# Cylindrical Manifolds and Tube Dynamics in the Restricted Three-Body Problem

Thesis by
Shane David Ross

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2004

(Defended April 7, 2004)

*The heavens declare the glory of God;*
*And the firmament shows His handiwork.*

*Psalm 19, of David*

# Preface

This thesis was submitted at the *California Institute of Technology* on March 23rd, 2004, as a partial fulfillment of the requirements for the degree of Doctor of Philosophy in Control and Dynamical Systems.

The thesis is intended to be a complete research report and an effort was made to maintain a coherent structure, but Chapter 6 was kept independent of the others in order to facilitate its publication. Certain chapters have been published or will be submitted for publication in slightly modified form:

- Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [2000] Heteroclinic connections between periodic orbits and resonance transitions in celestial mechanics. *Chaos* **10**(2), 427–469.

- Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [2004] *Dynamical Systems, the Three-Body Problem, and Space Mission Design*, in preparation.

- Ross, S. D. [2003] Statistical theory of interior-exterior transition and collision probabilities for minor bodies in the solar system, in *Libration Point Orbits and Applications* (G. Gomez, M.W. Lo and J.J. Masdemont, eds.), World Scientific, 2003, pp. 637-652.

# Acknowledgements

Firstly, I would like to thank Martin Lo, who first introduced me to dynamical systems theory and its application to celestial mechanics. He has been a good teacher and friend since the days of the SURF project nearly a decade ago when invariant manifolds and Smale horseshoes were but a glimmer in my young undergraduate mind.

I would also like to thank my adviser, Jerry Marsden, who has been a joy to learn from and has opened my eyes to the geometric way of viewing the mechanical world. Another "unofficial" adviser I want to thank is Wang Sang Koon for his friendship and assistance over the years.

I thank my other collaborators for being an integral part of my education: Michael Dellnitz, David Farrelly, Frederic Gabern, Gerard Gómez, Charles Jaffé, Oliver Junge, Francois Lekien, Josep Masdemont, Kathrin Padberg, Linda Petzold, Robert Preis, Radu Serban, Dan Scheeres, Bianca Thiere, Turgay Uzer, and Roby Wilson. I am honored to be numbered among such colleagues. I wish to extend a special thanks to Drs. Gómez and Masdemont for their early involvement in my three-body problem research, by engaging in helpful discussions and sharing their wonderful exploratory software tools.

I also wish to thank the following colleagues for helpful discussions and comments: Kyle Alfriend, Brian Barden, Julia Bell, Dong Eui Chang, Peter Goldreich, Kathleen Howell, Àngel Jorba, Andrew Lange, Melvin Leok, Belinda Marchand, Mark Milam, Richard Murray, Marci Paskowitz, Shawn Shadden, Carles Simó, Scott Tremaine, Benjamin Villac, and Stephen Wiggins.

Finally, I thank my parents, L. C. and Diane Ross, and sister, Natalie, for their constant support of my endeavors, and my wonderful fiancée Jessica Lozoya, for her support during the final leg of this race. This thesis is dedicated to "y'all."

# Cylindrical Manifolds and Tube Dynamics in the Restricted Three-Body Problem

by

Shane David Ross

In Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

## Abstract

Within the phase space of the planar circular restricted three-body problem, stable and unstable manifolds of periodic orbits with a $S^1 \times \mathbb{R}^1$ (cylindrical) geometry are shown to exist. The periodic orbits considered reside in bottleneck regions of the energy manifold, separating large zones associated with motion about one mass, the other mass, or both masses.

The cylinders have the physical property that all motion through the bottleneck in which the periodic orbit resides must occur through the interior of these surfaces. The cylinders thus mediate the global transport of test particles between large zones of the energy surface which are separated by the bottlenecks.

By elucidating the structuring role of the cylinders, we provide a new language for discussing some important problems in celestial mechanics. Furthermore, we propose that these cylindrical structures are the natural objects of study for the design of space mission trajectories which take advantage of three-body effects.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

The restricted three-body problem is a classic problem of celestial mechanics, wherein one is interested in the motion of a particle of negligible mass in the presence of two massive bodies. Attempts at its solution laid the foundation for dynamical systems theory and alerted Poincaré in the 1890s to the existence of deterministic chaos within Newtonian mechanics (Poincaré [1890, 1892-1899]).

In this work, we consider global solutions to the planar circular restricted three-body problem (PCR3BP) from a geometric point of view which takes full advantage of the autonomous Hamiltonian structure of the problem.

In Chapter 2, we develop some basic terminology for the problem and describe the local dynamics near saddle-center equilibrium points, building on the work of Conley [1968]. As is well known, the PCR3BP admits the five Euler-Lagrange equilibria, $L_i, i = 1, \ldots, 5$. We focus our attention on two of these equilibrium points, $L_1$ and $L_2$, also known as *libration points* in the literature.[1] Periodic orbits about these points are shown to exist. Furthermore, stable and unstable manifolds of these periodic orbits in phase space with a $S^1 \times \mathbb{R}^1$ (cylindrical) geometry are shown to exist. The periodic orbits considered reside in bottleneck regions of the energy manifold, separating large zones, otherwise known as

---

[1]The Euler-Lagrange points are named after Leonhard Euler (1707-1783) and Joseph Lagrange (1736-1813), the two mathematicians who showed that there are five positions of equilibrium in a rotating two-body gravity field, as is the case in the PCR3BP. Euler [1767] discovered $L_1, L_2$, and $L_3$ just a few years before Lagrange [1772] discovered $L_4$ and $L_5$, but it is common to refer to $L_1, L_2, L_3$ as the "Lagrange" or "Lagrangian points" as well, despite being historically inaccurate. The Euler-Lagrange points are also known as "libration points," from the word *librate* means "to vibrate slightly" and is derived from the Latin word *librare* meaning "to balance." An object that librates is poised between two competing forces in a state of equilibrium. An object near a libration point will be in a state of libration.

*realms*, associated with motion about one mass, the other mass, or both masses. The cylinders have the physical property that all motion through the bottleneck in which the periodic orbit resides must occur through the interior of these surfaces. The cylinders thus mediate the global transport of test particles between large zones of the energy surface which are separated by the bottlenecks.

In Chapter 3, the local picture is extended to the global phase space. The cylindrical stable and unstable invariant manifolds of $L_1$ and $L_2$, referred to as *tubes*, are shown to play a crucial role in our understanding of the global connectivity of the phase space. In particular, the existence of a heteroclinic connection between pairs of periodic orbits is numerically demonstrated, one around $L_1$ and the other around $L_2$, with the two periodic orbits having the same energy. This heteroclinic connection, along with previously known homoclinic connections, allows us to prove a theorem on the global orbit structure of the PCR3BP. This theorem can also be taken as a proof of "horseshoe-like" chaos in the system.

In Chapter 4, we use the tubes and their intersections to compute orbits of desired itineraries with respect to the aforementioned realms of the phase space and make the connection to the design of space mission trajectories.

In Chapter 5, we construct solutions of the restricted four-body problem using solutions of the restricted three-body problem as building blocks. This approach, which is appropriate for some low energy space mission trajectories, is referred to as the *patched three-body approximation*. We demonstrate the approach with the numerical construction of a low energy Earth-to-Moon trajectory which uses the Sun's perturbation.

In Chapter 6, we explore the use of tubes to describe phase space transport in the restricted three-body problem and similar Hamiltonian systems.

In Chapter 7, we make some concluding remarks.

# Chapter 2

# Motion near the Collinear Equilibrium Points

**In collaboration with W. Koon, M. Lo, and J. Marsden.**

## 2.1 Introduction

In this chapter and the next, we begin the detailed analysis of a simplification of the general three-body problem, one in which we study the motion of a test particle with negligible mass compared to the other two.

Suppose the two more massive bodies move in circular orbits about their common center of mass and the mass of the third body is too small to affect the motion of the two more massive bodies. The problem of the motion of the third body is called the circular, restricted, three-body problem, henceforth referred to as the CR3BP. If we further restrict the motion of the third body to be in the orbital plane of the other two bodies, the problem is called the **planar circular restricted three-body problem**, or the PCR3BP.

At first glance this problem may seem to have little application to motion in the solar system. After all, the observed orbits of solar system objects are non-circular, albeit with small eccentricities. However, the hierarchy of orbits and masses in the solar system (e.g., sun, planet, satellite, ring particle) means that the CR3BP provides a good approximation for certain systems, especially the *qualitative* behavior of those systems. In fact, we need consider only the lower dimensional PCR3BP to understand a wide range of dynamical behaviors for the third body.

This chapter is roughly divided into two parts. In the first part (§2.2-2.5), we describe the equations for the problem with particular reference to a constant of the motion, the Hamiltonian energy. We demonstrate the relationship between curves defined by the

Hamiltonian energy and the orbital path of the particle. We discuss the location and stability of equilibrium points.

In the second part (§2.6-2.9), we consider the motion of particles near two important equilibria, the *libration points* $L_1$ and $L_2$, which will be explained shortly. By considering the motion near these two equilibria in detail, we lay the foundation for understanding the global picture of the PCR3BP phase space, to be covered in Chapter 3.

We will be focusing on particular aspects of the three-body problem which are important for the discussion in later chapters. The books by Szebehely [1967] and Marchal [1990] provide authoritative coverage of the literature on the subject. For other general introductions to the three-body problem, see Abraham and Marsden [1978], Meyer and Hall [1992], Holmes [1990], or Simó [1999].

## 2.2 Planar Circular Restricted Three-Body Problem

**Problem Description.** Consider the motion of a ***particle*** $P$ of negligible mass moving under the gravitational influence of two masses $m_1$ and $m_2$, referred to as the primary masses, or simply the ***primaries***. In some cases, we refer to $m_1$ as the primary and $m_2$ as the ***secondary***. Assume that $m_1$ and $m_2$ have circular orbits about their common center of mass. The particle $P$ is free to move in the plane defined by the circular orbits of the primaries, but cannot affect their motion.

In the context of this chapter, we can imagine that $m_1$ represents the sun and $m_2$ represents a planet, and we are concerned with the motion of $P$, a comet or spacecraft of much smaller mass.

The system is made nondimensional by the following choice of units: the unit of mass is taken to be $m_1 + m_2$; the unit of length is chosen to be the constant separation between $m_1$ and $m_2$ (e.g., the distance between the centers of the sun and planet); the unit of time is chosen such that the orbital period of $m_1$ and $m_2$ about their center of mass is $2\pi$. The universal constant of gravitation then becomes $G = 1$. It then follows that the common mean motion, $n$, of the primaries is also unity. We will refer to this system of units as ***nondimensional*** or ***normalized*** units throughout the thesis.

We will use the normalized units for nearly all the discussions in this thesis. When appropriate, we can convert to dimensional units (e.g., km, km/s, s) to scale a problem. The conversion from units of distance, velocity, and time in the unprimed, normalized

system to the primed, dimensionalized system is where $L$ is the distance between the

$$\text{distance} \quad d' = Ld,$$
$$\text{velocity} \quad s' = Vs,$$
$$\text{time} \quad t' = \frac{T}{2\pi}t,$$

centers of $m_1$ and $m_2$, $V$ is the orbital velocity of $m_1$, $T$ is the orbital period of $m_1$ and $m_2$.

The only parameter of the system is the **mass parameter**,

$$\mu = \frac{m_2}{m_1 + m_2}.$$

If we assume that $m_1 > m_2$, then the masses of $m_1$ and $m_2$ in this system of units are, respectively,

$$\mu_1 = 1 - \mu \qquad \text{and} \qquad \mu_2 = \mu,$$

where $\mu \in [0, \frac{1}{2}]$, and thus $\mu_1 \geq \mu_2$. The phase space of the system is highly dependent on the mass parameter. A table of mass parameters and dimensional values $L$, $V$, and $T$, for several pairs of masses in the solar system is provided in Appendix A.

## 2.3  Equations of Motion

In this section, we familiarize the reader with some of the terminology of the PCR3BP and the all important concept of viewing the motion in the rotating frame.

There are several ways to derive the equations of motion for $P$ in the field of $m_1$ and $m_2$. We will go over a few of the ways, emphasizing the Hamiltonian structure for this system.[1] A simple technique is to use the covariance of the Lagrangian formulation and use the Lagrangian written using coordinates in a moving frame, as given in Marsden and Ratiu [1999]. This method directly gives the equations in Lagrangian form and the associated Hamiltonian form is given by the Legendre transformation. We shall discuss this approach later in this section, but we begin with the transformation between the inertial and rotating frames.

---

[1]For example, Whittaker [1927] and Abraham and Marsden [1978] use time dependent canonical transformation theory to transform the problem from an inertial frame to a rotating frame.

**Transformation between the Inertial and Rotating Frames.** Let $X$-$Y$ be an inertial frame with origin at the $m_1$-$m_2$ center of mass, as in Figure 2.1. Consider the set of axes $x$ and $y$ depicted in Figure 2.1. The $x$-axis lies along the line from $m_1$ to $m_2$ with the $y$-axis perpendicular to it, completing a right-handed coordinate system. The $x$-$y$ frame rotates with respect to the $X$-$Y$ inertial frame with an angular velocity equal to the mean
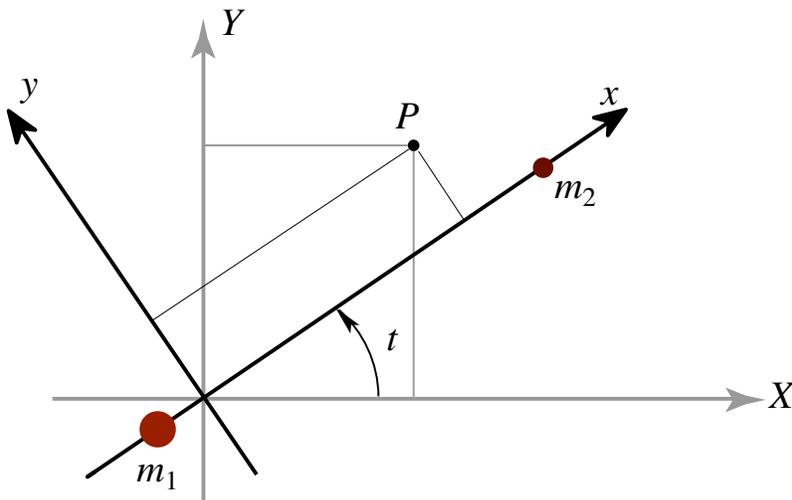


Figure 2.1: **Inertial and rotating frames.** The rotating coordinate system with coordinates $x$ and $y$ moves counterclockwise with unit angular velocity relative to the inertial frame with coordinates $X$ and $Y$.

motion, $n$, of either mass (unity in the normalized units). We will refer to this coordinate frame throughout the thesis as the ***rotating frame*** or the $m_1$-$m_2$ ***rotating frame***. Assume that the two frames coincide at $t = 0$. Let $(X, Y)$ and $(x, y)$ be the position of $P$ in the inertial and rotating frames, respectively. In normalized units, we have the following transformation of the particle's position between the two frames:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = A_t \begin{pmatrix} x \\ y \end{pmatrix}, \tag{2.1}$$

where

$$A_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \tag{2.2}$$

Differentiating gives us the transformation of velocity components from the rotating to the inertial frame:

$$
\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = \dot{A}_t \begin{pmatrix} x \\ y \end{pmatrix} + A_t \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix},
$$

$$
= -A_t J \begin{pmatrix} x \\ y \end{pmatrix} + A_t \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix},
$$

$$
= A_t \begin{pmatrix} \dot{x} - y \\ \dot{y} + x \end{pmatrix}, \tag{2.3}
$$

where

$$
J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
$$

**Rotating Frame.** The rotating frame is shown in Figure 2.2. The larger mass, $m_1$, is located at $(-\mu_2, 0)$ and the smaller mass, $m_2$, at $(\mu_1, 0)$. This is also true in the inertial frame when $t = 0$. At general times $t$,

$$
\begin{aligned}
(X_1, Y_1) &= (-\mu_2 \cos t, -\mu_2 \sin t), \\
(X_2, Y_2) &= (\mu_1 \cos t, \mu_1 \sin t),
\end{aligned}
$$

are the inertial frame positions of $m_1$ and $m_2$, respectively.

The rotating system of coordinates takes some getting used to, especially for those comortable with conic section orbits in the inertial frame from the two-body problem. We emphasize that in Figure 2.2, $(x, y)$ are the position coordinates of $P$ relative to the positions of the $m_1$ and $m_2$, *not relative to an inertial frame.*

**Gravitational Potential.** The gravitational potential which the particle experiences due to $m_1$ and $m_2$ (in normalized units) is

$$
\mathcal{U} = -\frac{\mu_1}{r_1} - \frac{\mu_2}{r_2} - \frac{1}{2}\mu_1\mu_2 \tag{2.4}
$$
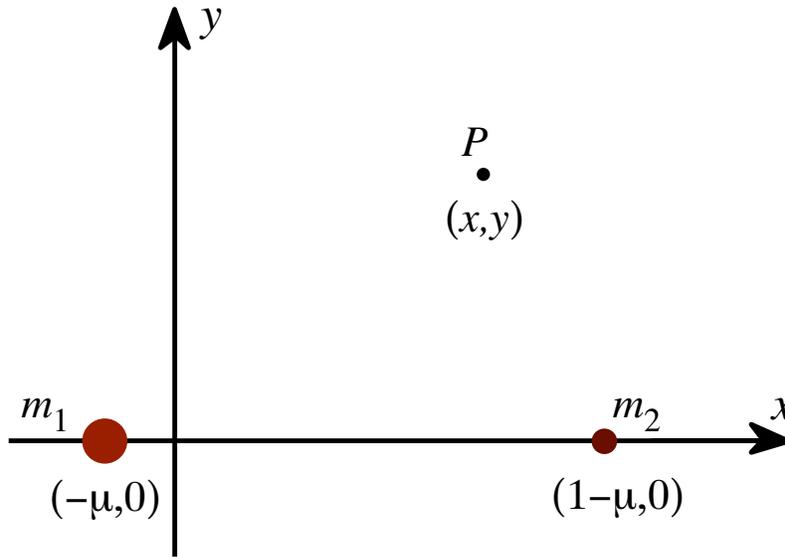
Figure 2.2: **Rotating coordinate frame.** The planar circular restricted three-body problem as viewed, not in any inertial frame, but in the rotating frame, where $m_1$ and $m_2$ are at fixed positions along the $x$-axis.

where $r_1$ and $r_2$ are the distances of $P$ from $m_1$ and $m_2$, respectively, given by

$$
\begin{aligned}
r_1^2 &= (X + \mu_2 \cos t)^2 + (Y + \mu_2 \sin t)^2, \\
r_2^2 &= (X - \mu_1 \cos t)^2 + (Y - \mu_1 \sin t)^2.
\end{aligned}
$$

The constant last term in the expression for $\mathcal{U}$ is added by convention (see, e.g., Llibre, Martinez and Simó [1985]), and will not affect the equations of motion.

**Lagrange Approach: Inertial Frame.** Recall the general form of the Euler-Lagrange equations:

$$
\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0,
$$

where the mechanical system is described by generalized coordinates $(q^1, \ldots, q^n)$. One usually chooses the Lagrangian $L$ to be of the form kinetic minus potential energy. See Marsden and Ratiu [1999] or other books on mechanics for a discussion.

In the inertial frame, the Lagrange $\mathcal{L}$ is kinetic minus potential energies and is given by

$$
\mathcal{L}(X, Y, \dot{X}, \dot{Y}, t) = \frac{1}{2}(\dot{X}^2 + \dot{Y}^2) - \mathcal{U}(X, Y, t).
$$

**Lagrange Approach: Rotating Frame.** In the rotating frame, the Lagrangian $L$ is given by

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\left((\dot{x} - y)^2 + (\dot{y} + x)^2\right) - U(x, y).$$

It is now *time-independent*, simplifying the analysis of solutions.

We obtain this formula for $L$ by simply rewriting the kinetic and potential energy of the inertial frame Lagrangian $\mathcal{L}$ in rotating coordinates. From Eq. (2.3), the potential energy is $\frac{1}{2}(\dot{X}^2 + \dot{Y}^2) = \frac{1}{2}\left((\dot{x} - y)^2 + (\dot{y} + x)^2\right)$. Also, since both the distances $r_1$ and $r_2$ are invariant under rotation, the gravitational potential is

$$U(x, y) = -\frac{\mu_1}{r_1} - \frac{\mu_2}{r_2} - \frac{1}{2}\mu_1\mu_2, \tag{2.5}$$

where $r_1$ and $r_2$ are expressed in rotating coordinates as

$$
\begin{aligned}
r_1^2 &= (x + \mu_2)^2 + y^2, \\
r_2^2 &= (x - \mu_1)^2 + y^2.
\end{aligned}
$$

The theory of moving systems says that one can simply write down the Euler-Lagrange equations in the rotating frame and one will get the correct equations. It is a very efficient generic method for computing equations for either moving systems or for systems seen from moving frames. See Marsden and Ratiu [1999] for more information.

In the present case, the Euler-Lagrange equations are given by

$$
\begin{aligned}
\frac{d}{dt}(\dot{x} - y) &= \dot{y} + x - U_x, \\
\frac{d}{dt}(\dot{y} + x) &= -(\dot{x} - y) - U_y.
\end{aligned}
$$

After simplification, we have

$$
\begin{aligned}
\ddot{x} - 2\dot{y} &= -\bar{U}_x, \\
\ddot{y} + 2\dot{x} &= -\bar{U}_y, \tag{2.6}
\end{aligned}
$$

where

$$\begin{aligned}
\bar{U}(x, y) &= -\frac{1}{2}(x^2 + y^2) + U(x, y), \\
&= -\frac{1}{2}(x^2 + y^2) - \frac{\mu_1}{r_1} - \frac{\mu_2}{r_2} - \frac{1}{2}\mu_1\mu_2, \quad (2.7) \\
&= -\frac{1}{2}(\mu_1 r_1^2 + \mu_2 r_2^2) - \frac{\mu_1}{r_1} - \frac{\mu_2}{r_2}, \quad (2.8)
\end{aligned}$$

is the augmented or **effective potential** and the subscripts denote its partial derivatives. This form of the equations has been studied in detail in Szebehely [1967].

**Hamiltonian Approach: Rotating Frame.** Whenever one has a Lagrangian system, one can transform it to Hamiltonian form by means of the Legendre transformation:

$$p_i = \frac{\partial L}{\partial \dot{q}^i}; \qquad H(q^i, p_i) = \sum_{i=1}^{n} p_i \dot{q}^i - L(q^i, p_i),$$

to get the equations in Hamiltonian form

$$\dot{q}^i = \frac{\partial H}{\partial p_i}; \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

In our case, the Legendre transformation is given by

$$\begin{aligned}
p_x &= \frac{\partial L}{\partial \dot{x}} = \dot{x} - y, \\
p_y &= \frac{\partial L}{\partial \dot{y}} = \dot{y} + x,
\end{aligned}$$

and so we obtain the Hamiltonian function

$$\begin{aligned}
H(x, y, p_x, p_y) &= p_x \dot{x} + p_y \dot{y} - L \\
&= \frac{1}{2}((p_x + y)^2 + (p_y - x)^2) + \bar{U}(x, y), \quad (2.9)
\end{aligned}$$

where $p_x$ and $p_y$ are the conjugate momenta.

Hence the Hamiltonian equations are given by

$$
\begin{aligned}
\dot{x} &= \frac{\partial H}{\partial p_x} = p_x + y, \\
\dot{y} &= \frac{\partial H}{\partial p_y} = p_y - x, \\
\dot{p}_x &= -\frac{\partial H}{\partial x} = p_y - x - \bar{U}_x, \\
\dot{p}_y &= -\frac{\partial H}{\partial y} = -p_x - y - \bar{U}_y.
\end{aligned}
\tag{2.10}
$$

Notice that both the Lagrangian and the Hamiltonian form of the equations in rotating coordinates $(x, y)$ give a time-independent system. Viewed as a dynamical system, it is a four dimensional dynamical system in either $(x, y, \dot{x}, \dot{y})$ or $(x, y, p_x, p_y)$ space.

**Energy Integral and Jacobi Constant.** Since the equations of motion of the PCR3BP (2.10) are Hamiltonian and independent of time, they have an *energy integral* of motion. We use the symbol $H$ when we regard the energy as a function of positions and momenta and $E$ when we regard it as a function of the positions and velocities,

$$
E(x, y, \dot{x}, \dot{y}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \bar{U}(x, y).
\tag{2.11}
$$

Physically, the measurement of the particle's position and velocity in either the inertial or rotating frames determines the value of the energy associated with the particle's motion.

The astronomy and astrodynamics communities use $-2E$, which is called the *Jacobi integral* and is given by

$$
C(x, y, \dot{x}, \dot{y}) = -(\dot{x}^2 + \dot{y}^2) - 2\bar{U}.
\tag{2.12}
$$

Usually in those communities, the existence of the Jacobi integral is derived directly from the equations of motion. The computation is straightforward:

$$
\begin{aligned}
\frac{d}{dt}(\dot{x}^2 + \dot{y}^2) &= 2(\dot{x}\ddot{x} + \dot{y}\ddot{y}) \\
&= 2[\dot{x}(2\dot{y} - \bar{U}_x) + \dot{y}(-2\dot{x} - \bar{U}_y)] = 2\frac{d}{dt}(-\bar{U}),
\end{aligned}
$$

so we get

$$
\frac{d}{dt}C = \frac{d}{dt}\left(-(\dot{x}^2 + \dot{y}^2) - 2\bar{U}\right) = 0.
$$

Throughout the thesis, we will use the terms "energy," "energy integral," "Jacobi integral," and "Jacobi constant" to refer to the same concept—the most important integral determining the motion of the particle. As they differ in sign, we will make it clear from the context when we are referring to increasing energy (decreasing Jacobi constant), etc. In general, there are no other integrals constraining the motion of the particle, making the PCR3BP a *non-integrable* problem.

## 2.4   Energy Surface and Regions of Possible Motion

In the two-body Kepler problem, one may divide the phase space into two major categories, based on the values of the Keplerian energy, $E_{Kep} = -\frac{1}{2a}$, where $a$ is the semimajor axis of the test particle's orbit around the central massive body. The following two cases divide the phase space into two major categories of possible motion for the test particle.

- (i) $E_{Kep} < 0$ : Negative Keplerian energies correspond to bound motion of the test particle about the single massive body, i.e., elliptical and circular orbits.

- (ii) $E_{Kep} > 0$ : Positive Keplerian energies correspond to unbound motion, i.e., hyperbolic orbits coming from and going to infinity.

The critical case of zero energy orbits between these two are the unbound parabolic orbits. If we restrict ourselves to the *planar* Kepler problem, we have a four-dimensional phase space, which we can view as $\mathbb{R}^4$: two position coordinates and their two corresponding velocities. For each value, $e$, the equation, $E_{Kep} = e$, describes a three-dimensional set in the four-dimensional phase space, termed the *energy surface corresponding to energy e*. The phase space can be viewed as a many layered "onion," each layer or leaf corresponding to a value of the energy. One says that the energy surfaces foliate the phase space.

In the three-body problem, the picture is more complicated, but we can follow a similar strategy of categorizing the possible motion of the test particle by energy, this time the three-body energy given in (2.11).

**Energy Surface.**   Let $\mathcal{M}$ be the **energy manifold** or **energy surface** given by setting the energy integral (2.11) equal to a constant, i.e.,

$$\mathcal{M}(\mu, e) = \{(x, y, \dot{x}, \dot{y}) \mid E(x, y, \dot{x}, \dot{y}) = e\}, \tag{2.13}$$

where $e$ is a constant. For a fixed $\mu$ and energy $e$, one can consider the surface $\mathcal{M}(\mu, e)$ as a three-dimensional surface embedded in the four-dimensional phase space.

**Hill's Region: the Region of Possible Motion.** The projection of this surface onto position space in the rotating frame, the $x$-$y$ plane, is the region of possible motion for a particle of energy $e$ in the field of two masses with mass parameter $\mu$. Let $M(\mu, e)$ denote this projection,[2]

$$M(\mu, e) = \{(x, y) \mid \bar{U}(x, y) \le e\}, \tag{2.14}$$

known historically as the **Hill's region**. The boundary of $M(\mu, e)$ is known as the **zero velocity curve**, and plays an important role in placing bounds on the motion of the particle.

**Zero Velocity Curves: the Boundaries of the Hill's Region.** The zero velocity curves are the locus of points in the $x$-$y$ plane where the kinetic energy, and hence the velocity, $v = \sqrt{\dot{x}^2 + \dot{y}^2}$, vanishes, i.e., $\frac{1}{2}v^2(x, y) = e - \bar{U}(x, y) = 0$. From (2.14), it is clear that the particle is only able to move on the side of this curve for which the kinetic energy is positive. The other side of the curve, where the kinetic energy is negative and motion is not possible, is known as the *forbidden realm.*

Recall that the energy $E$ is given by (2.11). Fixing the energy function to be a constant, i.e., $E(x, y, \dot{x}, \dot{y}) = e$, is like fixing a height in the plot of the effective potential, $\bar{U}(x, y)$. Consider the surface of the effective potential in Figure 2.3 and note the following features.

- Near either $m_1$ or $m_2$, we have a potential well.

- Far away from either $m_1$ or $m_2$, the term that corresponds to the centrifugal force dominates $\bar{U}$ in (2.7), i.e., $||\frac{1}{2}(x^2 + y^2)||/||\frac{1-\mu}{r_1} + \frac{\mu}{r_2}|| \gg 1$, and we have another potential well.

- By multivariable calculus, one finds that there are five critical points where the slope is zero: three saddle points along the $x$ axis and two symmetric points off the $x$ axis. As will be covered in the next section, these points are the $x$-$y$ locations of the

---

[2]Note that our convention is to use script letters for a region in the energy surface (including the energy surface itself, $\mathcal{M}$) and italicized capital letters for that region's projection onto the position space (e.g., $M$).
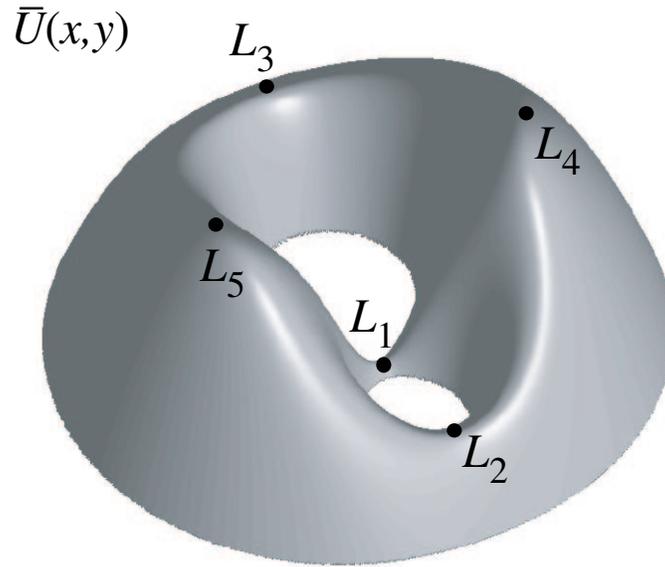
Figure 2.3: The plot of the effective potential $\bar{U}(x,y)$ for $\mu = 0.3$. The critical points are at the locations of the five equilibrium points, $L_i, i = 1, \ldots, 5$.

equilibrium points for a particle in the rotating frame, i.e., a particle placed here at rest (zero initial velocity), will stay at rest for all time (zero acceleration). We label these points $L_i, i = 1, \ldots, 5$, as in Figure 2.3.

- Let $E_i$ be the energy of a particle at rest at $L_i$, then $E_5 = E_4 > E_3 > E_2 > E_1$. Thus, $L_1$ is the location of the lowest energy equilibrium point and $L_4$ and $L_5$ are the highest energy equilibrium points. Since the energy is measured in a rotating frame, we cannot determine the stability properties of all the equilibrium points from their ordering by energy (e.g., $L_4$ and $L_5$ are spectrally stable for small $\mu$, despite being energy maxima, as covered in Szebehely [1967]).

**The Five Cases of the Hill's Region.** For a given $\mu$ there are five basic configurations for the Hill's region, corresponding to five intervals of energy value, $e$, in (2.13). We refer to these basic configurations as *cases*. The first four cases are shown in Figure 2.4. In the fifth case, motion over the entire $x$-$y$ plane is possible. We will show how to compute the energy intervals corresponding to these cases.

Contour plots of the effective potential give the five cases of Hill's region. The white areas in Figure 2.4 are the Hill's region and the shaded areas are the forbidden realm.

(a) Case 1 : $E<E_1$

(b) Case 2 : $E_1<E<E_2$

(c) Case 3 : $E_2<E<E_3$

(d) Case 4 : $E_3<E<E_4=E_5$
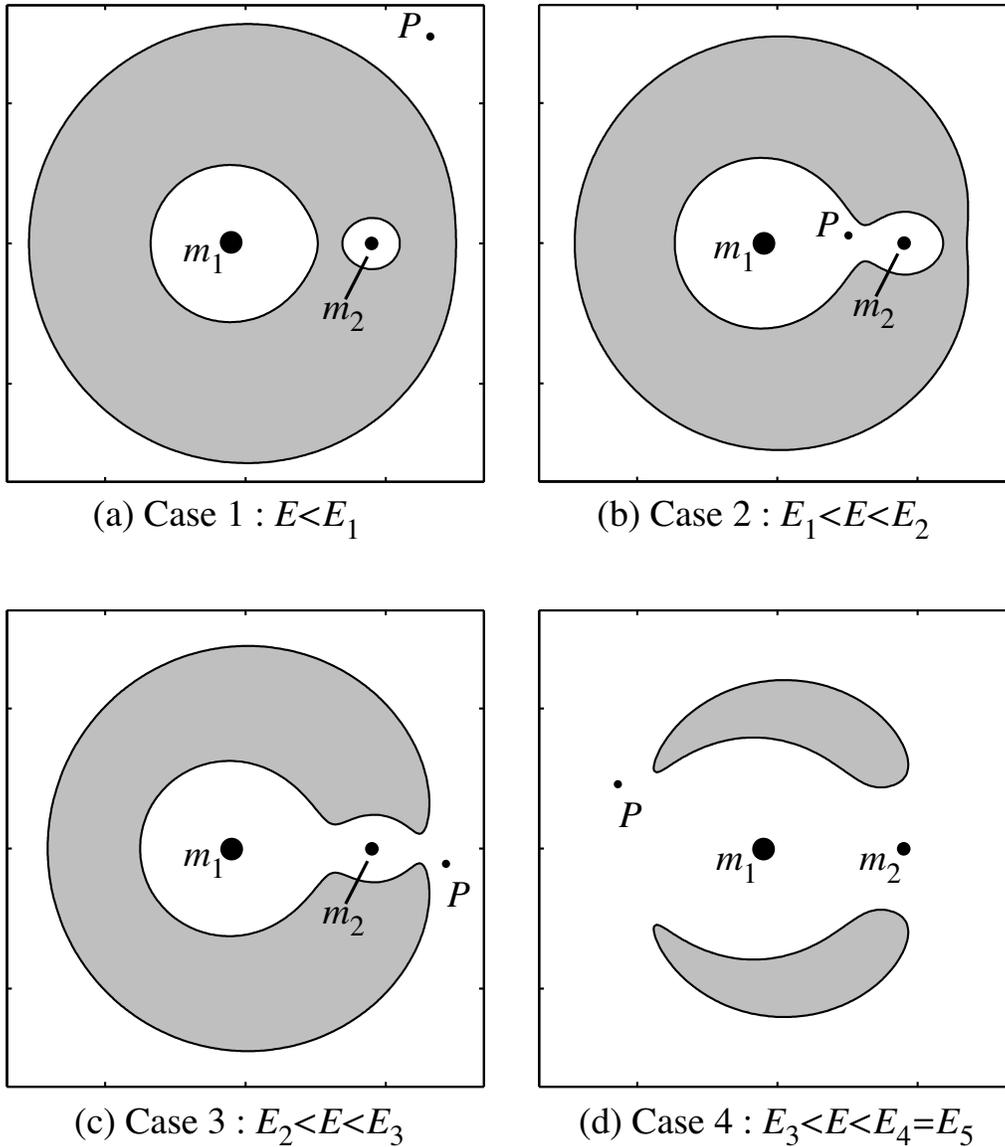
Figure 2.4: **Realms of possible motion.** Zero velocity curves for four values of the energy, one in each of the first four cases as described in the text, are shown on the $x$-$y$ plane for $\mu = 0.3$. These curves bound the zone, in white, accessible by the particle, $P$, for a given energy value, $E = e$. The part of the $x$-$y$ plane which is shaded is inaccessible for a given energy, and known as the *forbidden realm*. The outermost accessible realm, known as the *exterior realm*, extends to infinity.

- **Case 1,** $E < E_1$ : If the energy of the particle is below $E_1$, the particle cannot move between the regions around $m_1$ and $m_2$.

- **Case 2,** $E_1 < E < E_2$ : If the energy is just above $E_1$, a "neck" between the regions

around $m_1$ and $m_2$ opens up, permitting the particle to move between the two regions. The $L_1$ point is in this neck. We will see in §2.6 that the transport between the two adjacent regions is controlled by invariant manifold structures associated to $L_1$. The particle is still barred from moving between these two regions and the exterior region extending to infinity.

