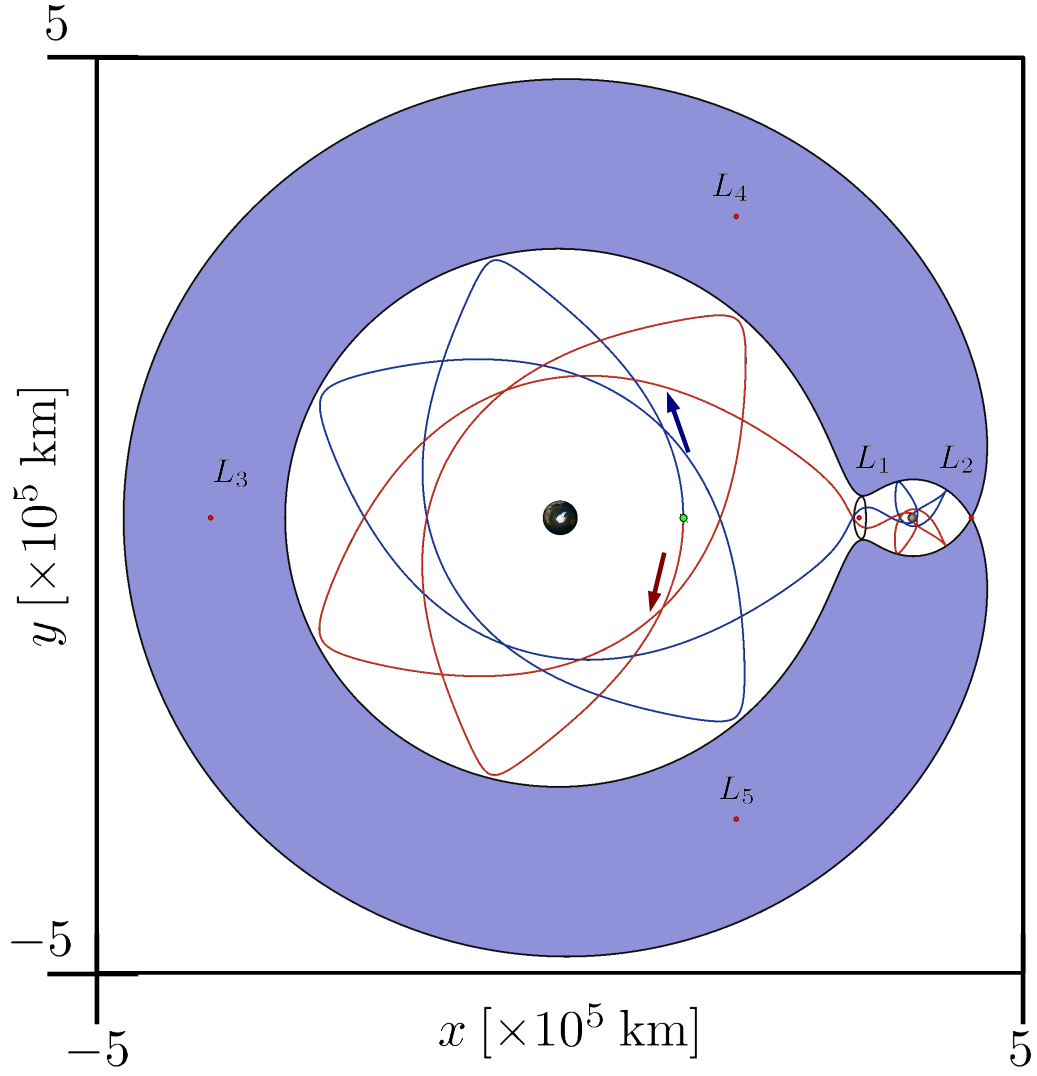


Figure 5.18. Trajectory arcs associated with Point 3 from Table 5.1

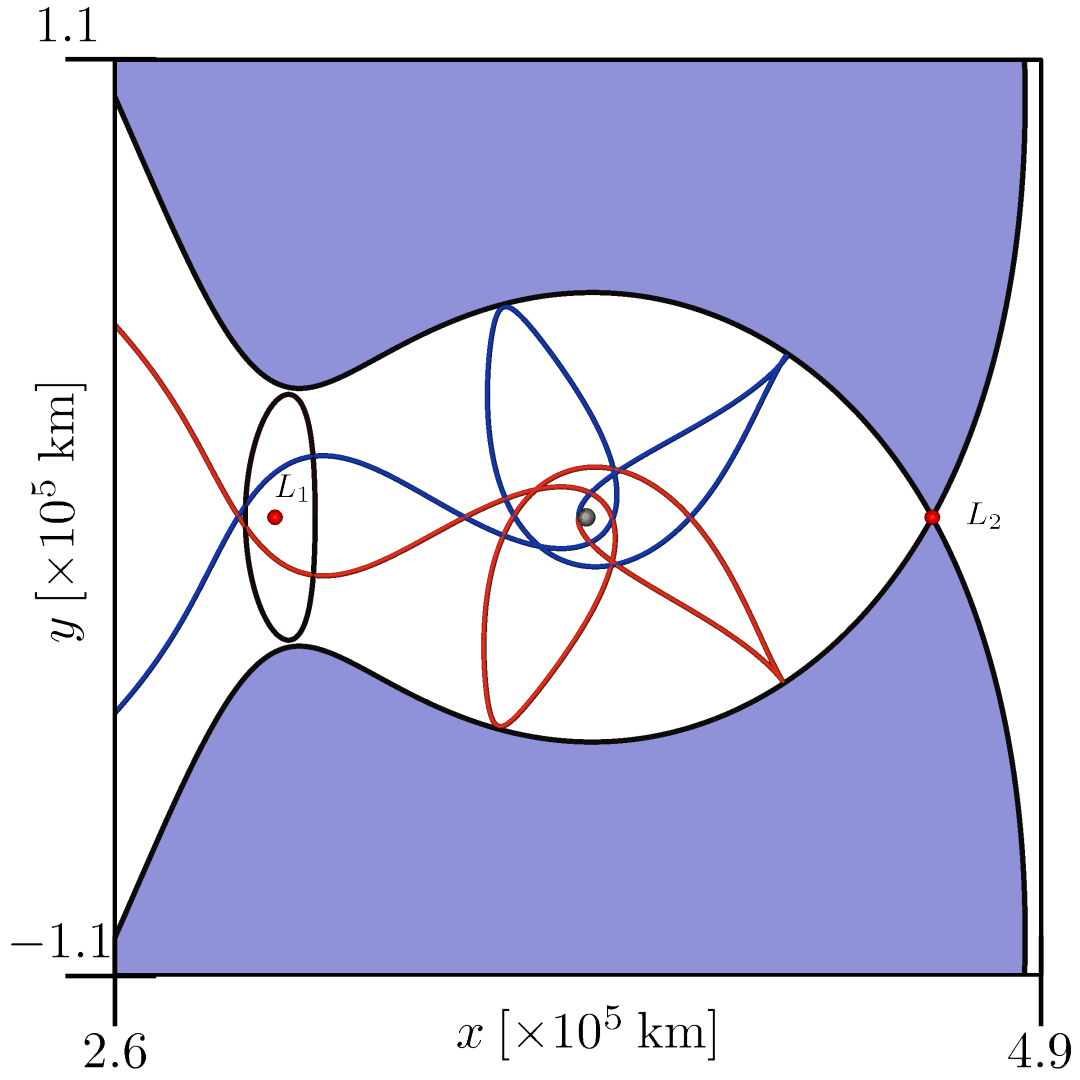


Figure 5.19. Zoom of Point 3 trajectory arcs focused in the region near the Moon

5.1.2 Structures In the L_4 Region with $C_{L_4} < C < C_{L_2}$

The collinear libration points have recently been a focus of numerous investigations in various systems, but the equilateral or triangular libration points in the EM system are also generating new attention. While the previous example highlighted all of the concepts embodied in this analysis, the following discussion focuses simply on presenting structures evident in FTLE maps near the triangular points. This discussion also serves to introduce the concept of periapse Poincaré maps and, in the process, bridges the gap between the previous and subsequent results. Unlike the characteristic roots computed for the collinear points, that include both real and imaginary values, the roots associated with the triangular points are pure imaginary, indicating that these equilibrium points are linearly stable. Analysis further indicates that periodic orbits near the equilateral points maintain this stability characteristic [67]. A consequence of the linear stability of these orbits is an obvious lack of stable and unstable manifolds. In recent years, some proposed mission scenarios included the possibility of transfers between orbits near the collinear points in the vicinity of the Moon and orbits associated with the triangular points. Despite the lack of flow that asymptotically approaches and departs the periodic orbits in the neighborhood of the triangular points, other transfer arcs may still approach these L_4 orbits. Inspection of FTLE maps in the vicinity of the L_4 and L_5 points supplies an immediate view of any potential arcs that may pass near the equilateral points, perhaps originating near the collinear points. If such options exist, the map structures can be exploited for transfer design. A brief survey of such maps in the vicinity of L_4 is summarized below.

To develop FTLE maps, an area near L_4 in the x - y plane is defined for examination. The specific region associated with the following FTLE maps is defined by an x - y bounding box centered on L_4 of width $x = x_{L_4} \pm 1.106 \times 10^5$ km and height $y = y_{L_4} \pm 1.106 \times 10^5$ km, as indicated in Figure 5.20. Since the map is defined in configuration space, the x and y coordinates are directly available on the map. The FTLE maps then represent the evolution through a time duration of 10 nondimensional time steps ($T \approx 43$ days), unlike previous maps defined on a Poincaré section. This time-duration mapping, coupled with a specific value of Jacobi constant does not constrain the two velocity states. Therefore, constraints are enforced on \dot{x} and \dot{y} such that, given x , y , and C , the associated initial

velocities represent a periapse point with respect to P_2 . Alternatively, either \dot{x} or \dot{y} is assigned a specific value, thereby constraining the remaining velocity state by the Jacobi constant value. The periapse condition is selected to produce maps that highlight orbits near the libration points in the vicinity of the Moon. The specific conditions to define such a periapse constraint are based on applications by Davis and Howell [68] as well as Haapala [69]. If the position and velocity of the spacecraft relative to P_2 are defined as,

$$\bar{q} = \begin{bmatrix} x - 1 + \mu & y & z \end{bmatrix}^T, \quad (5.1)$$

$$\dot{\bar{q}} = \begin{bmatrix} \dot{x} & \dot{y} & \dot{z} \end{bmatrix}^T, \quad (5.2)$$

then the conditions for periapses are,

$$\bar{q}\dot{\bar{q}}^T = (x + 1 - \mu)\dot{x} + y\dot{y} + z\dot{z} = 0, \quad (5.3)$$

$$v^2 + \bar{q}\ddot{\bar{q}}^T = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + (x - 1 + \mu)\ddot{x} + y\ddot{y} + z\ddot{z} \geq 0, \quad (5.4)$$

where, from Equation (2.27), $v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$.

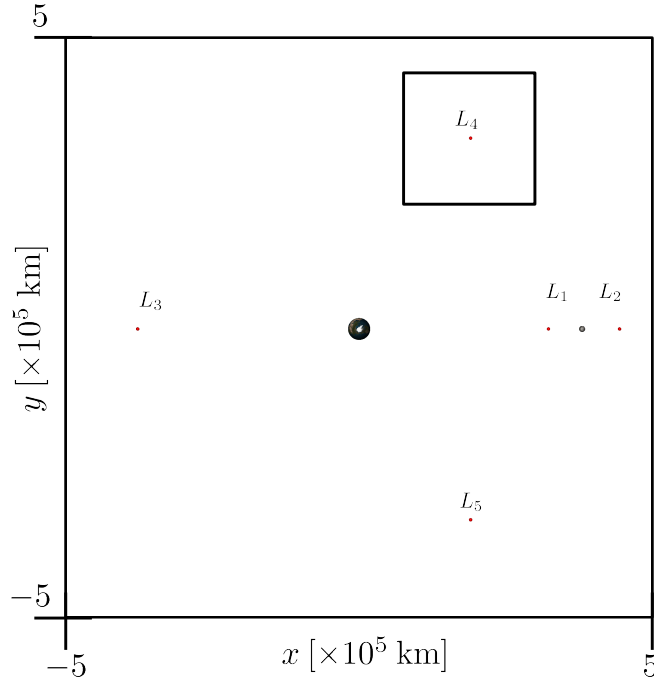


Figure 5.20. The L_4 region of interest explored in Section 5.1.2

A series of FTLE maps are constructed within the region near L_4 defined in Figure 5.20. The resulting maps appear in Figures 5.21 and 5.22. In each case, forward FTLE are represented in shades of blue and backward FTLE are colored red. The sequence begins in Figure 5.21 (a) and the maps are initially associated with a Jacobi constant value of $C \approx C_{L_2}$. The series of images concludes with Figure 5.22 (e) corresponding to a C value that is slightly greater than C_{L_4} . It is apparent from these images that some structure is available and the associated features represent potential trajectory options for exploration missions.