- **Case 3, $E_2 < E < E_3$** : This is the case that concerns us the most; when the energy is just above $E_2$. The particle can move between the vicinity of $m_1$ and $m_2$ and the exterior region via a neck around $L_2$.

- **Case 4, $E_3 < E < -\frac{3}{2} = E_4 = E_5$** : In this case the energy is above $E_3$ but below that of $E_4$ and $E_5$, which is always $-\frac{3}{2}$. The particle can pass directly from the vicinity of $m_1$ to the exterior region via a neck around $L_3$.

- **Case 5, $-\frac{3}{2} < E$** : If the energy is above $E_4 = E_5 = -\frac{3}{2}$, the forbidden realm disappears. Case 5 is where the particle is free to move in the entire $x$-$y$ plane.

**Realms of Possible Motion.** A glance at Figure 2.4 reveals that, beginning in case 1, there are three main regions of possible motion. We refer to these regions as **realms**. Considering, for example, Figure 2.4(a), the large region surrounding $m_1$ is referred to as the $m_1$ **realm**, sometimes referred to as the **interior realm**. The small region surrounding $m_2$ is the $m_2$ **realm**. The realm which lies outside both the $m_1$ and $m_2$ realms, and extends to infinity, is the **exterior realm**. For case 1, the realms are separated. Moving up in energy to case 2, a neck around $L_1$ opens up between the $m_1$ and $m_2$ realms, permitting the particle to pass between the two. An additional neck opens up around $L_2$ when we move up in energy to case 3, permitting travel between all three realms. Our main interest in this thesis will be case 3; but for comparison we shall occasionally bring up case 2.

The critical values of $E$ which separate these five cases are the values $E_i, i = 1, \ldots, 4$ previously mentioned, corresponding to the equilibrium points $L_i, i = 1, \ldots, 4$. These values can be easily calculated for small $\mu$ as will be shown in the following section. The graphs of the $E_i$ as a function of $\mu$ are shown in Figure 2.5. For case 3, the energy value lies between $E_2$ and $E_3$ which are the energy values of the libration points $L_2$ and $L_3$, respectively.
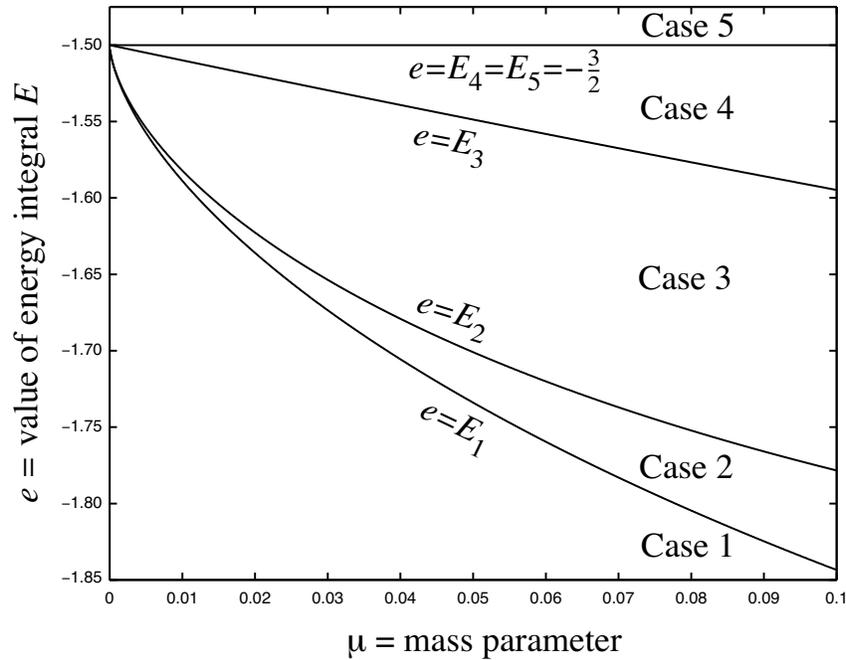
Figure 2.5: **The graphs of the $E_i$ as a function of** $\mu$ partition the $\mu$-$e$ plane into the five cases of possible motion. The Hill's regions for cases 1 through 4 are shown in Figure 2.4.

## 2.5  Location of the Equilibrium Points

Written in first-order form, the equations of motion for the PCR3BP are

$$
\begin{aligned}
\dot{x} &= v_x, \\
\dot{y} &= v_y, \\
\dot{v}_x &= 2v_y - \bar{U}_x, \\
\dot{v}_y &= -2v_x - \bar{U}_y,
\end{aligned}
\tag{2.15}
$$

To find equilibrium points, we set the right-hand sides of the system equal to zero. We see that equlibria in $(x, y, v_x, v_y)$ space are of the form $(x_e, y_e, 0, 0)$, where $(x_e, y_e)$ are critical points of the effective potential function $\bar{U}(x, y)$ shown in Figure 2.3.

As described in the previous section, the PCR3BP admits five equilibrium point solutions, which are shown in Figure 2.6:

- three collinear equilibria on the $x$-axis, called $L_1, L_2, L_3$; and

- two equilateral points called $L_4, L_5$.
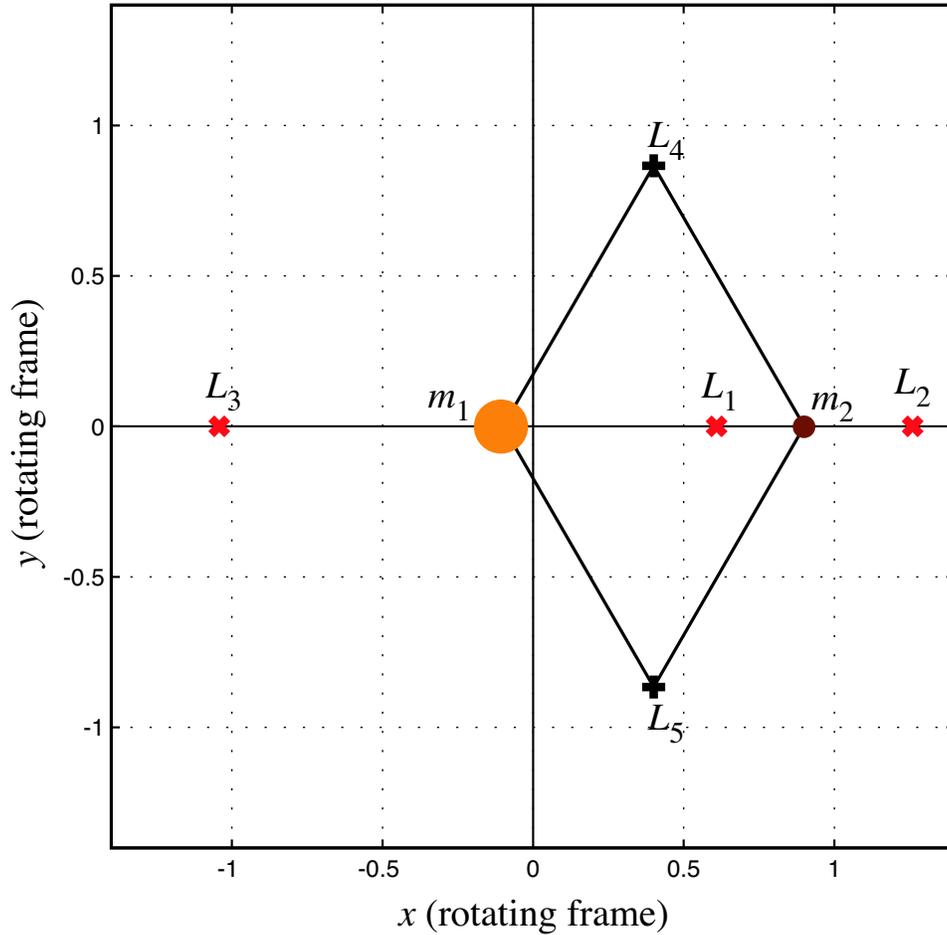
Figure 2.6: **Equilibrium points of the circular restricted three-body problem** in the $x$-$y$ plane of the frame rotating with the mean motion of the orbit of $m_1$ and $m_2$. A particle placed at rest at such a point will remain at rest for all time. The points marked with an 'x' are linearly unstable. Those marked with a '+' are unstable for $\mu \geq \mu_0 \simeq 0.038521$ and spectrally stable otherwise (see Szebehely [1967] for details). The points shown here are for $\mu = 0.3$.

The points $L_1, L_2$, and $L_3$ were discovered by Euler before Lagrange discovered the points, $L_4$ and $L_5$.[3] We are concerned with the collinear points, which can be found as follows.

---

[3]Euler [1767] discovered $L_1, L_2$, and $L_3$ just a few years before Lagrange [1772] discovered $L_4$ and $L_5$, but it is common in the literature to refer to $L_1, L_2, L_3$ as the "Lagrange" or "Lagrangian points" as well, despite being historically inaccurate.

**The Collinear Points.** Consider equilibria along the line of primaries where $y = 0$. In this case the effective potential function has the form

$$\bar{U}(x,0) = -\frac{1}{2}x^2 - \frac{1-\mu}{|x+\mu|} - \frac{\mu}{|x-1+\mu|}.$$

It can be determined that $\bar{U}(x,0)$ has precisely one critical point in each of the following three intervals along the $x$-axis: (i) $(-\infty, -\mu)$, (ii) $(-\mu, 1-\mu)$ and (iii) $(1-\mu, \infty)$.

This is because $\bar{U}(x,0) \to -\infty$ as $x \to \pm\infty$, as $x \to -\mu$, or as $x \to 1-\mu$. So $\bar{U}$ has at least one critical point on each of these three intervals. Also,

$$\frac{d^2\bar{U}}{dx^2} = -1 - \frac{1-\mu}{|x+\mu|^3} - \frac{\mu}{|x-1+\mu|^3},$$

is always negative, so $\bar{U}$ is concave. Therefore, $\bar{U}$ has precisely one critical point in each of these three intervals. A sketch of the graph of $\bar{U}(x,0)$ is given in Figure 2.7.
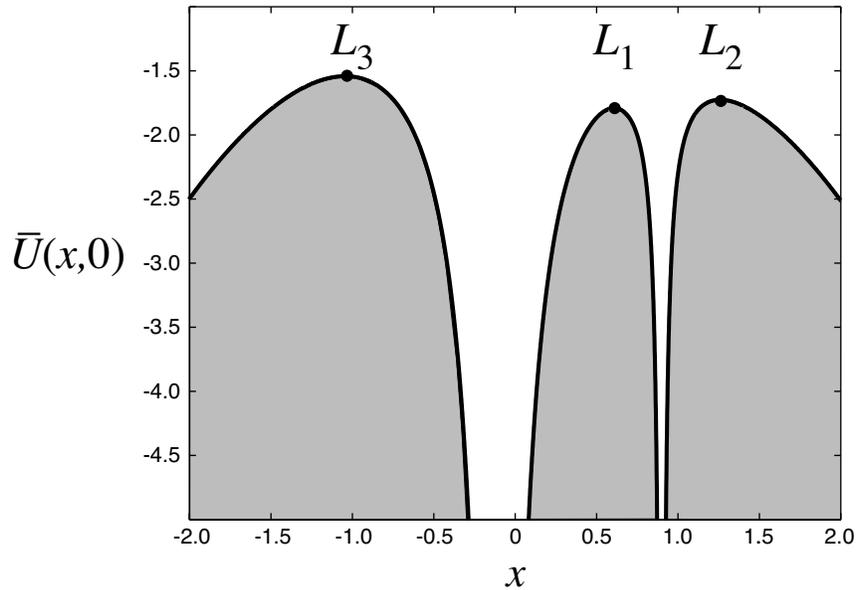


Figure 2.7: **The graph of** $\bar{U}(x,0)$ for $\mu = 0.1$ is shown. The solid line is the intersection of $\bar{U}(x,y)$ in Figure 2.3 with the plane defined by $y = 0$. At the $x$ locations of $m_1$ and $m_2$, the function plunges to $-\infty$. The maxima of $\bar{U}(x,0)$ correspond to the unstable collinear equilibrium points $L_1, L_2$, and $L_3$.

**Locating the Collinear Equilibria.** Computation of the $x$ values of the collinear points requires finging the maxima of the function $\bar{U}(x,0)$, i.e., the solutions of $\frac{d}{dx}\bar{U}(x,0) =$

0 which is a quintic equation after simplification. The distance from $L_i, i = 1, 2$ to the smaller primary is given by the unique postive solution $\gamma_i$ of the following equation:

$$\gamma^5 \mp (3 - \mu)\gamma^4 + (3 - 2\mu)\gamma^3 - \mu\gamma^2 \pm 2\mu\gamma - \mu = 0 \tag{2.16}$$

where the upper sign is for $L_1$ and the lower one for $L_2$ (see Szebehely [1967]). A similar equation can be found for $\gamma_3$, the distance between $L_3$ to the larger primary.

Historically, a lot of effort has been spent finding the series expansion for such solutions. Here, we will write down two of those, from Szebehely [1967], that are most useful for us:

$$\gamma_1 = r_h(1 - \frac{1}{3}r_h - \frac{1}{9}r_h^2 + \ldots), \tag{2.17}$$

$$\gamma_2 = r_h(1 + \frac{1}{3}r_h - \frac{1}{9}r_h^2 + \ldots), \tag{2.18}$$

where $r_h = (\frac{\mu}{3})^{\frac{1}{3}}$, the *Hill radius*, is the radius of the Hill sphere in the spatial problem. The Hill sphere is the 'bubble' in 3-D position space surrounding $m_2$ inside of which the gravitational field of $m_2$ has a dominant effect on the particle's motion over the field of $m_1$. It can also be thought of as a *sphere of influence* of $m_2$.

Locating these points to a higher accuracy numerically is straightforward. $\gamma_i, i = 1, 2$ can be solved by the Newton method, using $r_h$ as an initial solution for the above quintic equation (2.16).

As an example calculation, consider the motion of a particle in the Sun-Jupiter system ($\mu = 9.537 \times 10^{-4}$ from Appendix A). The Hill radius is $r_h = 6.825 \times 10^{-2}$, and $\gamma_1 = 6.666 \times 10^{-2}$ to third-order in $r_h$ via Eq. (2.17). Solving the quintic equation (2.16) numerically yields $\gamma_1 = 6.668 \times 10^{-2}$, and thus the $x$ position of $L_1$ is $x_{L_1} = x_{m_1} - \gamma_1 = (1 - \mu) - \gamma_1 = 0.9324$.

**A Note on Terminology.** Throughout the literature covering the equilibrium points in the PCR3BP, the points are given various names, such as libration points, Lagrange points, and Lagrangian points. In this thesis, we will restrict ourselves to the terms *libration* and *equilibrium* point.

## 2.6    Linearization near the Collinear Equilibria

In this section, we begin the study of the behavior of particle trajectories near the two libration points $L_1$ and $L_2$. As will become clear, we are particularly interested in particle's whose energy is just above that of the critical point $L_2$, that is, $E > E_2$ in case 3.

In this chapter we study the behavior of orbits near the two libration points $L_1$ and $L_2$ and particularly those orbits whose energy value $E$ is just above that of the critical point $L_2$, that is, $E > E_2$. As shown in Figure 2.4(c), the region of possible motion for case 3 contains a neck *about each libration point*. Thus, a particle starting in the exterior realm may pass through the neck around $L_2$ to the $m_2$ realm, and subsequently pass through the neck around $L_1$ to the $m_1$ realm, and so on.

The aim in the next few sections is to describe the geometry of trajectories in the neck regions. We begin by considering the equations of motion linearized near the equilibrium point inside the neck region. By virtue of Moser's generalization of a theorem of Lyapunov *all the qualitative results of such a discussion carry over to the full nonlinear equations* (see Moser [1958]). See Appendix B for more details on this theorem.

In the following sections, we will use $L$ to denote either $L_1$ or $L_2$. Furthermore, for a fixed energy $E$, consider a neighborhood of $L$ in the energy surface, whose position space projections are the neck regions described previously. We refer to this neighborhood as the **equilibrium region** and denote it by $\mathcal{R}$ on the energy surface. Following our adopted convention, $\mathcal{R}$ has the position space projection $R$.

**Hamiltonian Approach.**    To find the linearized equations around the collinear libration point $L$ with coordinates $(x_e, y_e, \dot{x}_e, \dot{y}_e) = (x_e, 0, 0, 0)$, we need the quadratic terms of the Hamiltonian $H$ in equation (2.9) as expanded about $(x_e, y_e, p_{xe}, p_{ye} = (x_e, 0, 0, x_e)$. After making a coordinate change with $(x_e, 0, 0, x_e)$ as the origin, these quadratic terms form the Hamiltonian function for the linearized equations, which we shall call $H_l$.

$$H_l = \frac{1}{2} \left[ (p_x + y)^2 + (p_y - x)^2 - ax^2 + by^2 \right],  \tag{2.19}$$

where $a$ and $b$ are defined by $a = 2\rho + 1$, and $b = \rho - 1$ and where

$$c = \mu |x_e - 1 + \mu|^{-3} + (1 - \mu)|x_e + \mu|^{-3}.  \tag{2.20}$$

It can be shown that both $a$ and $b$ are positive constants.

A short computation gives the linearized equations in the Hamiltonian form

$$
\begin{aligned}
\dot{x} &= \frac{\partial H_l}{\partial p_x} = p_x + y, \\
\dot{y} &= \frac{\partial H_l}{\partial p_y} = p_y - x, \\
\dot{p}_x &= -\frac{\partial H_l}{\partial x} = p_y - x + ax, \\
\dot{p}_y &= -\frac{\partial H_l}{\partial y} = -p_x - y - by.
\end{aligned} \tag{2.21}
$$

**Lagrangian Approach.** Using the inverse Legendre transformation: $v_x = p_x + y, v_y = p_y - x$, where $v_x, v_y$ correspond to velocity in the rotating coordinate system, we obtain the linearized equations in the Lagrangian form

$$
\begin{aligned}
\dot{x} &= v_x, \\
\dot{y} &= v_y, \\
\dot{v}_x &= 2v_y + ax, \\
\dot{v}_y &= -2v_x - by,
\end{aligned} \tag{2.22}
$$

which is the linearization of the equations (2.15) around the equilibrium point $(x_e, 0, 0, 0)$.

The integral $H_l$ of (2.19) now appears as

$$
E_l = \frac{1}{2}(v_x^2 + v_y^2 - ax^2 + by^2), \tag{2.23}
$$

which corresponds to the energy integral $E$ of the restricted problem. Notice that the zero-surface of the integral $E_l$ corresponds to the energy surface which passes through the libration point. We shall therefore study solutions of equations (2.22) on the surface $E_l = \varepsilon > 0$ which corresponds to the case where the Hill's region contains a neck about the libration point.

## 2.7 Geometry of Solutions near the Equilibria

Now we analyze the linearized equations (2.22). It is straightforward to find that the eigenvalues of this linear system have the form $\pm\lambda$ and $\pm i\nu$, where $\lambda$ and $\nu$ are positive

constants. The corresponding eigenvectors are

$$u_1 = (1, -\sigma, \lambda, -\lambda\sigma),$$

$$u_2 = (1, \sigma, -\lambda, -\lambda\sigma),$$

$$w_1 = (1, -i\tau, i\nu, \nu\tau),$$

$$w_2 = (1, i\tau, -i\nu, \nu\tau),$$

where $\sigma$ and $\tau$ are constants with $\sigma > 0$ and $\tau < 0$.

**Eigenvalues.** It can be shown that the characteristic polynomial for the linearized equations (2.22) written in matrix form

$$
\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{v}_x \\ \dot{v}_y \end{pmatrix} =
\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a & 0 & 0 & 2 \\ 0 & -b & -2 & 0 \end{pmatrix}
\begin{pmatrix} x \\ y \\ v_x \\ v_y \end{pmatrix} = A
\begin{pmatrix} x \\ y \\ v_x \\ v_y \end{pmatrix},
$$

is given by

$$p(\beta) = \beta^4 + (2 - c)\beta^2 + (1 + c - 2c^2).$$

Let $\alpha = \beta^2$, then the roots of $p(\alpha) = 0$ are as follows

$$\alpha_1 = \frac{c - 2 + \sqrt{9c^2 - 8c}}{2}, \qquad \alpha_2 = \frac{c - 2 - \sqrt{9c^2 - 8c}}{2}.$$

Since the last term of $p(\alpha) = 0$ is equal to $-ab$ which is negative, this quadratic equation must have one positive and one negative root. So, we have $\alpha_1 > 0$ and $\alpha_2 < 0$. Therefore, the eigenvalues of the linearized equations are of the form $\pm\lambda$ and $\pm i\nu$.

**Eigenvectors.** Let $v = (k_1, k_2, k_3, k_4)$ be an eigenvector of the linearized equations. If $\beta$ is an eigenvalue, then $Av = \beta v$ and we have the following relations

$$
\begin{aligned}
k_3 &= \beta k_1, & ak_1 + 2k_4 &= \beta k_3, \\
k_4 &= \beta k_2, & -bk_2 - 2k_3 &= \beta k_4.
\end{aligned}
$$

Notice that $k_1 \neq 0$, otherwise $k_1 = k_3 = k_4 = 0$ and $v = 0$. Thus, $k_1$ may be taken to be 1 and the equations relating the components of $v$ indicate that $v$ may have the form

$$v = (1, k_2, \beta, \beta k_2),$$

and that

$$a + 2\beta k_2 = \beta^2,$$
$$-bk_2 - 2\beta = \beta^2 k_2.$$

First let $\beta = \lambda$ and then $\beta = -\lambda$ to obtain

$$u_1 = (1, k_2, \lambda, \lambda k_2),$$
$$u_2 = (1, k_2', -\lambda, -\lambda k_2'),$$

where

$$a + 2\lambda k_2 = \lambda^2,$$
$$-bk_2 - 2\lambda = \lambda^2 k_2,$$
$$a - 2\lambda k_2' = \lambda^2,$$
$$-bk_2' + 2\lambda = \lambda^2 k_2'.$$

The first and the third equations show that $k_2 = -k_2'$, and, denoting $k_2' = \sigma$, the second and fourth give

$$\sigma = \frac{2\lambda}{\lambda^2 + b} > 0.$$

Similarly, taking $\beta = i\nu$, then $\beta = -i\nu$, we obtain

$$w_1 = (1, -i\tau, i\nu, \nu\tau),$$
$$w_2 = (1, i\tau, -i\nu, \nu\tau),$$

where

$$\tau = -\left(\frac{\nu^2 + a}{2\nu}\right) < 0.$$

**Eigenvectors as Axes for New Coordinate System.** To better understand the orbit structure on the phase space, we make a linear change of coordinates with the eigenvectors, $u_1, u_2, w_1, w_2$, as the axes of the new system. Using the corresponding new coordinates $(\xi, \eta, \zeta_1, \zeta_2)$, the differential equations assume the simple form

$$
\begin{aligned}
\dot{\xi} &= \lambda \xi, \\
\dot{\eta} &= -\lambda \eta, \\
\dot{\zeta_1} &= \nu \zeta_2, \\
\dot{\zeta_2} &= -\nu \zeta_1,
\end{aligned}
\tag{2.24}
$$

and the energy function (2.23) becomes

$$
E_l = \lambda \xi \eta + \frac{\nu}{2}(\zeta_1^2 + \zeta_2^2).
\tag{2.25}
$$

Solutions of the equations (2.24) can be conveniently written as

$$
\begin{aligned}
\xi(t) &= \xi^0 e^{\lambda t}, \\
\eta(t) &= \eta^0 e^{-\lambda t}, \\
\zeta(t) &= \zeta_1(t) + i\zeta_2(t) = \zeta^0 e^{-i\nu t},
\end{aligned}
\tag{2.26}
$$

where the constants $\xi^0, \eta^0$ and $\zeta^0 = \zeta_1^0 + i\zeta_2^0$ are the initial conditions. These linearized equations admit integrals in addition to the energy function (2.25); namely, the functions $\eta \xi$ and $|\zeta|^2 = \zeta_1^2 + \zeta_2^2$ are both constant along solutions.

**Phase Space of the Equilibrium Region.** For positive $\varepsilon$ and $c$, the region $\mathcal{R}$, which is determined by

$$
E_l = \varepsilon, \quad \text{and} \quad |\eta - \xi| \leq c,
\tag{2.27}
$$

which is homeomorphic to the product of a two-sphere and an interval; namely, for each fixed value of $\eta - \xi$ on the interval $I = [-c, c]$, the equation $E_l = \varepsilon$ determines the two-sphere

$$
\frac{\lambda}{4}(\eta + \xi)^2 + \frac{\nu}{2}(\zeta_1^2 + \zeta_2^2) = \varepsilon + \frac{\lambda}{4}(\eta - \xi)^2.
$$

The bounding sphere of $\mathcal{R}$ for which $\eta - \xi = -c$ will be called $n_1$, and that where $\eta - \xi = c$, $n_2$ (see Figure 2.8). We shall call the set of points on each bounding sphere where $\eta + \xi = 0$ the *equator*, and the sets where $\eta + \xi > 0$ or $\eta + \xi < 0$ will be called the *north and south hemispheres*, respectively.

**Flow in the Equilibrium Region.** To analyze the flow in $\mathcal{R}$ one simply considers the projections on the $\eta$-$\xi$ plane and $\zeta$ planes, respectively. In the first case we see the standard picture of an unstable critical point, and in the second, of a center. Figure 2.8 schematically illustrates the flow in the $\eta$-$\xi$ plane. The coordinate axes have been tilted by $45°$ in order to correspond to the direction of the flow in later figures. In Figure 2.8(a), $\mathcal{R}$ itself projects to a set bounded on two sides by the hyperbola $\eta\xi = \varepsilon/\lambda$, the thick solid hyperbolic segments on the top and bottom, (corresponding to $|\zeta|^2 = 0$, see (2.25)). $\mathcal{R}$ is bounded on two other sides by the line segments $\eta - \xi = \pm c$, the dotted vertical lines at left and right in Figure 2.8(a), which correspond to the bounding spheres, $n_1$ and $n_2$, respectively.

Since $\eta\xi$ is an integral of the equations in $\mathcal{R}$, the projections of orbits in the $\eta$-$\xi$ plane move on the branches of the corresponding hyperbolas $\eta\xi = $ constant, except in the case $\eta\xi = 0$ (where $\eta = 0$ or $\xi = 0$). If $\eta\xi > 0$, the branches connect the bounding line segments $\eta - \xi = \pm c$ and if $\eta\xi < 0$, they have both end points on the same segment. A check of equation (2.26) shows that the orbits move as indicated by the arrows in Figure 2.8.

To interpret Figure 2.8(b) as a flow in $\mathcal{R}$, notice that each point in the projection corresponds to a circle in $\mathcal{R}$ given by the "radius" variable $\rho = |\zeta|^2 = $ constant. Recall from (2.25) that $|\zeta|^2 = \frac{2}{\nu}(\varepsilon - \lambda\eta\xi)$. Of course, for points on the bounding hyperbolic segments ($\eta\xi = \varepsilon/\lambda$), the constant is zero so that the circle collapses to a point. Thus, the segments of the lines $\eta - \xi = \pm c$ in the projection correspond to the two-spheres bounding $\mathcal{R}$. This is because each corresponds to a circle crossed with an interval where the two end circles are pinched to a point.

We distinguish nine classes of orbits grouped into the following four categories:

1. The point at the origin in Figure 2.8(b), $\xi = \eta = 0$, corresponds to a ***periodic*** orbit in $\mathcal{R}$, known as the ***Lyapunov orbit*** (in, e.g., Szebehely [1967]).

2. The four half-open segments on the axes, $\eta\xi = 0$ (or equivalently $|\zeta|^2 = \rho^*$ where $\rho^* = 2\varepsilon/\nu$), correspond to four cylinders of orbits asymptotic to this periodic solution
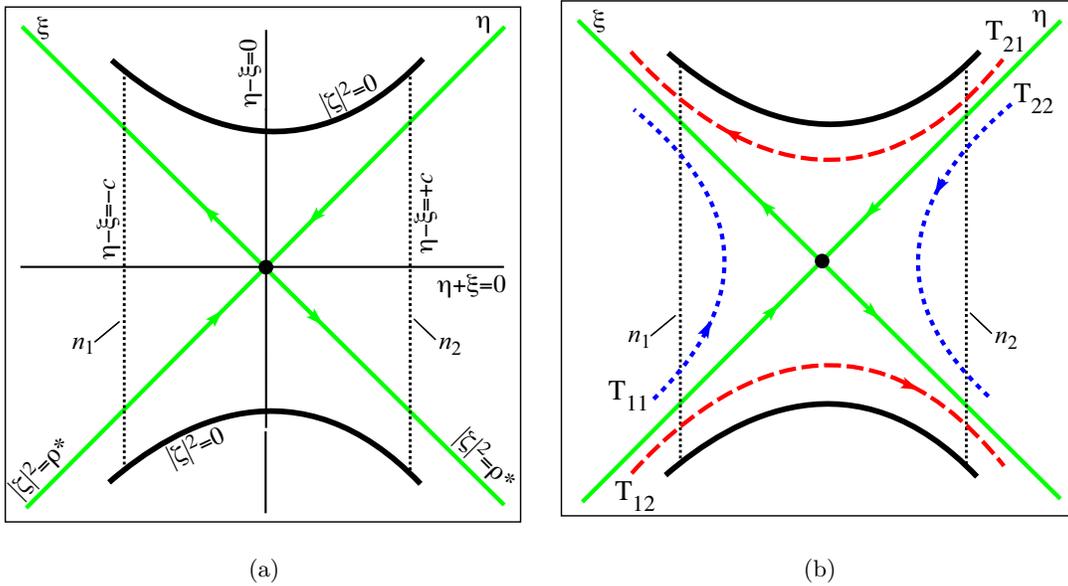
Figure 2.8: **The projection onto the $\eta$-$\xi$ plane** of orbits near the equilibrium point (note, axes tilted 45°). (a) The equilibrium region, $\mathcal{R}$, is bounded by the thick hyperbolic segments at top and bottom and the dotted vertical segments at left and right. At the origin is the periodic orbit in $\mathcal{R}$. The thick lines with arrows pointing toward or away from the origin are trajectories asympototically winding onto the periodic orbit. See the text for further descriptions. (b) Four additional trajectories are shown. The labeling $T_{ij}$ denotes the path of a particle which entered $\mathcal{R}$ through $n_i$ and exited through $n_j$. Two transit orbits, $T_{12}$ and $T_{21}$, and two non-transit orbits, $T_{11}$ and $T_{22}$, are shown.

either as time increases ($\xi = 0$) or as time decreases ($\eta = 0$). These are called **asymptotic** orbits. The are drawn as the thick lines with arrows pointing toward or away from the origin in Figures 2.8(a) and (b).

3. The hyperbolic segments determined by $\eta\xi = $ constant $> 0$ (or equivalently $|\zeta|^2 < \rho^*$) correspond to two cylinders which cross $\mathcal{R}$ from one bounding sphere to the other, meeting both in the same hemisphere; the north one if they go from $\eta - \xi = +c$ to $\eta - \xi = -c$, the south one in the other case. Since these orbits transit from one region to another, we call them **transit** orbits. The two trajectories labeled $T_{12}$ and $T_{21}$ in Figure 2.8(b) are transit orbits.

4. Finally the hyperbolic segments determined by $\eta\xi = $ constant $< 0$ ($|\zeta|^2 > \rho^*$) correspond to two cylinders of orbits in $\mathcal{R}$ each of which runs from one hemisphere to the other hemisphere on the same bounding sphere. Thus if $\xi > 0$, the sphere is

$n_1$ ($\eta - \xi = -c$) and orbits run from the south ($\eta + \xi < 0$) to the north ($\eta + \xi > 0$) hemisphere while the converse holds if $\xi < 0$, where the sphere is $n_2$. Since these orbits return to the same region, we call them **non-transit** orbits. See the two trajectories labeled $\text{T}_{11}$ and $\text{T}_{22}$ in Figure 2.8(b).

**McGehee Representation of the Equilibrium Region.** McGehee [1969], building on the work of Conley [1968], proposed a representation which makes it easier to visualize the region $\mathcal{R}$. Recall that $\mathcal{R}$ is homeomorphic to $S^2 \times I$. In McGehee [1969], it is represented by a spherical annulus, as shown in Figure 2.9(b).
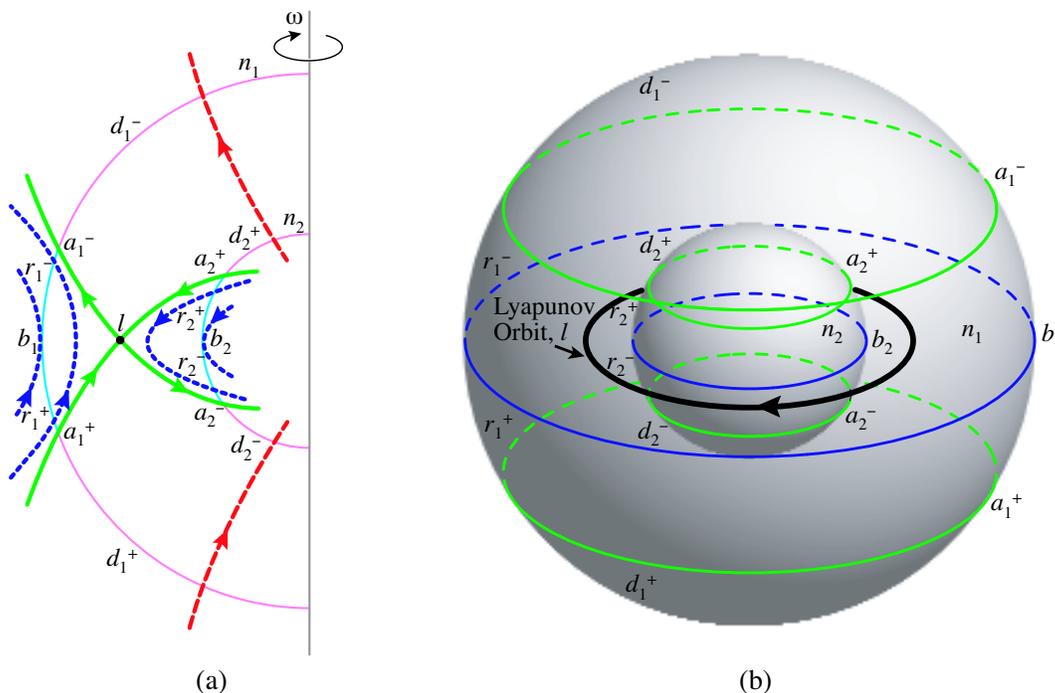


Figure 2.9: **McGehee representation of the equilibrium region.** (a) The cross section of the flow in the $\mathcal{R}$ region of the energy surface. (b) The McGehee representation of the flow in the region $\mathcal{R}$. See the text for details.

Figure 2.9(a) is a cross section of $\mathcal{R}$. Notice that this cross section is qualitatively the same as the illustration in Figure 2.8. The full picture (Figure 2.9(b)) is obtained by rotating this cross section, about the indicated axis $\omega$. The following classifications of orbits correspond to the previous four categories:

1. There is an unstable *periodic* orbit $l$, the Lyapunov orbit in the region $\mathcal{R}$.

2. Again let $n_1, n_2$ be the bounding spheres of region $\mathcal{R}$, and let $n$ denote either $n_1$ or $n_2$. We can divide $n$ into two hemispheres: $n^+$, where the flow enters $\mathcal{R}$, and $n^-$, where the flow leaves $\mathcal{R}$. We let $a^+$ and $a^-$ (where $|\zeta|^2 = \rho^*$) be the intersections with $n$ of the cylinders of orbits *asymptotic* to the unstable periodic orbit $l$. As shown in Figure 2.9(b), $a^+$ appears as a circle in $n^+$, and $a^-$ appears as a circle in $n^-$.

3. If we let $d^+$ be the spherical cap (where $|\zeta|^2 < \rho^*$) in $n^+$ bounded by $a^+$, shown in Figure 2.9(b), then the *transit* orbits entering $\mathcal{R}$ on $d^+$ exit on $d^-$ of the other bounding sphere. Similarly, letting $d^-$ ($|\zeta|^2 < \rho^*$) be the spherical cap in $n^-$ bounded by $a^-$, the transit orbits leaving on $d^-$ have come from $d^+$ on the other bounding sphere.

4. Note that the intersection $b$ of $n^+$ and $n^-$ is a circle of tangency points. Orbits tangent at this circle "bounce off," i.e., do not enter $\mathcal{R}$ locally. Moreover, if we let $r^+$ be a spherical zone which is bounded by $a^+$ and $b$, then *non-transit* orbits entering $\mathcal{R}$ on $r^+$ (where $|\zeta|^2 > \rho^*$) exit on the same bounding sphere through $r^-$ (where $|\zeta|^2 > \rho^*$) which is bounded by $a^-$ and $b$.

**Invariant Manifold Tubes as Separatrices.** The key observation here is that the asymptotic orbits are pieces of the stable and unstable manifold tubes of the Lyapunov orbit and they separate two distinct types of motion: transit orbits and non-transit orbits. The transit orbits, passing from one region to another, are those inside the cylindrical manifold, or **tube**. The non-transit orbits, which bounce back to their region of origin, are those outside the tube. This observation will be important for the numerical construction of interesting orbits in Chapters 4 and 5.

## 2.8 Flow Mappings in the Equilibrium Region

We now observe that on the two bounding spheres, each of the hemispheres $n^\pm$ is transverse to the flow. It follows that the flow in $\mathcal{R}$ defines four mappings—two between pairs of spherical caps $d^\pm$ and two between pairs of spherical zones $r^\pm$ (as in Llibre, Martinez and

Simó [1985]):

$$\psi_1 : d_1^+ \to d_2^-, \qquad \psi_2 : d_2^+ \to d_1^-, \tag{2.28}$$

$$\psi_3 : r_1^+ \to r_1^-, \qquad \psi_4 : r_2^+ \to r_2^-. \tag{2.29}$$

The four mappings are diffeomorphisms. Furthermore, all these mappings preserve the radius variable $\rho = |\zeta|^2$ since this is an integral in $\mathcal{R}$.

**The Infinite Twisting of the Mappings.** After computing from the solution (2.26) that

$$\frac{d}{dt} \ \mathrm{arg} \ \zeta = -\nu, \tag{2.30}$$

we see that the change in the argument of $\zeta$ for each of these mappings $\psi_i$ is approximately proportional to the negative of the time required to go from domain to range.

Also, this time approaches infinity as the flow approaches the circle $a^+$ ($|\zeta|^2 \to \rho^*$), since on the circle $a^+$ (where $|\zeta|^2 = \rho^*$) the orbits are asymptotic to the unstable periodic solution $l$. The proof is quite straightforward. Take $\psi_2$ as an exmaple. According to equations (2.26), we have $\xi(0) = \xi^0, \eta(0) = \eta^0$ on $d_2^+$ where $\eta^0$ and $\xi^0$ are both positive and

$$\eta^0 - \xi^0 = +c.$$

Similarly, if $T$ is the time required to go from domain to range, then $\xi(T) = \xi^0 e^{\lambda T}$ and $\eta(T) = \eta^0 e^{-\lambda T}$ on $d_1^-$ where

$$\eta(T) - \xi(T) = \eta^0 e^{-\lambda T} - \xi^0 e^{\lambda T} = -c.$$

Eliminating $c$ from the two above equations and solving for $T$, we obtain

$$T = \frac{1}{\lambda} \ln \frac{\eta^0}{\xi^0}.$$

Moreover, the energy integral (2.25) gives

$$\xi^0 \eta^0 = \frac{\varepsilon}{\lambda} - \frac{\nu}{2\lambda}|\zeta|^2 = \frac{\nu}{2\lambda}\left(\frac{2\varepsilon}{\nu} - |\zeta|^2\right) = \frac{\nu}{2\lambda}(\rho^* - \rho).$$

Hence,

$$T = \frac{1}{\lambda} \left( \ln \frac{2\lambda(\eta^0)^2}{\nu(\rho^* - \rho)} \right) = \frac{1}{\lambda} \left( \ln \frac{2\lambda(\eta^0)^2}{\nu} - \ln(\rho^* - \rho) \right)$$

where the last term determines the order of the required transit time.

These facts imply that arbitrary circles with radius variable $\rho = |\zeta|^2$ in the domain of the mappings are rotated by an amount that decreases to minus infinity as $\rho \to \rho^*$. Hence, the behavior of the flow in $\mathcal{R}$ should be obtained by adding some spiraling to the arrows given in Figure 2.9(a).

In Chapters 3 and 5, we will need a simple geometric consequence of the above observation on spiraling stated in terms of "abutting arcs" in the domain, or range of $\psi_i$. Namely, an arc lying in the closure of one of these sets ($d^\pm$ and $r^\pm$) is called an abutting arc if it is in the set itself except for one end point in the circle $a^\pm$. See Figure 2.10. For
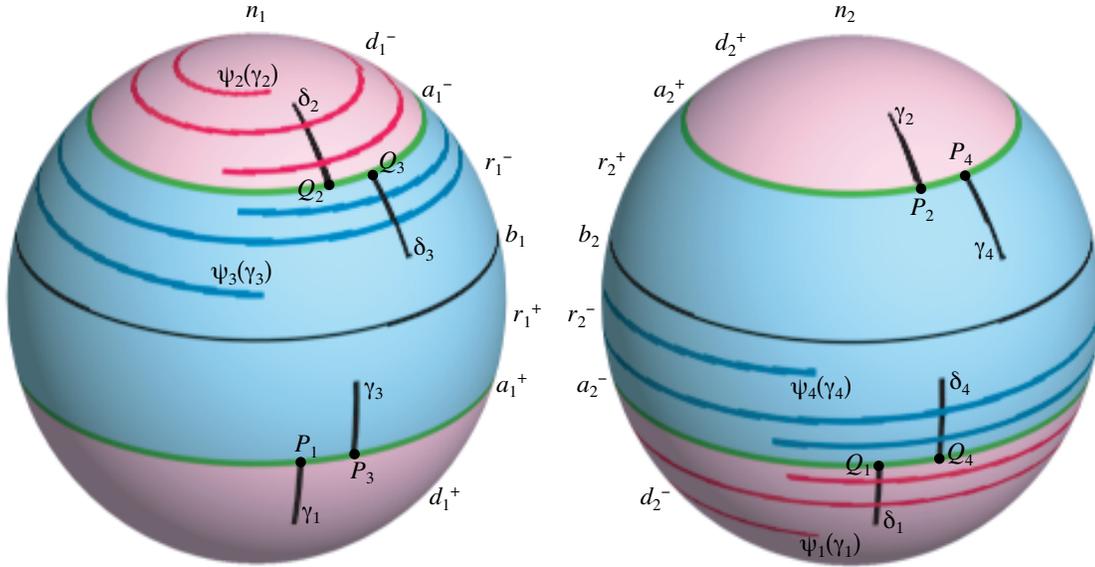


Figure 2.10: Spiraling of the images of arcs $\gamma_i$.

example, let $\gamma_1$ be an abutting arc in the domain $d_1^+$ of $\psi_1$ with one end point $P_1$ in $a_1^+$. Let $\delta_1$ be another abutting arc in the range $d_2^-$ of $\psi_1$ such that one of its end point $Q_1$ is in $a_2^-$. Then $\psi_1(\gamma_1)$ is an arc spiraling towards $a_2^-$ and cutting $\delta_1$ an infinite number of times in any neighborhood of the point of abutment $Q_1$.

This follows directly from the infinite twisting of the mappings $\psi_1$; namely, the image of $\gamma_1$ spirals infinitely many times around and down to $a_2^-$ in the range.

Similarly, let $\gamma_i$ be an abutting arc in the domain of $\psi_i$ with one end point $P_i$ in $a_2^+, a_1^+, a_2^+$ for $i = 2, 3, 4$, respectively. Let $\delta_i$ be another abutting arc in the range of $\psi_i$ such that one of its end points $Q_i$ is in $a_1^-, a_1^-, a_2^-$ respectively. Then $\psi_i(\gamma_i)$ is an arc spiraling towards $a_1^-, a_1^-, a_2^-$, respectively and cutting $\delta_i$ an infinite number of times in any neighborhood of the point of abutment $Q_i$.

## 2.9  Trajectories in the Neck Region

Having studied the orbit structure in the equilibrium region $\mathcal{R}$ and its projection on the $\eta$-$\xi$ plane, we now examine briefly the appearance of orbits in position space, that is, in the $x$-$y$ plane. In position space, $R$ appears as the neck region connecting two realms, so trajectories in $\mathcal{R}$ will be projected to tarjectories in the neck region.

Recall from §2.6 that the $\xi$ and $\eta$ coordinate axes are the eigenvectors $u_1 = (1, -\sigma, \lambda, -\lambda\sigma)$ and $u_2 = (1, \sigma, -\lambda, -\lambda\sigma)$, respectively. Their projection on the $x$-$y$ plane, $\bar{u}_1 = (1, -\sigma)$ and $\bar{u}_2 = (1, \sigma)$, plays an important role in the study of the appearance of orbits on the position space.

The image of a tilted projection of $\mathcal{R}$ on the $x$-$y$ plane provides the right mental picture. To build physical intuition regarding the flow in the equilibrium region, it is important to study the projection of the different classes of orbits on the $x$-$y$ plane. Here, we summarize the main results of Conley [1968].

Recall from §2.6 that the eigenvalues of the linear system (2.22) are $\pm\lambda$ and $\pm i\nu$ with corresponding eigenvectors $u_1, u_2, w_1, w_2$. Thus, the general (real) solution has the form

$$v(t) = (x(t), y(t), \dot{x}(t), \dot{y}(t)) = \alpha_1 e^{\lambda t} u_1 + \alpha_2 e^{-\lambda t} u_2 + 2\text{Re}(\beta e^{i\nu t} w_1), \qquad (2.31)$$

where $\alpha_1, \alpha_2$ are real and $\beta = \beta_1 + i\beta_2$ is complex. Notice that (2.31), while slightly more complicated, is essentially the same as (2.26).

Upon inspecting this general solution, we see that the solutions on the energy surface fall into different classes depending upon the limiting behavior of $x(t)$ (the $x$ coordinate of $v(t)$) as $t$ tends to plus or minus infinity. Notice that

$$x(t) = \alpha_1 e^{\lambda t} + \alpha_2 e^{-\lambda t} + 2(\beta_1 \cos \nu t - \beta_2 \sin \nu t). \qquad (2.32)$$

Thus, if $t \to +\infty$, then $x(t)$ is dominated by its $\alpha_1$ term. Hence, $x(t)$ tends to minus infinity

(staying on the left-hand side), is bounded (staying around the equilibrium point), or tends to plus infinity (staying on the right-hand side) according to $\alpha_1 < 0, \alpha_1 = 0, \alpha_1 > 0$. See Figure 2.11. The same statement holds if $t \to -\infty$ and $\alpha_2$ replaces $\alpha_1$. Different combinations of the signs of $\alpha_1$ and $\alpha_2$ will give us again the same nine classes of orbits which can be grouped into the same four categories:

1. If $\alpha_1 = \alpha_2 = 0$, we obtain a *periodic* solution which is a Lyapunov orbit. It has been proven in Conley [1968] that this periodic orbit, shown in Figure 2.11, projects onto the $x$-$y$ plane as an ellipse with major axis of length $2\tau\sqrt{\varepsilon/\kappa}$ in the direction of the $y$-axis, and minor axis of length $2\sqrt{\varepsilon/\kappa}$ in the direction of the $x$-axis. The orientation of the orbit is clockwise. Here $\kappa$ $(= -a + b\tau^2 + \nu^2 + \nu^2\tau^2)$ is a constant. Note that the size of the ellipse goes to zero with $\varepsilon$.

2. Orbits with $\alpha_1\alpha_2 = 0$ are *asymptotic* orbits. They are asymptotic to the periodic Lyapunov orbit. It has been proven in Conley [1968] that the asymptotic orbits with $\alpha_1 = 0$ project into the strip $S_1$ in the $x$-$y$ plane centering around $\bar{u}_2$ and bounded by the lines

$$y = \sigma x \pm 2\sqrt{\varepsilon(\sigma^2 + \tau^2)/\kappa}. \tag{2.33}$$

Similarly, asymptotic orbits with $\alpha_2 = 0$ project into the strip $S_2$ centering around $\bar{u}_1$ and bounded by the lines

$$y = -\sigma x \pm 2\sqrt{\varepsilon(\sigma^2 + \tau^2)/\kappa}. \tag{2.34}$$

Notice that the width of the strips goes to zero with $\varepsilon$.

3. Orbits with $\alpha_1\alpha_2 < 0$ are *transit* orbits because they cross the equilibrium region $\mathcal{R}$ from $-\infty$ (the left-hand side) to $+\infty$ (the right-hand side) or vice versa.

4. Orbits with $\alpha_1\alpha_2 > 0$ are *non-transit* orbits.

To study the projection of these last two categories of orbits, Conley [1968] proved a couple of propositions which allows one to determine at each point $(x, y)$ the "wedge" of velocities (if any) in which $\alpha_1\alpha_2 < 0$. See the shaded wedges in Figure 2.11. Since a detailed study will draw us too far afield, we simply state some of the main observations.

In Figure 2.11, $S_1$ and $S_2$ are the two strips mentioned above. Outside of each strip $S_i, i = 1, 2$, the sign of $\alpha_i$ is independent of the direction of the velocity. These signs
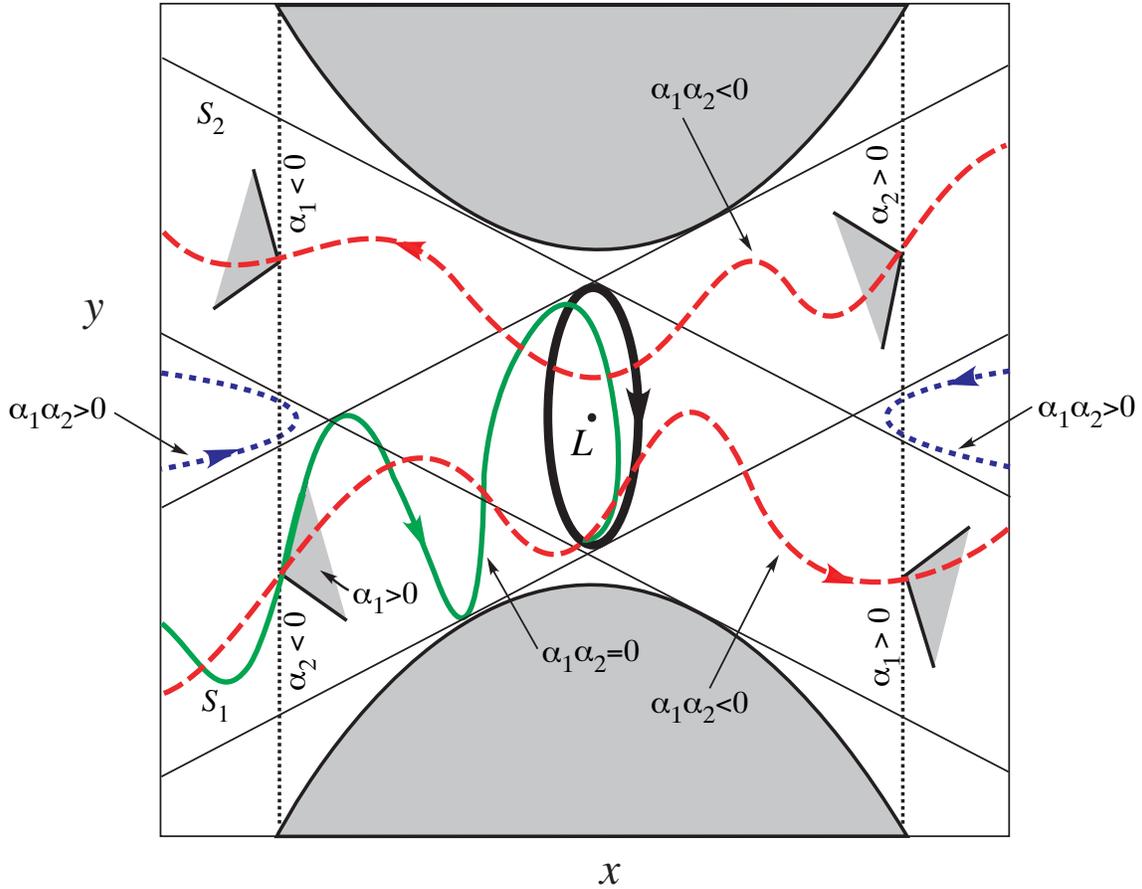
Figure 2.11: **The flow in the equilibrium region $R$ of position space.** Shown are the periodic orbit (ellipse); a typical asymptotic orbit winding onto the periodic orbit; two transit orbits; and two non-transit orbits. See the text for an explanation of the labeling.

can be determined in each of the components of the equilibrium region $R$ complementary to both strips. For example, in the left-most central components, both $\alpha$'s are negative, while in the right-most central components both $\alpha$'s are positive. Therefore, $\alpha_1 \alpha_2 > 0$ in both components and only non-transit orbits project onto these two components.

Inside the strips the situation is more complicated since in $S_i, i = 1, 2$, the signs of $\alpha_i$ depends on the direction of the velocity. For simplicity we have indicated this dependence only on the two vertical bounding line segments in Figure 2.11. For example, consider the intersection of strip $S_1$ with the left-most vertical line. On the subsegment so obtained there is at each point a wedge of velocity in which $\alpha_1$ is positive. The sign of $\alpha_2$ is always negative on this subsegment, so that orbits with velocity interior to the wedge are transit orbits ($\alpha_1 \alpha_2 < 0$). Of course, orbits with velocity on the boundary of the wedge are

asymptotic ($\alpha_1 \alpha_2 = 0$), while orbits with velocity outside of the wedge are non-transit. Here, only a transit and asymptotic orbit are illustrated. The situation on the remaining three subsegments is similar.

**The Flow in the Equilibrium Region.**   In summary, the phase space in the equilibrium region can be partitioned into four categories of distinctly different kinds of motion (see Figures 2.8, 2.9 and 2.11): the periodic Lyapunov orbits, asymptotic orbits, transit orbits, and, finally, non-transit orbits.

# Chapter 3

# Heteroclinic Connection and Global Orbit Structure

**In collaboration with W. Koon, M. Lo, and J. Marsden.**

As mentioned in Chapter 2, near one of the equilibrium points $L_1$ or $L_2$, there is a family of unstable periodic orbits called Lyapunov orbits. For appropriate energy values, the energy surface contains exactly one of these periodic solutions near each libration point. As dynamical systems theory suggests (see, for example, Wiggins [1990]), to understand fully the global dynamics of the flow, one should examine structures like homoclinic and heteroclinic orbits connecting these $L_1$ and $L_2$ Lyapunov orbits to themselves.

In §2.7, the local orbit structure near the libration points was shown to give (i) periodic orbits (the Lyapunov orbits), (ii) pieces of the stable and unstable manifolds of these periodic orbits, (iii) transit and (iv) non-transit orbits. In this chapter, we explore how these local structures are connected globally. Our goal is to show how homoclinic orbits in the interior realm are connected to the homoclinic orbits in the exterior realm by a heteroclinic cycle in the Jupiter realm. We refer to the union of these three structures as a ***chain***. An example is given in Figure 3.1.

The story is completed later in the chapter when this dynamical chain structure is used to show the existence of complex and interesting trajectories, some of which have been observed in actual comet trajectories, as discussed in Koon, Lo, Marsden and Ross [2000]. Guided by ideas laid down in this chapter, we numerically construct some interesting trajectories in Chapters 4 and 5.
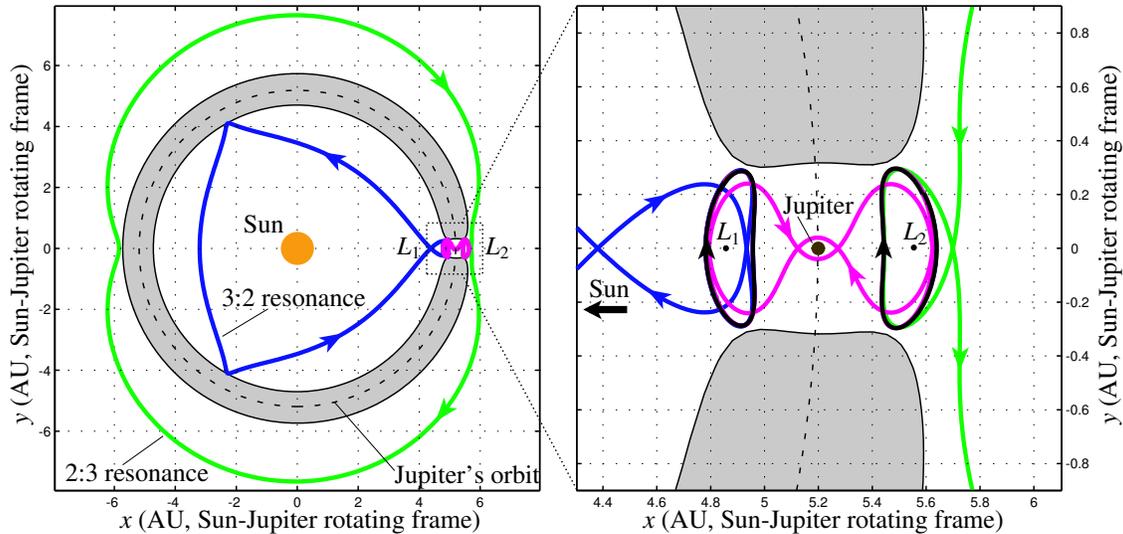
Figure 3.1: **A homoclinic-heteroclinic chain** corresponding to the Jupiter comet P/Oterma. The p.o.'s about $L_1$ and $L_2$ are black. Their homoclinic orbits are labeled with the mean motion resonances with which they are associated. The orbit homoclinic to the $L_1$ p.o. in the interior realm is labeled as the "3:2 resonance," since the particle (or comet) goes around the Sun 3 times while Jupiter goes around the sun 2 times as seen in an inertial frame. The orbit homoclinic to the $L_2$ p.o. in the exterior realm is similarly labeled as the "3:2 resonance." The pair of heteroclinic orbits connecting the $L_1$ and $L_2$ p.o.'s is also shown. These four structures together form a *chain*. We say this chain corresponds to the comet P/Oterma because this comet follows a trajectory close to this chain, as will be elaborated upon in this chapter and in Chapter 6. Distances are given in Astronomical Units (AU, about 150 million km).

**Organization of the Chapter.** In more detail, this chapter discusses the following topics. In §3.1, we discuss some of the results from Conley [1968] and McGehee [1969], which prove the existence of ***homoclinic*** orbits in both the interior and exterior realms. These are the orbits which are both forward and backward asymptotic to an unstable Lyapunov orbit. The heart of the proof is the construction of a function which counts the number of times an orbit segment with endpoints near the Lyapunov orbit winds around a solid torus.

We discuss in §3.2 the main results in Llibre, Martinez and Simó [1985] on the transversality of the invariant manifolds for the $L_1$ Lyapunov orbit. In dynamical systems theory, the property of being doubly asymptotic to a periodic orbit is described (and more quantitatively handled) by saying that the orbit is in both the *stable* and *unstable* manifold

of the periodic orbit, or that the homoclinic orbit is in the intersection of the stable and unstable manifolds of the periodic orbit. One of the most important issues which arises in this context is the transversality of the intersection. The presence of transversality will allow us to draw many profound conclusions about the orbit structure of the system under study.

Since neither Conley [1968] nor McGehee [1969] was able to settle this issue, Llibre, Martinez and Simó [1985] spent their major effort in proving analytically that the intersection is indeed transversal under appropriate conditions, at least in the interior realm. We summarize their results.

The theorems given in §3.2 are cited only for guidance on how to construct the transversal homoclinic orbits numerically. In §3.3 we compute intersections of stable and unstable manifolds of $L_1$ and $L_2$ p.o.'s to numerically demonstrate the existence of transversal homoclinic orbits in both the interior and exterior realms.

In §3.4 we use similar computational methods to numerically demonstrate the existence of transversal **heteroclinic** orbits in the Jupiter realm which connect asymptotically the $L_1$ and $L_2$ Lyapunov orbits. A heteroclinic orbit, also known as a *heteroclinic connection*, is an orbit lying in the intersection of the stable manifold of one periodic orbit and the unstable manifold of another periodic orbit. As discussed in Chapter 2, since the PCR3BP is a Hamiltonian system with two degree of freedom, its energy manifold is three-dimensional. From the work of Conley [1968], it was known that both the stable and unstable manifolds of the p.o.'s around $L_1$ and $L_2$ are two-dimensional. Hence, a dimension count suggests, but does not prove, the existence of such a heteroclinic connection. Careful numerical investigations allow us to show this connection is indeed present, as well as to isolate and study it.

It is worth noting that, inspired by these numerical demonstrations which were first reported in Koon, Lo, Marsden and Ross [2000], two teams of authors have rigorously proven the existence of transversal homoclinic and heteroclinic orbits (Wilczak and Zgliczyński [2003] and Kirchgraber and Stoffer [2004]). These are "computer-assisted proofs" using interval analysis methods, and they further verify the claims we make here.

In §3.5, we numerically demonstrate that, within an appropriate range of energy values, there exist chains of two homoclinic orbits and a symmetric heteroclinic cycle, as in Figure 3.1. The existence of these chains will be used in §3.6 to construct a suitable Poincaré

map which will allow us to classify as well as organize distinctively different types of global motions of the PCR3BP in terms of ultimate behavior with respect to the equilibrium points.

In §3.7 and §3.8, we extend the symbolic dynamics results of Llibre, Martinez and Simó [1985] to our situation and construct a set of bi-infinite sequences with two families of symbols.

In §3.9, we state the main theorem of this chapter and discuss its implications. The theorem gives the global orbit structure of the PCR3BP in a neighborhood of a chain of homoclinic orbits and a symmetric heteroclinic cycle.

## 3.1 Existence of Orbits Homoclinic to the Lyapunov Orbit

**Energy Manifold and Hill's Region.** Recall from §2.4 that the motion of a comet with energy $e$ is restricted to a three-dimensional energy surface $\mathcal{M}(\mu, e)$. The projection of the energy surface onto position space, $M(\mu, e)$, is the region of possible motion known as the Hill's region. The first four cases of possible motion, depending on energy, are given in Figure 2.4. For case 3, the region of possible motion contains necks around both $L_1$ and $L_2$ and the comet can transit from the interior realm to the exterior realm and vice versa. This is the case of most interest to us.

**Orbit Segments Winding around a Solid Torus.** From McGehee [1969], we know that the energy surface is broken up further into regions bounded by invariant tori. These invariant tori $\mathcal{A}_1$ and $\mathcal{A}_2$ project onto the darkly shaded annuli $A_1$ and $A_2$, respectively, shown for case 3 in Figure 3.2(a).

These annuli separate the Hill's region into sets corresponding to the invariant sets in the energy surface. It is interesting to note that for all of the cases, the singularities corresponding to the center of the Sun and Jupiter are separated from each other by an invariant torus (although we show only case 3), thus making it impossible for the comet to pass from a region arbitrarily close to the Sun to a region arbitrarily close to Jupiter. Similarly, Jupiter is separated from infinity by an invariant torus. We consider the regions of the energy surface projecting to the area between the two darkly shaded annuli, $A_1$ and $A_2$, i.e., the region containing Jupiter. The theorems of McGehee given below show that all orbits leaving the vicinity of one of the unstable periodic orbits proceed around one of

(a)  (b)
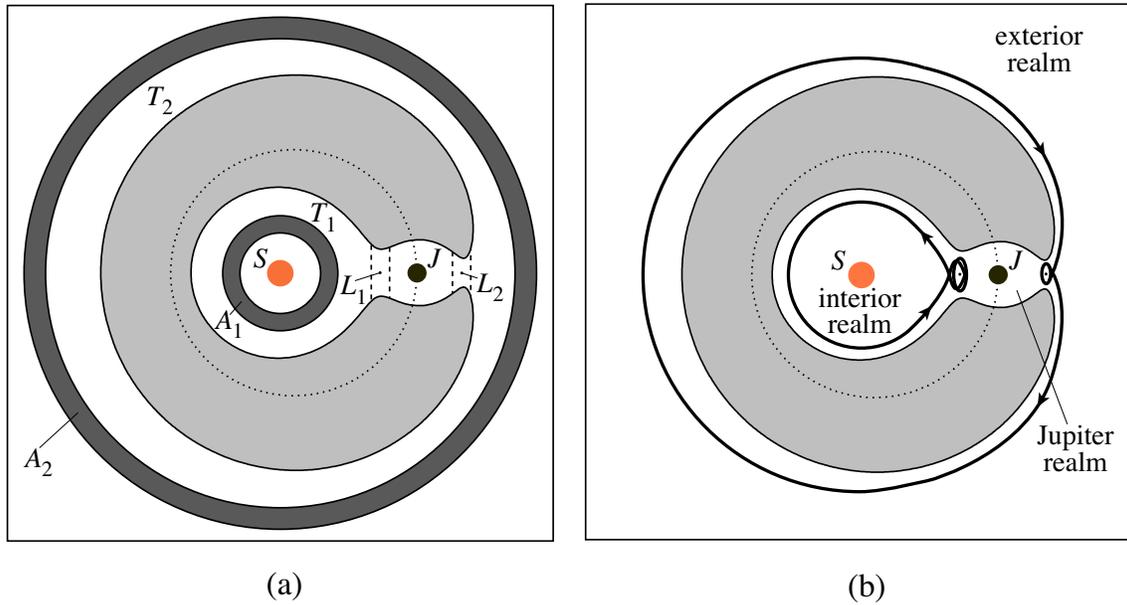
Figure 3.2: (a) The projection of invariant tori (darkly shaded) on position space for case 3. (b) Homoclinic orbits in the interior and exterior realms.

the annuli $T_1$ or $T_2$, projections of *solid tori* $\mathcal{T}_1$ or $\mathcal{T}_2$, before returning to that vicinity. The direction of motion is the same for all orbits, counterclockwise in the interior realm and clockwise in the exterior realm.

In Chapter 2, we studied the regions near the unstable periodic orbits to obtain a qualitative picture of the asymptotic orbits. Here we combine this picture of asymptotic orbits with the fact that orbits in $T_1$ or $T_2$ wind around in one direction *to construct homoclinic orbits* in both the interior and exterior realms, shown schematically in Figure 3.2(b).

We note that the following theorems do not literally apply to the system parameters ($\mu$) and energies ($e$) of interest to us. But they are a useful guide and numerical experiments reveal that the qualitative results they suggest hold over a larger set of system parameters and energies than is proven.

**Theorems of McGehee.**  To precisely state the theorems, we must first divide up the Hill's region and the energy surface. From §2.5, for small $\mu$ the two equilibrium points occur at an approximate distance $r_h = \left(\frac{\mu}{3}\right)^{\frac{1}{3}}$ on either side of Jupiter. We isolate these points by drawing vertical lines on each side of them, i.e., lines at $(1 - \mu \pm c_1 r_h, 0)$ and

$(1 - \mu \pm b_1 r_h, 0)$, where $b_1 < 1 < c_1$. This divides the Hill's region into five sets as shown in Figure 3.3.
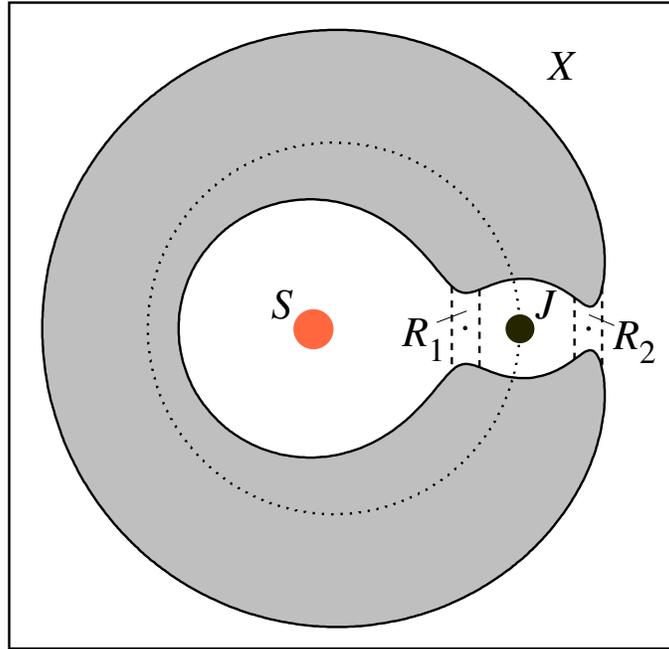


Figure 3.3: Division of Hill's region into five sets.

Let $S$ and $J$ be the large sets, i.e., realms, that contain the Sun and Jupiter, respectively; let region $R_1$ and region $R_2$ be those sets that contain the two equilibrium points $L_1$ and $L_2$, respectively; and let $X$ be the realm that lies exterior to the orbit of Jupiter. We also divide the energy surface $\mathcal{M}$ into sets projecting onto the sets shown in Figure 3.3. We will adopt the notation convention introduced in §2.4 where a set in the energy surface will be denoted by a script letter and its position space projection denoted by an italicized letter: e.g., region $\mathcal{R}_1$ for the set in the energy surface whose projection is the region $R_1$ in the position space. Theorem 3.1.1 leads to the assertion that *one can choose the division described above so that we simultaneously have sufficient control of the flow in both sets $\mathcal{S}$ and $\mathcal{R}_1$ to construct a homoclinic orbit. Theorem 3.1.2 makes the same assertion for sets $\mathcal{X}$ and $\mathcal{R}_2$.*

Let $\mathcal{R}$ denote either $\mathcal{R}_1$ or $\mathcal{R}_2$. As $\mathcal{R}$ is a function not only of $b_1$ and $c_1$, but also $\mu$ and $e$, we sometimes write it as $\mathcal{R}(\mu, e)$.

The analysis of $\mathcal{R}(\mu, e)$ is of a local nature. In fact, we limit ourselves to those energy values $e$ for which the linearized equations about the equilibrium point give us the qual-

itative picture of the flow. The flow for the linearized equations was already analyzed in some detail in Chapter 2.

We know that for $b_1$ and $c_1$ close to 1, i.e., for the region $\mathcal{R}$ close to the p.o., the flow for $\mathcal{R}$ is that shown in Figure 2.11. But we also know that we cannot make $c_1$ arbitrarily large without disturbing this qualitative picture. On the other hand, we would like to make $c_1$ large enough to obtain accurate estimates on the behavior of the flow in $\mathcal{S}$ and $\mathcal{X}$. The following theorems show that there exists a $c_1$ which allows us to balance these two factors. In the theorem, $n_{i,j}$ is the $n_j$ bounding sphere in region $\mathcal{R}_i$.

**Theorem 3.1.1.** **(McGehee)** *There exist constants $b_1$ and $c_1$ and an open set $O_1$ in the $(\mu, e)$-plane (see Figure 3.4(a)) containing the graph of $e = E_1(\mu)$ for small $\mu > 0$ such that, for $(\mu, e) \in O_1$:*

1. *The energy surface $\mathcal{M}(\mu, e)$ contains an invariant torus $\mathcal{A}_1$ separating the Sun from Jupiter.*

2. *For $e > E_1(\mu)$, the flow in $\mathcal{R}_1(\mu, e)$ is qualitatively the same as the flow for the linearized equations. (See Figure 2.11)*

3. *If we let $\mathcal{T}_1$ be that submanifold of $\mathcal{M}$ co-bounded by the invariant torus $\mathcal{A}_1$ and the left bounding sphere $n_{1,1}$ of the equilibrium region $\mathcal{R}_1$ (see Figure 3.4(b)), then there exists a function*

$$\theta : \mathcal{T}_1 \to \mathbb{R}$$

    *such that*

    (a) *$\theta$ is a meridional angular coordinate for $\mathcal{T}_1$;*

    (b) *$\theta$ is strictly increasing along orbits.*

**Theorem 3.1.2.** **(McGehee)** *There exist constants $b_1$ and $c_1$ and an open set $O_2$ in the $(\mu, e)$-plane containing the graph of $e = E_2(\mu)$ for small $\mu > 0$ such that, for $(\mu, e) \in O_2$:*

1. *The energy surface $\mathcal{M}(\mu, e)$ contains an invariant torus $\mathcal{A}_2$ separating the Sun and Jupiter from infinity.*

2. *For $e > E_2(\mu)$, the flow in $\mathcal{R}_2(\mu, e)$ is qualitatively the same as the flow for the linearized equations. (See Figure 2.11)*
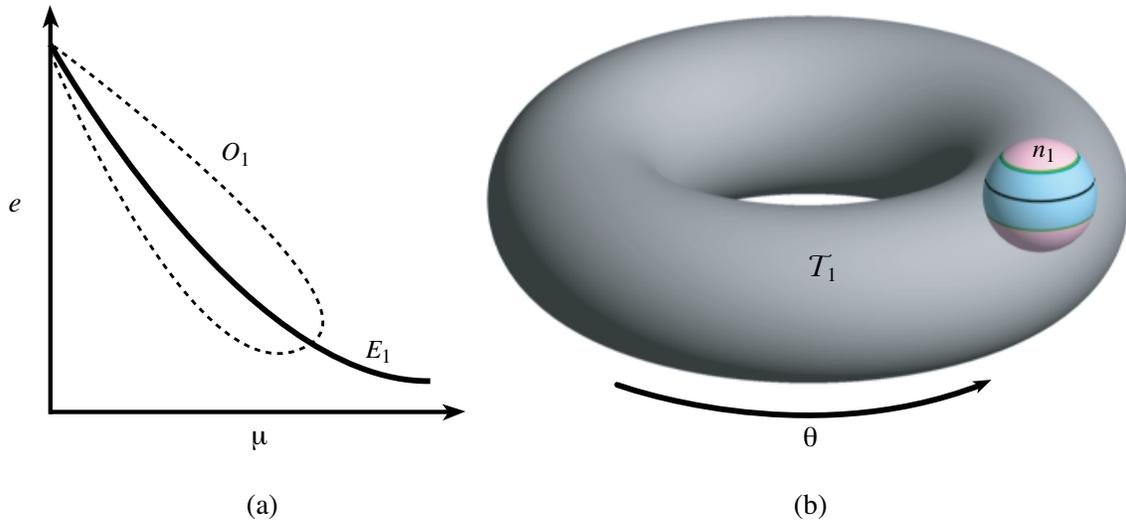
Figure 3.4: (a) Open set $O_1$ in $(\mu, e)$-plane. (b) The region $\mathcal{T}_1$ with meridional angular coordinate $\theta$.