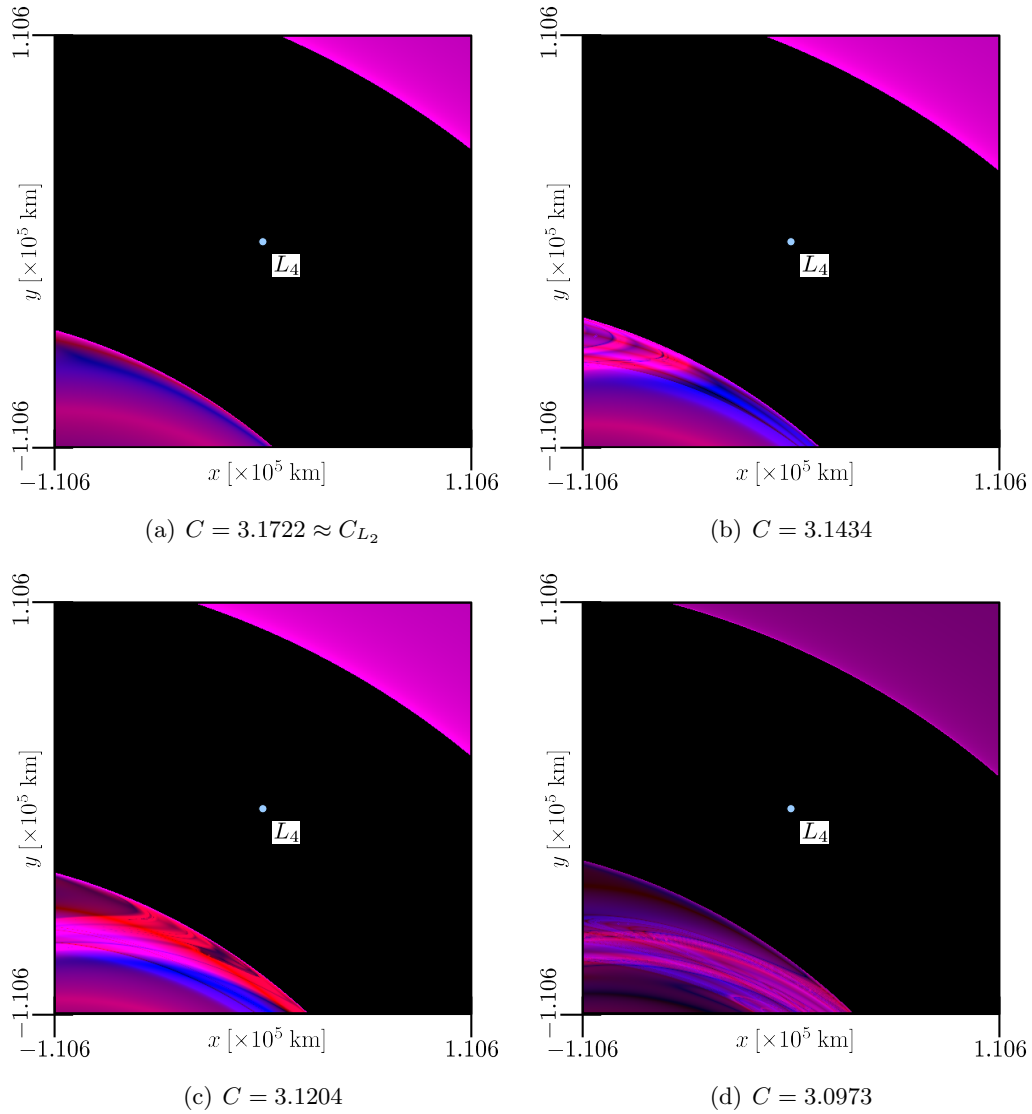


Figure 5.21. FTLE in the EM L_4 region for various C values (I)

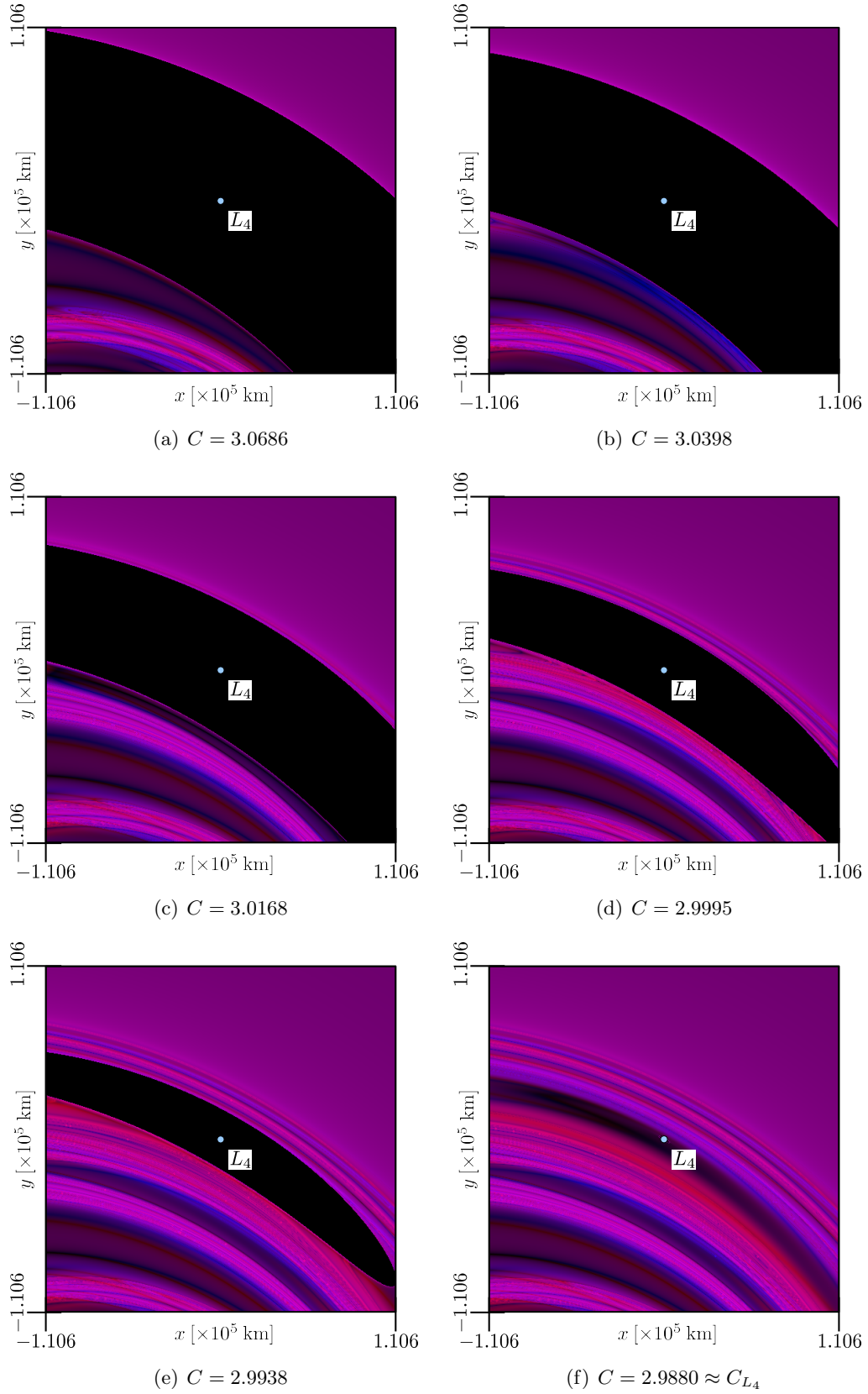


Figure 5.22. FTLE in the EM L_4 region for various C values (II)

5.2 Selected Lagrangian Coherent Structures in the Sun-Saturn System

Investigations combining the periodic nature of the CR3BP and the powerful analytical framework supplied by Poincaré maps are particularly insightful. Consider defining the hyperplane of the Poincaré map consistent with trajectory apsides. These maps are defined in configuration space and are, therefore, sometimes more intuitive than maps defined in a phase space. But, periapse Poincaré maps can allow for a naturally convenient setting for examination. The combination of such maps with invariant manifold theory to identify potential transit trajectories, homoclinic and heteroclinic connections, as well as cometary paths is the focus in Haapala [69]. Characterizing long-term trajectory behavior in a three-body system is the subject of additional analysis by Davis and Howell [68, 70]. In both cases, specific results are analyzed in the Sun-Saturn CR3BP ($\mu \approx 2.85804 \times 10^{-4}$). The discussion in this section elaborates (1) the potential advantages of FTLE/LCS analysis with their relationship to Sun-Saturn transit trajectories, and (2) the identification of features associated with both long-term periapse maps and FTLE maps.

5.2.1 Potential Sun-Saturn Transit Trajectories

Haapala explores trajectory arcs in the Sun-Saturn system [69]. The general approach in Haapala's analysis is the identification of invariant manifold crossings on a periapse Poincaré map. These crossings are generally focused in the P_2 region of the Sun-Saturn system. Transits, that is, pathways in and out of the P_2 (Saturn) region can be identified by inspecting the various intersecting regions of the map that are associated with the manifolds of L_1/L_2 Lyapunov orbits. The invariant manifolds associated with the L_1/L_2 Lyapunov orbits in the planar CR3BP act as separatrices in the flow for a specific C value [71]. Thus, the manifolds generate isolated regions in the map, each with certain characteristics. Regions defined by stable manifolds might overlap regions determined by the unstable manifolds. State vectors originating in such an isolated region must pass through, or transit, through the libration point gateways as they evolve. So, selecting a particular type of intersecting region supplies an initial state that propagates and subsequently enters or exits the P_2 region in a predictable way. This approach is generalized such that a single L_1/L_2 trajectory passes

through specified regions of the map, in a pre-determined order, to produce a predictable path.

As detailed by Haapala in [69], transit trajectory information is observed via periapse Poincaré maps. Understanding of the dynamical behavior is further expanded through exploration of the corresponding FTLE data. While intersecting regions are identified from a periapse map, differing behavior within these regions is only identifiable by blanket inspection of the region. An FTLE map supplies the requisite density, with every pixel in the map containing qualitative information. A periapse Poincaré map appears in Figure 5.23. The features in this map and the subsequent analysis are characteristic of a specific Jacobi constant value, $C = 3.01740$. For this C value, the ZVC are slightly open at L_2 . In the figure, stable manifold crossings of the map are indicated as blue dots while the first crossing of an unstable manifold, enclosing a closed lobe, is colored in red. This overlap of the stable

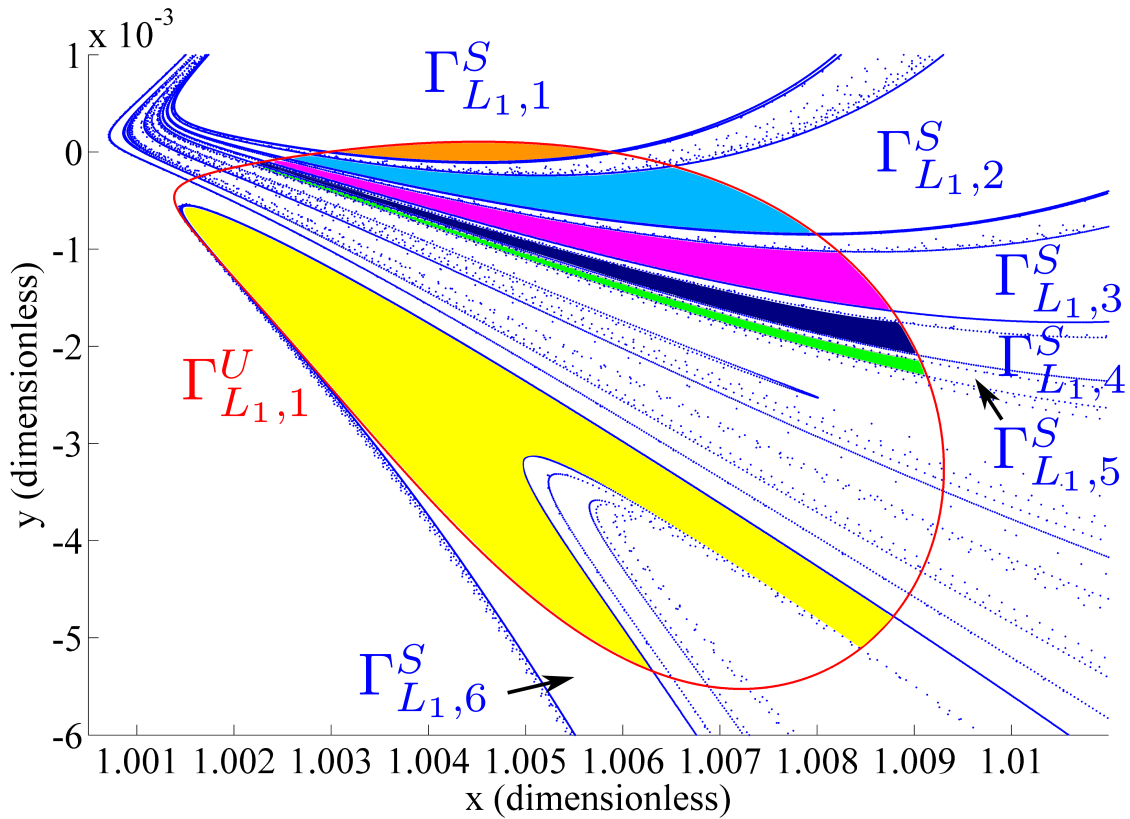


Figure 5.23. Periapse Poincaré map of manifolds (with permission: A. Haapala [69])