3. *If we let $\mathcal{T}_2$ be that submanifold of $\mathcal{M}$ co-bounded by the invariant torus $\mathcal{A}_2$ and the right bounding sphere $n_{2,2}$ of the equilibrium region $\mathcal{R}_2$, then there exists a function*

$$\theta : \mathcal{T}_2 \to \mathbb{R}$$

*such that*

(a) *$\theta$ is a meridional angular coordinate for $\mathcal{T}_2$;*

(b) *$\theta$ is strictly increasing along orbits.*

See McGehee [1969] for the proofs of these theorems.

Part 3 of the above theorems gives us the following properties for the flow in $\mathcal{T}_1$ and $\mathcal{T}_2$. The increase in $\theta$ along an orbit segment in $\mathcal{T}_1$ (or $\mathcal{T}_2$) with endpoints in the bounding sphere $n_{1,1}$ (or $n_{2,2}$, respectively) is close to a non-zero integer multiple of $2\pi$. The increase in $\theta$ along any other orbit segment which can be deformed to the first, keeping both endpoints in the bounding sphere $n$, is close to the same integer multiple of $2\pi$. Furthermore, the increase of $\theta$ along any orbit segment remaining for an arbitrarily long time in $\mathcal{T}_1$ or $\mathcal{T}_2$ is arbitrary large. As will be shown, these are precisely the properties we need to carry out the proof of the existence of a homoclinic orbit.

**The Existence of Orbits Homoclinic to a Lyapunov Periodic Orbit.** Recall from Chapter 2 that for an energy $e > E_i$, there is a periodic orbit (p.o.) around $L_i$, $i = 1$ or 2, with two-dimensional invariant unstable, $W^u_{L_i,p.o.}$ and stable $W^s_{L_i,p.o.}$, manifolds, the asymptotic orbits which are locally diffeomorphic to cylinders. We recall that a homoclinic orbit related to a periodic orbit $l$ is an orbit that tends to $l$ as $t \to \pm\infty$. Therefore, it is on the stable and unstable invariant manifolds of $l$. A homoclinic orbit is called a **_transversal homoclinic orbit_** if at some point of the orbit the tangent spaces to the stable and unstable manifolds at that point span the full tangent space to $\mathcal{M}(\mu, e)$ at the same point.

We assert that in our problem *either a transversal homoclinic orbit exists, or "total degeneracy" occurs.* Total degeneracy is the case when every orbit asymptotic to the unstable periodic orbit at one end is also asymptotic to the same periodic orbit at the other end and hence is a homoclinic orbit. In other words, the total degeneracy situation occurs when the stable and unstable manifolds of the Lyapunov orbit coincide with each other. In either event we conclude the existence of a homoclinic orbit. We shall sketch the proof below for completeness. For more details, see Conley [1968] and McGehee [1969].

Assume that total degeneracy does not occur. The first step of the proof of the preceding assertion is to find an orbit segment in $\mathcal{T}_1$ connecting either $d_1^-$ to $a_1^+$ or $a_1^-$ to $d_1^+$ as follows. Consider Figure 3.5, where we show an example of the latter, and where $n_1$ denotes $n_{1,1}$. Since $\mathcal{T}_1$ is compact and our flow, which is Hamiltonian, preserves
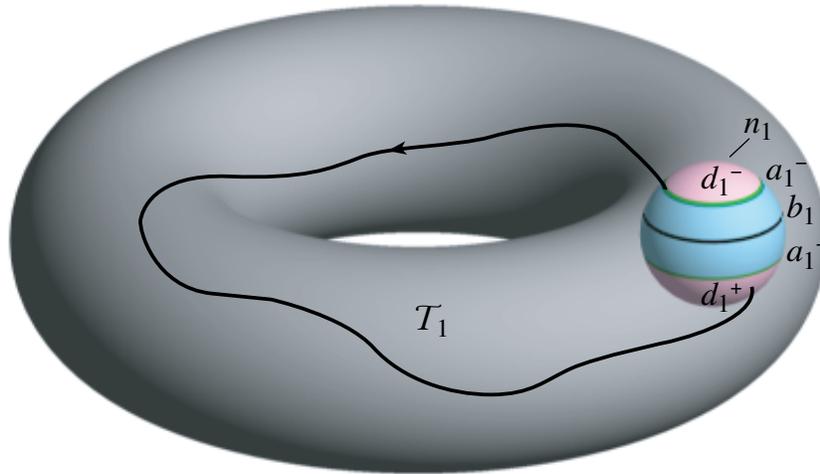


Figure 3.5: The existence of orbits homoclinic to the Lyapunov orbit.

a nondegenerate area element, we can conclude that some orbit which leaves $\mathcal{R}_1$ (and crosses the bounding sphere $n_1$) and so enters $\mathcal{T}_1$ must also leave $\mathcal{T}_1$ and re-enter $\mathcal{R}_1$ (and recross $n_1$). Therefore, for some point $p \in d_1^-$ of $n_1$, there is an orbit segment connecting $p$ to a point $q \in d_1^+$ of $n_1$. Recall from §2.7 and Figure 2.9 that in $\mathcal{R}_1$, the spherical caps $d_1^-$ and $d_1^+$ are where the flow crosses $n_1$.

Starting with this orbit segment connecting $p$ to $q$, we can find an orbit segment connecting either $d_1^-$ to $a_1^+$ or $a_1^-$ to $d_1^+$ as follows. Let $\gamma$ be an arc in $d_1^-$ linking $p$ to $a_1^-$ (where $\gamma \cap a_1^-$ is not on a homoclinic orbit). If all of $\gamma$ is carried by the flow to the spherical cap $d_1^+$, then we shall have an orbit segment with one endpoint in $a_1^-$ and the other in $d_1^+$. Otherwise, starting from $p$, there is some maximal initial half-open subarc $\gamma'$ of $\gamma$ which is carried by the flow to $d_1^+$. Let $r$ be the first point of $\gamma$ not in $\gamma'$, then the orbit segment with one endpoint at $r$ must become arbitrarily long. But the only way this orbit segment can become arbitrarily long is to approach the asymptotic set, since the number of times it can wind around $\mathcal{T}_1$ is finite and therefore must contain an arbitrarily long subsegment in $\mathcal{R}_1$. Because of our knowledge of the flow in $\mathcal{R}_1$, we know that long orbit segments in $\mathcal{R}_1$ must lie close to the cylinders of asymptotic orbits and therefore $r$ must be carried to $a_1^+$. Hence, in either case we conclude that there is an orbit segment connecting the set $d_1^\pm$ in one hemisphere to the set of asymptotic orbits in the other.

Without loss of generality, we can suppose that we have found an orbit segment with one endpoint, called $\alpha$, in $a_1^-$ and the other in $d_1^+$. We now choose for $\gamma$ the whole set $a_1^-$. Using arguments similar to the above, we can conclude that either all of $a_1^-$ is carried by the flow inside $d_1^+$, or there exists a point $\beta \in a_1^-$ such that the orbit segment with $\beta$ as an endpoint becomes asymptotic at the other end. If the first possibility holds, we would have a map of $d_1^-$ to the interior of $d_1^+$, contradicting area preservation of Hamiltonian flow. Thus we have proven that either transversal homoclinic orbits exist or total degeneracy occurs for the interior realm. The same proof also works for the exterior realm.

## 3.2 Existence of Transversal Homoclinic Orbits in the Interior Realm

Conley [1968] and McGehee [1969] did not settle the issue of when one has transversality of the homoclinic orbit families for the PCR3BP, since total degeneracy was a possibility they

could not rule out analytically. Subsequently, Llibre, Martinez and Simó [1985] devoted their major effort to show that under appropriate conditions, the invariant manifolds of the $L_1$ Lyapunov orbits do meet transversely. In this section, we summarize their analytical results. Moreover, in §3.4 we explore numerically the existence of transversal homoclinic orbits in both the interior and exterior realms.

To state the major analytical results of Llibre, Martinez and Simó [1985], we first need to set up some notation. As mentioned earlier, near $L_1$ and for energy values $e$ where $E_1 < e < E_2$ (case 2) there is a family of unstable Lyapunov orbits. When $e$ approaches $E_1$ from above, the periodic orbit tends to $L_1$. There are one-dimensional invariant stable, $W^s_{L_1}$, and unstable, $W^u_{L_1}$, manifolds associated to $L_1$.

Notice that equations (2.15) have the following symmetry

$$s : (x, y, v_x, v_y, t) \rightarrow (x, -y, -v_x, v_y, -t). \tag{3.1}$$

Therefore, if we know the unstable manifold of $L_1$ of the Lyapunov orbit (which is a symmetrical periodic orbit) the corresponding stable manifold is obtained through the use of the stated symmetry. This observation will be used in later sections to find the transversal homoclinic orbits.

**Analytical Results for $L_1$ Lyapunov Orbit in Interior Realm.** Using the basic framework developed in McGehee [1969], Llibre, Martinez and Simó [1985] were able to prove the following two theorems. Together these two theorems imply that for sufficiently small $\mu$ and for an appropriate range of $\Delta E = e - E_1$, the invariant manifolds $W^{s,\mathcal{S}}_{L_1,\text{p.o.}}$ and $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ in the interior realm $\mathcal{S}$ intersect transversely.

**Theorem 3.2.1.   (Llibre-Martinez-Simó)** *For $\mu$ sufficiently small, the branch $W^{u,\mathcal{S}}_{L_1}$ of $W^u_{L_1}$ in the interior realm $\mathcal{S}$ has a projection on position space (see Figure 3.6(a)) given by*

$$\begin{aligned} d &= \mu^{1/3}\left(\frac{2}{3}N - 3^{1/6} + M\cos t + o(1)\right), \\ \alpha &= -\pi + \mu^{1/3}(Nt + 2M\sin t + o(1)), \end{aligned}$$

*where $d$ is the distance to the zero velocity curve, $\alpha$ is the angular coordinate and $N$ and $M$ are constants.*

*In particular, for a sequence of values of μ which have the following asymptotic expression:*

$$\mu_k = \frac{1}{N^3 k^3}(1 + o(1)), \tag{3.2}$$

*the first intersection of this projection with the x-axis is orthogonal to that axis, giving a symmetric (1,1)-homoclinic orbit for $L_1$. The prefix (1,1) refers to the first intersection (with the Poincaré section defined by the plane $y = 0, x < 0$) of both the stable and unstable manifolds of $L_1$.*
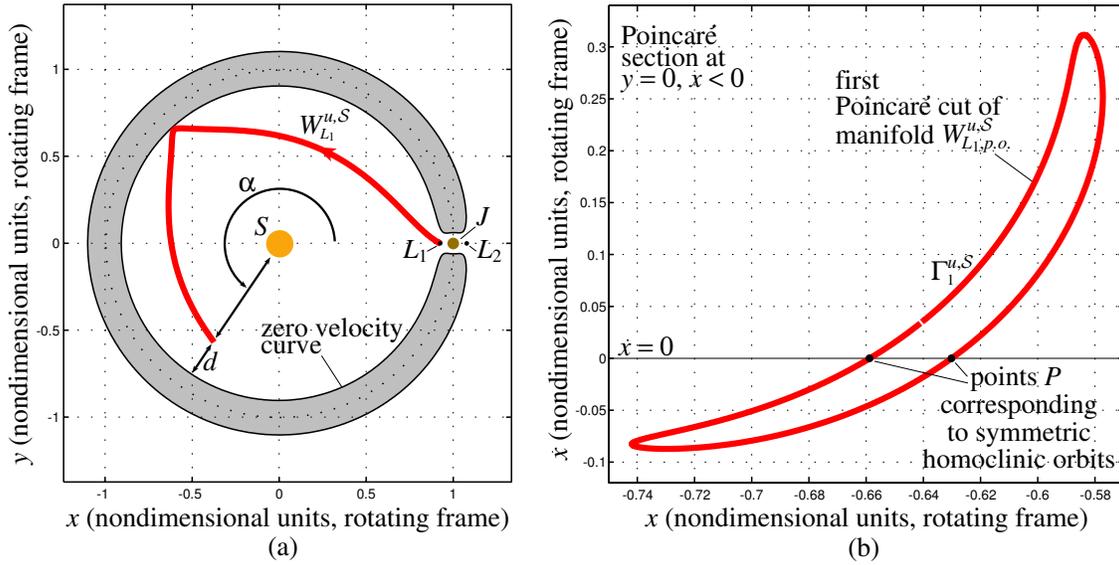


Figure 3.6: (a) Projection of the interior branch of the manifold $W^u_{L_1}$ on the position space. (b) First intersection (Poincaré "cut") $\Gamma^{u,\mathcal{S}}_1$ of the interior branch of $W^u_{L_1,\text{p.o.}}$ with the plane $y = 0, x < 0$.

**Theorem 3.2.2. (Llibre-Martinez-Simó)** *For μ and $\Delta E = e - E_1$ sufficiently small, the branch $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ of $W^u_{L_1,\text{p.o.}}$ contained initially in the interior realm $\mathcal{S}$ of the energy surface intersects the plane $y = 0$ for $x < 0$ in a curve diffeomorphic to a circle (see Figure 3.6(b)).*

*In particular, for points in the $(\mu, e)$ plane such that there is a $\mu_k$ of Theorem 3.2.1 for which*

$$\Delta E > L\mu_k^{4/3}(\mu - \mu_k)^2 \tag{3.3}$$

*holds (where L is a constant), there exist symmetric transversal (1,1)-homoclinic orbits.*

For details of the proofs, see Llibre, Martinez and Simó [1985]. We would like to make

a few comments about these results which are pertinent to the main thrust of this chapter.

1. The main objective of both theorems (3.2.1 and 3.2.2) is to study the transversality of the invariant manifolds for the $L_1$ Lyapunov orbit on the energy surface whose energy $e$ is slightly greater than $E_1(\mu)$ as one varies $\mu$ and $e$. The main step is to obtain an expression for the first intersection of the unstable manifold $W_{L_1,\text{p.o.}}^{u,\mathcal{S}}$ with the plane $y = 0$ in the region $x < 0$, which we label $\Gamma_1^{u,\mathcal{S}}$. While formulas were provided in Llibre, Martinez and Simó [1985] for this closed curve as a function of $\mu$ and $\Delta E$ in the variables $x, \dot{x}$, they are quite complicated and difficult to interpret and hence are not included here. But the key point is the following. According to Theorem 3.2.1, the set of values of $\mu$ for which we have a symmetric (1,1)-homoclinic orbit associated to $L_1$ is discrete and is given by equation (3.2). Then for any other value of $\mu$ the unstable manifold $W_{L_1}^{u,\mathcal{S}}$ of $L_1$ reaches the $(x, \dot{x})$-plane in a point $(x_1, \dot{x}_1)$ outside $\dot{x} = 0$. Therefore, if $\Delta E$ is too small, $\Gamma_1^{u,\mathcal{S}}$ does not cut the $x$-axis and hence (by symmetry), the intersection $\Gamma_1^{s,\mathcal{S}}$ of the stable manifold $W_{L_1,\text{p.o.}}^{s,\mathcal{S}}$ with the plane $y = 0, x < 0$ does not cut the $x$-axis either. Therefore the first intersections of the invariant manifolds do not meet and there is no symmetric (1,1)-homoclinic orbit.

However, for a fixed value of $\mu$, if we increase $\Delta E$, we hope that $\Gamma_1^{u,\mathcal{S}}$ of the unstable manifold will become large. Therefore we can look for some value of $\Delta E$ such that $\Gamma_1^{u,\mathcal{S}}$ becomes tangent to the $x$-axis or even intersects it at more than one point. Then, due to the symmetry of the PCR3BP (3.1), $\Gamma_1^{s,\mathcal{S}}$ of the stable manifold also intersects the $x$-axis at the same points. Points $P$ on the $x$-axis where $\Gamma_1^{u,\mathcal{S}}$ and $\Gamma_1^{s,\mathcal{S}}$ intersect correspond to (symmetric) orbits homoclinic to the Lyapunov orbit (see Figure 3.6(b)). If $\Gamma_1^{u,\mathcal{S}}$ is transversal to $\Gamma_1^{s,\mathcal{S}}$ at $P$ then the homoclinic orbit is transversal. The results of Theorem 3.2.2 say that the above phenomenon occurs if $\Delta E > L\mu_k^{4/3}(\mu - \mu_k)^2$ holds.

2. Using the results of Theorem 3.2.2, Llibre, Martinez and Simó [1985] were able to draw the mesh of homoclinic tangencies for the $(\mu, \Delta E)$-plane. The numbers in Figure 3.7 show the number of symmetric (1,1)-homoclinic points found in the first intersection of $W_{L_1,\text{p.o.}}^{u,\mathcal{S}}$ with the plane $y = 0, x < 0$ when one varies $\mu$ and $\Delta E$. For us, the key point of the theorems is that for the wide range of $\mu$ which exist in the solar system, the invariant manifolds of the $L_1$ Lyapunov orbit intersect transversely for sufficiently large $\Delta E$.

3. The heart of the proofs of these two theorems is to obtain expressions for $W_{L_1}^{u,\mathcal{S}}$ as a
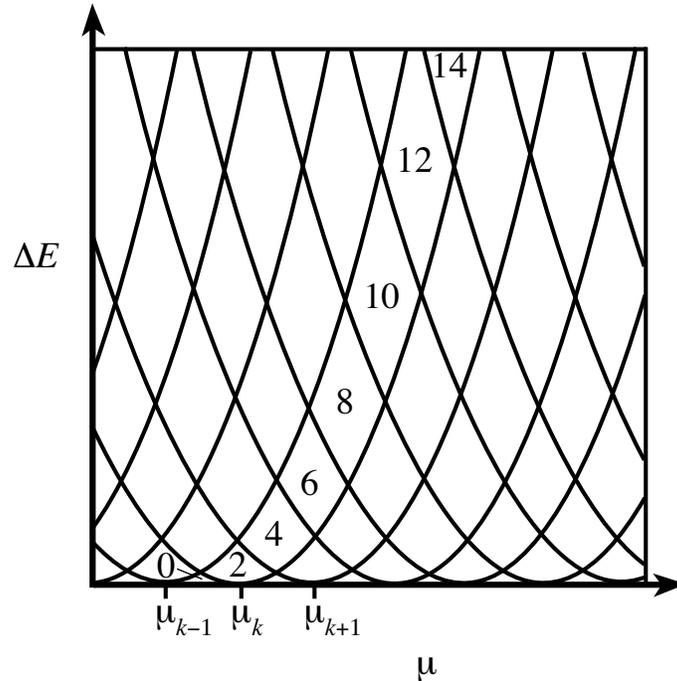
Figure 3.7: Partition of the $(\mu, \Delta E)$-plane according to the number of symmetric (1,1)-homoclinic points found in the first intersection of $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ with the plane $y = 0, x < 0$.

function of $\mu$ and for $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ as a function of $\mu$ and $\Delta E$. By using the basic framework of McGehee [1969], Llibre, Martinez and Simó [1985] divided the annulus $T_1$ in the interior realm $S$ into two sets: a small neighborhood $H$ near $R_1$ and the rest of the realm outside this small neighborhood. In the neighborhood $H$, the PCR3BP can be considered as a perturbation of the Hill's problem. In celestial mechanics, it is well known that Hill's problem studies the behavior near the small mass of the PCR3BP in the limit when $\mu$ approaches zero. In the rest of the realm away from the small mass, the PCR3BP can be approximated by the two-body problem in a rotating frame. Through a number of careful estimations, Llibre, Martinez and Simó [1985] were able to obtain these analytical results.

**Summary.** Conley [1968] and McGehee [1969] proved the existence of homoclinic orbits for both the interior and exterior realm, and Llibre, Martinez and Simó [1985] showed analytically the existence of transversal symmetric (1,1)-homoclinic orbits in the interior realm under appropriate conditions. For our problem, we need to find transversal homoclinic orbits in both interior and exterior realms as well as transversal heteroclinic cycles for the $L_1$ and $L_2$ Lyapunov orbits. In the following sections, we perform some numerical

explorations using the methods described in Chapter 4. For more details on finding invariant manifolds numerically, see Gómez, Jorba, Masdemont and Simó [1991] and references therein.

## 3.3 Existence of Transversal Homoclinic Orbits in the Exterior Realm

We turn our attention now to numerical explorations of the problem, and in particular, to the existence of transversal homoclinic orbits for the $L_2$ Lyapunov orbit in the exterior realm. Though there are no analytical results proving the existence of transversal homoclinic orbits in the $\mathcal{X}$ realm, we can construct them numerically by finding an intersection of the manifolds $W^s_{L_2,\text{p.o.}}$ and $W^u_{L_2,\text{p.o.}}$ on an appropriately chosen Poincaré section.

Numerical experiments guided by geometrical insight suggest that we cut the flow by the plane $y = 0$, the line passing through the two masses in the rotating frame. The branch of the manifold $W^u_{L_2,\text{p.o.}}$ which enters the $\mathcal{X}$ realm flows clockwise in the position space, as shown in See Figure 3.8(a). We refer to this exterior branch of the manifold as $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$. Outside of a neighborhood of $n_2$ in the $\mathcal{X}$ realm, this two-dimensional manifold tube $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ first intersects the plane $y = 0$ on the part of $\mathcal{T}_2$ which is opposite to $L_2$ with respect to the Sun (i.e., $x < 0$). The intersection shown in Figure 3.8(b) is a curve diffeomorphic to a circle, as one would expect geometrically. We call this intersection the first *cut* of the tube $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ with $y = 0$. Some arcs of this curve produce successive intersections without leaving the $\mathcal{X}$ realm. The $q$-th of these intersections of $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ with $y = 0$ will be referred to as $\Gamma^{u,\mathcal{X}}_q$. In a similar manner we call $\Gamma^{s,\mathcal{X}}_p$ the corresponding $p$-th intersection with $y = 0$ of $W^{s,\mathcal{X}}_{L_2,\text{p.o.}}$.

A point in $y = 0$ belonging to $\Gamma^{u,\mathcal{X}}_q \cap \Gamma^{s,\mathcal{X}}_p$ (if not empty) will be called a $(q,p)$-***homoclinic point***. The existence of $(q,p)$-homoclinic points for certain $q$ and $p$ is shown in McGehee [1969].

Our goal is to obtain the first such transversal intersection of $\Gamma^{u,\mathcal{X}}_q$ with $\Gamma^{s,\mathcal{X}}_p$ and so obtain a *transversal $(q,p)$-homoclinic point*. The $(q,p)$-homoclinic point $P$ is transversal if $\Gamma^{u,\mathcal{X}}_q$ and $\Gamma^{s,\mathcal{X}}_p$, which necessarily intersect at $P$, do so transversely: that is, their tangent spaces span the $(x,\dot{x})$-plane at $P$. Other intersections (for larger $q$ and $p$) may exist, but we will restrict ourselves for now to the first. Suppose that the unstable manifold
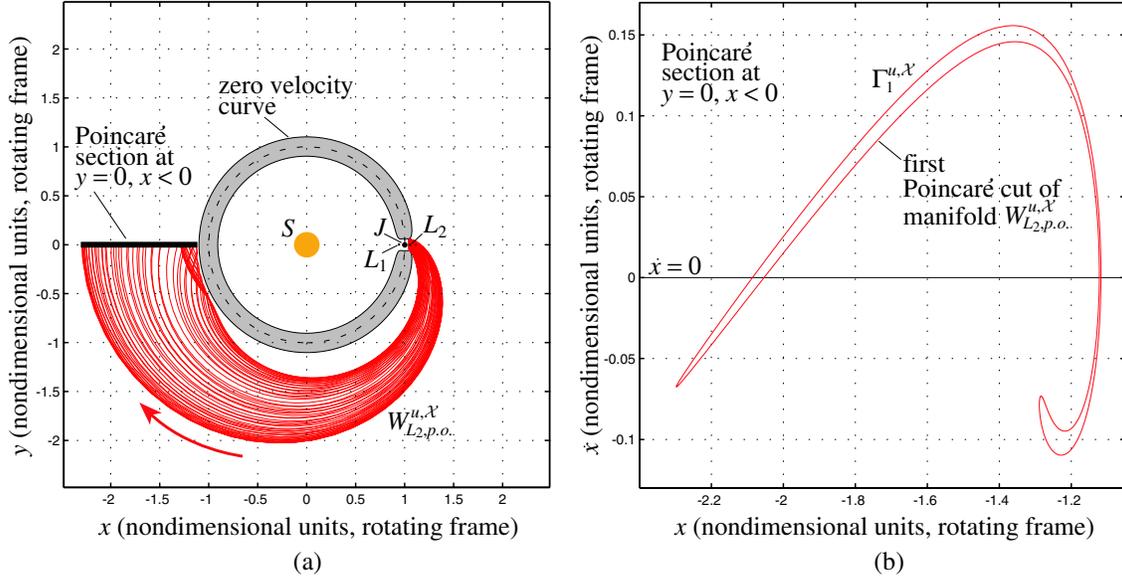
Figure 3.8: (a) The position space projection of the unstable manifold "tube" $W_{L_2,\text{p.o.}}^{u,\mathcal{X}}$ until the first intersection with the Poincaré section at $y = 0, x < 0$. (b) The first Poincaré cut $\Gamma_1^{u,\mathcal{X}}$ of the manifold $W_{L_2,\text{p.o.}}^{u,\mathcal{X}}$ on the plane $y = 0, x < 0$.

intersection $\Gamma_q^{u,\mathcal{X}}$ is a closed curve $\gamma$ in the variables $x, \dot{x}$. Let $s_x$ be the symmetry with respect to the $x$-axis on this plane. Then due to the symmetry of the PCR3BP (3.1), the $q$-th intersection $\Gamma_q^{s,\mathcal{X}}$ of the stable manifold $W_{L_2,\text{p.o.}}^{s,\mathcal{X}}$ with $y = 0$ is $s_x\gamma$. For some minimum $q$, the closed curve $\gamma$ intersects the $\dot{x} = 0$ line of the $(x, \dot{x})$-plane. Points $P$ along the curve $\gamma$ which intersect the $\dot{x} = 0$ line are $(q, q)$-homoclinic points, corresponding to (symmetric) orbits homoclinic to the Lyapunov orbit. If the curve $\gamma$ is transversal to the curve $s_x\gamma$ at the point $P$ then the homoclinic orbit corresponding to $P$ is transversal. If intersections between the curves $\gamma$ and $s_x\gamma$ exist off the line $\dot{x} = 0$ (i.e., if the set $(\gamma \cap s_x\gamma)\backslash\{\dot{x} = 0\}$ is nonempty), then nonsymmetric homoclinic orbits exist.

Consider Figure 3.8(b), where we used the values $\mu = 9.537 \times 10^{-4}$ and $\Delta E = e - E_2 = 0.005$ to compute the unstable Poincaré cut $\Gamma_1^{u,\mathcal{X}}$. If we also plotted the stable cut $\Gamma_1^{s,\mathcal{X}}$, which is the mirror image of unstable cut $\Gamma_1^{u,\mathcal{X}}$ (i.e., $s_x\Gamma_1^{s,\mathcal{X}}$), we would find several points of intersection. In Figure 3.9(a), we focus on the left-most group of points, centered at about $x = -2.07$. We find two $\dot{x} = 0$ intersections which are transversal homoclinic points in the $\mathcal{X}$ realm. The transversal symmetric $(1, 1)$-homoclinic orbit corresponding to the left $\dot{x} = 0$ intersection is shown in Figure 3.9(b).
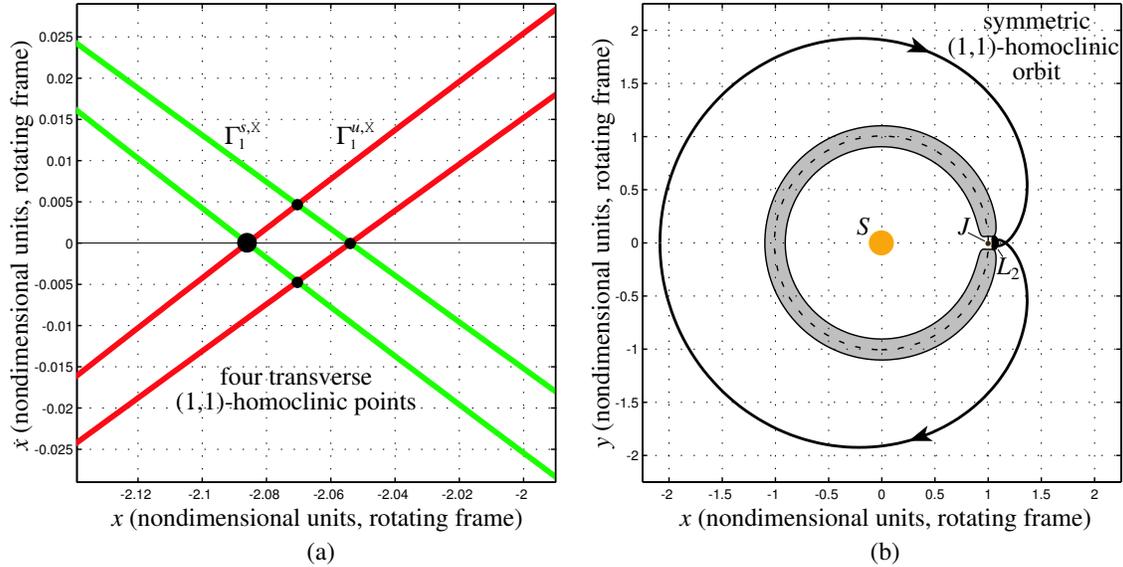
Figure 3.9: (a) A group of four transverse $(1, 1)$-homoclinic points. (b) The symmetric $(1, 1)$-homoclinic orbit corresponding to the left $\dot{x} = 0$ $(1, 1)$-homoclinic point (the large black dot in (a)).

We also notice two off-axis intersections in Figure 3.9(a), completing the local transversal intersection of two closed loops in the $(x, \dot{x})$-plane. As these two intersections occur near the line $\dot{x} = 0$, the appearance of the corresponding homoclinic orbits in position space will be nearly symmetric. A more pronounced case of nonsymmetry occurs for the other group of intersection points centered near $x = -1.15$ on the right side of Figure 3.8(b), for which we have the nonsymmetric $(1, 1)$-homoclinic orbit given in Figure 3.10.

**Homoclinic Orbits in the Exterior and Jupiter Realms.** A similar procedure can numerically produce homoclinic orbits in the interior realm as well as in the Jupiter realm. We can even look at cuts beyond the first and large values of $\mu$ and $\Delta E$, such as shown in Figure 3.11(a). For example, in Figure 3.11(b) we show an interior realm $(1, 3)$-homoclinic orbit (note, also $(2, 2)$ and $(3, 1)$, using $\bar{q} + \bar{p} = q + p$) associated to an $L_1$ Lyapunov orbit for $\mu = 0.1, \Delta E = e - E_1 = 0.03715$.
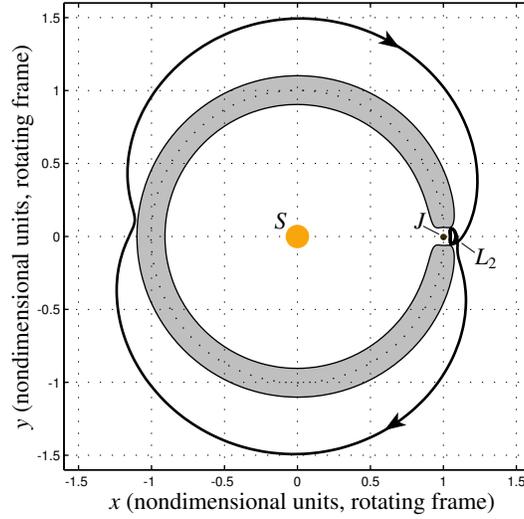
Figure 3.10: A nonsymmetric $(1, 1)$-homoclinic orbit.



(a)  (b)
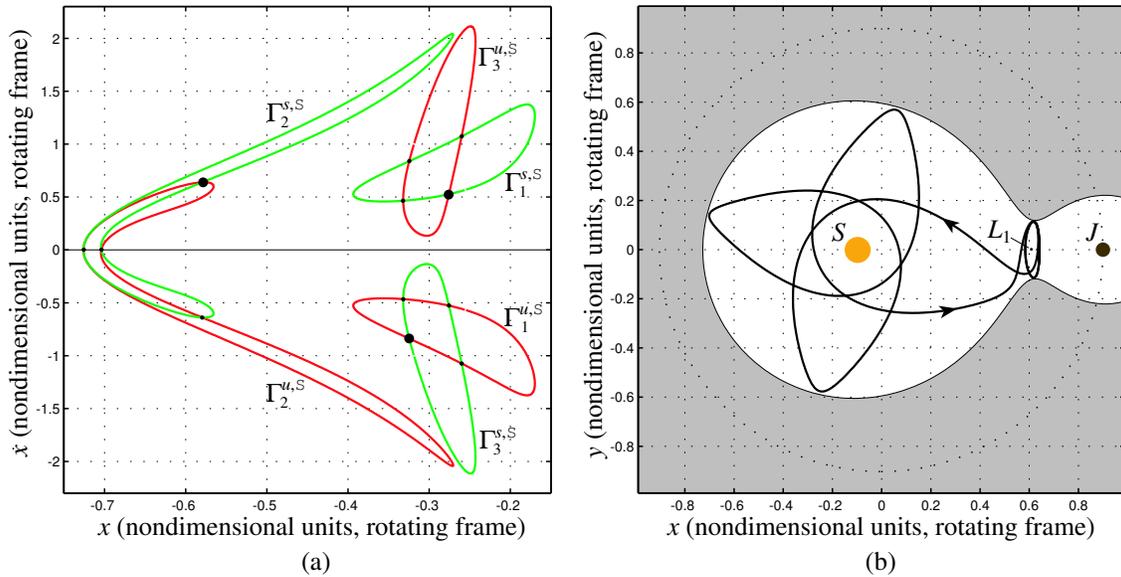
Figure 3.11: (a) The first three Poincaré cuts of the unstable ($W_{L_1,\text{p.o.}}^{u,\mathcal{S}}$) and stable ($W_{L_1,\text{p.o.}}^{s,\mathcal{S}}$) manifolds with the plane $y = 0$. (b) A nonsymmetric $(1, 3)$-homoclinic orbit in the interior realm (corresponding to the three large dots in (a)).

## 3.4 Existence of Heteroclinic Connections between Lyapunov Orbits

We construct a heteroclinic connection between Lyapunov orbits of $L_1$ and $L_2$ by finding an intersection of their respective invariant manifolds in the $\mathcal{J}$ realm. To do so, we seek

points of intersection on a suitably chosen Poincaré section. For instance, to generate a heteroclinic orbit which goes from an $L_1$ Lyapunov orbit (as $t \to -\infty$) to an $L_2$ Lyapunov orbit (as $t \to +\infty$), we proceed as follows.

We restrict ourselves for now to case 3 ($e \in (E_2, E_3)$; see Figure 2.4), for which the Hill's region opens enough to permit Lyapunov orbits about both $L_1$ and $L_2$ to exist. Let the branch of the unstable manifold of the $L_1$ Lyapunov orbit which enters the $\mathcal{J}$ realm be denoted $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$. On the same energy surface there is an $L_2$ Lyapunov orbit, whose stable manifold in the $\mathcal{J}$ realm we similarly denote $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$. The projection of the two-dimensional manifold tubes onto the position space is shown in Figure 3.12(a).
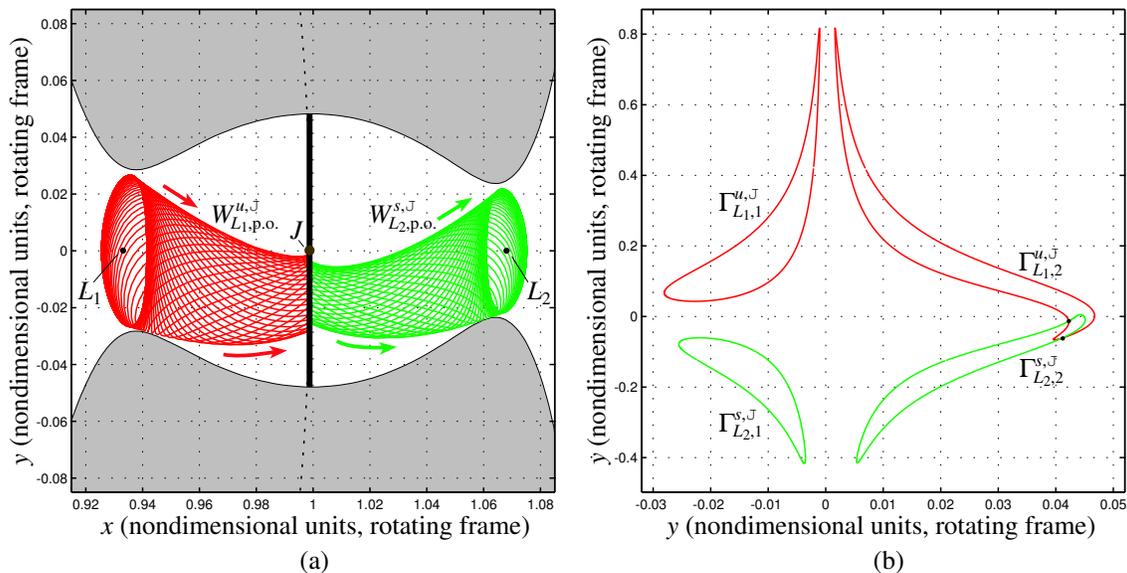


Figure 3.12: (a) The projection of invariant manifolds $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$ and $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$ in the realm $J$ of the position space. (b) The first two Poincaré cuts of the invariant manifolds with the plane $x = 1 - \mu$.

To find intersections between these two tubes, we cut the flow by the plane $x = 1 - \mu$, denoted by the thick black line in Figure 3.12(a). The cuts on this plane are shown in Figure 3.12(b).

This convenient plane maximizes the number of intersections for values of $\mu, e$ which produce manifolds making a limited number of revolutions around Jupiter before escaping from the $\mathcal{J}$ realm. The $q$-th intersection of $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$ with the plane $x = 1 - \mu$ will be labeled $\Gamma^{u,\mathcal{J}}_{L_1,q}$. Similarly, we will call $\Gamma^{s,\mathcal{J}}_{L_2,p}$ the $p$-th intersection of $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$ with $x = 1 - \mu$.

Numerical experiments show that the $L_1$ Lyapunov orbit unstable manifold $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$ does not coincide with the $L_2$ Lyapunov orbit stable manifold $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$. Moreover, for a

wide range of $\mu$ and $e$ values (where $e \in (E_2, E_3)$), numerical explorations demonstrate that they do intersect transversely. While we recognize that for certain values of $\mu$ and $e$, there are tangencies between the stable and unstable manifold, we will not deal with this interesting case in this thesis. Hence, from now on, we will concentrate our numerical explorations only on the cases where the stable and unstable manifold intersect transversely.

Now, suppose that $\Gamma_{L_1,q}^{u,\mathcal{J}}$ and $\Gamma_{L_2,p}^{s,\mathcal{J}}$ are each closed curves in the variables $y, \dot{y}$. A point in the plane $x = 1 - \mu$ belonging to the intersection of the two closed curves (i.e., $\Gamma_{L_1,q}^{u,\mathcal{J}} \cap \Gamma_{L_2,p}^{s,\mathcal{J}}$) will be called a $(q, p)$-**heteroclinic point** because such a point corresponds to a heteroclinic orbit going from the $L_1$ Lyapunov orbit to the $L_2$ Lyapunov orbit. Moreover, since we restrict ourselves to the case where $W_{L_1,\text{p.o.}}^{u,\mathcal{J}}$ and $W_{L_2,\text{p.o.}}^{s,\mathcal{J}}$ intersect transversely, the $(q, p)$-heteroclinic point will be a *transversal heteroclinic point*. Our objective is to obtain the first intersection point (or group of points) of the curve $\Gamma_{L_1,q}^{u,\mathcal{J}}$ with the curve $\Gamma_{L_2,p}^{s,\mathcal{J}}$ and so obtain the minimum values of $q$ and $p$ such that we have a transversal $(q, p)$-heteroclinic point. Other intersections may exist, but we will restrict ourselves for now to the first. For some minimum $q$ and $p$, we have an intersection of the curves, and some number of $(q, p)$-heteroclinic points, depending on the geometry of the intersection. Note that the sum $q + p$ must be an even positive integer.

As we are interested in heteroclinic points for the Sun-Jupiter system ($\mu = 9.537 \times 10^{-4}$), we take $e = -1.5185$ and numerically obtain the intersections of the invariant manifolds $W_{L_1,\text{p.o.}}^{u,\mathcal{J}}$ and $W_{L_2,\text{p.o.}}^{s,\mathcal{J}}$ with the plane $x = 1 - \mu$. In Figure 3.12(b) we show the curves $\Gamma_{L_1,q}^{u,\mathcal{J}}$ for $q = 1, 2$ and $\Gamma_{L_2,p}^{s,\mathcal{J}}$ for $p = 1, 2$. Notice that $\Gamma_{L_1,2}^{u,\mathcal{J}}$ and $\Gamma_{L_2,2}^{s,\mathcal{J}}$ intersect in two points (the black dots in Figure 3.12(b) near $y = 0.042$). Thus, the minimum $q$ and $p$ for a heteroclinic point to appear for these particular values of $\mu, e$ are $q = 2$ and $p = 2$. The $(2, 2)$-heteroclinic points can each be forward and backward integrated to produce heteroclinic orbits going from the $L_1$ Lyapunov orbit to the $L_2$ Lyapunov orbit, otherwise known as a *heteroclinic connection*. We show one of the heteroclinic orbits in Figure 3.13. Notice that the number of revolutions around Jupiter is given by $(q+p-1)/2$. The reverse orbit, going from the $L_2$ Lyapunov orbit to the $L_1$ Lyapunov orbit, is easily given by the symmetry $s$ (3.1). It is the mirror image (about the $x$-axis) of the trajectory in Figure 3.13, with the direction arrows reversed. These two heteroclinic connections together form a symmetric **heteroclinic cycle**.
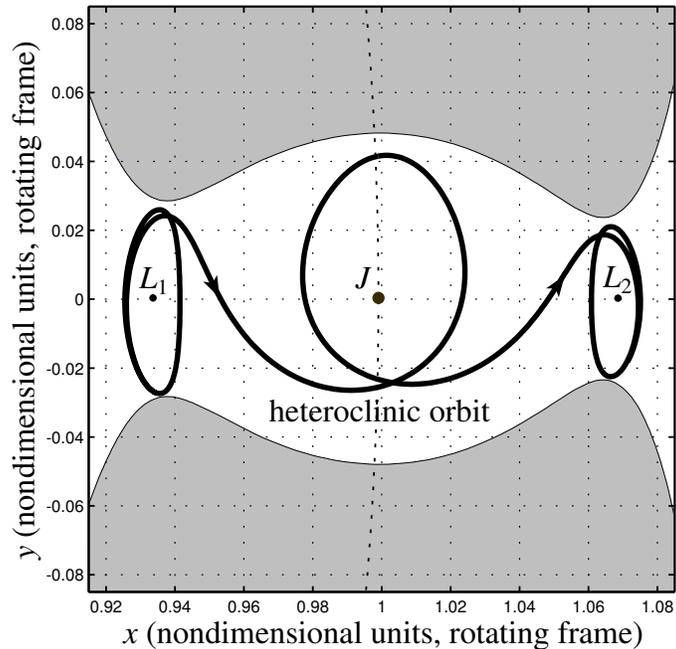
Figure 3.13: The existence of a transversal $(2,2)$-heteroclinic orbit in the $J$ realm.

## 3.5 Existence of Chains of Homoclinic Orbits and Heteroclinic Cycles

We have numerically demonstrated the existence of homoclinic and heteroclinic orbits associated to the $L_1$ and $L_2$ Lyapunov orbits for case 3. We now take the final step, combining homoclinic and heteroclinic orbits of the same energy value to generate what is called a ***homoclinic-heteroclinic chain*** of orbits, which connect asymptotically the $L_1$ and $L_2$ Lyapunov orbits to each other. As will be seen, these chains imply a complicated dynamics connecting the interior, exterior, and Jupiter realms.

As an example, we again choose the Sun-Jupiter system ($\mu = 9.537 \times 10^{-4}$), but now an energy value similar to that of comet P/Oterma during its Jupiter encounters ($e = -1.515$). Using the described methodologies, we obtain an interior realm orbit homoclinic to the $L_1$ Lyapunov orbit, an exterior realm orbit homoclinic to the $L_2$ Lyapunov orbit, and a heteroclinic cycle connecting the $L_1$ and $L_2$ Lyapunov orbits in the Jupiter realm. The union of these orbits is the homoclinic-heteroclinic chain shown in Figure 3.1. The existence of homoclinic-heteroclinic chains has important consequences, which will be expanded upon further in the following sections.

## 3.6    Construction of a Suitable Poincaré Map

The idea of reducing the study of the global orbit structure of a system of differential equations to the study of an associated discrete map is due to Poincaré [1890], who first utilized the method in his studies of the restricted three-body problem. In this section we use the chain of two homoclinic orbits and one symmetric heteroclinic cycle (such as the one shown in Figure 3.1) to construct a suitable Poincaré map. Our choice of Poincaré map will allow us to study the complex global orbit structure near the chain. We find an invariant set for this map near some transversal homoclinic and heteroclinic points along the chain where "Smale horseshoe"-like dynamics exist. We then use symbolic dynamics to characterize the chaotic motion of a comet in a neighborhood of the chain as it transitions intermittently through the interior, Jupiter and exterior realms. Not only do we prove the existence of the invariant set, but we also numerically approximate it in Chapter 4, gaining further insight into the complex global dynamics associated with the chains.

Here is additional detail about how we proceed: In this section, we construct a Poincaré map $P$ transversal to the flow around a chain whose domain $U$ consists of four different squares $U_i, i = 1, 2, 3, 4$, located in different parts of phase space in the neighborhood of the chain, as shown schematically in Figure 3.14.
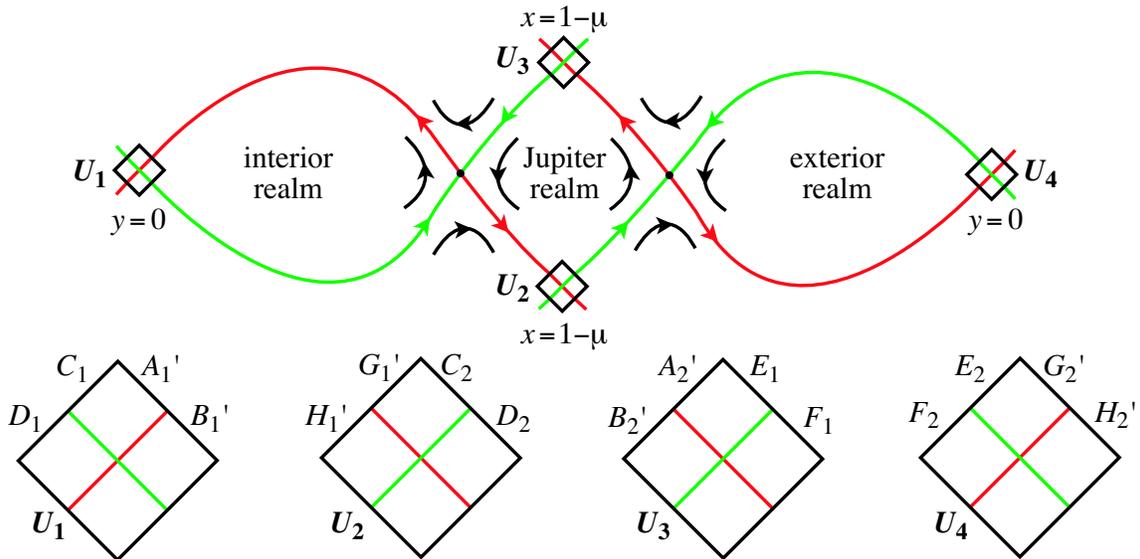


Figure 3.14:  The construction of a suitable Poincaré map. The labeling $D_1$, etc., is described in the text.

Squares $U_1$ and $U_4$ are contained in the surface $y = 0$ and each centers around a transversal homoclinic point in the interior and the exterior realm, respectively. Squares $U_2$ and $U_3$ are contained in the surface $x = 1 - \mu$ ($y < 0$ and $y > 0$, respectively) and center around transversal heteroclinic points in the Jupiter realm which are symmetric with respect to each other. Clearly, for any orbit which passes through a point $q$ in one of the squares and whose images and pre-images ($P^n(q), n = 0, \pm 1, \pm 2, \ldots$) all remain in the domain $U$, the whereabouts of $P^n(q)$ (as $n$ increases or decreases) can provide some of the essential information about the history of the particular orbit. We record this history with a bi-infinite sequence. This well-known technique of studying only the set of points that forever remain in the domain $U$ (the ***invariant set***) provides us with all the periodic solutions as well as the recurrent solutions in the neighborhood of the chain.

The technique of characterizing the orbit structure of a dynamical system via a set of bi-infinite sequences of "symbols" is known as ***symbolic dynamics***.

In §3.7 and §3.8, we extend the symbolic dynamics results of Llibre, Martinez and Simó [1985] to our situation and construct a set of bi-infinite sequences with two families of symbols. The first family is a *subshift of finite type* with four symbols $\{u_1, u_2, u_3, u_4\}$. It is used to keep track of the whereabouts of an orbit with respect to the four squares $U_1, U_2, U_3, U_4$. The symbol $u_i$ is recorded every time the $U_i$ square is pierced by the orbit. Subshift here means that among the set of all bi-infinite sequences of four symbols, (i.e., $(\ldots, u_{i_{-1}}; u_{i_0}, u_{i_1}, u_{i_2}, \ldots)$ where $i_j$ ranges from 1 to 4), certain sequences where the adjacent entries in the sequence violate certain relations are not allowed. For example, from $U_1$, the (forward) flow cannot get to $U_4$ without passing through other squares. Hence, in the bi-infinite sequence, the symbol $u_1$ cannot be followed by $u_4$. The relations can be defined by a matrix $A$ called the *transition matrix*. In our case,

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

It is constructed by the following rule: $(A)_{kl} = 1$ if the ordered pair of symbols $u_k, u_l$ may appear as adjacent entries in the symbolic sequence, and $(A)_{kl} = 0$ if the ordered pair of symbols $u_k, u_l$ may not appear as adjacent entries. For example, since $u_1$ cannot be

followed by $u_4$, we have $(A)_{14} = 0$.

The second family is a *full shift of infinite type* with symbols of positive integers greater than a fixed integer $m$. This set of bi-infinite sequences of positive integers is used to keep track of the number of integer revolutions that the projection of an orbit winds around either $L_1$ or $L_2$ when the orbit enters the equilibrium regions $\mathcal{R}_1$ or $\mathcal{R}_2$, respectively.

In §3.9, we state the main theorem of this chaoter and discuss its implications. The theorem gives the global orbit structure of the PCR3BP in a neighborhood of a chain of homoclinic orbits and a symmetric heteroclinic cycle. It says essentially that given any bi-infinite sequence

$$\alpha = (u, r) = (\ldots, (u_{i_{-1}}, r_{-1}); (u_{i_0}, r_0), (u_{i_1}, r_1), (u_{i_2}, r_2) \ldots),$$

there exist initial conditions near the transversal homoclinic and heteroclinic points (the intersection of the chain with $U$) such that an orbit corresponding to such initial conditions starts at $U_{i_0}$ and goes to $U_{i_1}$ (provided $(A)_{i_0 i_1} = 1$). This orbit passes through either the equilibrium region $\mathcal{R}_1$ or $\mathcal{R}_2$ depending on whether the initial index ($i_0$ in the current case) is $1, 3$ or $2, 4$ (see Figure 3.14 for reference). For example if $i_0 = 1$, then the projection of the orbit winds around $L_1$ for $r_0$ revolutions inside the region $\mathcal{R}_1$ before leaving for $U_{i_1}$. After that, the same process begins with $(u_{i_1}, r_1)$ in place of $(u_{i_0}, r_0)$ and $(u_{i_2}, r_2)$ in place of $(u_{i_1}, r_1)$, etc. For negative time, a similar behavior is described for $(u_{i_{-1}}, r_{-1}), (u_{i_0}, r_0)$, etc. While the formalism involved in the proof is fairly standard, there are a few new features which may be worth pointing out. While most of these comments will be made earlier, we provide a sketch of the proof in §3.9 both for completeness and for the convenience of the reader. For more details, one can consult Koon, Lo, Marsden and Ross [2000], Moser [1973], Llibre, Martinez and Simó [1985], Wiggins [1990], and Wiggins [1993].

In Chapter 4 we numerically construct sets of orbits with prescribed itineraries. By successive application of the Poincaré map $P$ to a transversal plane in a (rather large) neighborhood of a chain, we can theoretically generate regions of orbits with itineraries of any size.

**Construction of a Suitable Poincaré Map.** In §3.5, we showed that with an appropriate energy value, there exists a chain of two homoclinic orbits and one symmetric heteroclinic cycle. For simplicity of exposition, let us suppose that the chain $\mathcal{C}$ consists

of $(1,1)$-transversal homoclinic orbits in the interior and exterior realms and a symmetric $(1,1)$-transversal heteroclinic cycle in the Jupiter realm. A similar study can be done for other cases.

Now we are ready to construct a Poincaré map. The first step is to construct the transversal maps on the bounding spheres of the equilibrium regions $\mathcal{R}_1$ and $\mathcal{R}_2$. Let $\epsilon_1$ and $\epsilon_2$ be small positive quantities. For the bounding spheres $n_{1,1}$ and $n_{1,2}$ of the equilibrium region $\mathcal{R}_1$, we define $\{A_1, B_1, C_1, D_1\}$, $\{E_1, F_1, G_1, H_1\}$ as the set of points of $\{d_{1,1}^-, r_{1,1}^-, r_{1,1}^+, d_{1,1}^+\}$, $\{d_{1,2}^+, r_{1,2}^+, r_{1,2}^-, d_{1,2}^-\}$, respectively, such that $||\zeta|^2 - \rho^*| < \epsilon_1$. These sets correspond to thin strips on the bounding sphere centered on the asymptotic sets $\{a_{1,1}^-, a_{1,1}^+\}$, $\{a_{1,2}^+, a_{1,2}^-\}$, respectively, as shown in Figure 3.15. Similarly, given $\epsilon_2$ small, we can define corresponding strips for the bounding spheres $n_{2,1}$ and $n_{2,2}$ of the equilibrium region $\mathcal{R}_2$.
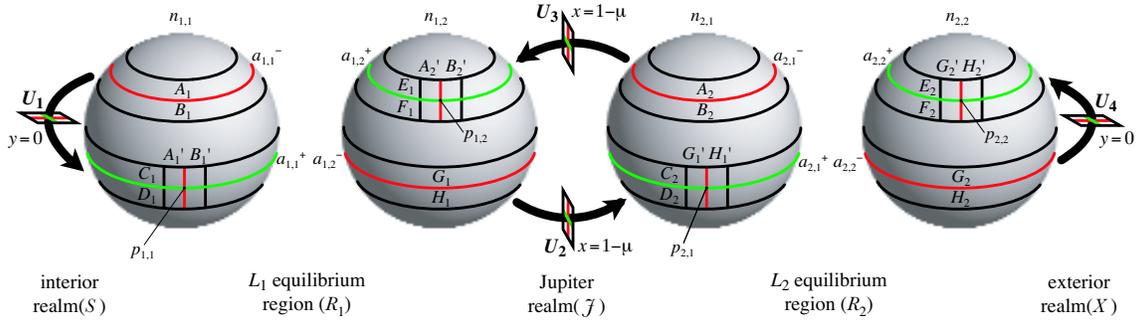


Figure 3.15: The strips near the asymptotic sets on the spheres $n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}$.

If $\epsilon_1$ and $\epsilon_2$ are small enough, the flow is transversal to the surfaces just defined. Recall from §2.8 that orbits entering $\mathcal{R}_1$ through $C_1, D_1, E_1, F_1$ leave it through $B_1, H_1, A_1, G_1$, respectively, because $|\zeta|^2$ is a first integral in $\mathcal{R}_1$. Therefore the diffeomorphisms $\psi_{1,i}$ send $D_1, E_1, C_1, F_1$ into $H_1, A_1, B_1, G_1$ respectively, for $i = 1, 2, 3, 4$, where the $\psi_{1,i}$ are the mappings given in (2.28) and (2.29). Similar results hold for orbits entering $\mathcal{R}_2$ and the corresponding diffeomorphisms $\psi_{2,i}$ send $D_2, E_2, C_2, F_2$ into $H_2, A_2, B_2, G_2$ respectively, for $i = 1, 2, 3, 4$.

The second step is to construct transversal maps outside of the equilibrium regions. Let $p_{1,1} \in a_{1,1}^+$ (resp. $p_{2,2} \in a_{2,2}^+$) be a point of the transversal homoclinic orbit of $\mathcal{C}$ in the interior (resp. exterior) realm. Let $A_1'$ and $B_1'$ (resp. $G_2'$ and $H_2'$) be the first images of $A_1$ and $B_1$ (resp. $G_2$ and $H_2$) in $n_{1,1}$ (resp. $n_{2,2}$) sent by the forward flow outside $\mathcal{R}_1$

(resp. $\mathcal{R}_2$). The maps sending $A_1, B_1, G_2, H_2$ onto $A_1', B_1', G_2', H_2'$ are diffeomorphisms. In a neighborhood of $p_{1,1}$ (resp. $p_{2,2}$) the qualitative picture of $A_1'$ and $B_1'$ (resp. $G_2'$ and $H_2'$) is shown in Figure 3.15 provided $\epsilon_1$ and $\epsilon_2$ are sufficiently small.

Similarly, let $p_{1,2} \in a_{1,2}^+$ and $p_{2,1} \in a_{2,1}^+$ be points of the transversal heteroclinic cycle of $\mathcal{C}$ in the Jupiter realm. Let $A_2'$ and $B_2'$ (resp. $G_1'$ and $H_1'$) be the first images of $A_2$ and $B_2$ (resp. $G_1$ and $H_1$) in $n_{1,2}$ (resp. $n_{2,1}$) sent by the flow outside $\mathcal{R}_1$ and $\mathcal{R}_2$. The maps sending $A_2, B_2, G_1, H_1$ onto $A_2', B_2', G_1', H_1'$ are diffeomorphisms. In a neighborhood of $p_{1,2}$ (resp. $p_{2,1}$) the qualitative picture of $A_2'$ and $B_2'$ (resp. $G_1'$ and $H_1'$) is also shown in Figure 3.15.

Now let $U_1$ (resp. $U_4$) be the sets diffeomorphic to $(C_1 \cup D_1) \cap (A_1' \cup B_1')$ (resp. $(E_2 \cup F_2) \cap (G_2' \cup H_2')$) defined by following the flow backwards up to the first crossing with the surface $y = 0$. Similarly, let $U_2$ (resp. $U_3$) be the sets diffeomorphic to $(C_2 \cup D_2) \cap (G_1' \cup H_1')$ (resp. $(E_1 \cup F_1) \cap (A_2' \cup B_2')$) defined by following the flow backwards up to the first crossing with the surface $x = 1 - \mu$. See Figures 3.14 and 3.15. Since each of the sets $U_i$ are topologically a square, we shall refer to them loosely as squares in the rest of this section.

Let $U = U_1 \cup U_2 \cup U_3 \cup U_4$. We define the Poincaré map

$$P : U \to U \tag{3.4}$$

in the following way: To each point $q \in U$ we assign the corresponding first intersection point with $U$ of the orbit passing through $q$, if such an intersection exists. For simplicity of notation, we loosely refer to $U_1$ as $(C_1 \cup D_1) \cap (A_1' \cup B_1')$ even though $U_1$ actually lies in the surface $y = 0$. Similar convention will be used for the other $U_i$'s.

Now we consider the invariant set of points, $\Lambda$, which remain in $U$ under all forward and backward iterations by $P$. Thus $\Lambda$ is defined as

$$\Lambda = \cap_{n=-\infty}^{\infty} P^n(U). \tag{3.5}$$

This invariant set contains all the periodic solutions as well as the recurrent solutions near the chain and provides insight into the global dynamics in a neighborhood of the chain.

**Horseshoe-type Map and Conley-Moser Conditions.** We review a standard text-book example to introduce the next section. For a horseshoe-type map $h : Q \to Q$ of a square $Q$ into itself, which satisfies the *Conley-Moser conditions*, the invariant set of all iterations

$$\Lambda_h = \cap_{n=-\infty}^{\infty} h^n(Q), \tag{3.6}$$

can be constructed and visualized in a standard way. The Conley-Moser conditions are the following.

- **Strip condition**: $h$ maps "horizontal strips" $H_0, H_1$ to "vertical strips" $V_0, V_1$, (with horizontal boundaries to horizontal boundaries and vertical boundaries to vertical boundaries).

- **Hyperbolicity condition**: $h$ has uniform contraction in horizontal direction and expansion in vertical direction.

The invariant set of first iterations

$$\Lambda_h^1 = h^{-1}(Q) \cap Q \cap h^1(Q), \tag{3.7}$$

has 4 squares, with addresses $(0;0), (1;0), (1;1), (0;1)$. Invariant set of second iterations has 16 squares contained in 4 squares of first stage. This process can be repeated ad infinitum due to the Conley-Moser conditions. What remains is invariant set of points $\Lambda_h$ which are in 1-to-1 correspondence with set of bi-infinite sequences of 2 symbols $(\ldots, 0; 1, \ldots)$.

## 3.7  Horseshoe-like Dynamics

Compared with the standard textbook example above which studies the chaotic dynamics in a neighborhood of a transversal homoclinic point of a two-dimensional map $\bar{f}$, the Poincaré map $P$ constructed in this chapter has a number of special properties.

**Domain of the Poincaré Map $P$.** Instead of studying the first return map $\bar{f}$ (induced by the flow $f$) on a (small) topological square $Q$, the domain $U$ of the Poincaré map $P$ consists of four squares $U_i, i = 1, 2, 3, 4$ which center around $p_{1,1}, p_{2,1}, p_{1,2}, p_{2,2}$, respectively, as shown in Figures 3.16 and 3.17.
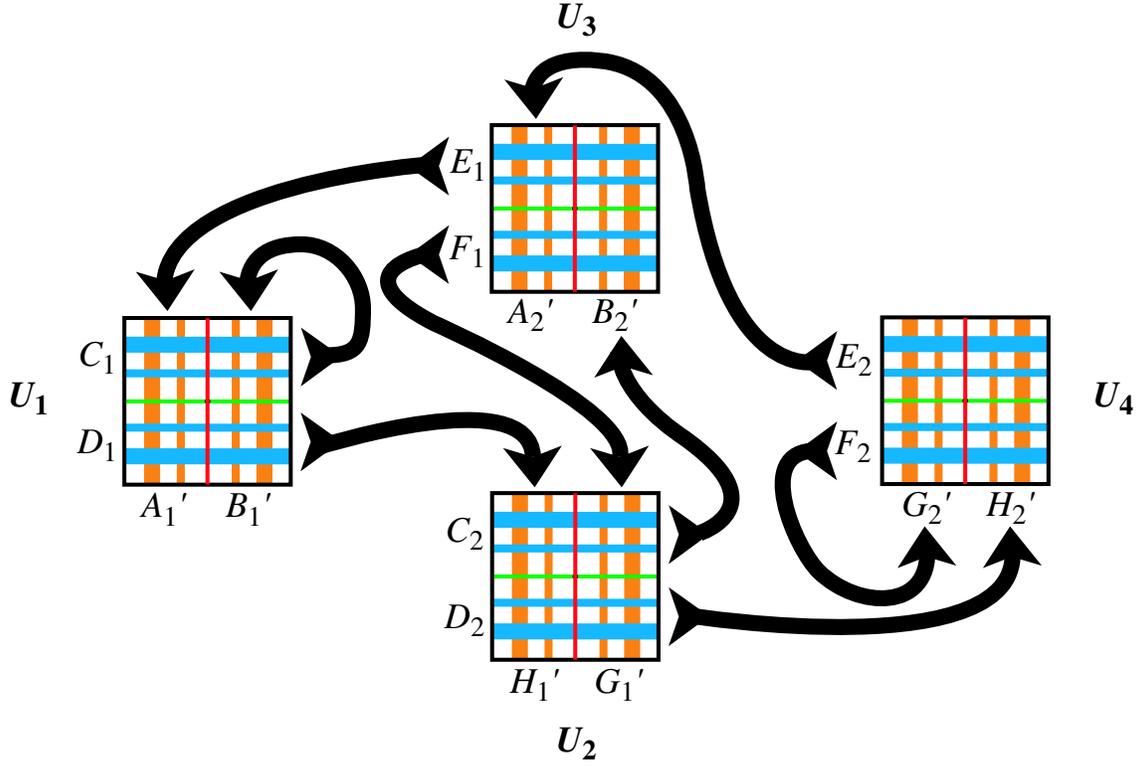
Figure 3.16: The families of horizontal strips and their images under $P$.

Moreover, the map $P$ is not defined on points in $U$ belonging to the invariant manifolds of the $L_1$ and $L_2$ Lyapunov orbits. Take $U_1$ as an example. On the curves $\Gamma_{L_1,1}^{u,\mathcal{S}}$ and $\Gamma_{L_1,1}^{s,\mathcal{S}}$ which are the first intersections of the unstable and stable invariant manifolds of the $L_1$ Lyapunov orbit with the surface $y = 0$ in the interior (Sun) realm, the Poincaré map is singular because any point on those curves will be carried by the flow asymptotically backward or forward towards the $L_1$ Lyapunov orbit. Hence, we have a kind of singular Poincaré map as it has been considered by Devaney [1981]. We return to this point at the end of §3.8.

Therefore, we must consider in fact four small (open) squares in $U_1$, namely:

$$(C_1 \cap A_1'), (C_1 \cap B_1'), (D_1 \cap A_1') \quad \text{and} \quad (D_1 \cap B_1').$$

Similar consideration is also needed for the other $U_i$'s which add up to sixteen small squares in total, as shown in Figure 3.17.
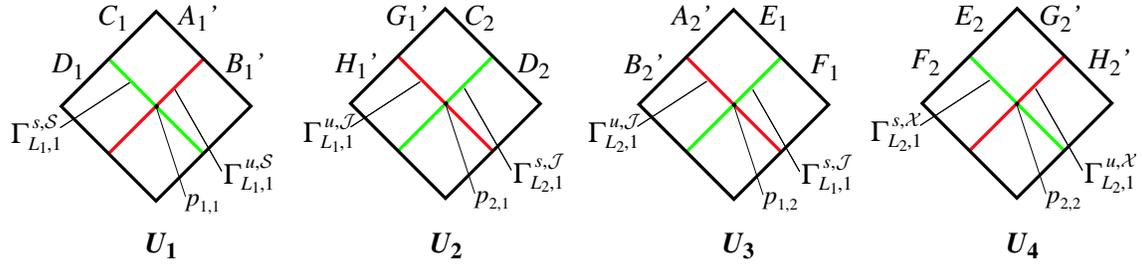
Figure 3.17: The domain $U = U_1 \cup U_2 \cup U_3 \cup U_4$ of the Poincaré map $P$.

**Horizontal and Vertical Strips.** For the standard textbook example, the first return map $\bar{f}$ (induced by the flow $f$) on the square $Q$ qualitatively looks like a Smale horseshoe map. Conley and Moser found conditions for the map $\bar{f}$ to satisfy in order for it to have an invariant subset $\Lambda_{\bar{f}}$ of $Q$ on which it has chaotic dynamics. These conditions are a combination of geometrical and analytical conditions.

1. The geometrical part consists of generalizing the notion of horizontal and vertical rectangles to horizontal and vertical strips in $Q$ by allowing the boundaries to be Lipschitz curves, as shown in Figure 3.18, rather than straight lines. With this generalization in hand one then requires "horizontal" strips to map to "vertical" strips with horizontal boundaries mapping to horizontal boundaries and vertical boundaries mapping to vertical boundaries.

2. The analytical part comes from requiring uniform contraction in the horizontal directions and expansion in the vertical direction.

For the Poincaré map $P$ constructed in this chapter, the situation becomes more complicated in two ways. First, the number of strips in each family generated after one iteration is not two or even finite, but is instead infinite. Second, we need to use subshift to keep track of the image of each family of strips. Here, we discuss first the issue of each family having an infinite number of strips.

**First Iteration: 8 Families of Vertical Strips $V_n^{ji}$.** Let us consider $U \cap P(U)$. For simplicity of exposition, take $U_1$ as an example and consider the small squares $(D_1 \cap A_1')$ and $(D_1 \cap B_1')$, shown on the left side of Figure 3.19.

Recall the observation in §2.8 on the spiraling of an abutting arc with an endpoint in the asymptotic set of a bounding sphere. The image of the squares $(D_1 \cap A_1')$ and
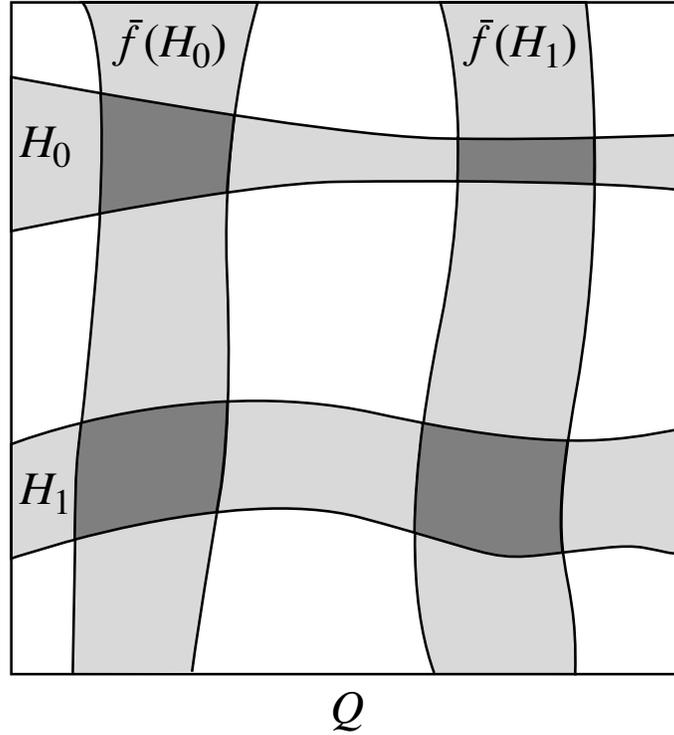
Figure 3.18: Generalization of the notion of horizontal and vertical rectangles for the Conley-Moser conditions.

$(D_1 \cap B_1')$ under $P$ is a strip contained in $H_1'$ of arbitrarily long length, cutting $U_2$ an infinite number of times and spiraling towards $\Gamma_{L_1,1}^{u,\mathcal{J}}$, becoming skinnier when approaching the limit. The intersection of this strip with $U$ (in fact only with $U_2$) forms an infinite number of components. All but perhaps one of the components are limited by the sides $e_6$ and $e_8$, shown in Figure 3.19. We call each of the components of

$$P((D_1 \cap A_1') \cup (D_1 \cap B_1')) \cap U \subset H_1'$$

a **vertical strip** of $H_1'$ (in $U_2$).

Now consider all the vertical strips in $H_1'$ and denote these by $\mathcal{V}H_{1,0}', \mathcal{V}H_{1,1}', \ldots$, beginning with the strips nearest to $e_5$. We have on $H_1'$ a family of vertical strips $\{\mathcal{V}H_{1,n}'\}$ bounded by the sides $e_6$ and $e_8$ (in $U_2$) and with the width of $\mathcal{V}H_{1,n}'$ tending to zero as $n$ tends to infinity. We define

$$\mathcal{V}H_{1,\infty}' = \lim_{n \to \infty} \mathcal{V}H_{1,n}'.$$
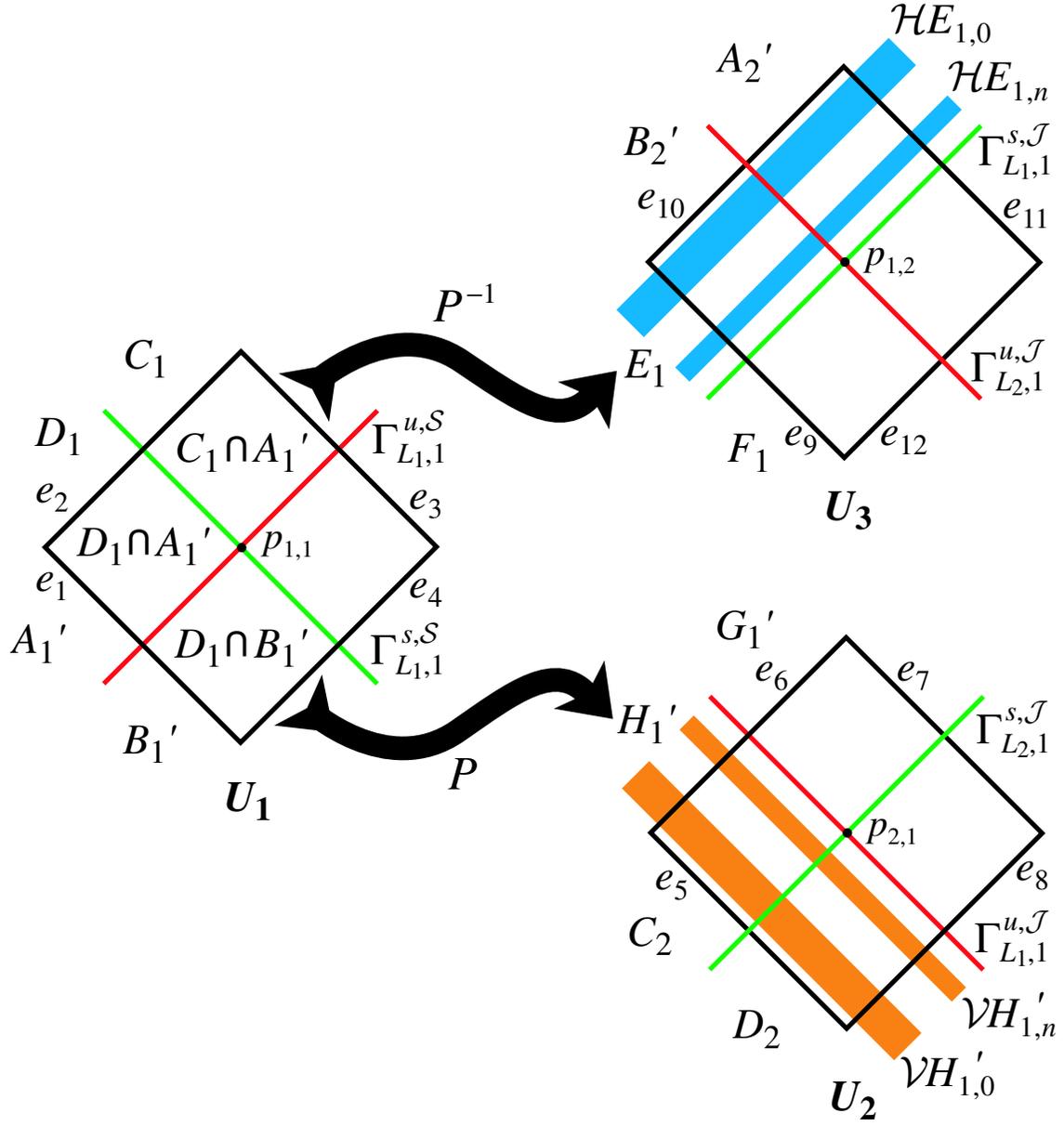
Figure 3.19: The topological squares and the images of some rectangles. We show schematically only two strips although there are an infinite number, getting increasingly slender as they approach the invariant manifold ($\Gamma_{L_1,1}^{u,\mathcal{J}}$ on $U_2$ and $\Gamma_{L_1,1}^{s,\mathcal{J}}$ on $U_3$).