(blue) and unstable (red) manifolds facilitates the design process. The lobe outlined in red contains intersections of states interior to the first crossing of an unstable manifold with states inside successive crossings of stable manifolds. The manifold structures are labeled with the symbol Γ with superscripts U and S indicating unstable and stable manifolds, respectively. The subscript, L_1 , indicates that these manifold structures are associated with a L_1 Lyapunov orbit in the Sun-Saturn system. Regions of intersection between the stable and unstable manifolds are colored by their associated crossing of the hyperplane, where different solid-colored regions denote subsequent intersections as indicated by subscript indices (1-6). The corresponding FTLE map is plotted in Figure 5.24. Consider the yellow region

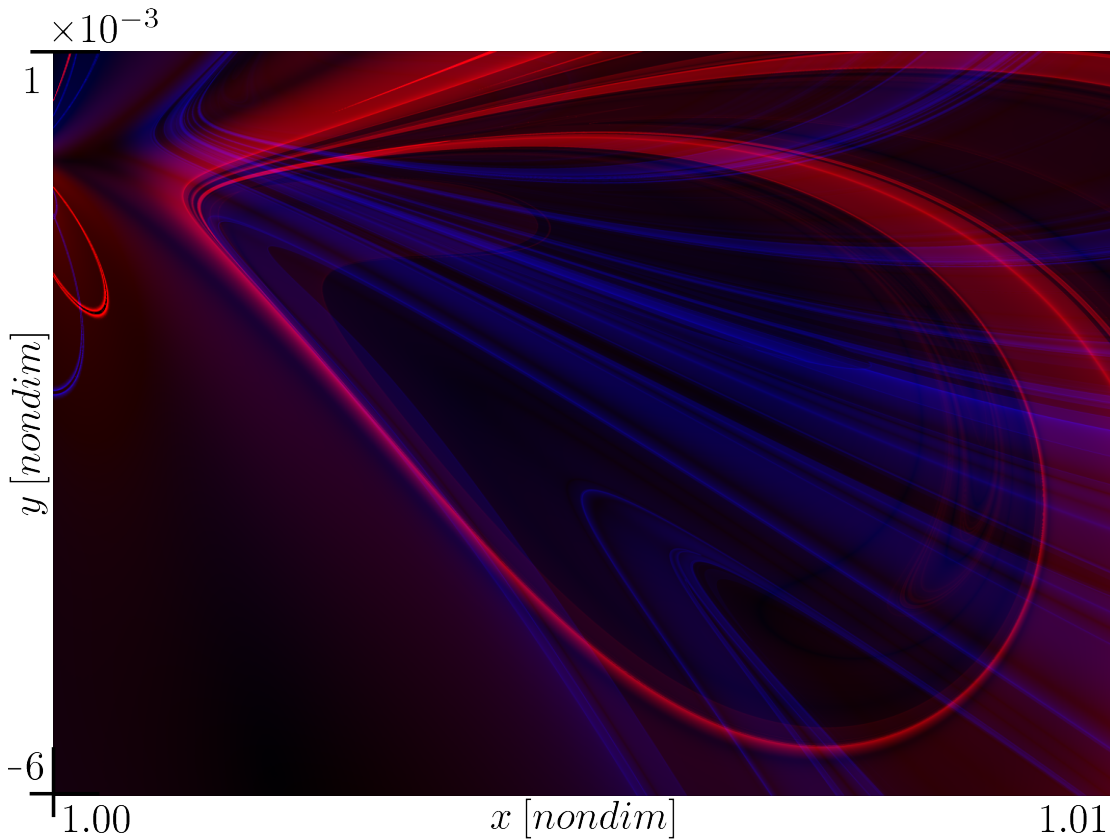


Figure 5.24. FTLE map corresponding to Figure 5.23

from Figure 5.23. This region is also identified in Figure 5.24. All of the points within the red lobe are first crossings of trajectories that enter the P_2 region through the L_1 gateway; this conclusion results from the fact that all such points lie within the red boundary, i.e.,

the first unstable crossing of the manifold associated with the L_1 Lyapunov orbit. Overlaid on the red lobe are backward crossings of the map by the stable manifold emanating from the same orbit. Recall that the colored ribbons identify a particular crossing number where yellow implies the 6th crossing. Within the area of the FTLE map (Figure 5.24) that corresponds to the yellow band from Figure 5.23, note the region of relatively larger FTLE values (lighter color) running diagonally along the edge of the lobe. The information associated with the FTLE map in this region is not captured in the associated Poincaré manifold plot. Other structures appear that originate with flow in this region that is not previously captured. Two zoomed views focus on specific areas of this region and appear as insets in Figure 5.25. These specific areas are identified due to the presence of intersecting

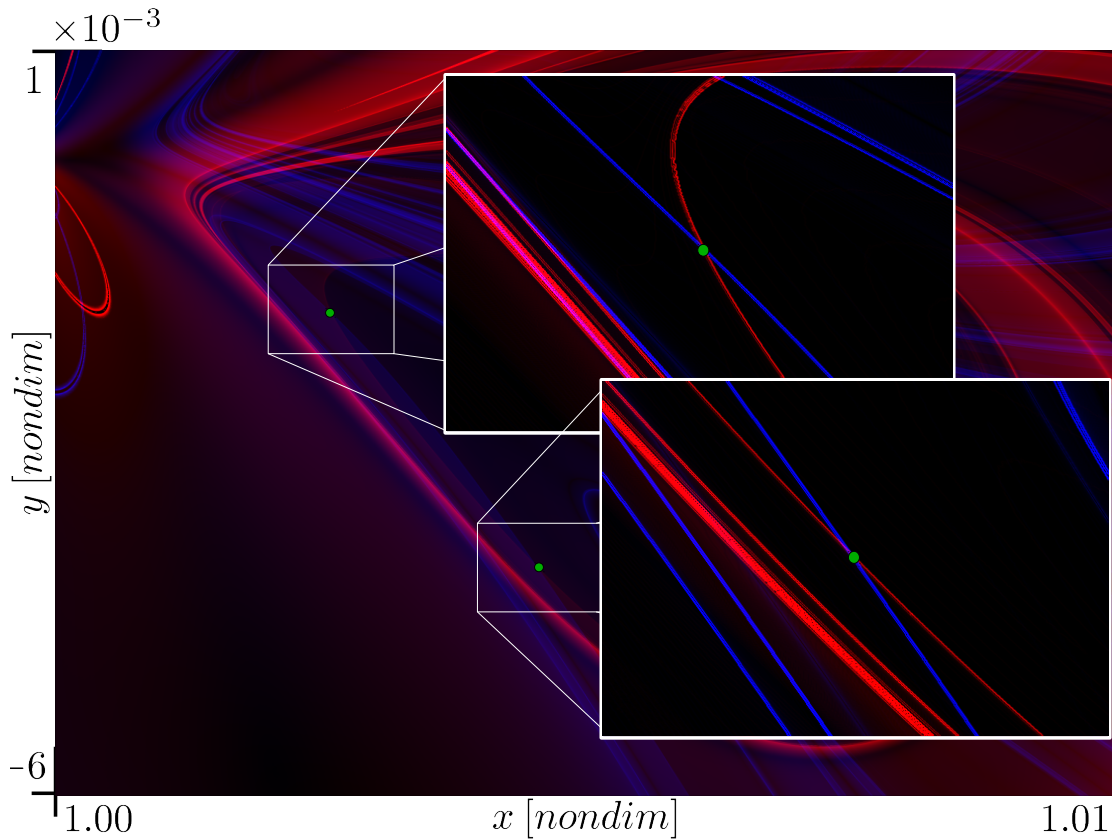


Figure 5.25. Zoomed insets focused on regions of interest from Figure 5.24

FTLE. Note that digital enhancement is performed on the insets to highlight the structures. The intersections are represented by green dots in Figure 5.25 and the associated states are

listed in Table 5.2. These states are evolved forward and backward in time until they effectively meet and the associated trajectories, along with zoomed views of the Saturn region, appear in Figures 5.26–5.29. Note that both the Sun and Saturn appear scaled $100\times$ and $20\times$, respectively, in these figures. Corrections are required to join the arcs at the point of discontinuity given the manual point identification from the associated image. Thus, these intersections, readily identified from the FTLE data and not immediately apparent from an associated manifold Poincaré plot, immediately produce interesting results in the form of potentially periodic orbits that pass back and forth between the two primary regions in the Sun-Saturn system. This data is now the focus of further examination elsewhere.

Table 5.2
Point Data Selected Manually from Figure 5.25 for Inspection

Inset	$[x\ y\ z\ \dot{x}\ \dot{y}\ \dot{z}]^T$	$[x\ y\ z\ \dot{y}\ \dot{y}\ \dot{z}]^T$	Figure No.
	(non-dim)	(km $\times 10^9$ km/s)	
Top	$[1.002602\ -0.001474\ 0.0\ \dots\ 0.180594\ 0.353733\ 0.0]^T$	$[1.430439\ -0.002104\ 0.0\ \dots\ 1.742012\ 3.412110\ 0.0]^T$	5.26, 5.27
Bottom	$[1.004517\ -0.003834\ 0.0\ \dots\ 0.170356\ 0.213345\ 0.0]^T$	$[1.433169\ -0.005471\ 0.0\ \dots\ 1.643254\ 2.057928\ 0.0]^T$	5.28, 5.29

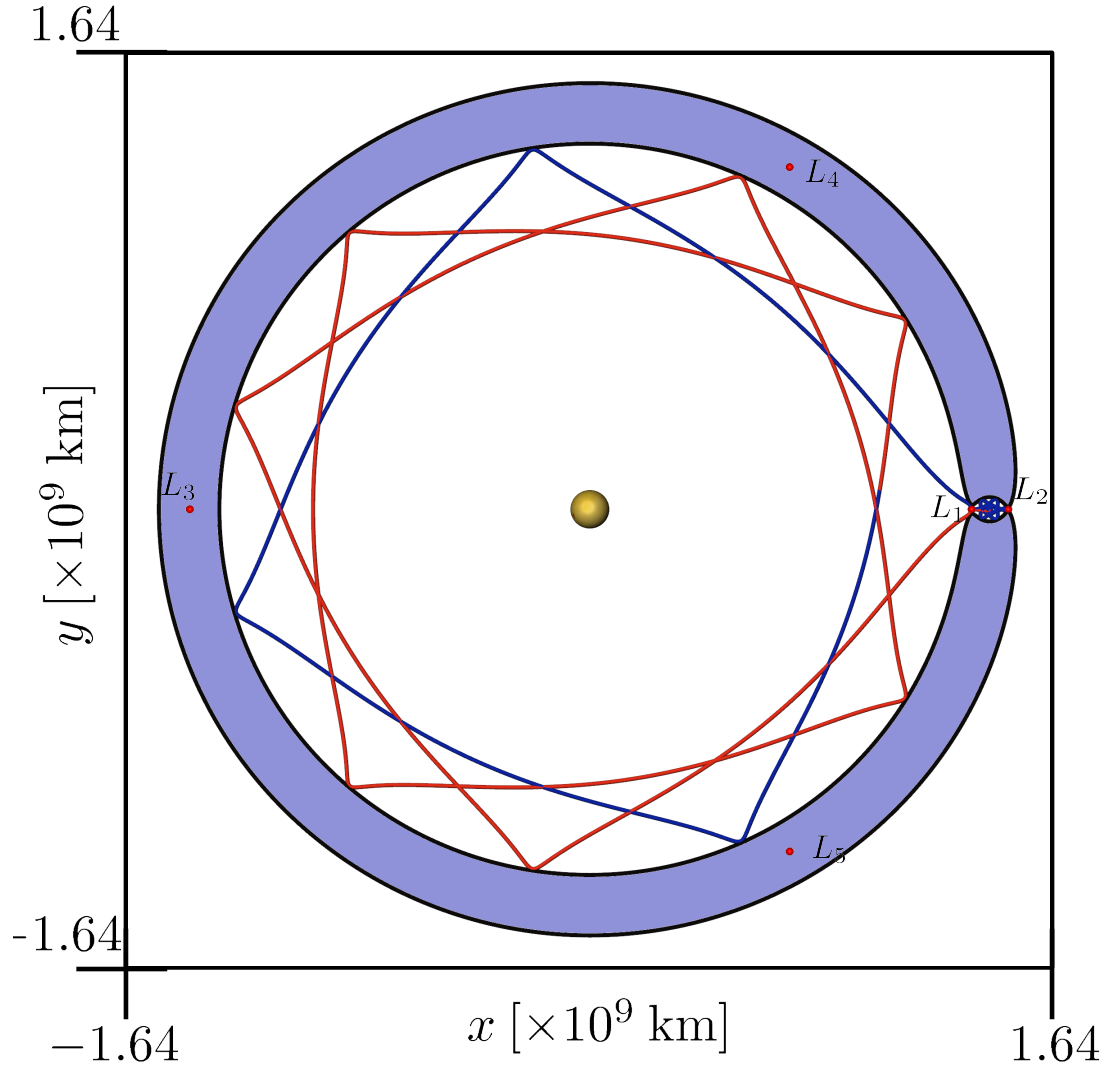


Figure 5.26. Trajectories from green point in top inset of Figure 5.25

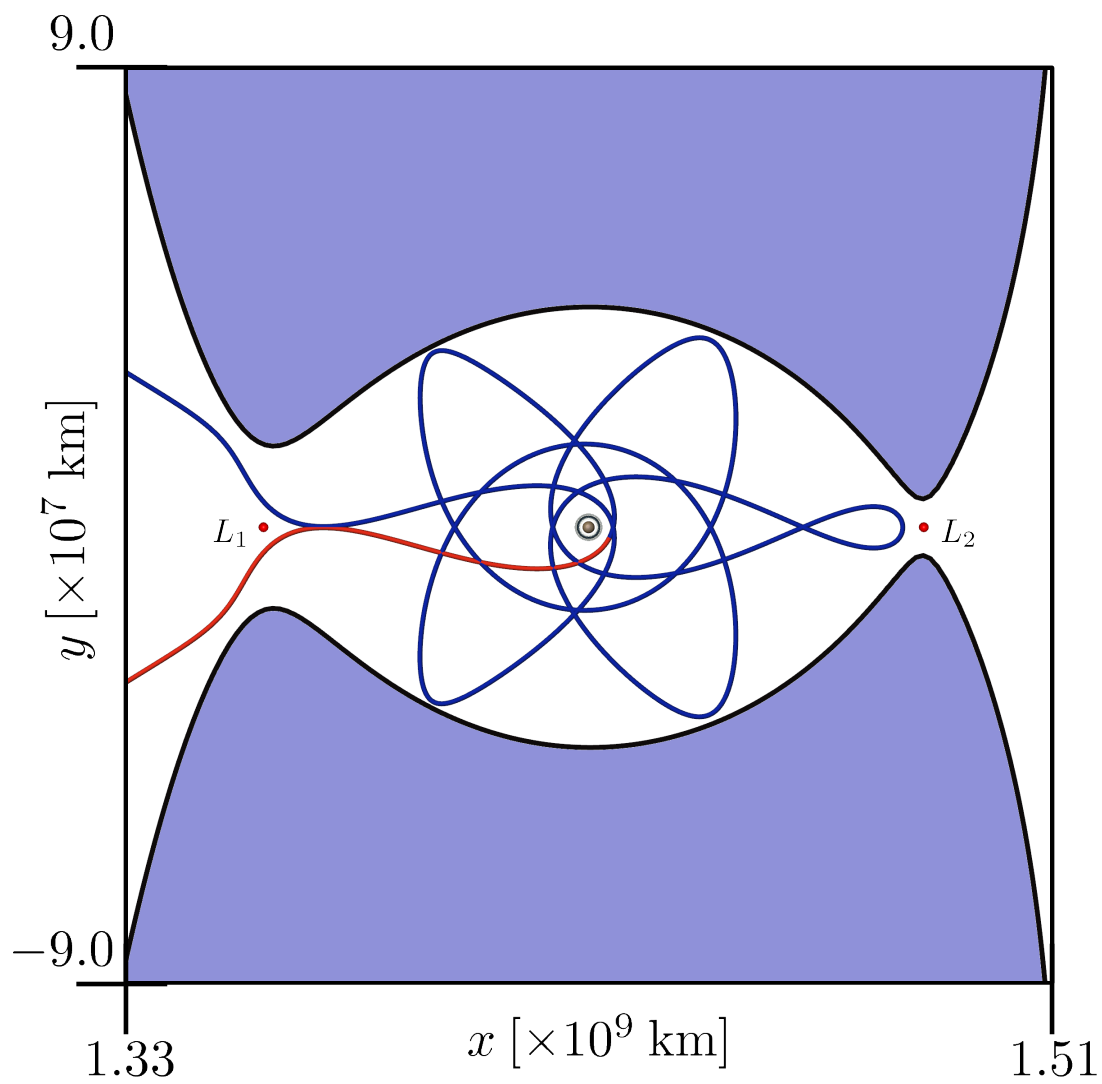


Figure 5.27. Zoom focused on Saturn region from Figure 5.26

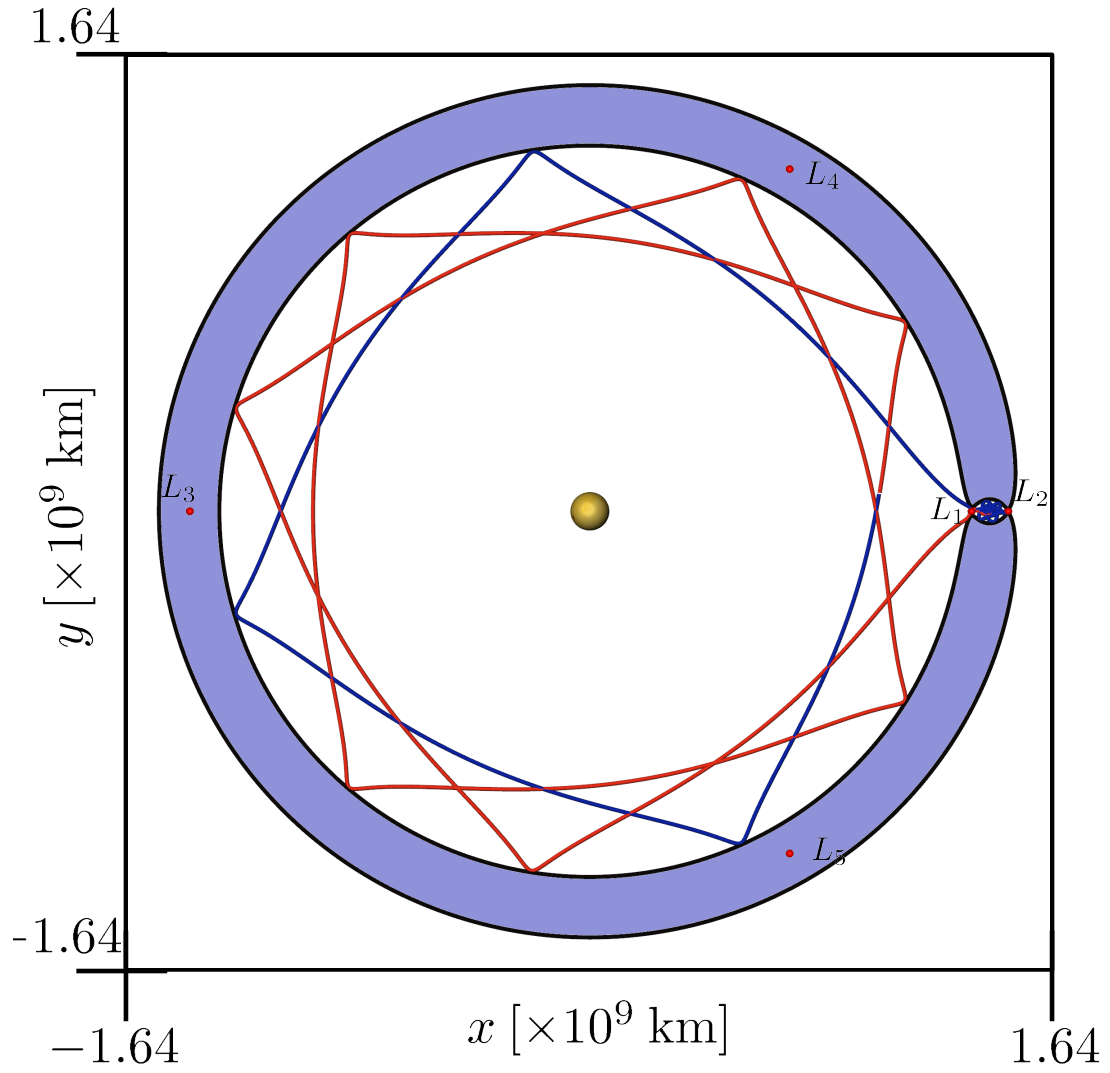


Figure 5.28. Trajectories from green point in bottom inset of Figure 5.25

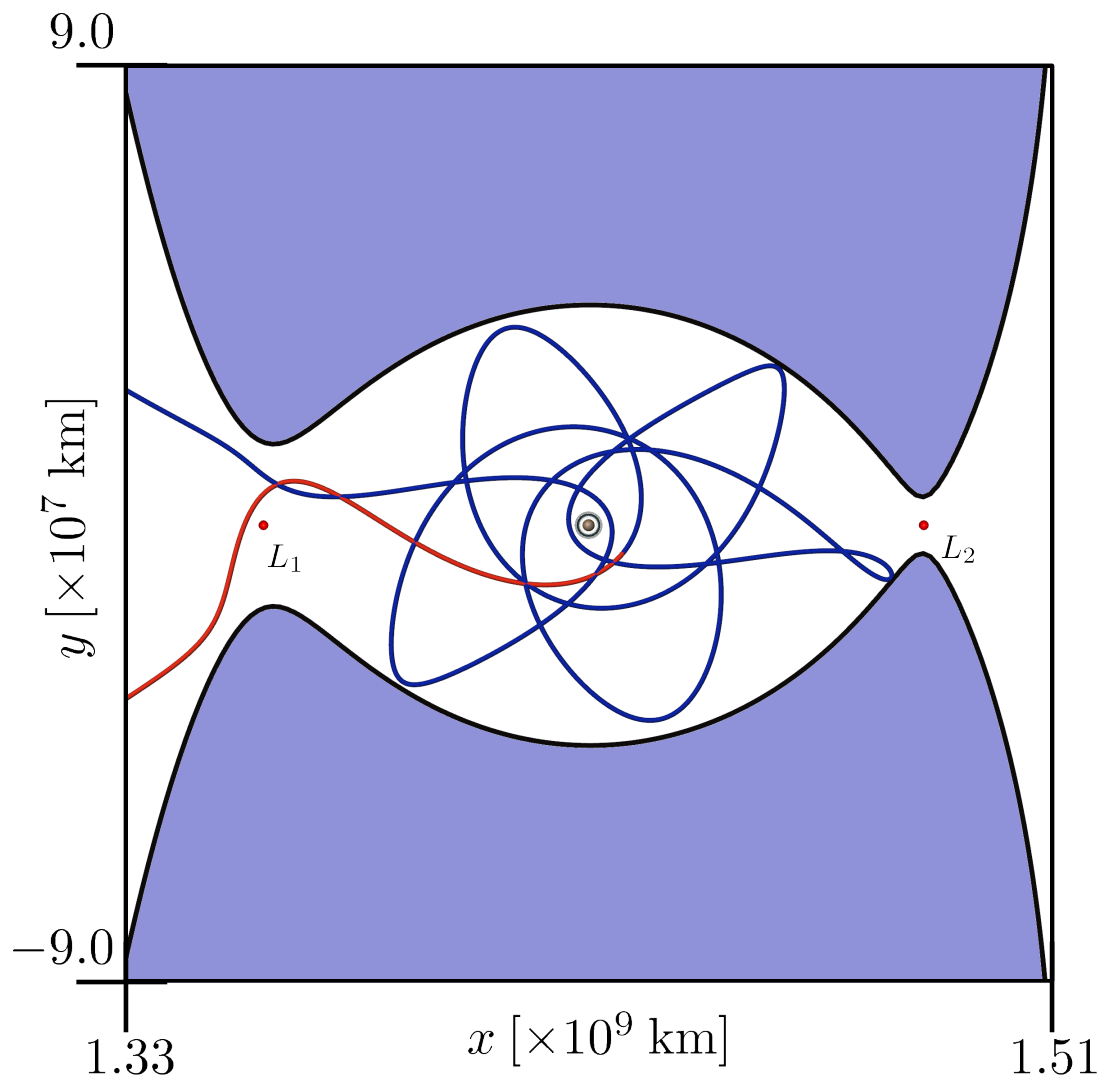


Figure 5.29. Zoom focused on Saturn region from Figure 5.28

5.2.2 Long-Term Trajectory Analysis in the Sun-Saturn System

Related to the analysis from Section 5.2.1 is an investigation of the long-term behavior near P_2 in the CR3BP. Specifically, the ongoing analysis of Davis and Howell [68,70] employs periapse Poincaré maps in the vicinity of Saturn to identify capture and escape orbits from the P_2 region. The Jacobi constant utilized here is slightly smaller than that employed in Section 5.2.1, namely, $C \approx 3.01730$, indicating that the ZVC are more widely open at L_2 .