Clearly, $\mathcal{V}H'_{1,\infty}$ is simply the vertical curve $\Gamma_{L_1,1}^{u,\mathcal{J}}$ which is on the Jupiter realm branch of the unstable invariant manifold of the $L_1$ Lyapunov orbit. Similar constructions can be carried out for the other small squares $(C_1 \cap A'_1)$ and $(C_1 \cap B'_1)$ of $U_1$ which yield a family of vertical strips in $B'_1$. In order to keep track of these families of vertical strips

more effectively, we shall rename $\{\mathcal{V}B'_{1,n}\}$ and $\{\mathcal{V}H'_{1,n}\}$ as $\{V_n^{11}\}$ and $\{V_n^{21}\}$, respectively. Notice that for $V_n^{ji}$, the index $ji$ indicates that the family is in the square $U_j$ and it came from the square $U_i$. For simplicity of illustration, we have used rectangles to represent strips in Figure 3.19. Similar representations will be used throughout the rest of this section.

Similarly, we can look at the first iterate by $P$ of the other $U_i$'s and obtain families of vertical strips in

$$B'_2(\{V_n^{32}\}), H'_2(\{V_n^{42}\}), A'_1(\{V_n^{13}\}), G'_1(\{V_n^{23}\}), A'_2(\{V_n^{34}\}), G'_2(\{V_n^{44}\}).$$

Therefore, $U \cap P(U)$ is the disjoint union of eight families of pairwise disjoint vertical strips.

**First Iteration: 8 Families of Horizontal Strips $H_m^{ij}$.** An analogous study can be done for $U \cap P^{-1}(U)$. Consider the small squares $(D_1 \cap A'_1)$ and $(C_1 \cap A'_1)$ of $U_1$ in Figure 3.19. Then $P^{-1}((D_1 \cap A'_1) \cup (C_1 \cap A'_1))$ is a strip contained in $E_1$ of arbitrarily long length, cutting $U_3$ an infinite number of times and spiraling towards $\Gamma_{L_1,1}^{s,\mathcal{J}}$, becoming thinner while approaching the limit. The intersection of this strip with $U$ (in fact only with $U_3$) forms an infinite number of components. All but perhaps one of the components are limited by the sides $e_9$ and $e_{11}$. We call each of the components of

$$P^{-1}((D_1 \cap A'_1) \cup (C_1 \cap A'_1)) \cap U \subset E_1$$

a **horizontal strip** of $E_1$ (in $U_3$).

Now consider all the horizontal strips in $E_1$ and denote these by $\mathcal{H}E_{1,0}, \mathcal{H}E_{1,1}, \ldots$, beginning with the strip nearest to $e_{10}$. We have on $E_1$ a family of horizontal strips $\{\mathcal{H}E_{1,n}\}$ bounded by the sides $e_9$ and $e_{11}$ (in $U_3$) and with the width of $\mathcal{H}E_{1,n}$ tending to zero as $n$ tends to infinity. We define

$$\mathcal{H}E_{1,\infty} = \lim_{n \to \infty} \mathcal{H}E_{1,n}.$$

Clearly, $\mathcal{H}E_{1,\infty}$ is simply the horizontal curve $\Gamma_{L_1,1}^{s,\mathcal{J}}$ which is on the stable invariant manifolds of the $L_1$ Lyapunov orbit.

Similar constructions can be carried out for the other small squares $(C_1 \cap B'_1)$ and

$(D_1 \cap B_1')$ of $U_1$ which yield a family of horizontal strips in $C_1$. We shall again rename $\{\mathcal{H}C_{1,n}\}$ and $\{\mathcal{H}E_{1,n}\}$ as $\{H_n^{11}\}$ and $\{H_n^{31}\}$, respectively. Notice that for $H_n^{ij}$, the index $ij$ indicates that the family is in the square $U_i$ and it will go to the square $U_j$.

Similarly, we can look at the first iterate by $P^{-1}$ of the other $U_i$'s and obtain families of horizontal strips in

$$D_1(\{H_n^{12}\}), F_1(\{H_n^{32}\}), C_2(\{H_n^{23}\}), E_2(\{H_n^{43}\}), D_2(\{H_n^{24}\}), F_2(\{H_n^{44}\}).$$

Therefore, $U \cap P^{-1}(U)$ is the disjoint union of eight families of pairwise disjoint horizontal strips.

**Meaning of the Label $V_n^{ji}$.** We discuss briefly the meaning of the subscript $n$ in the vertical strip $V_n^{ji}$. It can be used to keep track of the number of revolutions the projection of the associated orbits wind around $L_1$ or $L_2$. For example, the orbit which pierces the vertical strip $V_{k+1}^{21}$ has wound one more time around $L_1$ than the orbit which pierces the vertical strip $V_k^{21}$. Moreover, given any $\epsilon_1$ for the width of the strips $D_1$ and $H_1'$, there is a minimum number of integer revolutions $r_{\min}$ around $L_1$ an orbit will make in going from $D_1$ (in $U_1$) to $H_1'$ (in $U_2$). With this specific $\epsilon_1$, the orbit which pierces $V_n^{21}$ has wound around $L_1$ for $(n + r_{\min})$ times. In the rest of this chapter, we assume that we have adjusted the widths (the $\epsilon_j$'s) of all the other corresponding pairs of strips so that the minimum number of revolutions around $L_1$ or $L_2$ is the same for all the $U_i$'s. With this adjustment, any orbit which pierces $V_n^{ji}$ is now in $U_j$. It came from $U_i$ and has wound around $L_1$ (if $u_i = 1, 3$) or $L_2$ (if $u_i = 2, 4$) for $(n + r_{\min})$ times.

**The Generalized Conley-Moser Conditions.** For the standard textbook example introduced earlier about the dynamics near a transversal homoclinic point, it is well known that if the first return map $\bar{f}$ (induced by $f$) on the square $Q$ satisfies the following Conley-Moser conditions, then there exists an invariant set $\Lambda_{\bar{f}}$ of $Q$ on which $\bar{f}$ has chaotic dynamics.

**Condition** 1: There exist a finite (or possibly infinite) number of horizontal and vertical strips $\mathcal{H}_i$ and $\mathcal{V}_i$ with $i$ in an index set. The mapping $\bar{f}$ takes $\mathcal{H}_i$ homeomorphically onto $\mathcal{V}_i$, with horizontal boundaries mapped to horizontal boundaries and vertical boundaries mapped to vertical boundaries.

**Condition** 2: Suppose $\mathcal{V}$ is a vertical strip contained in $\bigcup_i \mathcal{V}_i$. Then $\bar{f}(\mathcal{V}) \cap \mathcal{V}_i = \bar{\mathcal{V}}_i$ is a

vertical strip for every $i$. Moreover, $w(\bar{\mathcal{V}}_i) \leq \nu_v w(\mathcal{V})$ for some $0 < \nu_v < 1$ where $w(\mathcal{V})$

is the width of strip $\mathcal{V}$. Similarly, suppose $\mathcal{H}$ is a horizontal strip contained in $\bigcup_i \mathcal{H}_i$.

Then $\bar{f}^{-1}(\mathcal{H}) \cap \mathcal{H}_i = \bar{\mathcal{H}}_i$ is a horizontal strip for every $i$. Moreover, $w(\bar{\mathcal{H}}_i) \leq \nu_h w(\mathcal{H})$

for some $0 < \nu_h < 1$.

In analogy with the conditions mentioned at the end of §3.6 for the horseshoe map, we call Condition 1 the **strip condition**. Similarly, since Condition 2 requires a uniform contraction in the horizontal direction and expansion in the vertical direction, it can be called the **hyperbolicity condition**.

For the Poincaré map $P$ constructed in §3.6, the situation is more complex. Now we have four squares $U_1$ through $U_4$ together with eight families of pairwise disjoint horizontal strips and eight families of pairwise disjoint vertical strips. We state below the theorem that the Poincaré map $P$ of the PCR3BP satisfies the **generalized Conley-Moser conditions**.

**Theorem 3.7.1.** *The Poincaré map $P$ satisfies the following generalized Conley-Moser conditions:*

**Generalized Condition 1:** *$P$ maps horizontal strips to vertical strips, i.e.,*

$$P(H_n^{11}) = V_n^{11} \quad P(H_n^{12}) = V_n^{21} \quad P(H_n^{23}) = V_n^{32} \quad P(H_n^{24}) = V_n^{42}$$
$$P(H_n^{31}) = V_n^{13} \quad P(H_n^{32}) = V_n^{23} \quad P(H_n^{43}) = V_n^{34} \quad P(H_n^{44}) = V_n^{44}$$

*for all positive integers $n$, with horizontal boundaries mapping to horizontal boundaries and vertical boundaries mapping to vertical boundaries.*

**Generalized Condition 2:** *Let $V$ be a vertical strip contained in $\bigcup_i V_i^{13}$. Then*

$$V_n' = P(V) \cap V_n^{11} \quad and \quad V_n'' = P(V) \cap V_n^{21}$$

*are two vertical strips for every $n$. Moreover,*

$$w(V_n') \leq \nu_v w(V) \quad and \quad w(V_n'') \leq \nu_v w(V)$$

*for some $0 < \nu_v < 1$, where $w(V)$ is the width of $V$. Similarly, let $H$ be a horizontal*

*strip contained in $\bigcup_i H_i^{11}$. Then*

$$H_n' = P^{-1}(H) \cap H_n^{31} \quad and \quad H_n'' = P^{-1}(H) \cap H_n^{11}$$

*are two horizontal strips for every n. Moreover,*

$$w(H_n') \leq \nu_h w(H) \quad and \quad w(H_n'') \leq \nu_h w(H)$$

*for some $0 < \nu_h < 1$. Similar assertions are true for the other families of vertical and horizontal strips.*

The proof is in Koon, Lo, Marsden and Ross [2000].

Recall that

$$\mathcal{H}C_{1,n} = H_n^{11} \quad \mathcal{H}D_{1,n} = H_n^{12} \quad \mathcal{H}E_{1,n} = H_n^{31} \quad \mathcal{H}F_{1,n} = H_n^{32}$$

$$\mathcal{H}C_{2,n} = H_n^{23} \quad \mathcal{H}D_{2,n} = H_n^{24} \quad \mathcal{H}E_{2,n} = H_n^{43} \quad \mathcal{H}F_{2,n} = H_n^{44}$$

$$\mathcal{V}A_{1,n}' = V_n^{13} \quad \mathcal{V}B_{1,n}' = V_n^{11} \quad \mathcal{V}G_{1,n}' = V_n^{23} \quad \mathcal{V}H_{1,n}' = V_n^{21}$$

$$\mathcal{V}A_{2,n}' = V_n^{34} \quad \mathcal{V}B_{2,n}' = V_n^{32} \quad \mathcal{V}G_{2,n}' = V_n^{44} \quad \mathcal{V}H_{2,n}' = V_n^{42},$$

where $\mathcal{H}C_{1,n}$ is the $n$-th horizontal strip of the horizontal rectangle $C_1$ and $\mathcal{V}A_{1,n}'$ is the $n$-th vertical strip of the vertical rectangle $A_1'$, etc. Moreover, the index $ij$ of $\{H_n^{ij}\}$ indicates that the family is in the square $U_i$ and it will go to the square $U_j$ and the index $ji$ of $\{V_n^{ji}\}$ indicates that the family is in the square $U_j$ and it came from the square $U_i$, as illustrated in Figure 3.19.

We use this result to sketch the proof of the main theorem on the global orbit structure of the PCR3BP given in §3.8 and §3.9.

## 3.8   Symbolic Dynamics

In §3.6 and §3.7, we have constructed a Poincaré map $P$ on $U$ whose domain consists of four topological squares $U_i, i = 1, 2, 3, 4$, each of which is further subdivided into four smaller squares by two curves that lie on the invariant manifolds of the Lyapunov orbits. Moreover, $P$ satisfies the generalized Conley-Moser conditions.

While we need to take stock of certain new features, the basic formalism developed by

Smale, Conley and Moser still holds with a few modifications.

For the horseshoe map $h$ which bends a square $Q$ into a horseshoe and intersects it with the square, one has an infinite Cantor set of trapped points in the invariant set $\Lambda_h$, given earlier in (3.7),

$$\Lambda_h = \cap_{n=-\infty}^{\infty} h^n(Q),$$

which is the set of points in the square $Q$ that remain in the square under all forward and backward iterations by $h$.

We can define an element of the invariant set by

$$p = \{q \in Q \mid h^i(q) \in H_{s_i}, i = 0, \pm 1, \pm 2, \ldots\},$$

where $s_i$ denotes one of the elements in $\Sigma^2 = \{0, 1\}$ and $H_0, H_1$ are the two original horizontal rectangles in $D$. Moreover, an address which is a bi-infinite sequence of two symbols $\{0, 1\}$ (in $\Sigma^2$) can be attached to every point $p$ in the invariant set $\Lambda_h$, which will not only describe its location, but also tell its whole history and future under iteration of the map. By this we mean that there is a map $\phi : \Lambda_h \to \Sigma^2$ defined by

$$\phi(p) = (\ldots, s_{-n}, \ldots, s_{-1}; s_0, s_1, \ldots, s_n, \ldots),$$

where $s_i = 0$ if $h^i(p) \in H_0$ and $s_i = 1$ if $h^i(p) \in H_1$.

One easy way to imagine the invariant set $\Lambda_h$ is to draw the regions that remain trapped for one forward and one backward iteration in the square $Q$. This is the intersection of the thickest vertical and horizontal strips, so it is four squares lying in the corners of the original square. The set trapped for two iterations forwards and two backwards is obtained by intersecting the thinner strips of these figures, yielding sixteen smaller squares contained in the four squares of the first stage, as shown in Figure 3.20. Notice the addresses that have been assigned to those squares. This process can be repeated ad infinitum. After infinitely many steps, what remains is a Cantor set of points which are in one-to-one correspondence with the set of bi-infinite sequences of two symbols $\{0, 1\}$.

For the Poincaré map $P$, we can use a similar technique to visualize the invariant set $\Lambda$ and its associated set of bi-infinite sequences. Instead of one square $Q$, we have four squares $U_i, 1 = 1, 2, 3, 4$. After one forward and one backward iteration, instead of the intersections of two vertical rectangles and two horizontal rectangles, we have the
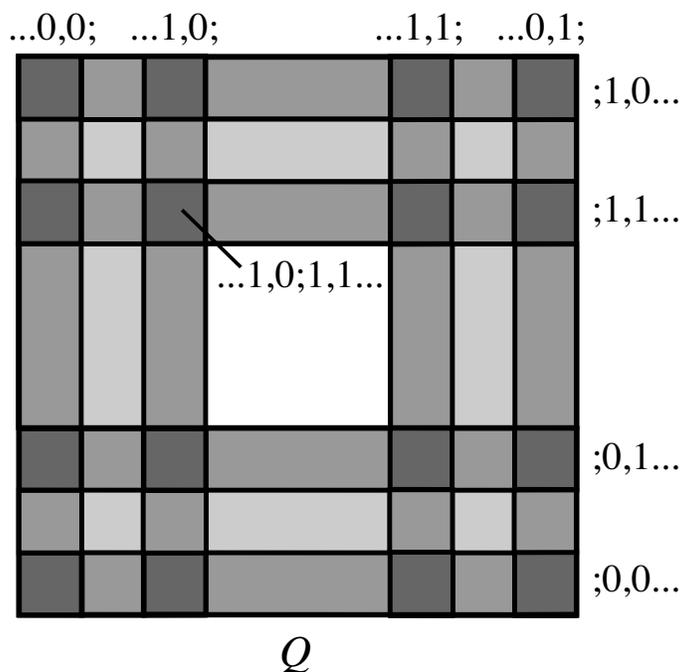
Figure 3.20: The invariant set $\Lambda_h$ of the horseshoe map $h$.

intersections of eight families of vertical strips $\{V_n^{ji}\}$ and eight families of horizontal strips $\{H_n^{ij}\}$, with the indices $ij$ corresponding to the nonzero entries of the transition matrix $A$. Using Figure 3.21 as a guide, recall from §3.6 that for $\{V_n^{ji}\}$, the index $ji$ indicates that the family is in the square $U_j$ and it came from the square $U_i$; for $\{H_n^{ij}\}$, the index $ij$ indicates that the family is in the square $U_i$ and it will go to the square $U_j$.

For simplicity of illustration, we draw Figure 3.21 schematically. Taking the family $\{H_n^{12}\}$ as an example, we draw two horizontal rectangles to represent the first and the $n$-th horizontal strips. This horizontal family is in the square $U_1$ and it will go to the square $U_2$. Similarly, for $\{V_m^{13}\}$, only the first and the $m$-th vertical rectangles are shown. This vertical family is in the square $U_1$ and it came from the square $U_3$. The same method has been used to illustrate all the other families of horizontal and vertical strips.

As for assigning the addresses for points remaining in $U$, take the "square" $S_{m;n}^{3;12}$ as an example. Since $S_{m;n}^{3;12}$ is the intersection of the horizontal strip $H_n^{12}$ and the vertical strip $V_m^{13}$, we can use $(\ldots, u_3, m; u_1, n, u_2, \ldots)$ to represent its location. As usual, the central block of this sequence also tells the history of the points in this "square" $(S_{m;n}^{3;12})$:

1. they are currently in $U_1$ and will go to $U_2$ and on their way their projection will
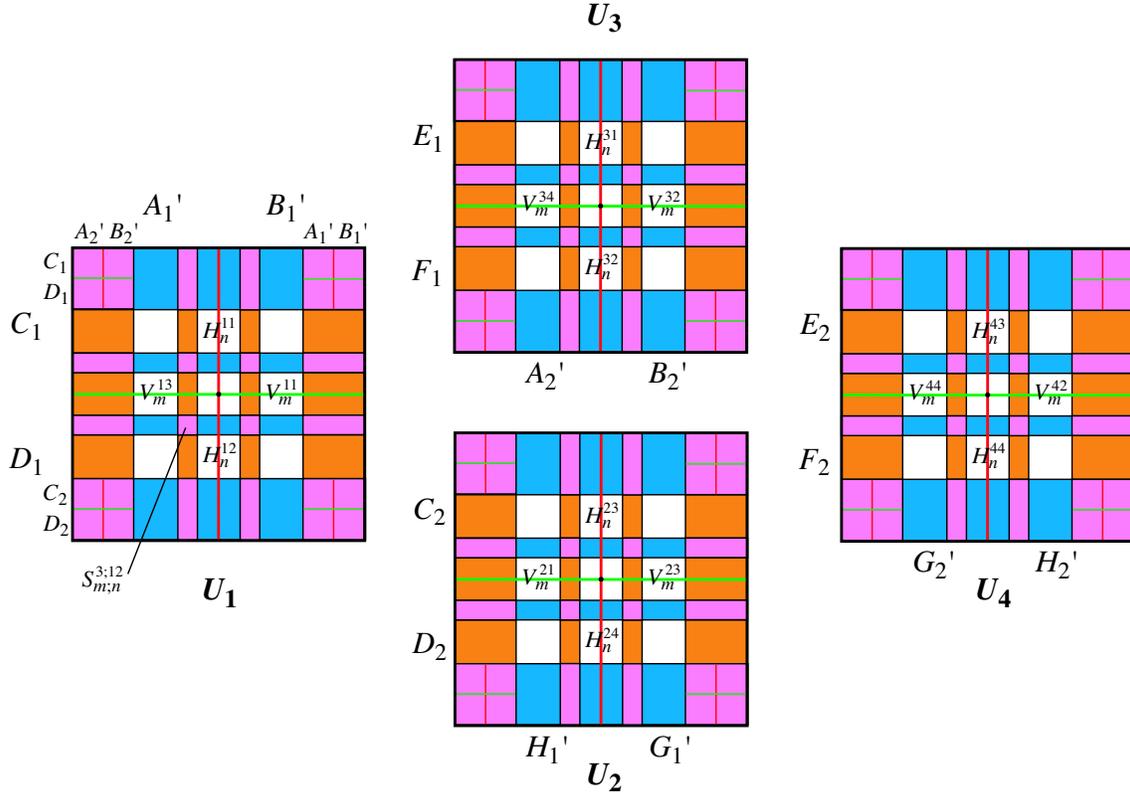
Figure 3.21: The invariant set $\Lambda$ of the Poincaré map $P$.

wind around $L_1$ for $(n + r_{\min})$ revolutions where $r_{\min}$ is the minimum number of revolutions discussed earlier in §3.6;

2. they came from $U_3$ and their position space projection has wound around $L_1$ for $(m + r_{\min})$ revolutions.

Similar sequences can be assigned to the other "squares" which are the intersections of all the other horizontal and vertical strips.

Moreover, since the Poincaré map $P$ satisfies the generalized Conley-Moser conditions, this process can be repeated ad infinitum as in the case of the horseshoe map. After an infinite number of steps, what remains in $U$ is a Cantor set of points which are in one-to-one correspondence with the set of bi-infinite sequences

$$(\ldots, (u_{i_{-1}}, n_{-1}); (u_{i_0}, n_0), (u_{i_1}, n_1), (u_{i_2}, n_2), \ldots).$$

Hence, we have shown that the invariant set $\Lambda$ for the Poincaré map $P$ corresponds to

a set of bi-infinite sequences with two families of symbols. The first family is a subshift of finite type with four symbols $\{u_1, u_2, u_3, u_4\}$ (with a transition matrix $A$ defined in §3.6). It is used to keep track of the history of the map $P$ with respect to the four squares $U_1, U_2, U_3, U_4$.

The second family is a full shift of infinite type with symbols of nonnegative integers. This set of integers is used to keep track of individual members of each vertical or horizontal family ($\{V_n^{ji}\}$ or $\{H_n^{ij}\}$). As mentioned at the end of §3.6, this set of integers also corresponds to the number of revolutions that the position space projection of an orbit winds around either $L_1$ and $L_2$.

**Singular Poincaré Map.** We discuss briefly the issue of the singular Poincaré map and how it relates to certain modifications of the space of symbol sequences $\Sigma$. Let $\Sigma = \{((u_{i_j}, n_j))\}$ be the set of bi-infinite sequences of elements of $S \times \mathbb{N}$ with a transition matrix $A$ defined on $S$. Here, $S = \{u_1, u_2, u_3, u_4\}$ and $\mathbb{N}$ is the set of non-negative integers. As usual, a compactification $\bar{\Sigma}$ of $\Sigma$ is obtained with the inclusion of sequences of the following types:

$$
\begin{aligned}
\beta &= (\ldots; (u_{i_0}, n_0), \ldots, (u_{i_k}, \infty)) \\
\gamma &= (\infty, (u_{i_{-l}}, n_{-l}), \ldots; (u_{i_0}, n_0), \ldots) \\
\delta &= (\infty, (u_{i_{-l}}, n_{-l}), \ldots; (u_{i_0}, n_0), \ldots, (u_{i_k}, \infty)).
\end{aligned}
$$

The elements of $\Sigma \subset \bar{\Sigma}$ will be called **type** $\alpha$ from now on. Moreover, the shift map $\sigma$ on $\Sigma$ defined by $\sigma((u_{i_j}, n_j)) = (u_{i_{j+1}}, n_{j+1})$ can be extended to a shift map $\bar{\sigma}$ in a natural way. The domain of $\bar{\sigma}$ is

$$
D(\bar{\sigma}) = \{(u, n) \in \bar{\Sigma} \mid n_0 \neq \infty\}
$$

and the range of $\bar{\sigma}$ is

$$
R(\bar{\sigma}) = \{(u, n) \in \bar{\Sigma} \mid n_1 \neq \infty\}.
$$

By studying Figure 3.21, it should be clear that $H_\infty^{12}$ (or $H_\infty^{11}$) is simply the horizontal curve $\Gamma_{L_1,1}^{s,\mathcal{S}}$ which is on the interior (Sun) realm branch of the stable invariant manifold of the $L_1$ Lyapunov orbit and any point on this curve will be carried forward asymptotically towards the $L_1$ Lyapunov orbit. Hence, any element of type $\beta$ corresponds to an orbit which tends to either the $L_1$ or $L_2$ Lyapunov orbit asymptotically after $k$ iterations.

Similarly, any element of type $\gamma$ corresponds to an orbit which is carried by the flow asymptotically backward towards one of the Lyapunov orbits after $l$ backward iterations. As for an element of type $\delta$, we have either a homoclinic or a heteroclinic orbit.

## 3.9　Global Orbit Structure

Now we are ready to put together all the results in §3.7 and §3.8 and to state the main theorem of this chapter which provides a symbolic dynamics description of the global orbit structure of the PCR3BP near a chain of homoclinic orbits and a symmetric heteroclinic cycle. For simplicity of exposition, we have assumed in the past that the chain consists of $(1,1)$-homoclinic orbits in the interior and exterior realms and a symmetric $(1,1)$-heteroclinic cycle in the Jupiter realm. Now we consider the general situation. Let us suppose from now on that the chain $\mathcal{C}$ is made up of a symmetric $(q_2, p_2)$-heteroclinic cycle in the Jupiter realm together with two homoclinic orbits, one of which is a $(q_1, p_1)$ orbit in the interior realm and the other is a $(q_3, p_3)$ orbit in the exterior realm.

**Theorem 3.9.1.** **(Global Orbit Structure)** *Consider an element* $(u, r) \in \bar{\Sigma}$ *with* $r_j \geq r_{\min}$ *for all* $j$. *Then there are initial conditions, unique in a neighborhood of the given chain of two homoclinic orbits and one symmetric heteroclinic cycle (associated with* $p_{1,1}, p_{2,2}, p_{1,2}, p_{2,1}$, *respectively), such that the following statements are true.*

1. *For an element of type*

$$\alpha = (\ldots, (u_{i_{-1}}, r_{-1}); (u_{i_0}, r_0), (u_{i_1}, r_1), (u_{i_2}, r_2), \ldots),$$

*the orbit corresponding to such conditions starts at* $U_{i_0}$ *and goes to* $U_{i_1}$ *if* $(A)_{i_0 i_1} = 1$. *This orbit passes through either the equilibrium region* $\mathcal{R}_1$ *or* $\mathcal{R}_2$ *depending on whether the initial index* $i_0$ *is* $1, 3$ *or* $2, 4$. *If* $i_0 = 1, 3$, *the projection of the orbit winds around* $L_1$ *for* $r_0$ *revolutions inside the region* $\mathcal{R}_1$ *before leaving for* $U_{i_1}$. *Otherwise, it winds around* $L_2$ *for* $r_0$ *revolutions before leaving for* $U_{i_1}$. *After that, the same process begins with* $(u_{i_1}, r_1)$ *in place of* $(u_{i_0}, r_0)$ *and* $(u_{i_2}, r_2)$ *in place of* $(u_{i_1}, r_1)$, *etc. For negative time a similar behavior is described for* $(u_{i_{-1}}, r_{-1}), (u_{i_0}, r_0)$, *etc.*

   *For this orbit, the number of revolutions that the comet winds around Jupiter or the Sun (in the interior or exterior realm) is a constant which depends on the realm and*

*the given chain of homoclinic orbits and heteroclinic cycle. For the Jupiter realm, the number is $(q_2 + p_2 - 1)/2$. For the interior and exterior realms, the number is $q_1 + p_1 - 1$ and $q_3 + p_3 - 1$, respectively. Note that $q_i$ and $p_i$ are positive integers.*

2. *For an element of type*

$$\beta = (\ldots ; (u_{i_0}, r_0), \ldots, (u_{i_k}, \infty)),$$

*the orbit tends asymptotically towards one of the Lyapunov orbits after $k$ iterations. If $u_{i_k} = 1, 3$, the orbit tends towards the $L_1$ orbit and stays in region $\mathcal{R}_1$. If $u_{i_k} = 2, 4$, it tends towards the $L_2$ orbit and stays in region $\mathcal{R}_2$.*

3. *For an element of type*

$$\gamma = (\infty, (u_{i_{-l}}, r_{-l}), \ldots ; (u_{i_0}, r_0), \ldots),$$

*the orbit tends asymptotically backward towards one of the Lyapunov orbits after $l$ backward iterations. If $u_{i_{-l}} = 1, 2$, the orbit tends towards the $L_1$ orbit and stays in region $\mathcal{R}_1$. If $u_{i_{-l}} = 3, 4$, it tends towards the $L_2$ orbit and stays in region $\mathcal{R}_2$.*

4. *For an element of type*

$$\delta = (\infty, (u_{i_{-l}}, r_{-l}), \ldots ; (u_{i_0}, r_0), \ldots, (u_{i_k}, \infty)),$$

*the orbit tends asymptotically towards the $L_1$ or $L_2$ Lyapunov orbit after $k$ iteration, depending on whether $u_{i_k} = 1, 3$ or $2, 4$. It also tends asymptotically backward towards the $L_1$ or $L_2$ orbit after $l$ iterations backwards, depending on whether $u_{i_l} = 1, 2$ or $3, 4$.*

We provide a sketch of the proof here, which makes use of the major results in §3.7 and §3.8. The proof itself is in Koon, Lo, Marsden and Ross [2000]. While we still need to fully establish the fact that the Poincaré map $P$ does satisfy the generalized Conley-Moser conditions as mentioned at the end of §3.7, we refer the reader to the proof in Koon, Lo, Marsden and Ross [2000] so that we can discuss the implications of this theorem.

**Sketch of Proof.** First construct a Poincaré map $P$ whose domain $U$ consists of four different squares $U_i, i = 1, 2, 3, 4$. Squares $U_1$ and $U_4$ are contained in the surface $y = 0$ and they center around $(q_1, p_1)$ and $(q_3, p_3)$-transversal homoclinic points in the interior and the exterior realms, respectively. Squares $U_2$ and $U_3$ are contained in the surface $x = 1 - \mu$ and center around $(q_2, p_2)$-transversal heteroclinic points in the Jupiter realm which are symmetric with respect to each other.

Adjust the widths of all the corresponding pairs of the thin strips on the bounding spheres so that the minimum number of revolutions $r_{\min}$ around $L_1$ or $L_2$ is the same for all the $U_i$'s. With this adjustment, any orbit which pierces $V_m^{ji}$ is now in $U_j$. It came from $U_i$ and has wound around $L_1$ (if $u_i = 1, 3$) or $L_2$ (if $u_i = 2, 4$) for $(m + r_{\min})$ times. A similar analysis holds for $H_n^{ji}$.

Assume that we have shown that the Poincaré map $P$ satisfies the generalized Conley-Moser conditions. Then our discussion in §3.8 on symbolic dynamics shows that for any bi-infinite sequence of type $\alpha$, $\alpha = (u, r)$, we can find initial conditions $(u, n)$ in $U$ such that the orbit with this initial condition has exactly the history of $(u, r)$. Here, $r_j = n_j + r_{\min}$. Similar arguments also hold for bi-infinite sequences of other types.

**Some Comments on the Implications of the Theorem.** Type $\alpha$ orbits include "oscillating," "capture" and "non-transit" orbits. Oscillating orbits are orbits which cross from one realm to the other infinitely many times, capture orbits are orbits which cross sometime but eventually stay in one realm, and non-transit orbits always stay in the same realm. Type $\beta$ and type $\gamma$ orbits are asymptotic orbits which wind to one of the Lyapunov orbits. Type $\delta$ orbits are homoclinic and heteroclinic orbits.

Similar to the standard textbook example, it is easy to verify that both the shift map $\bar{\sigma}$ and the Poincaré map $P$ have the following properties:

1. a countable infinity of periodic orbits of all periods,

2. an uncountable infinity of nonperiodic orbits, and

3. a "dense orbit."

Moreover, both $\bar{\sigma}$ and $P$ model the phenomenon that is called ***deterministic chaos*** in dynamical systems theory. Most notably, they exhibit the phenomenon of sensitive dependence on initial conditions, i.e., the distance between nearby initial conditions grows

under some fixed number of iterates. This phenomenon corresponds to the "random" jumping of the comets between the interior, the Jupiter and the exterior realms.

# Chapter 4

# Construction of Trajectories with Prescribed Itineraries

Let us summarize the major results of the thesis to this point. We have been developing a framework for understanding the motion of a particle in the gravity field of two massive bodies, $m_1$ and $m_2$. In particular, we have considered the planar circular restricted three-body model. In this model, there is a constant of the motion, the energy, which divides the phase space of the particles motion into five cases (see Figure 2.5).

In the first four cases, depicted in Figure 2.4, the energy surface is naturally divided into three large realms of motion:

1. the $m_1$ realm, surrounding $m_1$;

2. the $m_2$ realm, surrounding $m_2$;

3. the exterior realm, which includes neither $m_1$ nor $m_2$, and is exterior to them both.

The energy cases are defined according to which realms of motion are connected. The connections appear as necks surrounding the location of libration points. The necks increase their width with increasing energy, corresponding to "easier," or more probable, transport between realms with increasing energy. For example, in case 3, the particle has enough energy to move between all three realms: the $m_1$, $m_2$, and exterior realms. The particle moves between realms via necks surrounding $L_1$ (connecting the $m_1$ and $m_2$ realms) and $L_2$ (connecting the $m_2$ and exterior realms). According to the terminology developed in Chapters 2 and 3, the neck regions surrounding $L_1$ and $L_2$ are denoted $R_1$ and $R_2$, respectively, as in Figure 4.1.
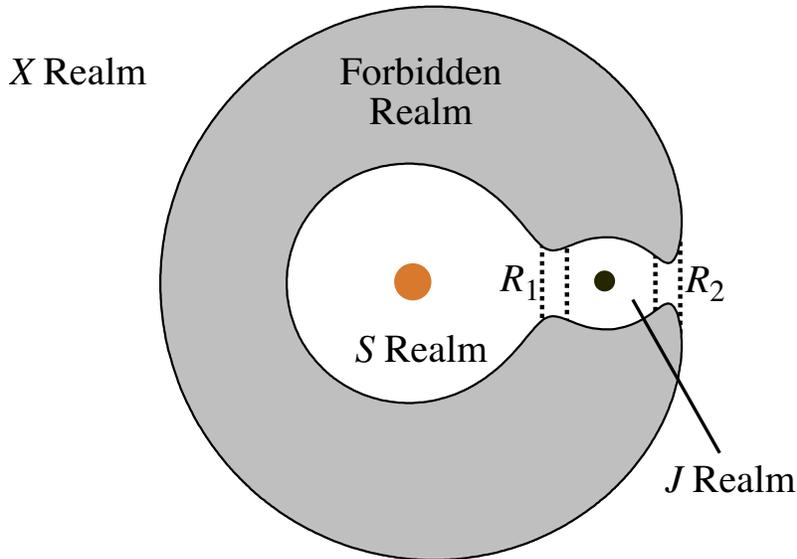
Figure 4.1: A schematic of the rotating frame for a particle in the gravitational field of the sun and Jupiter. Here, $m_1 = S$ and $m_2 = J$. Thus, the realm around the sun is the $S$ realm, the realm around Jupiter is the $J$ realm, and the realm not containing $S$ or $J$ is the $X$ realm. One can construct orbits which connect the three realms using the stable and unstable manifold tubes associated to libration orbits in the necks around $L_1$ and $L_2$, the equilibrium regions $R_1$ and $R_2$, respectively (discussed in Chapters 2 and 3).

The stable and unstable manifold tubes emanating from libration orbits in these necks are the objects governing the motion between realms. Since these are global objects, we can, in theory, compute them out to arbitrarily long times and distances from the neck. Particles with initial conditions interior to a stable (unstable) manifold tube are guaranteed to move from one realm to another when evolved forward (backward) in time. When one finds intersections between the regions interior to stable and unstable manifolds, one can pick any initial condition in the intersection region and integrate it both forward and backward. The resulting solution in the phase space corresponds to a desired particle path, i.e., a desired itinerary for the particle.

## 4.1 Trajectories with Prescribed Itineraries

In this section, we learn the basic ingredients for constructing orbits with prescribed itineraries. For simplicity of exposition in the discussion which follows, consider the planar

motion of a particle in the gravitational field of the sun and Jupiter ($\mu = 9.537 \times 10^{-4}$). We label the realm around the sun with an $S$, the realm around Jupiter with a $J$, and the exterior realm with an $X$, as in Figure 4.1. We will use the set of symbols $\{S, J, X\}$ to denote the location of the trajectory to construct finite itineraries of length $k$ of the form $(A_1, A_2, \ldots, A_k)$, where $A_i \in \{S, J, X\}, i = 1, \ldots, k$ . Using a conceptually simple procedure, trajectories with arbitrarily large itineraries can be constructed numerically.