A three-crossing FTLE map centered at Saturn is plotted in Figure 5.30. The coloring

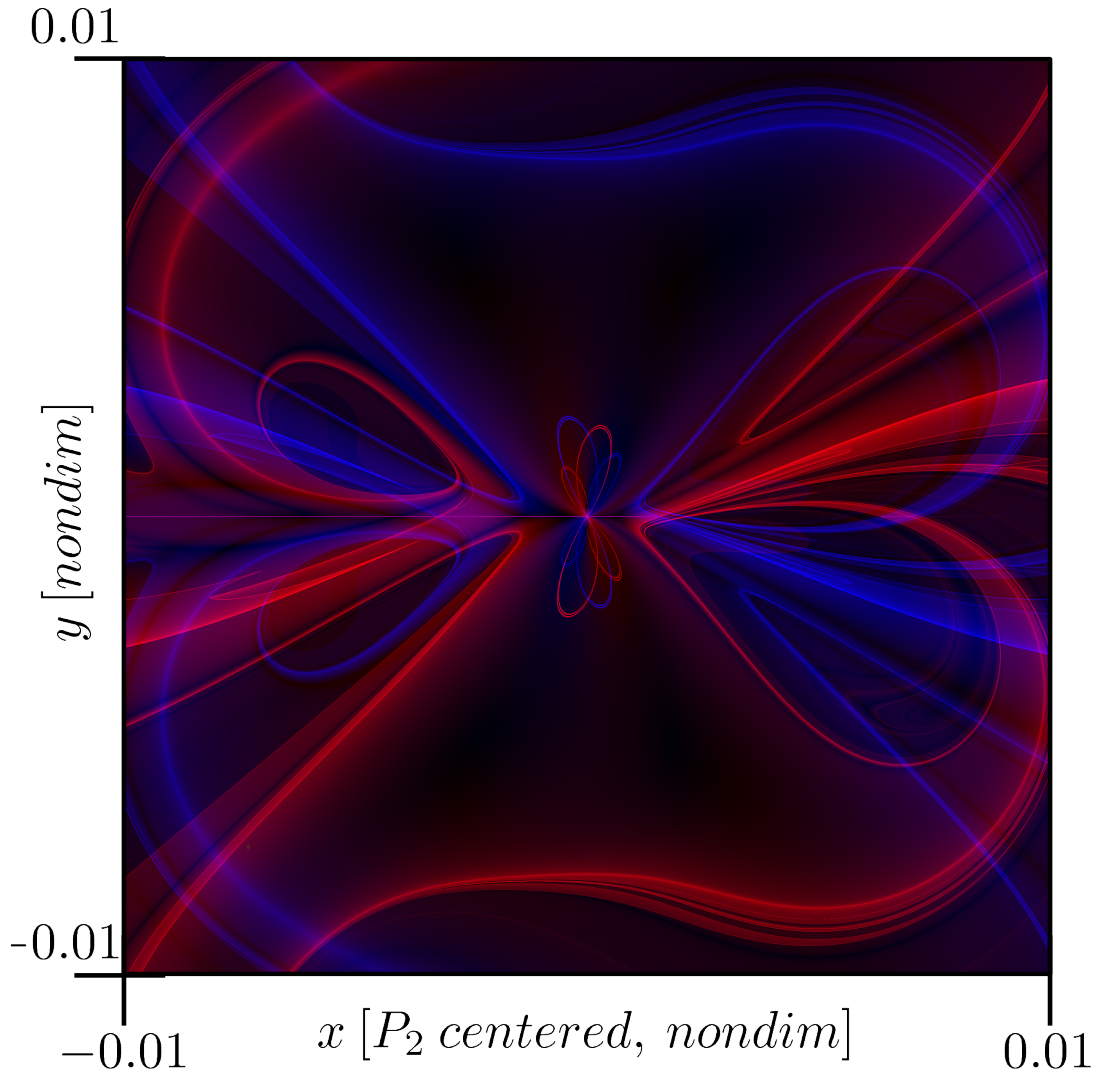


Figure 5.30. FTLE for $3\times$ of the periapse map in the vicinity of Saturn

in this figure is consistent with previous FTLE maps, i.e., forward FTLE is represented in blue and backward FTLE is colored red. Recall that LCS associated with forward FTLE correspond to stable manifold structures and unstable manifold structures are associated with backward FTLE¹. Immediately apparent in the figure is the region from the previous section where the first crossing of the unstable manifold from the orbit near L_1 appears as a red lobe on the right side of the figure generally below vertical center. Subsequent crossings of this same L_1 unstable manifold in red overlap a blue lobe above vertical center, also on the right side of the image. This blue lobe is the first stable manifold crossing associated with the L_1 Lyapunov orbit. Similarly, subsequent crossings of the stable manifold overlap the unstable lobe. On the left side of the figure similar, but smaller, lobes are apparent. These lobes are associated with the stable/unstable manifolds crossings associated with an L_2 Lyapunov orbit. Interesting features are observed in the form of sharp, bright loops of FTLE close to P_2 . All of these salient features appear, again, within three iterations of the periapse Poincaré map.

Since long-term behavior is a focus of related research efforts, a longer evolution is performed to better establish the associated structures and yield sharpened and more detailed features. This longer evolution involves ten crossings of the periapse map and appears in Figure 5.31. A similar periapse Poincaré map appears in Figure 5.32 with permission from Davis and Howell [70] (some simulation results to produce this map are from Haapala). In this figure, the same system parameters are employed to produce periapse crossings from stable manifolds associated with L_1 (red) and L_2 (blue) orbits. Note that the units in Figure 5.32 are Hill radii. Additionally, captured orbits are evolved over a significant time duration to produce the map crossings indicated in black – these are orbits that remain in the vicinity of Saturn for at least 1000 years. This map is included for general comparison. Specifically, in Figure 5.31, the “hourglass”² region in the center, similar to that defined in Figure 5.32, is apparent. This region of low relative FTLE values is indicative of uniform, organized motion and implies long-term stability. Also, note the similar narrow regions of stability on the x axis to the left of center termed the “arrowhead”² region for its distinctive shape on the map and the shape of the associated trajectories.

¹In the case of apse mapping, crossings actually occur at different times so the observed structures are not technically LCS, but they essentially supply the same qualitative information.

²as termed by Davis

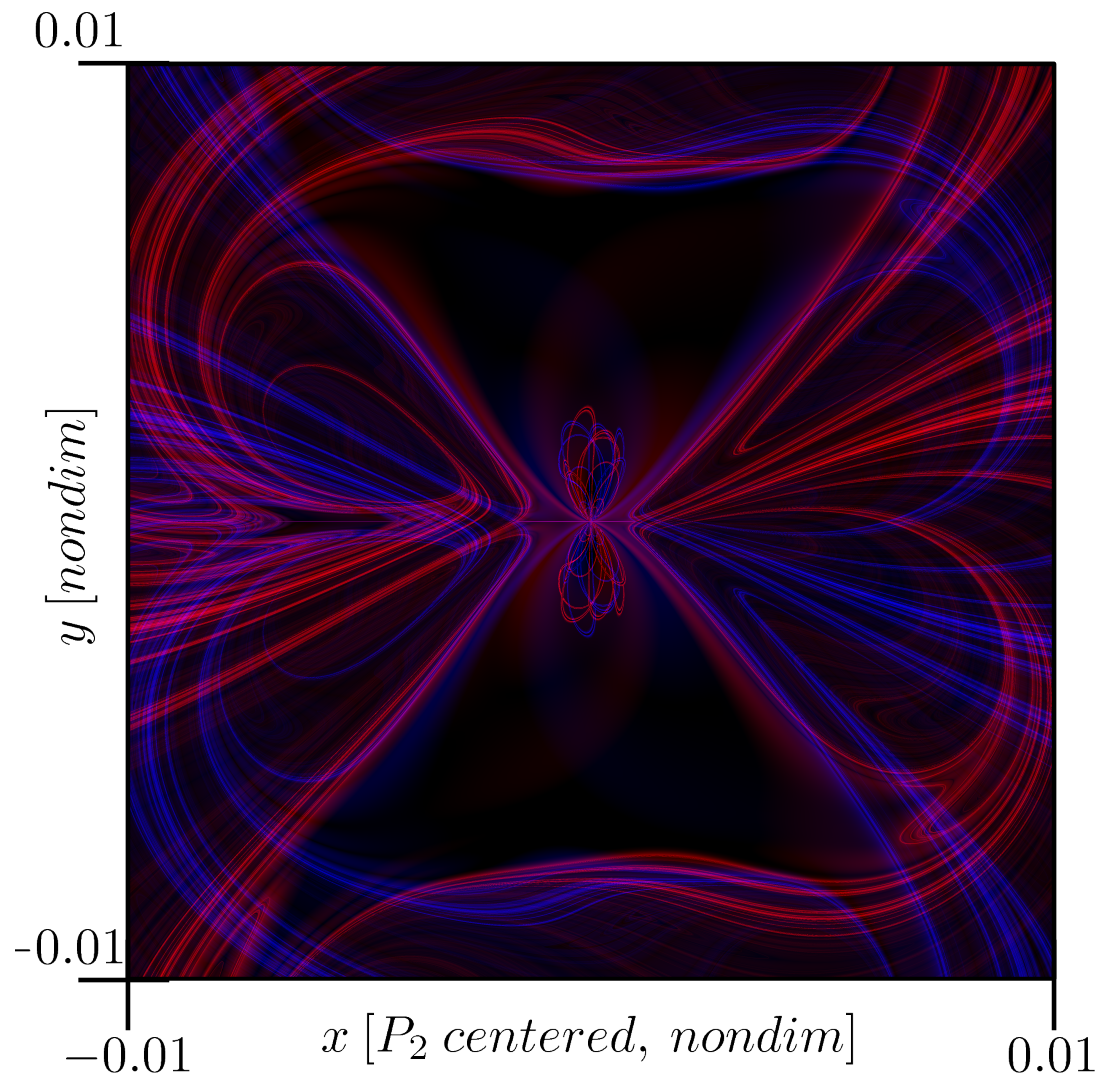


Figure 5.31. FTLE for $10\times$ of the periaapse map in the vicinity of Saturn

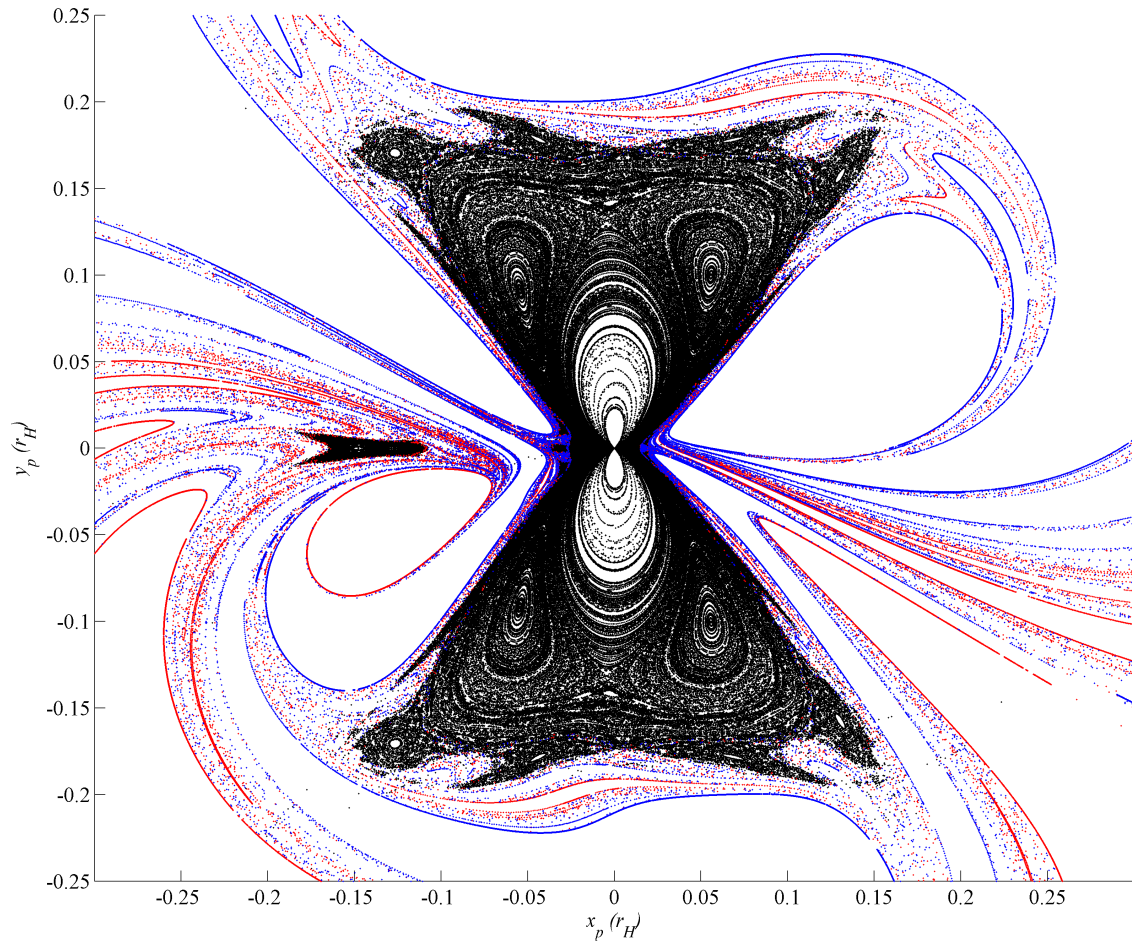


Figure 5.32. Periapse Poincaré map (with permission: D. C. Davis and K. C. Howell [70])

Beyond the general comparisons concerning the availability of similar information from both approaches, the sharp, bright structures appearing in Figures 5.30 and 5.31 near Saturn warrant investigation. A zoomed view of this region is provided in Figures 5.33 and 5.34, where focused views are offered for both the 3- and 10-crossing maps to indicate the evolution of the structures. Note that the actual diameters of Saturn and its rings are included, to scale, in the center of the figures. Several trajectories are investigated along the various

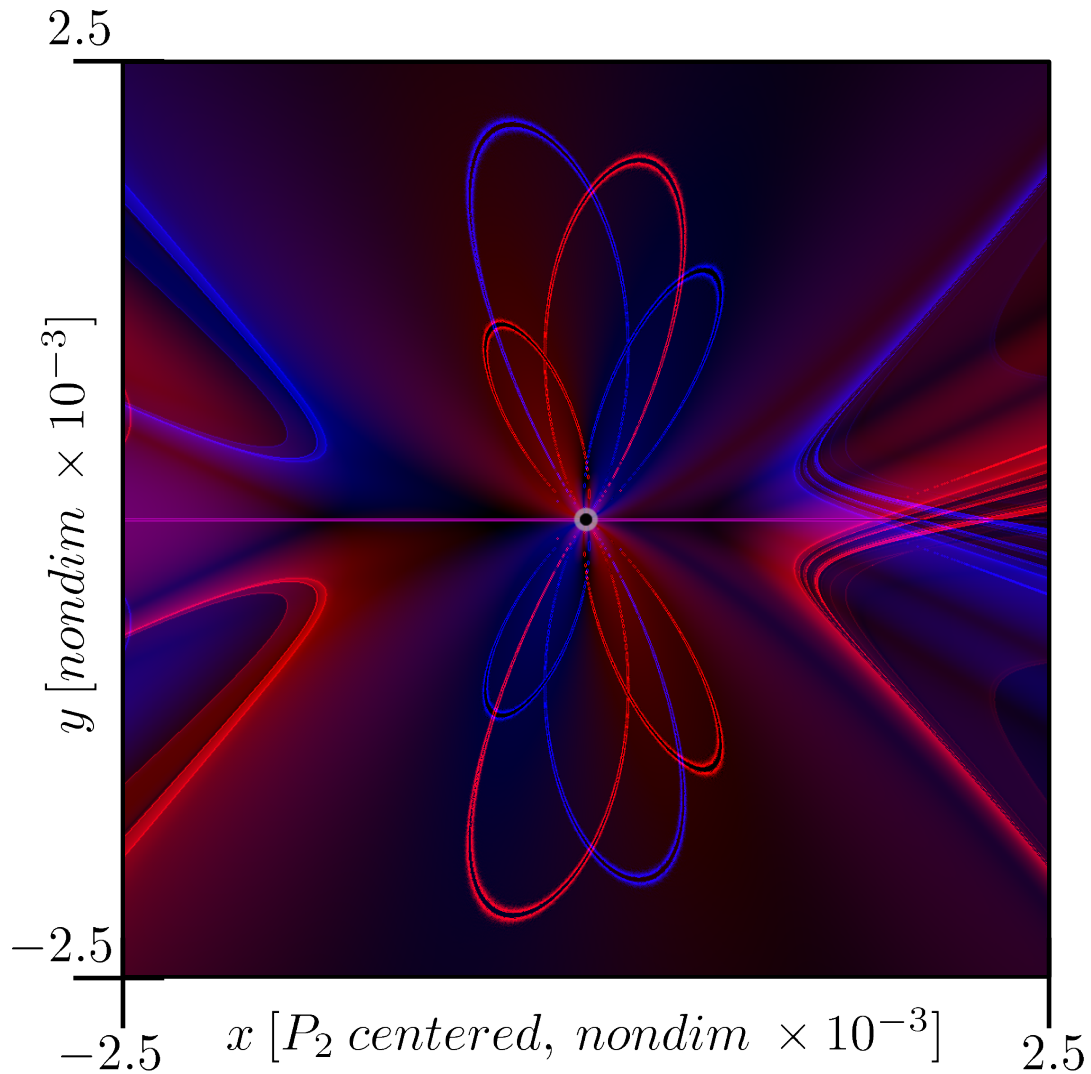


Figure 5.33. Zoom focused on distinctive structures of the 3-crossing FTLE map

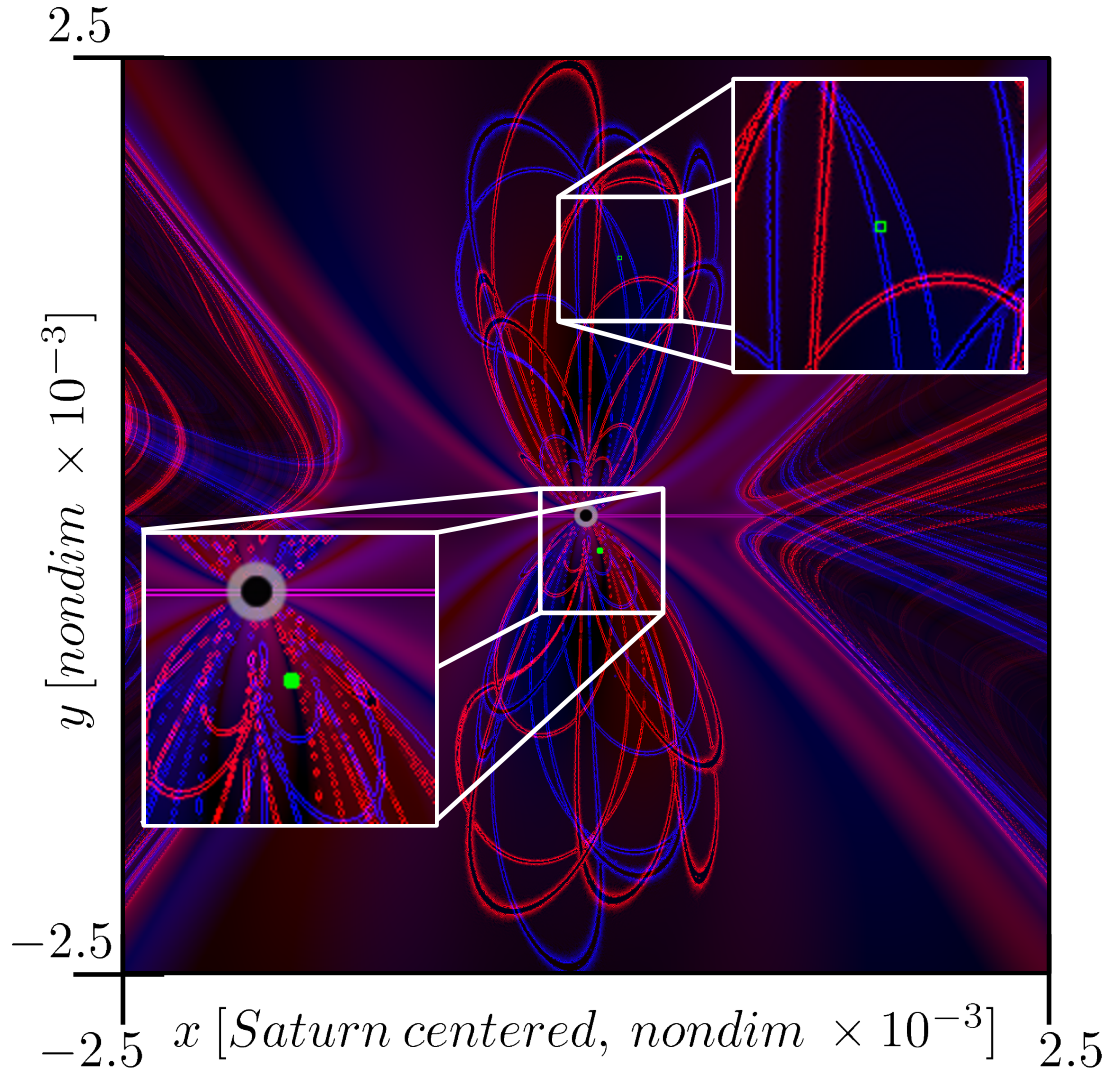


Figure 5.34. Zoom focused on distinctive structures of the 10-crossing FTLE map

structures. A few of these trajectories are specifically indicated in the upper inset zoom of Figure 5.34. Multiple states are selected and propagated within the indicated green box. The resulting information reveals two main points: (1) the structures result from numerical issues associated with proximity to Saturn and (2) while these structures do not represent LCS, they supply information regarding the fate of trajectories on or near the structures. Specifically, given an initial state along one of the features, subsequent evolution frequently

leads to impact with Saturn. Some illustration appears in Figure 5.35, where selected trajectories are plotted (note, unlike the zoomed FTLE images, the Sun and Saturn are scaled $100\times$ and $20\times$, respectively). Specifically, in Figure 5.35, a comparison trajectory (black) from the darker FTLE region near the structure indicated in the upper inset of Figure 5.34 is plotted along with an adjacent trajectory (green, originating on the ridge) that experiences extremely different behavior; this difference of behavior ultimately supplies the high FTLE value observed in the map. It is indicated that the trajectories pass within the radius of Saturn and, given the singular nature of the primaries in the CR3BP, continue to evolve after the “impact”. As the second (green) trajectory continues, it is apparent that the Jacobi integral is not conserved. This is perceivable as the trajectory violates the ZVC. Another state is selected centered at the *solid* green dot¹ in the lower inset, much closer to Saturn, in Figure 5.34. This state is evolved forward and backward in time for 9 crossings of the periaapse Poincaré map (~ 31 years in both forward and backward time, or ~ 62 years in total), and the resulting trajectories appear in Figures 5.36 and 5.37 where Saturn is now drawn to scale. Note that these trajectories, originating from a black region amidst the tangle of impacts, do not impact Saturn over the length of the integration and C is held constant to the order of 10^{-10} . Despite the ultimate decrease in numerical accuracy in close proximity to P_2 , the effectiveness of the the FTLE map to indicate subsequent collisions (or numerical difficulty), and perhaps more importantly, indicate regions near Saturn that do not impact within the duration of the simulation, is clear.

¹ $[x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z}]^T = [0.999797 \ -0.000217 \ 0.0 \ 1.456477 \ 0.560183 \ 0.0]^T$ (nondim) =
 $[1.426436 \ -0.000310 \ 0.0 \ 14.049186 \ 5.403533 \ 0.0]^T$ ($\times 10^9$ km | $\frac{\text{km}}{\text{s}}$)

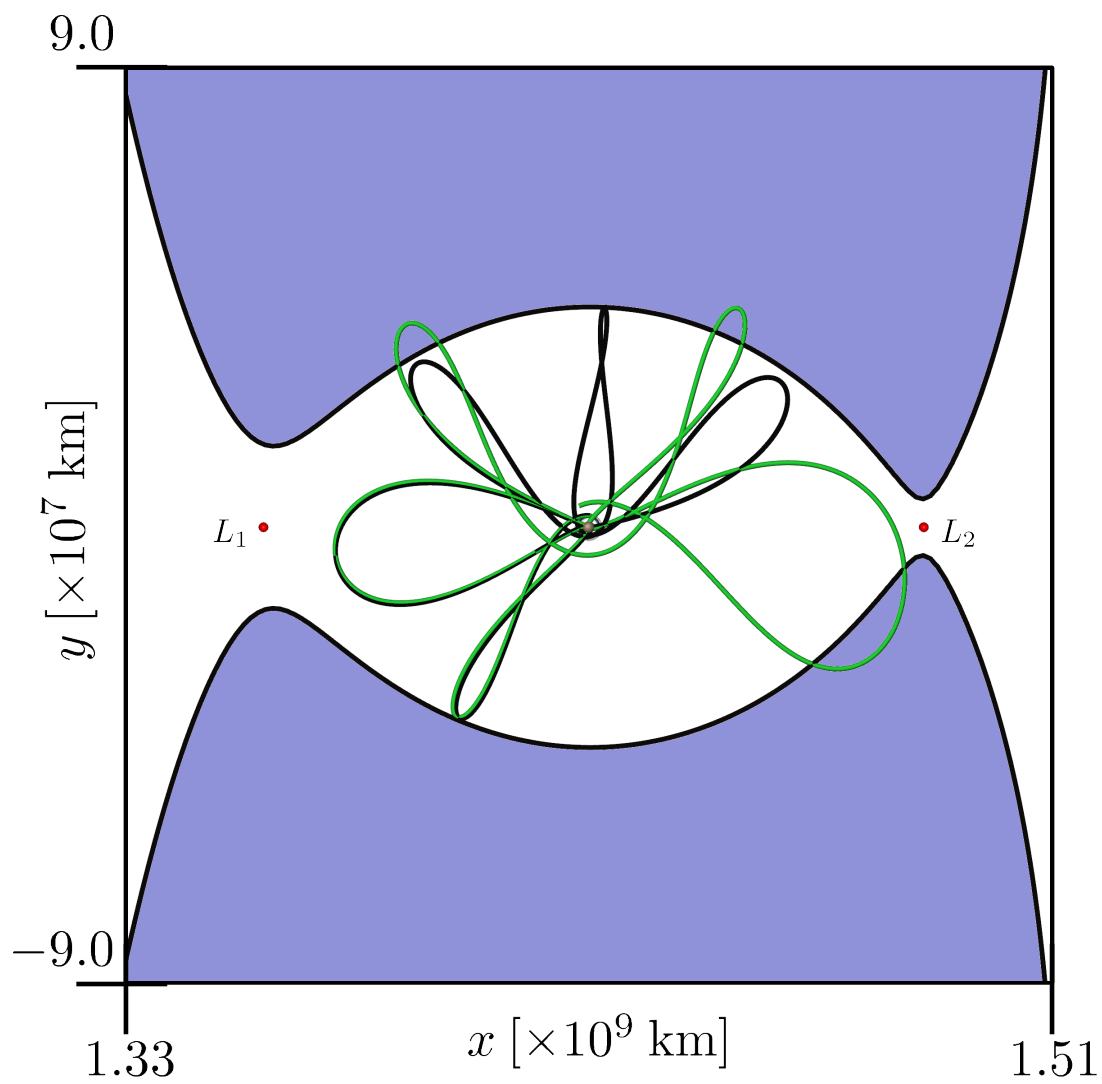


Figure 5.35. Wide numerical divergence induced by close proximity Saturn passages