**A Trajectory with Itinerary $(X, J, S)$.**   Suppose we want to find an initial condition corresponding to a particle which begins in the exterior realm and passes through the Jupiter realm to the sun realm. We transcribe this goal into a search for an initial condition with the itinerary $(X, J, S)$. In principle, we could start with a large number of initial points in the four-dimensional phase space and save only those whose orbits correspond with this itinerary. But we can simplify the search tremendously by using *tube dynamics* on an energy surface. Then our search becomes one of searching for an area on a two-dimensional Poincaré section for which all the points in that area correspond to an initial condition with this itinerary. We will use the following step by step procedure to find the itinerary region, or "itinerarea," which corresponds to an orbit with itinerary $(X, J, S)$.

**Algorithm for Finding An "Itinerarea."**   The reader may wish to reproduce the results of this section in order to gain familiarity with the method.

1. **Select an appropriate energy.** One first needs to set the energy to a value such that the itinerary we seek exists. We want the particle to go between all three realms, $X$, $J$, and $S$, so we need to be in energy case 3 as described in §2.4. For the given $\mu$, we compute the case 3 energy interval,

$$[E_2, E_3] \approx [-1.519, -1.501].$$

For illustrative purposes, we will take a value $e \in [E_2, E_3]$ near the lower end of the interval. This corresponds to necks around $L_1$ and $L_2$ which are slightly open. The value we will use for the construction is $e = -1.515$. A schematic of the realms of possible motion for this energy is shown in Figure 4.2(a), and notice the labeling of the $X$, $J$, and $S$ realms.

2. **Compute the $L_1$ and $L_2$ periodic orbits.** Consider the libration point $L_i$, stand-

(a)                                                    (b)
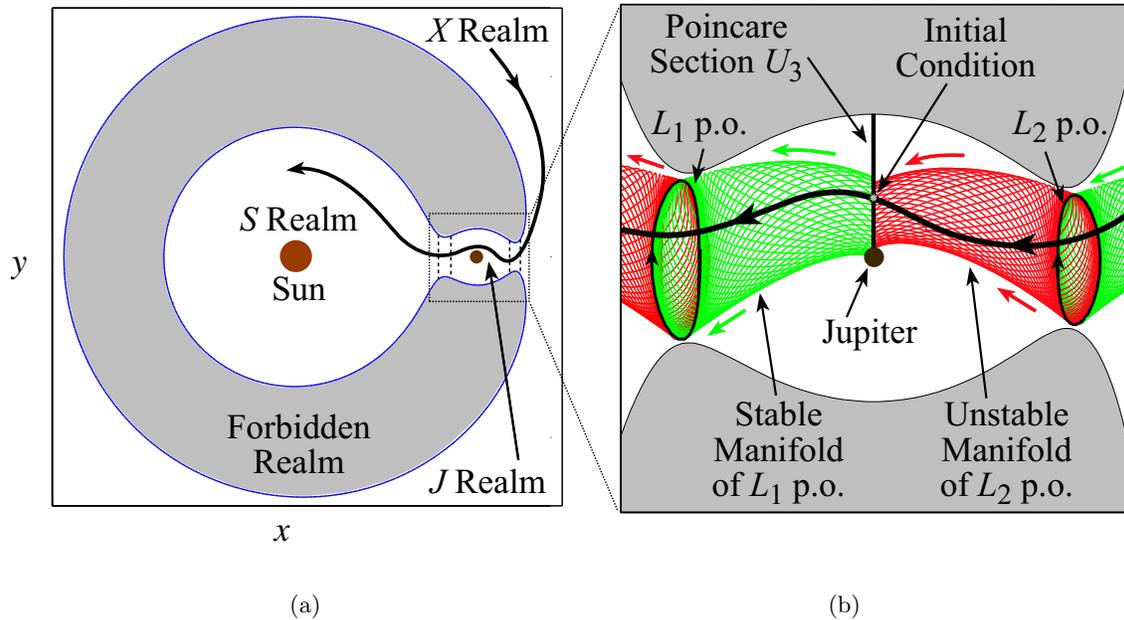
Figure 4.2: (a) A schematic of the realms of possible motion for a case 3 energy. The $X$, $J$, and $S$ realms are labeled. The trajectory shown as a heavy black line is a trajectory with an itinerary $(X, J, S)$. (b) A close-up around the $J$-realm. The position space projection of the stable and unstable manifold tubes of the $L_i, i = 1, 2$ periodic orbits (p.o.) are shown. The $J$-branch of the $L_1$ stable (resp., $L_2$ unstable) tubes are labeled. We seek the intersection of the interior of these two tubes on the Poincaré section $U_3$. Taking an initial condition from this intersection and numerically integrating it forward and backward in time produces the trajectory with itinerary $(X, J, S)$.

ing for either $L_1$ or $L_2$. We will consider one procedure which computes periodic orbits around $L_i$ in a relatively straightforward fashion. This procedure begins with small "seed" periodic orbits obtained from the linearized equations of motion near $L_i$, and uses differential correction and numerical continuation to generate the desired the periodic orbit corresponding to the chosen energy $e$. We will make a remark on other procedures.

(a). *Computing the Location of the Equilibrium Points.* Compute the location of $L_i$, ( $x_e$, 0, 0, 0), using the procedure in §2.5. Consider the linearized equations of motion in a coordinate system centered on $L_i$, Eq. (2.22). The eigenvalues and eigenvectors for the linearized system are given by explicit formulas in §2.7. One can then compute, using the general solution (2.31) to (2.22), the initial conditions for a periodic orbit (p.o.) of $x$ amplitude $A_x > 0$. In (2.31), let $t = \alpha_1 = \alpha_2 = 0$ and $\beta = -A_x/2$. When transformed back to the original coordinates, this yields an

initial condition

$$\begin{aligned}
\bar{x}_0 &= (\,x_e,\ 0,\ 0,\ 0\,) + 2\mathrm{Re}(\beta w_1),\\
&= (\,x_e - A_x,\ 0,\ 0,\ v_{y0}\,), \hspace{4cm} (4.1)
\end{aligned}$$

where $v_{y0} = -A_x \nu \tau > 0$, $\nu = \frac{1}{2}(c - 2 - \sqrt{9c^2 - 8c}) > 0$, and $\tau = -(\nu^2 + 2c + 1)/2\nu$, using relations from §2.7, where $c$ is given in (2.20).

(b). *Generating a Family of Periodic Orbits Using Differential Correction and Numerical Continuation.* The initial condition given in (4.1) will only yield a good approximation to a p.o. in the nonlinear equations (2.15) in the case of $A_x \ll 1$. But we want a p.o. of energy $e$, which may correspond to a large amplitude. We thus proceed as follows. Let $A_{x1}, A_{x2} \ll 1$ be two small amplitudes, with $A_{x1} < A_{x2}$, and corresponding initial conditions $\bar{x}_{0,g}^{(1)}$ and $\bar{x}_{0,g}^{(2)}$, respectively, where $g$ denotes that this is an initial guess to a true periodic solution of (2.15). We will use **differential correction**, a targeting procedure, to generate p.o.'s in the nonlinear equations which are accurate to some specified tolerance $d$. In other words, if $\bar{x}_{\mathrm{po}}(0) \equiv \bar{x}_0$ is an initial condition on a p.o., $x_{\mathrm{po}}(t)$, of period $T$, we want

$$|\bar{x}_{\mathrm{po}}(T) - \bar{x}_{\mathrm{po}}(0)| < d,$$

for a specified $d << 1$.

Differential correction uses the analytical approximation as the first guess in an iterative process which updates the initial conditions while keeping some values constant. In our case, we want to keep the $x$ value constant and update the $y$ velocity.

Given a reference solution $\bar{x}(t)$ going from $\bar{x}_0$ to $\bar{x}_1$ under

$$\dot{x} = f(x),$$

we want to adjust $\bar{x}_0$ by $\delta\bar{x}_0$ so trajectory will end at a desired phase point $x_d$, near $\bar{x}_1$. We thus need the sensitivity of $\delta\bar{x}_1$ w.r.t. $\delta\bar{x}_0$. The linear approximation to this sensitivity is given by the *state transition matrix*, discussed below.

Let trajectories of the differential equations $\dot{x} = f(x)$, e.g., (2.15), with $x(t_0) = x_0$

be denoted by $\phi(t, t_0; x_0)$. A trajectory that starts from the perturbed initial vector $\bar{x}_0 + \delta\bar{x}_0$, and evolves until $t + \delta t$, progresses with the distance

$$\delta\bar{x}(t + \delta t) = \phi(t + \delta t, t_0; \bar{x}_0 + \delta\bar{x}_0) - \phi(t, t_0; \bar{x}_0)$$

with respect to the reference solution $\bar{x}(t)$.

Measuring the distance at time $t_1 + \delta t_1$ gives

$$\delta\bar{x}(t_1 + \delta t_1) = \phi(t_1 + \delta t_1; t_0; \bar{x}_0 + \delta\bar{x}_0) - \phi(t_1, t_0; \bar{x}_0).$$

Taylor expansion yields

$$\begin{aligned}
\delta\bar{x}(t_1 + \delta t_1) &= \frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial x_0}\delta\bar{x}_0 + \frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial t_1}\delta t_1 + \mathrm{higher - order\ terms}, \\
&= \frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial x_0}\delta\bar{x}_0 + \dot{\bar{x}}_1\delta t_1 + \mathrm{h.o.t.},
\end{aligned}$$

where the first part of the second term comes from $\frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial t_1} = \frac{d\phi(t, t_0; \bar{x}_0)}{dt} = f(\phi(t, t_0; \bar{x}_0))$, evaluated at $t = t_1$. The matrix $\frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial x_0}$ which satisfies the above relation to first order (when $\delta t_1 = 0$) is called **state transition matrix**. Usually abbreviated as $\Phi(t_1, t_0)$, this matrix given by

$$\delta\bar{x}(t_1) = \Phi(t_1, t_0)\delta\bar{x}_0, \tag{4.2}$$

will play an important role in differential correction. Equation (4.2) can also be seen as the solution to the **variational equations** of (2.15) which are only linearized equations for variations $\delta\bar{x}$ given by

$$\delta\dot{\bar{x}}(t) = Df(\bar{x}(t))\delta\bar{x},$$

where the Jacobian matrix evaluated at $\bar{x}(t)$ is

$$Df(\bar{x}(t)) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\bar{U}_{xx} & -\bar{U}_{xy} & 0 & 2 \\ -\bar{U}_{yx} & -\bar{U}_{yy} & -2 & 0 \end{pmatrix}_{\bar{x}(t)},$$

and $\bar{U}_{ab}$ are the double partial derivatives of the effective potential (2.7).

Suppose we want to reach a desired endpoint, $x_d$, but

$$\bar{x}(t_1) = \phi(t_1, t_0; \bar{x}_0) = \bar{x}_1 = x_d - \delta\bar{x}_1,$$

is slightly off ($|\delta\bar{x}_1| > d$) and we need to correct it. Since

$$
\begin{aligned}
\phi(t_1, t_0; \bar{x}_0 + \delta\bar{x}_0) &= \phi(t_1, t_0; \bar{x}_0) + \frac{\partial\phi(t_1, t_0; \bar{x}_0)}{\partial x_0}\delta\bar{x}_0 + \text{h.o.t.,} \\
&= \phi(t_1, t_0; \bar{x}_0) + \Phi(t_1, t_0)\delta\bar{x}_0 + \text{h.o.t.,} \\
&= \bar{x}_1 + \delta\bar{x}_1 + \text{h.o.t.,} \\
&= x_d + \text{h.o.t.,}
\end{aligned}
$$

this implies that changing $\bar{x}_0$ by $\delta\bar{x}_0 = \Phi(t_1, t_0)^{-1}\delta\bar{x}_1$ will perform the correction to first order. By iteration, the process produces convergence:

$$|\phi(t_1, t_0; \bar{x}_0 + \Delta\bar{x}_0) - x_d| < d,$$

where $\Delta\bar{x}_0$ is the accumulation of corrections $\delta\bar{x}_0$ which yields $x_d$ within the desired tolerance $d$.

We seek periodic orbits which are symmetric w.r.t. the $x$-axis ($y = 0$), noting that $y \mapsto -y, t \mapsto -t$ leaves equations of motion (2.15) unchanged, i.e., the symmetry gives mirror image solution $\bar{x}(-t)$ for each $\bar{x}(t)$, completing the other half of the periodic orbit.

From the earlier step, we choose an approximate initial condition (at $t_0 = 0$) which intersects the $x$-axis perpendicularly,

$$\bar{x}(0) = (x_0 \ 0 \ 0 \ v_{y0})^T.$$

Using a standard Runga-Kutta 7(8) integration package with an appropriate tolerance (say, $10^{-14}$), we integrate this initial condition until the next $x$-axis crossing using the following procedure: (1) integrate until $y(t)$ changes sign; (2) then change the time step until, e.g., $|y(t)| < 10^{-11}$ which we refer to as the "crossing"; (3) at the crossing, $t_1 \equiv t$, $y_1 \equiv y(t_1)$.

This gives use $\bar{x}(t_1)$, so we can also compute $\Phi(t_1, 0)$ as well. For a p.o., the desired final state has the form

$$\bar{x}(t_1) = (x_1 \ 0 \ 0 \ v_{y1})^T,$$

where $t_1 = T/2$, the time of one half-cycle of the p.o. The actual value for $v_{x1}$ as a result from numerical integration may not be 0. For our purposes, we want $|v_{x1}| < d$, e.g., $d = 10^{-8}$. The state transition matrix after one half-cycle, $\Phi(t_1, 0)$, can be used to adjust the initial values to obtain a p.o. as

$$\delta\bar{x}_1 \approx \Phi(t_1, 0)\delta\bar{x}_0 + \dot{\bar{x}}_1\delta t_1.$$

Suppose $|v_{x1}| > d$ and we hold $x_0$ fixed. The correction to $v_{y0}$ can be calculated from

$$\begin{aligned}
\delta v_{x1} &= \Phi_{34}\delta v_{y0} + \dot{v}_{x1}\delta t_1 + \text{h.o.t.}, \\
0 = \delta y_1 &= \Phi_{24}\delta v_{y0} + v_{y1}\delta t_1 + \text{h.o.t.},
\end{aligned}$$

where $\Phi_{ij}$ is an element of the matrix $\Phi(t_1, 0)$. Here, $\delta v_{x1} = v_{x1}$ since we want $v_{x1} = 0$. Hence,

$$\delta v_{y0} \approx \left(\Phi_{34} - \frac{1}{v_{y1}}\Phi_{24}\right)^{-1} v_{x1},$$

can be used to cancel out $v_{x1}$ if we let

$$v_{y0} \mapsto v_{y0} - \delta v_{y0}.$$

This process converges to $|v_{x1}| < d$ within a few iterations typically.

The above procedure yields an accurate initial condition for a p.o. from a single initial guess. If our initial guess came from the linear approximation near the equilibrium point, it has been observed numerically that we can only use this procedure for small amplitude p.o.'s around $L_i$, say $\approx 10^{-4}$ for $\mu = 9.537 \times 10^{-4}$. But if we want an orbit of arbitrarily large amplitude (which is in one-to-one correspondence with the energy $e$), we need to use numerical continuation to generate a family of orbits which reaches the appropriate energy $e$.

We proceed as follows. Suppose we find two small nearby p.o. initial conditions,

$\bar{x}_0^{(1)}, \bar{x}_0^{(2)}$, correct to within the tolerance $d$, using the differential correction procedure described above. We can generate a family of p.o.'s with increasing amplitude around $L_i$ in the following way. Let

$$
\begin{aligned}
\Delta &= \bar{x}_0^{(2)} - \bar{x}_0^{(1)}, \\
&= (\Delta x_0 \ 0 \ 0 \ \Delta v_{y0})^T.
\end{aligned}
$$

Extrapolate to an initial guess for $\bar{x}_0^{(3)}$ via

$$
\begin{aligned}
\bar{x}_{0,g}^{(3)} &= \bar{x}_0^{(2)} + \Delta, \\
&= \left( (x_0^{(2)} + \Delta x_0) \ 0 \ 0 \ (v_{y0}^{(2)} + \Delta v_{y0}) \right)^T \\
&= \left( x_0^{(3)} \ 0 \ 0 \ v_{y0}^{(3)} \right)^T.
\end{aligned}
$$

Keeping $x_0^{(3)}$ fixed, we can use differential correction to compute an accurate solution $\bar{x}_0^{(3)}$ from the initial guess $\bar{x}_{0,g}^{(3)}$ and repeat the process until we have a family of solutions. We can keep track of the energy of each p.o. and when we have two solutions, $\bar{x}_0^{(k)}, \bar{x}_0^{(k+1)}$, whose energies bracket the desired energy $e$, we can refine our continuation until we find a p.o. of energy $e$ to within a desired amount.

(c). *Remark on Other Procedures.* One can expand the nonlinear equations of motion (2.15) to some truncation order $k$ using the Legendre polynomials $P_n, n = 3, \ldots, k$. First, assume a periodic orbit exists around $L_i$ for energy $e = E_i + \varepsilon_i$ where $\varepsilon_i > 0$ is small. An analytical approximation to a periodic solution of the $j$th order equations of motion can be obtained, beginning with a sinusoidal solution to the linearized equations of motion. Second, using the Lindstedt-Poincaré method, one can obtain higher-order approximations to the periodic solution, finally obtaining an approximation up to our truncation order $k$. This procedure is described in Gómez, Jorba, Masdemont and Simó [1991]. Analytical expansions to the periodic solution are extremely useful, but may require thousands of terms to obtain the desired accuracy (e.g., if $k > 10$). The accuracy of the periodic solution can be measured as the distance between the initial condition and final condition after one period. One typically wants this to be smaller than $10^{-8}$, but this depends on the application.

3. **Computation of invariant manifolds.** First, we find the local approximations to

the unstable and stable manifolds of the p.o. A simple way to compute an approximation of $W^u_{L_1,p.o.}$ (resp., $W^s_{L_1,p.o.}$) is provided by Parker and Chua [1989] and is based on Floquet theory. Once a periodic orbit of period $T$ has been obtained in the previous step, one computes the state transition matrix over one period $\Phi(T,0)$, otherwise known as the *monodromy matrix*. From the eigenvectors of the monodromy matrix, the local approximations of the unstable (resp., stable) manifolds of the periodic solution are obtained.

The linear approximation, in the form of a state vector, is integrated in the nonlinear equations of motion (2.15) to produce the approximation of the stable and unstable manifolds, a procedure known as *globalization of the manifolds*. The computation is well covered in Howell, Barden and Lo [1997] (see also Gómez, Masdemont and Simó [1993]). For periodic orbits in the three-body problem, this approach works well.

4. **Take a Poincaré surface of section of the globalized stable and unstable manifolds.** In Chapter 3, we defined the four Poincaré surfaces of section, $U_i, i = 1, \ldots, 4$, which were locally defined around heteroclinic points, with a map $P$ linking them. In order to link the present numerical construction with the earlier theoretical framework and terminology, we adopt the following convention. The $U_1$ and $U_4$ (Poincaré) sections will be defined by the following two-dimensional surfaces:

$U_1 = \{(x, \dot{x}) \mid y = 0, \ x < 0, \ \dot{y}(x, \dot{x}; e) < 0\}$, in the $S$ realm;

$U_4 = \{(x, \dot{x}) \mid y = 0, \ x < -1, \ \dot{y}(x, \dot{x}; e) > 0\}$, in the $X$ realm,

where $\dot{y}(x, \dot{x}; e)$ denotes that $\dot{y}$ is obtained from the energy equation (2.11). The $U_2$ and $U_3$ sections will be defined by the following:

$U_2 = \{(y, \dot{y}) \mid x = 1 - \mu, \ y < 0, \ \dot{x}(y, \dot{y}; e) > 0\}$, in the lower half of the $J$ realm;

$U_3 = \{(y, \dot{y}) \mid x = 1 - \mu, \ y > 0, \ \dot{x}(y, \dot{y}; e) < 0\}$, in the upper half of the $J$ realm.

Figure 4.3 depicts the locations of the Poincaré sections in the rotating frame.

The $U_i$ are at strategically placed locations, allowing us to get cross sections of the flow within the three-dimensional energy surface $\mathcal{M}(\mu, e)$. To pick the appropriate $U_i$ on which to find an $(X, J, S)$ itinerarea, we reason as follows. From our discussions regarding the $L_1$ and $L_2$ p.o. stable and unstable manifold tubes in Chapters 2 and 3, we know that, in a frame parallel to the rotating frame but centered on the point $L_i$, the two unstable manifold tube branches are locally heading in the second and fourth quadrants. Similarly,
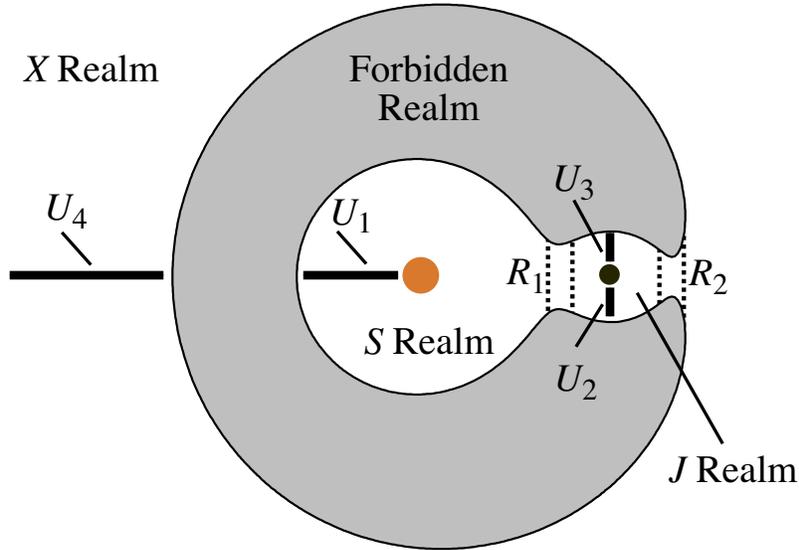
Figure 4.3: The location of the four Poincaré sections $U_1$, $U_2$, $U_3$, and $U_4$, with respect to the $S$, $J$, and $X$ realms and the neck regions, $R_1$ and $R_2$, connecting them.

the unstable manifold tube branches are locally coming from the first and third quadrants. To refresh your memory, review Figure 2.11 and §2.9.

---

**Aside: Why does this method work?** Recall from Chapters 2 and 3 the McGehee representation of the equilibrium region $\mathcal{R}$, which is in between two realms, e.g., the $S$ and $J$ realms. Emanating from the unstable p.o. are four cylinders of asymptotic orbits which form pieces of the stable and unstable manifold tubes of the p.o. They intersect the bounding spheres at asymptotic circles, separating spherical caps, which contain transit orbits, from spherical zones, which contain non-transit orbits. In order for an initial condition $s_0 \in \mathcal{R}$ to transit from one realm to another, it must be inside the tubes.

For a fixed energy in case 3, consider the spherical caps of transit orbits. These are building blocks from which we construct orbits of prescribed itineraries. Consider their images and pre-images on a suitable Poincaré section, e.g., the Poincaré section $U_3$ between $L_1$ and $L_2$. Spherical caps have the geometry of a disk, so we expect the images and pre-images will also appear as disk, or distorted disks. In Figure 4.4, the image of the cap on the left bounding sphere of the $L_2$ equilibrium region $\mathcal{R}_2$ is shown, containing trajectories leaving $\mathcal{R}_2$. On this same figure, we show the pre-image of the cap on the right bounding
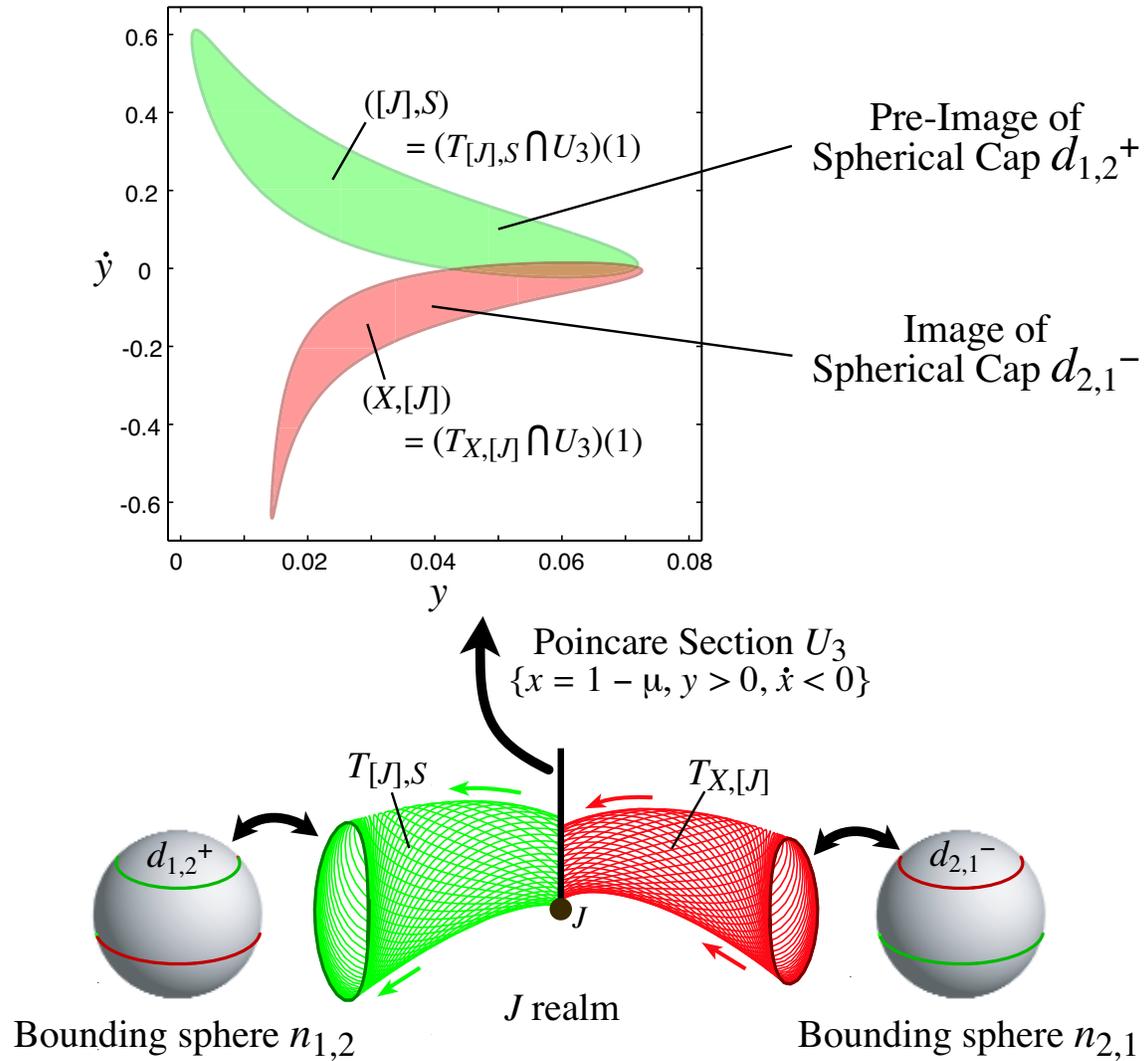
Figure 4.4: We seek transit orbits from the exterior to interior realm by looking at the intersections of images and pre-images of the "caps of transit orbits," introduced for the equilibrium regions in Chapters 2 and 3. See the text for an explanation.

sphere of $\mathcal{R}_1$ containing orbits entering $\mathcal{R}_1$.

The intersection of the unstable manifold tube of the $L_2$ p.o. with $U_3$ forms the boundary of the image of the cap containing transit orbits leaving $\mathcal{R}_2$. All of these orbits came from the $X$ realm and are now in the $J$ realm, so we label this region $(X, [J])$.

Similarly, the intersection of the stable manifold tube of the $L_1$ p.o. with $U_3$ forms the boundary of the pre-image of the cap of transit orbits entering $\mathcal{R}_1$. All of these orbits are now in the $J$ realm and are headed for the $S$ realm, so we label this region $([J], S)$.

Note that the regions $(X, [J])$ and $([J], S)$ intersect.

---

5. **Consider tube dynamics to compute the desired itinerarea.** From our discussion in Chapter 3, we know that the stable and unstable manifold tubes of the $L_1$ and $L_2$ p.o.'s bound regions in the energy surface exactly corresponding to motion between realms.

Key to our construction is the connectivity of the stable and unstable manifolds of the $L_1$ and $L_2$ p.o.'s. Consider Figure 4.5, where we show their projection onto position space which appear as strips of variable width. We show the tube projections up to their first intersection with the $U_i$.

For convenience in the discussion which follows, we introduce a new labeling convention. The set $T_{[A],B}$ is the *solid tube* of trajectories which are currently in the $A$ realm and heading toward the $B$ realm. The boundary of $T_{[A],B}$ is the stable manifold of the p.o. lying in the neck between the $A$ and $B$ realms. Similarly, the set $T_{A,[B]}$ is the solid tube of trajectories which came from the $A$ realm and are currently in the $B$ realm and its boundary is the unstable manifold of the p.o. lying in the neck between the $A$ and $B$ realms.

Consider the $J$ realm. Suppose the initial condition for the trajectory we want to construct with itinerary $(X, J, S)$ is in this realm. Then the itinerarea is $(X, [J], S)$. All particles with initial conditions in the itinerarea labeled $(X, [J], S)$ are such that when numerically integrated backward in time they transit to the $X$ realm, and when integrated forward in time they transit to the $S$ realm, like the trajectory shown in Figure 4.2. During the backward integration segment, the trajectory was within the $J$-branch of the $L_2$ p.o. unstable tube, labeled $T_{X,[J]}$ in Figure 4.5. Similarly, during the forward integration segment, the trajectory will be within the $J$-branch of the $L_1$ p.o. stable tube, labeled $T_{[J],S}$ in Figure 4.5. These two tubes are known from numerical experiments to intersect on the $U_3$ Poincaré section.

$T_{X,[J]}$ may conatin pieces which wind around Jupiter several times. Therefore, $T_{X,[J]}$ will intersect $U_3$ several times. We denote the $n$th intersection of $T_{X,[J]}$ with $U_3$ by $\left(T_{X,[J]} \bigcap U_3\right)(n)$. For the present, we will restrict ourselves to $n = 1$.

The set $\left(T_{X,[J]} \bigcap U_3\right)(1)$ is an itinerarea of particles which came from the $X$ realm
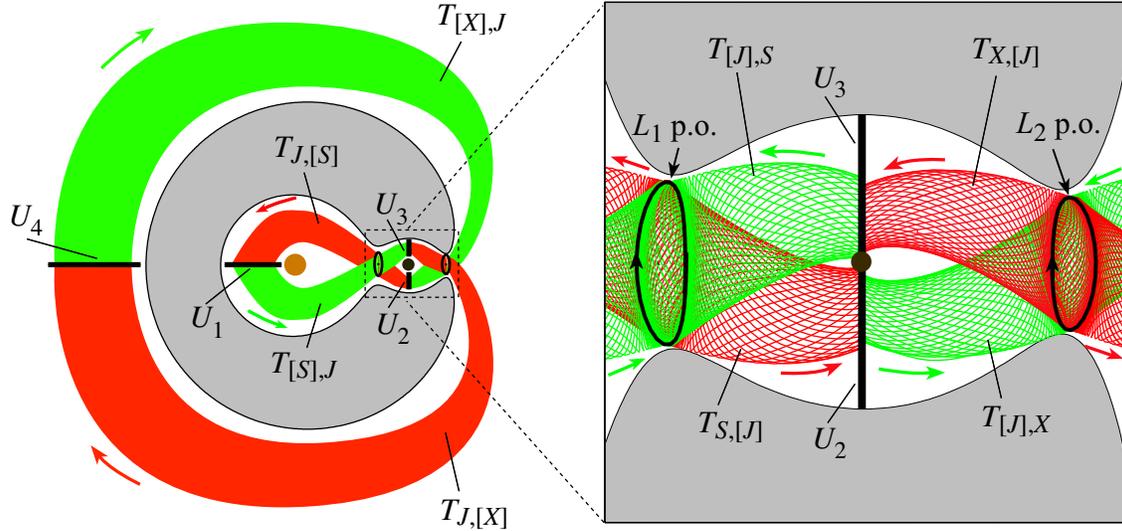
Figure 4.5: Position space projection of the $L_1$ and $L_2$ periodic orbit stable and unstable manifold tubes (schematic). The tubes are labeled according to the behavior of trajectories inside the boundaries defined by the stable and unstable manifolds. For example, $T_{[X],J}$ contains trajectories which are currently in the $X$ realm and heading toward the $J$ realm. Note the symmetry about the sun-Jupiter line. The location of the Poincaré surfaces of sections $U_i$ are also shown. Magnification of the $J$ realm is shown at right.

and are now in $J$ realm. Let us denote it by $I_{X,[J]}$, or simply $(X, [J])$, as in Figure 4.6. In Figure 4.6, we also plot $([J], S) = \left(T_{[J],S} \bigcap U_3\right)(1)$. We denote the intersection $(X, [J]) \bigcap ([J], S)$ by $(X, [J], S)$. This itinerarea contains initial conditions for orbits with itinerary $(X, J, S)$, like the one shown in Figure 4.2.

6. **Numerically integrate an initial condition in the appropriate itinerarea.** Once we have the itinerarea, the last step is forward and backward numerical integration of any initial condition within the itinerarea. Continuing the example, suppose we have obtained the set $(X, [J], S)$, a subset of the $y$-$\dot{y}$ plane in $U_3$. We desire an initial condition $s_0 = (x_0 \ y_0 \ \dot{x}_0 \ \dot{y}_0)^T$.

(a). First, we know from our choice of Poincaré section $(U_3)$ that $x_0 = 1 - \mu$.

(b). We then pick values $(y_0, \dot{y}_0) \in (X, [J], S) \subset U_3$

(c). Finally, $\dot{x}_0$ comes from the energy equation (2.11).

$$\dot{x}_0 = -\sqrt{-\dot{y}_0^2 - 2\bar{U}(x_0, y_0)} \tag{4.3}$$
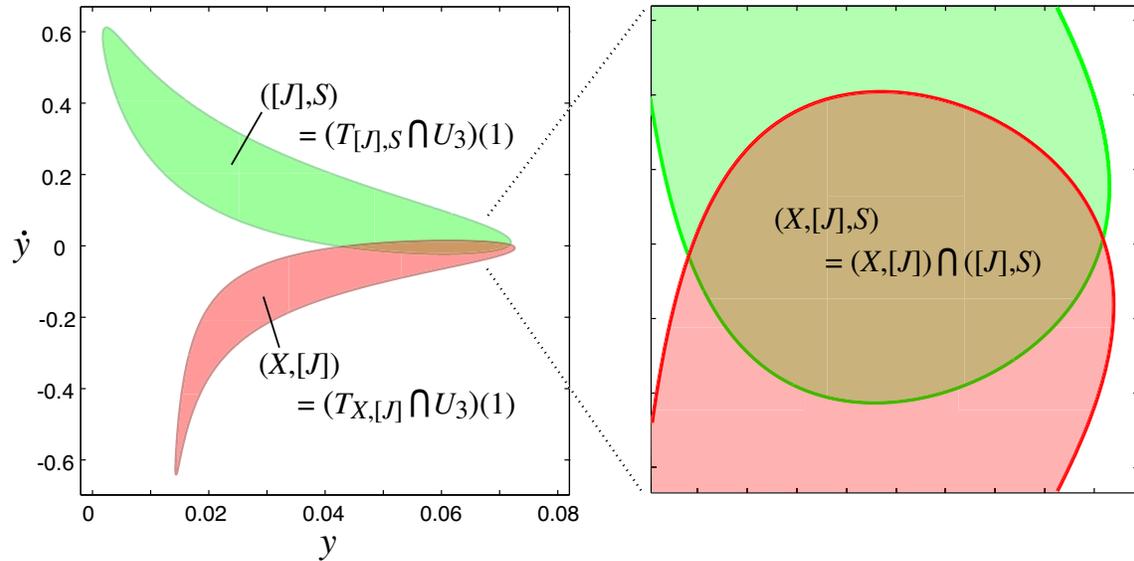
Figure 4.6: **An itinerarea with label (X,[J],S).** The itinerareas $(X, [J])$ and $([J], S)$ on $U_3$ are shown to intersect in the left panel. The right panel shows a close-up of the intersection region, the itinerarea $(X, [J], S)$, which contains initial conditons for orbits with itinerary $(X, J, S)$. See the text for details.

where the effective potential is given in Eq. (2.7). We take the negative sign of the square root by the definition of the $U_3$ Poincaré section.

We want the solution $s(t)$ which passes through $s_0$ at time $t = 0$, i.e., $s(0) = s_0$. Evolving our initial point $s_0$ forward and backward under the equations of motion (2.15) within some time interval $[-\tau, \tau]$ for $\tau > 0$ yields the desired solution. We are guided in our choice of $\tau$ by the integration times of the trajectories along the tube boundary, which yields an initial guess for $\tau$. Through simple trial and error starting from a reasonable guess, we find the $\tau$ which produces the appropriate trajectory, whose projection onto position space looks like that shown in Figure 4.2.