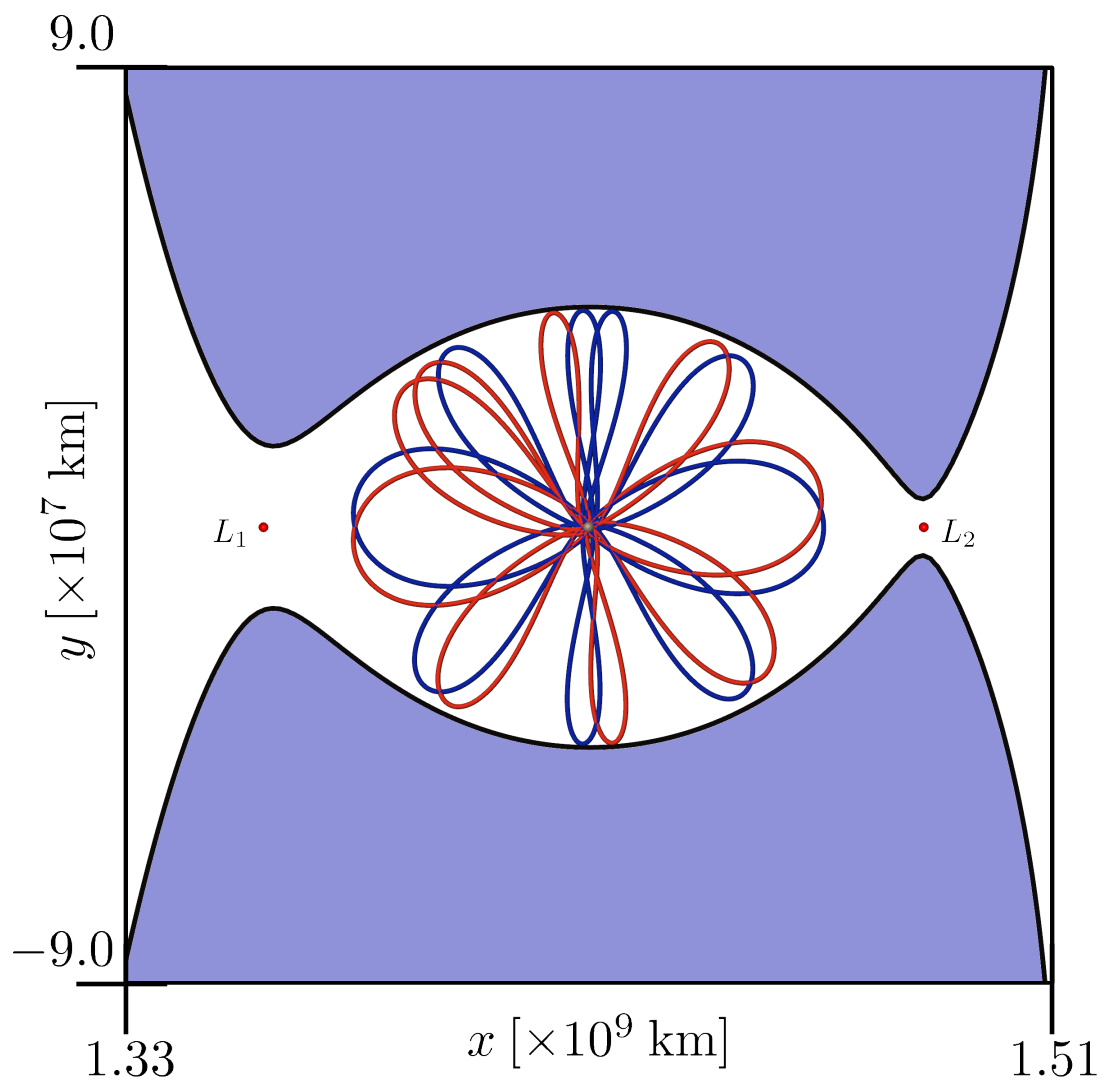


Figure 5.36. Trajectory from 9-crossing stable region indicated in Figure 5.34

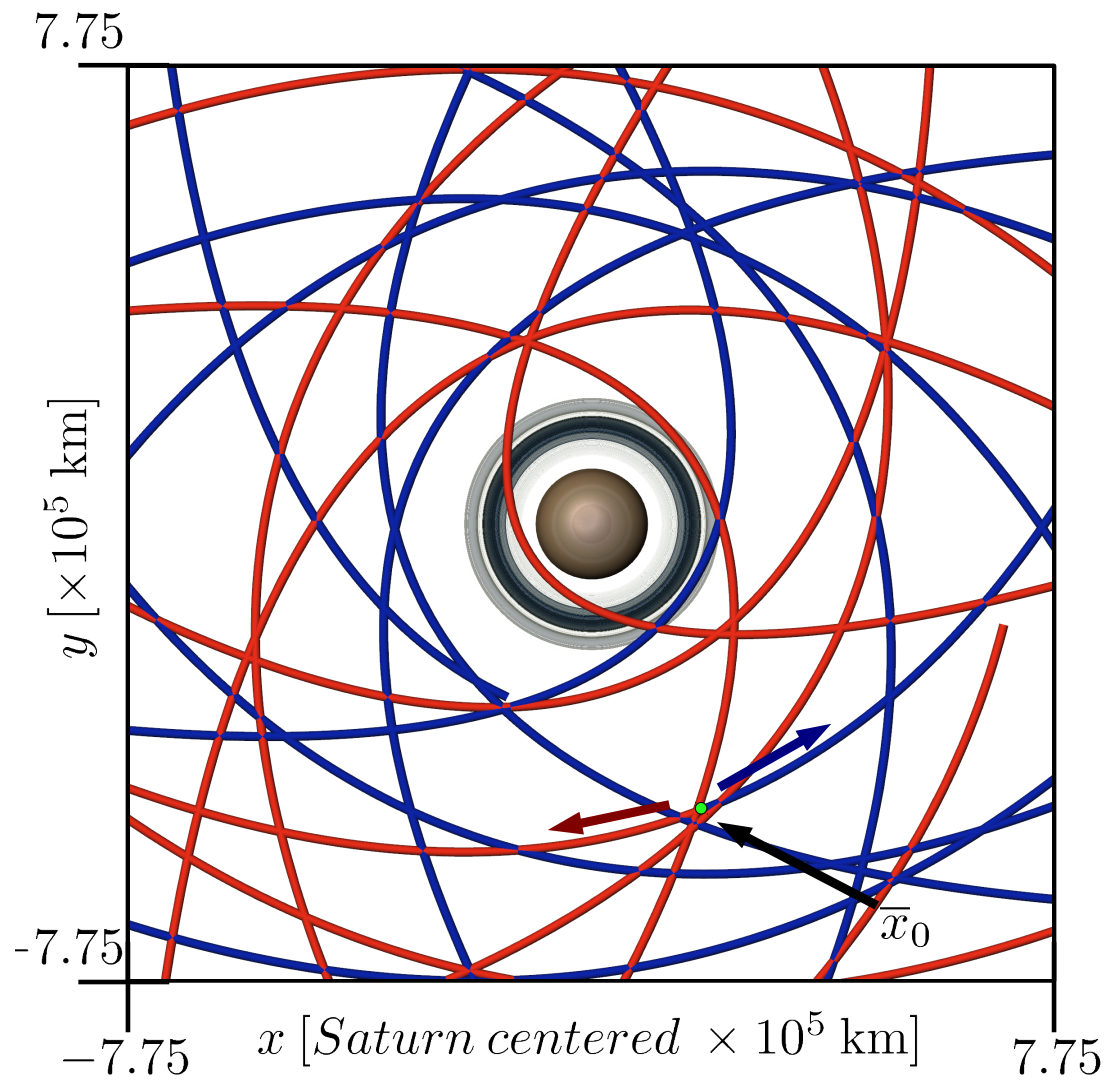


Figure 5.37. Saturn focused view of Figure 5.36

6. CONCLUSION

Lagrangian coherent structures represent a relatively untapped approach for revealing insight in mission design and trajectory analysis. In this chapter a summary of the investigation is provided. Additionally, the potential for further application of these concepts and techniques is elaborated. Finally, concluding remarks are provided.

6.1 Summary

The dual purpose of this analysis is to apply the fairly nascent concept of LCS to the CR3BP and to gather and organize the framework associated with these ideas. As such, the preliminary chapters establish the dynamical context and the ideas associated with Lagrangian coherent structures. Specifically, the model development of the circular restricted three-body problem is presented within the context of the general n -body problem and the general three-body problem. Dynamical systems concepts exhibited by the autonomous CR3BP are presented and summarized. Subsequently, additional concepts leading to the formulation of the characteristic Lyapunov exponent, its truncated version as the finite-time Lyapunov exponent, and foundations for Lagrangian coherent structures are introduced. General computational and computer visualization methods are described providing the means to realize the primary goal.

To establish the applicability of the described concepts, results are presented in two solar system three-body configurations, namely the Earth-Moon CR3BP and the Sun-Saturn CR3BP. Results in the Earth-Moon system that are associated with a specific periodic orbit as well as a sample FTLE exposition for a specific EM region with consideration for potential mission design applications contribute to an comprehensive overview of the concepts. Sun-Saturn results correlate with ongoing work by Haapala [69] and Davis and Howell [68, 70]. These results highlight the advantages of the application of FTLE and LCS in mission design investigations in conjunction with other strategies including general Poincaré mapping.

6.2 Future Work

Given the potential represented by Lagrangian coherent structures as applied to astrodynamics, there is large set of possibilities for investigation and application of the concepts compiled here. In fact, an ancillary goal of this analysis is to establish a foundation for the expansion of the methods to more detailed models of multi-body motion. However, additional investigation is possible in the planar CR3BP including expanded analysis near the triangular points. Further planar inspections introducing perturbing bodies, such as a four-body formulation, supplies a non-autonomous option while still exhibiting convenient constraints for FTLE definition on surfaces of section. These planar applications represent one of many possible research directions.

The application of FTLE and LCS methods to problems involving motion free to move in all spatial dimensions is a compelling possibility. Poincaré mapping is not generally amenable to spatial three-body analysis given the lack of sufficient constraints to define a map. While, FTLE information is conveniently displayed on a Poincaré section, it can be defined anywhere adjacent trajectories evolve. That is, the FTLE definition supplies the ability to take analysis efforts directly into three dimensions effectively. Specifically, extension of these methods to a dynamical regime focused more exclusively in three dimensions, such as the spatial CR3BP, is a possibility. This three-dimensional augmentation of the analysis involves added complexities as well as a natural extension to higher fidelity models including combinations of multi-body ephemerides.

6.3 Concluding Remarks

While LCS continue to draw significant attention as a new focus for research in general and their applicability has been demonstrated in various problems, applications to astrodynamical work has been limited. Therefore, a comprehensive development of the associated theory with respect to orbital mechanics and trajectory design is timely. The applications in this analysis serve as a foundation to establish the relevance of Lagrangian coherent structures in a mission design and analysis context and, in turn, expand their use.

LIST OF REFERENCES

LIST OF REFERENCES

- [1] G. Galilei, *Sidereus Nuncius (1610)*. Chicago: The University of Chicago Press, 1989. Translated, A. Van Helden.
- [2] I. Newton, *The Mathematical Principles of Natural Philosophy (1687)*. New York: Prometheus Books, 1995. Translated, A. Motte.
- [3] J. Kepler, *New Astronomy (1609)*. Cambridge: Cambridge University Press, 1992. Translated, W. H. Donahue.
- [4] S. Drake, *Discoveries and Opinions of Galileo*. New York: Anchor Books, 1957.
- [5] J. Barrow-Green, *Poincaré and the Three Body Problem*. New York: Cambridge University Press, 1997.
- [6] L. Euler, *Theoria Motuum Lunae*. Petropoli: Typis Academiae Imperialis Scientiarum, 1772. Reprinted in “Opera Omnia, Series 2” (L. Courvoisier, ed.), Vol. 22, Orell Füssli Turici, Lausanne, 1958.
- [7] V. Szebehely, *Theory of Orbits: The Restricted Problem of Three Bodies*. New York: Academic Press, 1967.
- [8] H. Poincaré, *New Methods of Celestial Mechanics (1893)*. New York: American Institute of Physics, 1993. Translated, D. Goroff.
- [9] K. F. Sundman, “Mémoire Sur le Problème des Trois Corps,” *Acta Mathematica*, vol. 36, pp. 105–179, 1912.
- [10] F. Diacu, “The Solution of the n -body Problem,” *The Mathematical Intelligencer*, vol. 18, no. 3, pp. 66–70, 1996.
- [11] Q. Wang, “The global solution of the n -body problem,” *Celestial Mechanics*, vol. 50, pp. 73–88, 1991.
- [12] F. R. Moulton, *Periodic Orbits*. Washington: The Carnegie Institution, 1920. Publication No. 161.
- [13] K. C. Howell, “Families of Orbits in the Vicinity of the Collinear Libration Points,” *Journal of the Astronautical Sciences*, vol. 49, no. 1, pp. 107–125, 2001.
- [14] M. Hénon, “Numerical Exploration of the Restricted Three-Body Problem,” in *Proceedings of International Astronomical Union Symposium No. 25*, (London), pp. 157–169, Academic Press, 1966.
- [15] J. F. Kennedy, “John F. Kennedy Moon Speech.” [Online Publication, retrieved 10 Nov 2010], <http://er.jsc.nasa.gov/seh/ricetalk.htm>, 1962.
- [16] W. E. Burrows, *This New Ocean: The Story of the First Space Age*. New York: Random House, Inc., 1998.

- [17] J. V. Breakwell, A. A. Kamel, and M. J. Ratner, "Station-Keeping for a Translunar Communications Station," *Celestial Mechanics*, vol. 10, no. 3, pp. 357–373, 1974.
- [18] R. Farquhar, "The Utilization of Halo Orbits in Advanced Lunar Operations," Tech. Rep. X-551-70-449, NASA Goddard, Greenbelt, MD, 1970.
- [19] K. C. Howell and J. V. Breakwell, "Almost Rectilinear Halo Orbits," *Celestial Mechanics*, vol. 32, pp. 29–52, January 1984.
- [20] K. C. Howell, "Three-Dimensional Periodic 'Halo' Orbits," *Celestial Mechanics*, vol. 32, pp. 53–72, January 1984.
- [21] R. W. Farquhar, "The Flight of ISEE-3/ICE: Origins, Mission History, and a Legacy," *Journal of the Astronautical Sciences*, vol. 49, pp. 23–73, January-March, 2001.
- [22] M. Lo, B. G. Williams, W. E. Bollman, D. Han, Y. Hahn, J. L. Bell, E. A. Hirst, R. A. Corwin, P. E. Hong, K. C. Howell, B. Barden, and R. Wilson, "Genesis Mission Design," in *AIAA/AAS Astrodynamics Specialist Conference*, (Boston, Massachusetts), August 10-12, 1998.
- [23] T. S. Brown, "Multi-Body Mission Design in the Saturnian System with Emphasis on Enceladus Accessibility," M.S. Thesis, Purdue University, West Lafayette, Indiana, 2008.
- [24] D. C. Folta, T. A. Pavlak, K. C. Howell, M. A. Woodard, and D. W. Woodfork, "Stationkeeping of Lissajous Trajectories in the Earth-Moon System with Applications to ARTEMIS," in *20th AAS/AIAA Space Flight Mechanics Meeting*, (San Diego, California), February, 2010. Paper No. AAS 10-113.
- [25] D. C. Folta, M. Woodard, K. Howell, C. Patterson, and W. Schlei, "Applications of Multi-Body Dynamical Environments: The ARTEMIS Transfer Trajectory Design," in *IAF 61st International Astronautical Congress*, (Prague, Czech Republic), September, 2010.
- [26] T. Peacock and J. Dabin, "Introduction to focus issue: Lagrangian coherent structures," *Chaos*, vol. 20, pp. 017501–1–3, 2010.
- [27] J. P. Richter, *The Literary Works of Leonardo da Vinci*. London: Phaidon Press Ltd., 1970.
- [28] G. Haller, "Finding finite-time invariant manifolds in two-dimensional velocity fields," *Chaos*, vol. 10, no. 1, pp. 99–108, 2000.
- [29] G. Haller, "Distinguished material surfaces and coherent structures in three-dimensional fluid flows," *Physica D*, vol. 149, pp. 248–277, 2001.
- [30] G. Haller and G. Yuan, "Lagrangian coherent structures and mixing in two-dimensional turbulence," *Physica D*, vol. 147, pp. 352–370, 2000.
- [31] G. Haller, "Lagrangian coherent structures from approximate velocity data," *Physics of Fluids*, vol. 14, no. 6, pp. 1851–1861, 2002.
- [32] S. C. Shadden, F. Lekien, and J. E. Marsden, "Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows," *Physica D*, vol. 212, pp. 271–304, 2005.

- [33] M. Mathur, G. Haller, T. Peacock, J. Ruppert-Felsot, and H. Swinney, "Uncovering the Lagrangian Skeleton of Turbulence," *Physical Review Letters*, vol. 98, pp. 144502–1–4, 2007.
- [34] F. Lekien, S. C. Shadden, and J. E. Marsden, "Lagrangian coherent structures in n-dimensional systems," *Journal of Mathematical Physics*, vol. 48, pp. 065404–1–19, 2007.
- [35] C. Garth, F. Gerhardt, X. Tricoche, and H. Hagen, "Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13(6), pp. 1464–1471, 2007.
- [36] B. F. Villac, "Using FLI maps for preliminary spacecraft trajectory design in multi-body environments," *Celestial Mechanics and Dynamical Astronomy*, vol. 102, no. 1-3, pp. 29–48, 2008.
- [37] B. Villac and S. Broschart, "Applications of Chaoticity Indicators to Stability Analysis Around Small Bodies," in *19th AAS/AIAA Space Flight Mechanics Meeting*, (Savannah, Georgia), February 9-12, 2009. Paper No. AAS 09-221.
- [38] E. S. Gawlik, J. E. Marsden, P. C. du Toit, and S. Campagnola, "Lagrangian coherent structures in the planar elliptic restricted three-body problem," *Celestial Mechanics and Dynamical Astronomy*, vol. 103, no. 3, pp. 227–249, 2009.
- [39] D. T. Greenwood, *Principles of Dynamics*. Upper Saddle River, New Jersey: Prentice Hall, second ed., 1988.
- [40] A. M. Lyapunov, *The General Problem of the Stability of Motion (1892)*. Bristol, Pennsylvania: Taylor and Francis, 1992. Trans. Russian to French by Édouard Davaux; French to English by A. T. Fuller.
- [41] A. E. Roy and M. W. Ovenden, "On the Occurrence of Commensurable Mean Motions in the Solar System: II. The Mirror Theorem," *Monthly Notices of the Royal Astronomical Society*, vol. 115, 1955.
- [42] A. Bennett, *Lagrangian Fluid Dynamics*. Providence, Rhode Island: American Mathematical Society, 1997.
- [43] K. P. Burr, "13.021 Marine Hydrodynamics Lecture 2." [Online Publication, retrieved 21 Jul 2010], http://web.mit.edu/fluids-modules/www/potential_flows/LecturesHTML/lec02/lecture2.html, 2003.
- [44] V. I. Oseledec, "A Multiplicative Ergodic Theorem. Ljapunov Characteristic Numbers for Dynamical Systems," *Transactions of the Moscow Mathematical Society*, vol. 19, pp. 197–231, 1968.
- [45] R. L. Anderson, *Low Thrust Trajectory Design for Resonance Flybys and Captures Using Invariant Manifolds*. Ph.D. Thesis, University of Colorado at Boulder, Boulder, Colorado, 2005.
- [46] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*. New York: Springer-Verlag, 1989.
- [47] D. Scheeres, D. Han, and Y. Hou, "Influence of Unstable Manifolds on Orbit Uncertainty," *Journal of Guidance Control and Dynamics*, vol. 24, no. 3, pp. 573–585, 2001.

- [48] R. L. Anderson, M. W. Lo, and G. H. Born, "Application of Local Lyapunov Exponents to Maneuver Design and Navigation in the Three-Body Problem," in *AAS/AIAA Astrodynamics Specialist Conference*, (Big Sky, Montana), August 3-7, 2003. Paper No. AAS 03-569.
- [49] H. D. I. Abarbanel, R. Brown, and M. B. Kennel, "Variation of Lyapunov Exponents on a Strange Attractor," *Journal of Nonlinear Science*, vol. 1, pp. 175–199, 1991.
- [50] S. C. Shadden, "Lagrangian Coherent Structures: Analysis of time-dependent dynamical systems using finite-time Lyapunov exponents." [Online Publication, retrieved 27 May 2009], <http://www.cds.caltech.edu/~shawn/LCS-tutorial/>, 2005.
- [51] G. Katsiaris and C. L. Goudas, "The Symplectic Property of Matrizants," *Astrophysics and Space Science*, vol. 8, pp. 243–250, 1970.
- [52] B. G. Marchand, "Temporary Satellite Capture of Short Period Jupiter Family Comets from the Perspective of Dynamical Systems," M.S. Thesis, Purdue University, West Lafayette, Indiana, 2000.
- [53] D. Eberly, R. Gardner, B. Morse, and S. Pizer, "Ridges for Image Analysis," *Journal of Mathematical Imaging and Vision*, vol. 4, pp. 355–371, 1994.
- [54] C. Garth. Personal communication, 2010.
- [55] "The GNU Scientific Library Manual." [Online Publication, retrieved 24 Aug 2010], http://www.gnu.org/software/gsl/manual/html_node/index.html, 2010.
- [56] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, New Jersey: Pearson Prentice Hall, third ed., 2008.
- [57] Z. P. Olikara, "Computation of Quasi-Periodic Tori in the Circular Restricted Three-Body Problem," M.S. Thesis, Purdue University, West Lafayette, Indiana, 2010.
- [58] M. Nixon and A. Aguado, *Feature Extraction & Image Processing*. Oxford: Academic Press, second ed., 2008.
- [59] R. van den Boomgaard and R. van der Weij, "Gaussian Convolutions. Numerical Approximations Based on Interpolation," in *Scale-Space '01: Proceedings of the Third International Conference on Scale-Space and Morphology in Computer Vision*, (London, UK), pp. 205–214, Springer-Verlag, 2001.
- [60] L. Yan, J. Ma, W. Wang, Q. Liu, and Q. Zhou, "A Novel Approach to Ocular Image Enhancement with Diffusion and Parallel AOS Algorithm," in *Advances in Intelligent Computing: International Conference on Intelligent Computing, ICIC 2005, Proceedings, Part 1*, (Berlin, DE), pp. 1–9, Springer-Verlag, 2005.
- [61] "Lecture 11: Differential Geometry." [Online Publication, retrieved 13 Oct 2010], http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/diffgeom.pdf, 2000.
- [62] "The OpenMP API Specification for Parallel Programming." [Online Publication, retrieved 4 Oct 2010], <http://openmp.org/wp/about-openmp/>, 2010.
- [63] "NVIDIA® CUDA™ Architecture, Version 1.1." [Online Publication, retrieved 8 Nov 2010], http://developer.download.nvidia.com/compute/cuda/docs/CUDA_Architecture_Overview.pdf, April, 2009.

- [64] “Teem: Tools to process and visualize scientific data and images.” [Online Publication, retrieved 12 Oct 2010], <http://teem.sourceforge.net/>, 2010.
- [65] “VTK: The Visualization Toolkit.” [Online Publication, retrieved 12 Oct 2010], <http://www.vtk.org/>, 2010.
- [66] “Qt: A cross-platform application and UI framework.” [Online Publication, retrieved 12 Oct 2010], <http://qt.nokia.com/products>, 2010.
- [67] D. J. Grebow, “Generating Periodic Orbits in the Circular Restricted Three-Body Problem with Applications to Lunar South Pole Coverage,” M.S. Thesis, Purdue University, West Lafayette, Indiana, 2006.
- [68] D. C. Davis and K. C. Howell, “Trajectory Evolution in the Multi-Body Problem with Applications in the Saturnian System,” in *IAF 61st International Astronautical Congress*, (Prague, Czech Republic), September, 2010.
- [69] A. F. Haapala, “Trajectory Design Using Periapse Maps and Invariant Manifolds,” M.S. Thesis, Purdue University, West Lafayette, Indiana, 2010.
- [70] D. C. Davis and K. C. Howell, “Long-Term Evolution of Trajectories Near the Smaller Primary in the Restricted Problem,” in *AAS/AIAA Astrodynamics Specialist Conference*, (San Diego, California), February, 2010. Paper No. AAS 10-184.
- [71] S. Wiggins, L. Wiesenfeld, C. Jaffé, and T. Uzer, “Impenetrable Barriers in the Phase-Space,” *Physical Review Letters*, vol. 86, pp. 5478–5481, 2001.