## 4.2   Example Itinerary: $(X, J, S, J, X)$

In what follows, we apply the numerical construction techniques discussed above to an example with a longer itinerary ($k = 5$). As our example, we construct a trajectory with itinerary $(X, J, S, J, X)$. This example is chosen because it roughly corresponds to the behavior of comet P/Oterma with respect to the sun-Jupiter system during the years 1910 to 1980 (see Koon, Lo, Marsden and Ross [2001b]).

We seek itinerareas with label $(X, J, S, J, X)$ on one of the Poincaré sections $\{U_i\}$. We use the energy, $e = -1.519$, which is in the range $[E_2, E_3]$ for the sun-Jupiter system ($\mu = 9.537 \times 10^{-4}$).

In Figure 4.7, we show the first few intersections of the $L_1$ and $L_2$ p.o. tubes with $U_3$. We need to introduce a change in notation. What we called $(X, [J])$ in the previous
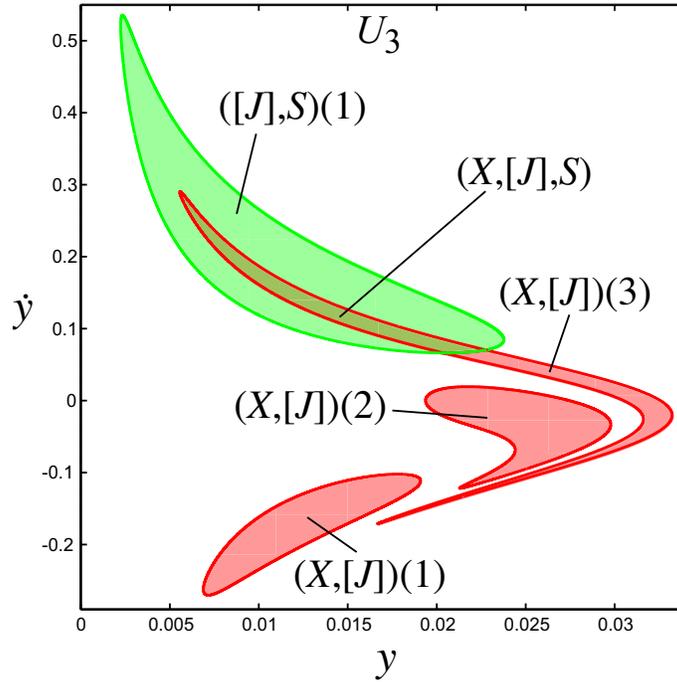


Figure 4.7: The first intersection of $T_{[J],S}$ and the first three intersections of $T_{X,[J]}$ with $U_3$ are shown. We use the notation $(X, [J])(n) = (T_{X,[J]} \bigcap U_3)(n)$ for the $n$th intersection of $T_{X,[J]}$ with $U_3$. The intersection $(X, [J], S) = (X, [J])(3) \bigcap ([J], S)(1)$ contains all solutions $s(t)$ which come from the $X$ realm, perform two full revolutions around Jupiter in the $J$ realm, and then exit to the $S$ realm.

section is now simply the first intersection of the tube, $T_{X,[J]}$, with the Poincaré section $U_3$, denoted $(X, [J])(1)$. Similarly, we use the notation $(X, [J])(n) = (T_{X,[J]} \bigcap U_3)(n)$ for the $n$th intersection of $T_{X,[J]}$ with $U_3$.

Let

$$
\begin{aligned}
f_{33} : U_3 &\rightarrow U_3, \\
(y, \dot{y}) &\mapsto (y', \dot{y}'),
\end{aligned}
$$

denote the Poincaré map from $U_3$ to $U_3$ (or at least defined on an appropriate restriction

of $U_3$). The map $f_{33}$ is area preserving owing to the Hamiltonian nature of the flow and the choice of Poincaré section. In particular, we have $(X, [J])(n) = f_{33}((X, [J])(1))$ and $m((X, [J])(n)) = m((X, [J])(n-1))$ where $m(A) \geq 0$ denotes the usual two-dimensional area of a set $A \subset \mathbb{R}^2$.

There is an intersection, $(X, [J], S) = (X, [J])(3) \bigcap ([J], S)(1)$. All initial conditions $s_0 \in (X, [J], S)$ correspond to solutions $s(t)$ which come from the $X$ realm, perform two full revolutions around Jupiter in the $J$ realm, and then exit to the $S$ realm.

To find itinerareas with the additional symbols, we take the $(X, [J], S)$ itinerarea and evolve it forward under the equations of motion (2.15) until it intersects the $U_1$ section in the $S$ realm, shown in Figure 4.8. Following our notation, this itinerarea is labeled $(X, J, [S])$. Notice that it lies entirely within the $(J, [S])$ itinerarea, as we would expect.
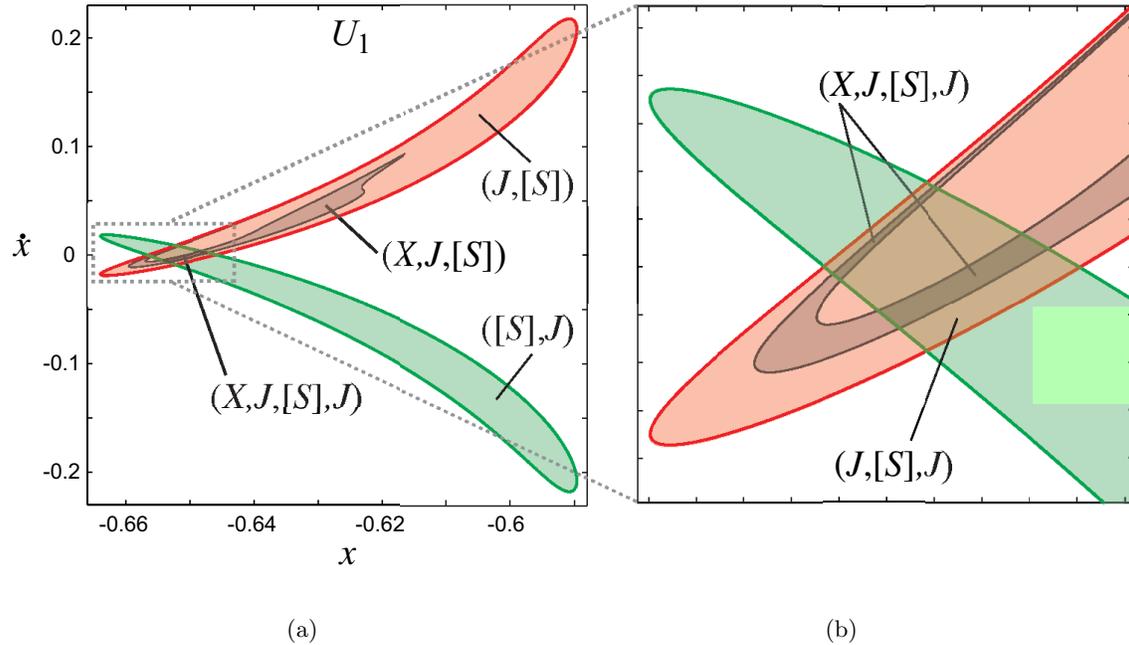


Figure 4.8: (a) The $U_1$ Poincaré section in the $S$ realm is shown. $(X, J, [S])$ is obtained by evolving the $(X, [J], S) \subset U_3$ itinerarea forward until it intersects $U_1$. (b) A close-up of the intersection of $(X, J, [S])$ with $([S], J)$, i.e., the $(X, J, [S], J)$ itinerarea, which consists of two disconnected large pieces. For this construction, we follow the evolution of the larger piece.

Furthermore, as seen in Figure 4.7, part of the boundary of $(X, [J], S)$ is on the boundary of $([J], S)(1)$, i.e., the boundary of the tubes connecting the $J$ and $S$ realms. Due to the infinite winding near the boundary of the tubes upon their approach to the $L_1$ p.o., this portion of the $(X, J, [S])$ set spirals around the boundary of $(J, [S])$ in $U_1$, as is suggested

in Figure 4.8.

The intersection $(X, J, [S]) \bigcap ([S], J)$ consists of a pair of large strips (and infinitely many smaller ones, due to the infinite winding described in Chapter 2), reminiscent of the strips around heteroclinic and homoclinic points which we encountered in Chapter 3. The pair of strips, shown close-up in Figure 4.8(b), carry the label $(X, J.[S], J)$ and bring us one symbol closer to our desired itinerarea. Taking the larger of the two strips, we evolve it forward in time until it re-enters the $J$ realm and intersects $U_2$, shown in Figure 4.9. Notice the symmetry between Figure 4.9 and Figure 4.7, i.e., $y \mapsto -y, t \mapsto -t$.



(a)                  (b)
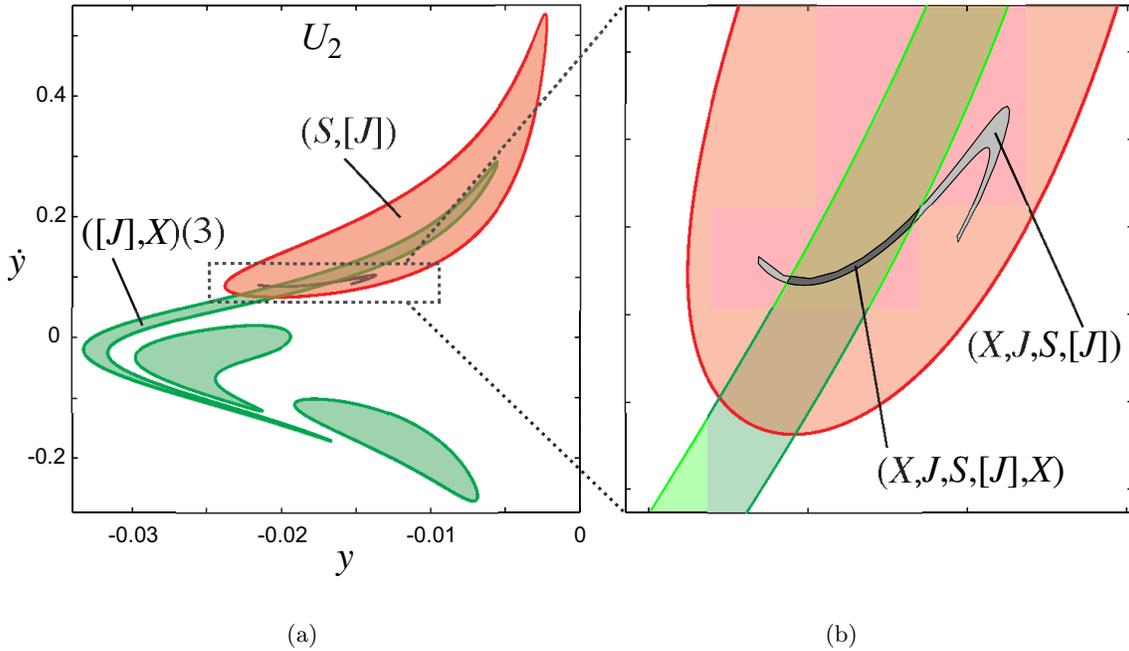
Figure 4.9: (a) We evolve the larger of the $(X, J, [S], J)$ pieces on $U_1$ until it intersects $U_2$. (b) A close-up of the intersection of the $(X, J, S, [J])$ and $([J], X)$ itinerareas is shown, yeilding the desired $(X, J, S, [J], X)$ itinerarea.

The $(X, J, S, [J])$ and $(S, [J])$ itinerareas intersect in a thin strip, the desired $(X, J, S, [J], X)$ itinerarea. Any trajectory passing through this strip will escape from the $J$ to the $X$ realm in forward time, and will perform a $S \rightarrow J \rightarrow X$ journey in backward time.

Taking any initial condition in this itinerarea strip and numerically integrating it forward and backward in time yields a trajectory with the desired itinerary. We give an example in Figure 4.10. Orbits in the region are considered *robust* because nearby orbits have the same finite itinerary. Regions corresponding to other allowable itineraries of any length can also theoretically be generated with this same systematic procedure. Not

only do we know such orbits exist from Theorem 3.9.1, but we have a relatively simple method for producing them.



(a)

(b)
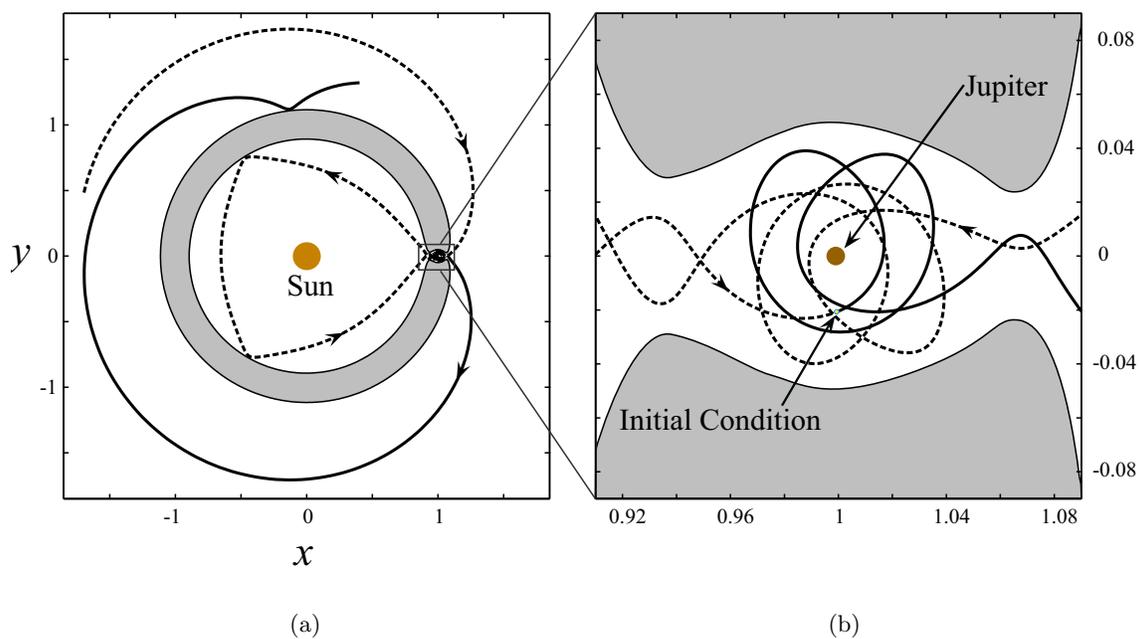
Figure 4.10: (a) **A trajectory with the itinerary** $(X, J, S, J, X)$ computed using an initial condition inside the $(X, J, S, [J], X)$ itinerarea of the $U_2$ Poincaré section, shown in Figure 4.9. (b) A close-up of this trajectory in the $J$ realm is shown and the initial condition is labeled. The backward (respectively, forward) integrated portion of the trajectory is a dashed (respectively, solid) curve.

# Chapter 5

# Trajectories in the Four-Body Problem

In this chapter, we describe a procedure to construct trajectories for a spacecraft in the four-body problem using solutions from the three-body problem covered in chapter 4. We illustrate the procedure in the construction of an important example mission: a "low-energy transfer" to the moon which uses ballistic capture.[1] The term *low-energy* is used to refer to the low fuel and therefore energy required to control the trajectory from a given starting condition to a targeted final condition.

**The Patched Conic Approximation.** For many years, trajectory designers for spacecraft on interplanetary missions have obtained good initial trajectory solutions in the $N$-body problem by dividing the spacecraft's motion into pieces in which the influence of *only one body at a time is considered*. This patched two-body, or "patched conic," approximation has worked well for missions such as the Voyager probes which have high relative velocity encounters with the bodies they visit. The criterion for switching from the influence of one body to another involves a dividing surface in the configuration space known as the sphere of influence, related to the Hill radius first encountered in chapter 2.

The patched conic approximation breaks down when we consider low relative velocity encounters, which are critical for low energy trajectories. In this regime, two bodies (e.g., the Earth and Moon) both influence the motion of the spacecraft with the same order of magnitude, and the restricted three-body problem must be used to model the motion of the spacecraft. Furthermore, the criterion for switching between the influence of one pair of bodies to another pair involves a dividing surface in the full phase space.

---

[1] *Ballistic capture* means that no propulsion is necessary (i.e., no $\Delta V$) to achieve a capture orbit at the destination body. In general, this "capture" is temporary.

**Spacecraft Trajectory Design Using the Patched Three-Body Approximation.**
In this chapter, we describe how to construct trajectories in the four-body problem using
invariant manifold tubes from multiple three-body systems using the **patched three-body approximation**. This is important for taking full advantage of $N$-body dynamics to
reduce fuel consumption, and is especially useful in the design of interplanetary trajectories
which *visit multiple bodies.* These may include mission trajectories such as a low energy
mission to orbit multiple moons of Jupiter or a low energy transfer from the Earth to
the Moon. For instance, using the phase space tubes in each three-body system, we are
able to construct a transfer trajectory from the Earth which executes an unpropelled (i.e.,
ballistic) capture at the Moon. An Earth-to-Moon trajectory of this type, which utilizes
the perturbation by the Sun, requires less fuel than the usual Hohmann transfer, such as
those used by the Apollo missions of the 1960s.

To design, for instance, a spacecraft trajectory from the Earth to the moon which
also takes advantage of the sun's gravity, we would want to model the trajectory as two
pieces: the first piece being a solution of the sun-Earth-spacecraft system (where the
moon's gravitational influence is unimportant) and the second piece being a solution of
the Earth-moon-spacecraft system (where the sun's influence is important). The two
pieces are connected by two initial conditions, $s_{pp}^-$ and $s_{pp}^+$, which together form the *patch
point* between two three-body solution arcs. Both $s_{pp}^-$ and $s_{pp}^+$ are at the same location in
position space, but we permit them to have differing velocities. The velocity discontinuity,
of norm $\Delta V$, corresponds to the impulsive rocket maneuver which will be necessary to
effect the transition between the two three-body solutions. Evolving $s_{pp}^-$ backward in
time gives the first piece; a solution in one three-body system, e.g., sun-Earth-spacecraft.
Evolving $s_{pp}^+$ forward in time gives the second piece; a solution in the other three-body
system, e.g., Earth-moon-spacecraft. We will discuss how to find appropriate patch points
pairs, $s_{pp}^\pm$.

## 5.1    Modeling the Four-Body Problem

Consider a particle $P$ in field of three massive bodies, $M_0$, $M_1$, and $M_2$. We suppose that
the massive bodies are in one of the two prescribed motions about one another:

(a). **Concentric Circular Model (CCM).** $M_0$ is a central body about which $M_1$ and
$M_2$ move in circular orbits of radii $d_1$ and $d_2$, respectively, where $d_2 > d_1$. In general,

we suppose $M_1, M_2 \ll M_0$. This is a model of, e.g., the Jupiter-Ganymede-Callisto system (as $M_0$, $M_1$, and $M_2$, respectively).

(b). **Bicircular Model (BCM).** $M_1$ and $M_2$ are in circular motion about their barycenter, with mutual separation $d_1$. Considering all the mass in the $M_1$-$M_2$ system to be concentrated at its barycenter, we suppose $M_0$ and the $M_1$-$M_2$ barycenter are in a circular orbit of radius $d_2 > d_1$ about their common center of mass. In general, we suppose $M_2 \ll M_1 \ll M_0$. This is a model of, e.g., the sun-Earth-Moon system (as $M_0$, $M_1$, and $M_2$, respectively).
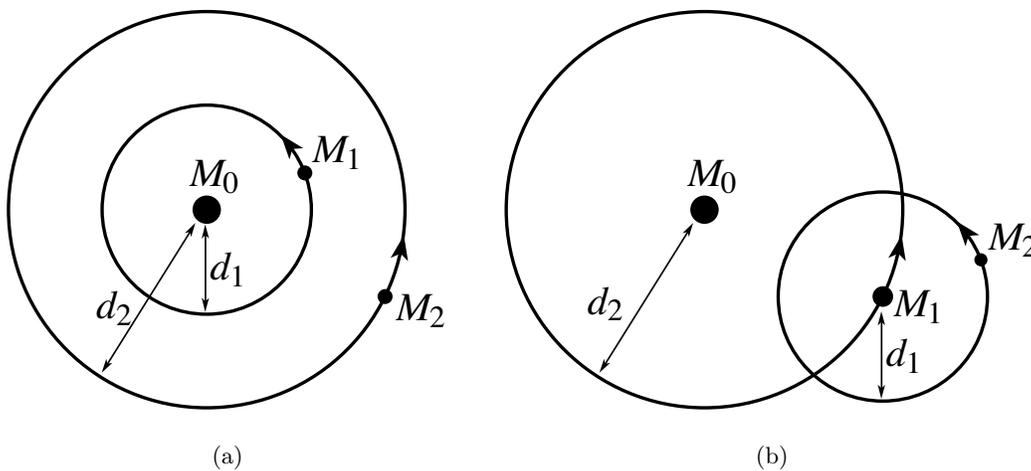


Figure 5.1: **Models of motion for the four-body problem.** (a) *Concentric Circular Model.* $M_0$ is a central body about which $M_1$ and $M_2$ move in circular orbits of radii $d_1$ and $d_2$, respectively, where $d_2 > d_1$. (b) *Bicircular Model.* $M_1$ and $M_2$ are in circular motion about their barycenter, with mutual separation $d_1$. Considering all the mass in the $M_1$-$M_2$ system to be concentrated at its barycenter, we suppose $M_0$ and the $M_1$-$M_2$ barycenter are in a circular orbit of radius $d_2 > d_1$ about their common center of mass.

## 5.2   Bicircular Model

In this thesis, we will only consider the BCM, which can be used to compute a low-energy Earth-to-Moon trajectory. The CCM has been used to produce a "Petit Grand Tour" of two of Jupiter's moons, as reported in Koon, Lo, Marsden and Ross [1999] and Gómez, Koon, Lo, Marsden, Masdemont and Ross [2001]. More recently, an extension of the CCM

has been used to design a "Multi-Moon Orbiter" of three of Jupiter's moons (see Ross, Koon, Lo and Marsden [2003] and Ross, Koon, Lo and Marsden [2004]).

**Equations of Motion in Earth-Moon Rotating Frame**   As mentioned earlier, we use the equations of motion derived under the BCM assumptions as the underlying dynamical model. The bicircular problem is a simplified version of the restricted four-body problem. The objective is to describe the motion of a spacecraft of negligible mass under the gravitational attraction of the Earth, Moon, and Sun. "Negligible mass" means that the spacecraft does not influence the motion of the Earth, Moon, and Sun. This description follows that of Simó, Gómez, Jorba and Masdemont [1995].

In this model we suppose that the Earth and Moon are revolving in circular orbits around their center of mass (barycenter) and the Earth-Moon barycenter is moving in a circular orbit around the center of mass of the Sun-Earth-Moon system. The orbits of all four bodies are in the same plane. We remark that, with these assumptions, the motion of these three bodies is not coherent. That is, the assumed motions do not satisfy Newton's equations. However, numerical simulation shows that, in some regions of phase space, this model gives the same qualitative behavior as the real system. Thus, the model is extremely useful for the study of some kinds of orbits, in particular the "Shoot the Moon" trajectory of Koon, Lo, Marsden and Ross [2001a].

To simplify the equations, the units of length, time, and mass are chosen such that the angular velocity of rotation of the Earth and Moon (around their barycenter), the sum of the masses of the Earth and Moon, and the gravitational constant are all equal to one. With these normalized units, the Earth-Moon distance is also one. Let $\mu$ be the mass of the Moon in these units. Then $1 - \mu$ the mass of the Earth. Let $m_S$ the mass of the Sun. Let the semimajor axis of the Sun be $a_S$.

We use a synodic (rotating) coordinates with respect to the Earth-Moon system. The origin is taken at the center of mass of the Earth-Moon system. The $x$-axis is given by the line that goes from the Earth to the Moon, and the $y$-axis is taken such that the system is orthogonal and positive oriented. Note that, in this synodic (non-inertial) frame, the Earth and Moon have fixed positions and the Sun is rotating clockwise around the barycenter of the Earth-Moon system. The positions of the Earth and Moon are fixed at $(-\mu, 0)$ and $(1 - \mu, 0)$, respectively. The angular velocity of the Sun in these synodic coordinates is denoted by $\omega_S$ and the phase of the Sun at $t = 0$ is $\theta_{S0}$. See Figure 5.2.
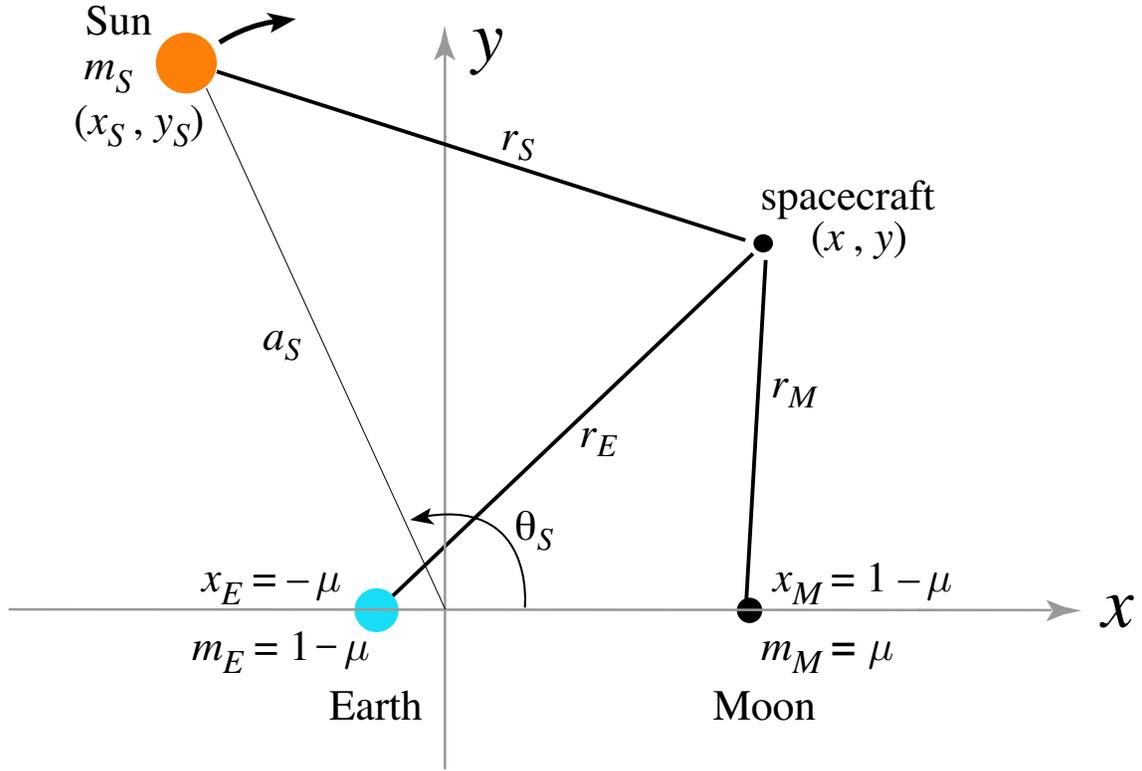
Figure 5.2:  Rotating coordinate frame in the BCM approximation with Earth and Moon fixed on the $x$-axis. As seen in this frame, the Sun rotates clockwise around the Earth-Moon barycenter (the origin) with angular frequency $\omega_S$.

Using nondimensional units, the equations of motion in the BCM are

$$
\begin{aligned}
\dot{x} &= u, \\
\dot{y} &= v, \\
\dot{u} &= x + 2v - c_E(x + \mu_M) - c_M(x - \mu_E) - c_S(x - x_S) - \alpha_S x_S, \\
\dot{v} &= y - 2u - c_E y - c_M y - c_S(y - y_S) - \alpha_S y_S,
\end{aligned}
\tag{5.1}
$$

where

$$
\begin{aligned}
c_i &= \frac{\mu_i}{r_i^3}, \quad \text{for} \quad i = E, M, S, \tag{5.2} \\
\alpha_S &= \frac{m_S}{a_S^3}, \tag{5.3}
\end{aligned}
$$

and

$$r_E = \sqrt{(x + \mu_M)^2 + y^2},$$

$$r_M = \sqrt{(x - \mu_E)^2 + y^2},$$

$$r_S = \sqrt{(x - x_S)^2 + (y - y_S)^2},$$

with

$$\mu_E = 1 - \mu,$$

$$\mu_M = \mu,$$

$$x_S = a_S \cos(\theta_S),$$

$$y_S = a_S \sin(\theta_S),$$

$$\theta_S = -\omega_S t + \theta_{S0}.$$

The values of the parameters are as follows:

$$\mu = \frac{m_M}{m_M + m_E} = 0.01215, \tag{5.4}$$

from Appendix A where $m_M$ and $m_E$ denote the mass of the Moon and Earth, respectively. The Sun's mass (1 unit = Earth + Moon mass) is

$$m_S = 328900.54, \tag{5.5}$$

the Sun's distance (1 unit = Earth-Moon distance) is

$$a_S = 388.81114, \tag{5.6}$$

and the Sun's angular velocity in synodic coordinates is

$$\omega_S = 0.925195985520347. \tag{5.7}$$

In the above equations, time is scaled by the period of the Earth and Moon around their center of mass ($T/2\pi$, where $T = 2.361 \times 10^6$ s), positions are scaled by the average Earth-Moon distance ($L = 3.850 \times 10^5$ km), and velocities are scaled by the Moon's

average orbital speed around the Earth ($2\pi L/T = 1.025$ km/s). One can find these values in Appendix A.

**Equations of Motion in Sun-Earth Rotating Frame**  In this model we suppose that the Sun and Earth are revolving in circular orbits around their barycenter and the Moon is moving in a circular orbit around the center of the Earth. The orbits of all four bodies are in the same plane. We remark that, with these assumptions, the motion of these three bodies is not coherent. That is, the assumed motions do not satisfy Newton's equations. However, the model is extremely useful for the study of some kinds of orbits, in particular the "Shoot the Moon" trajectory.

Let $\mu$ be the mass of the Earth, $1 - \mu$ the mass of the Sun and $m_M$ the mass of the Moon. Let the distance between the Sun and the Earth be taken as unity. Let the orbit of the Sun and Earth around the Sun-Earth barycenter also be taken as unity. The distance from the Earth to the Moon is $a_M$. We use rotating coordinates with respect to the Sun-Earth system, so that the positions of the Sun and Earth are fixed at $(-\mu, 0)$ and $(1 - \mu, 0)$, respectively. The angular velocity of the Moon in these synodic coordinates is denoted by $\omega_M$ and the phase of the Moon at $t = 0$ is $\theta_{M0}$.

In the rotating frame just defined and using nondimensional units, the equations of motion in the Sun-Earth rotating frame are

$$
\begin{aligned}
\dot{x} &= u, \\
\dot{y} &= v, \\
\dot{u} &= x + 2v - c_S(x + \mu_E) - c_E(x - \mu_S) - c_M(x - x_M), \\
\dot{v} &= y - 2u - c_S y - c_E y - c_M(y - y_M),
\end{aligned} \tag{5.8}
$$

where

$$
c_i = \frac{\mu_i}{r_i^3}, \quad \text{for} \quad i = S, E, M \tag{5.9}
$$

$$
\alpha_S = \frac{m_S}{a_S^3}, \tag{5.10}
$$

and

$$r_S = \sqrt{(x + \mu_E)^2 + y^2},$$
$$r_E = \sqrt{(x - \mu_S)^2 + y^2},$$
$$r_M = \sqrt{(x - x_M)^2 + (y - y_M)^2},$$

with

$$\mu_S = 1 - \mu,$$
$$\mu_E = \mu,$$
$$x_M = a_M \cos(\theta_M),$$
$$y_M = a_M \sin(\theta_M),$$
$$\theta_M = \omega_M t + \theta_{M0}.$$

The values of the parameters are as follows:

$$\mu = \frac{m_E}{m_E + m_S} = 3.036 \times 10^{-6}, \tag{5.11}$$

where $m_E$ and $m_S$ denote the mass of the Earth and Sun, respectively. The Moon's mass (1 unit = Sun + Earth mass) is

$$m_M = 3.733998734625702 \times 10^{-8}. \tag{5.12}$$

the Earth-Moon distance (1 unit = Sun-Earth) is

$$a_M = 2.573565073532068 \times 10^{-3}, \tag{5.13}$$

and the Moon's angular velocity in synodic coordinates is

$$\omega_M = 12.36886949284508. \tag{5.14}$$

In the above equations, time is scaled by the period of the Sun and Earth around their center of mass ($T/2\pi$, where $T = 3.156 \times 10^7$ s), positions are scaled by the average Sun-Earth distance ($L = 1.496 \times 10^8$ km), and velocities are scaled by the Earth's average

orbital speed around the Sun ($2\pi L/T = 29.7840$ km/s), according to Appendix A.

**Transforming Coordinates between Rotating Frames.** A necessary algorithm is the transformation between the two rotating coordinate frames: the Earth-Moon rotating frame and the Sun-Earth rotating frame.

Let the phase space trajectory in rotating coordinate system $A$ be denoted by $\mathbf{x}_A^{\mathrm{ro}}(t_A)$ where $\mathbf{x}_A^{\mathrm{ro}} = [x, y, u, v]^T$ is in the nondimensional position and velocity units associated with system $A$ and $t_A$ is in the corresponding time units of system $A$.

We first transform to inertial coordinates centered on the primary $m_i$, $i = 1$ or $2$, via

$$\mathbf{x}_A^{\mathrm{in}} = \mathbf{R}(\mathbf{x}_A^{\mathrm{ro}} - \mathbf{d}_A), \tag{5.15}$$

where

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}, \tag{5.16}$$

$$R_{11} = R_{22} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}, R_{21} = \begin{pmatrix} -s & -c \\ c & -s \end{pmatrix}, R_{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.17}$$

$$c = \cos(\theta(t_A)), s = \sin(\theta(t_A)),$$

$$\theta(t_A) = t_A + \theta_{A0},$$

and $\mathbf{d}_A = [x_A^0, 0, 0, 0]^T$ and $x_A^0$ is $-\mu_A$ or $1 - \mu_A$ depending on whether the $A$ system inertial frame is $m_1$- or $m_2$-centered, respectively.

We then change from the units of system $A$ to the units of another system, $B$. Let $L_{AB} = \frac{L_A}{L_B}$ be the ratio of the length scales and $T_{AB} = \frac{T_A}{T_B}$ be the ratio of the time scales. The inertial frame position, velocity, and time coordinates in the $B$ system are then

$$x_B^{\mathrm{in,pos}} = L_{AB} x_A^{\mathrm{in,pos}}, \tag{5.18}$$

$$x_B^{\mathrm{in,vel}} = \frac{L_{AB}}{T_{AB}} x_A^{\mathrm{in,vel}}, \tag{5.19}$$

$$t_B = T_{AB} t_A, \tag{5.20}$$

respectively.

The primary $m_i$ of system $A$ is the primary $m_j$ of system $B$. Thus, the trajectory

$\mathbf{x}_B^{\text{in}}(t_B)$ is in $m_j$-centered inertial coordinates in the units of system $B$. To transform back to rotating coordinates, we use

$$\mathbf{x}_B^{\text{ro}} = \mathbf{R}^{-1}\mathbf{x}_B^{\text{in}} + \mathbf{d}_B, \tag{5.21}$$

where $\mathbf{d}_B = [x_B^0, 0, 0, 0]^T$ and $x_B^0$ is $-\mu_B$ or $1 - \mu_B$ depending on whether the $B$ system inertial frame is $m_1$- or $m_2$-centered, respectively.

## 5.3  Example Mission: Low-Energy Transfer to the Moon

**Hiten Mission.**  The traditional approach to construct a spacecraft transfer trajectory to the moon from the Earth is by Hohmann transfer. This type of transfer uses only two-body dynamics. It is constructed by determining a two-body Keplerian ellipse from an Earth parking orbit to the orbit of the moon, illustrated schematically in Figure 5.3(a). The two bodies involved are the Earth and a spacecraft. Such a transfer requires a large $\Delta V$ for the spacecraft to get captured by the moon.

In 1991, the failed Japanese mission, Muses-A, whose propellant budget did not permit it to transfer to the moon via the usual method was given a new life with an innovative trajectory design, based on the work of Belbruno and Miller [1993]. Its re-incarnation, renamed Hiten, used a low-energy transfer with a ballistic capture at the moon. An Earth-to-Moon trajectory of this type, shown in Figures 5.3(b) and (c), which utilizes the perturbation by the Sun, requires less fuel than the usual Hohmann transfer.

**Using the Patched Three-Body Approximation to Systematically Design Earth-to-Moon Trajectories with Ballistic Capture.**  In this section, we present an approach to the problem of the orbital dynamics of this interesting trajectory by implementing *in a systematic way* the view that the Sun-Earth-Moon-spacecraft four-body system can be approximated as two three-body systems. Figure 5.4(a) shows a schematic of this trajectory in the Sun-Earth rotating frame, showing the two legs of the trajectory: (1) the Sun-Earth Lagrange point portion and (2) the lunar capture portion.

Within each three-body system, using our understanding of the invariant manifold structures associated with the Lagrange points $L_1$ and $L_2$, we transfer from a 200 km altitude Earth orbit into the region where the invariant manifold structure of the Sun-
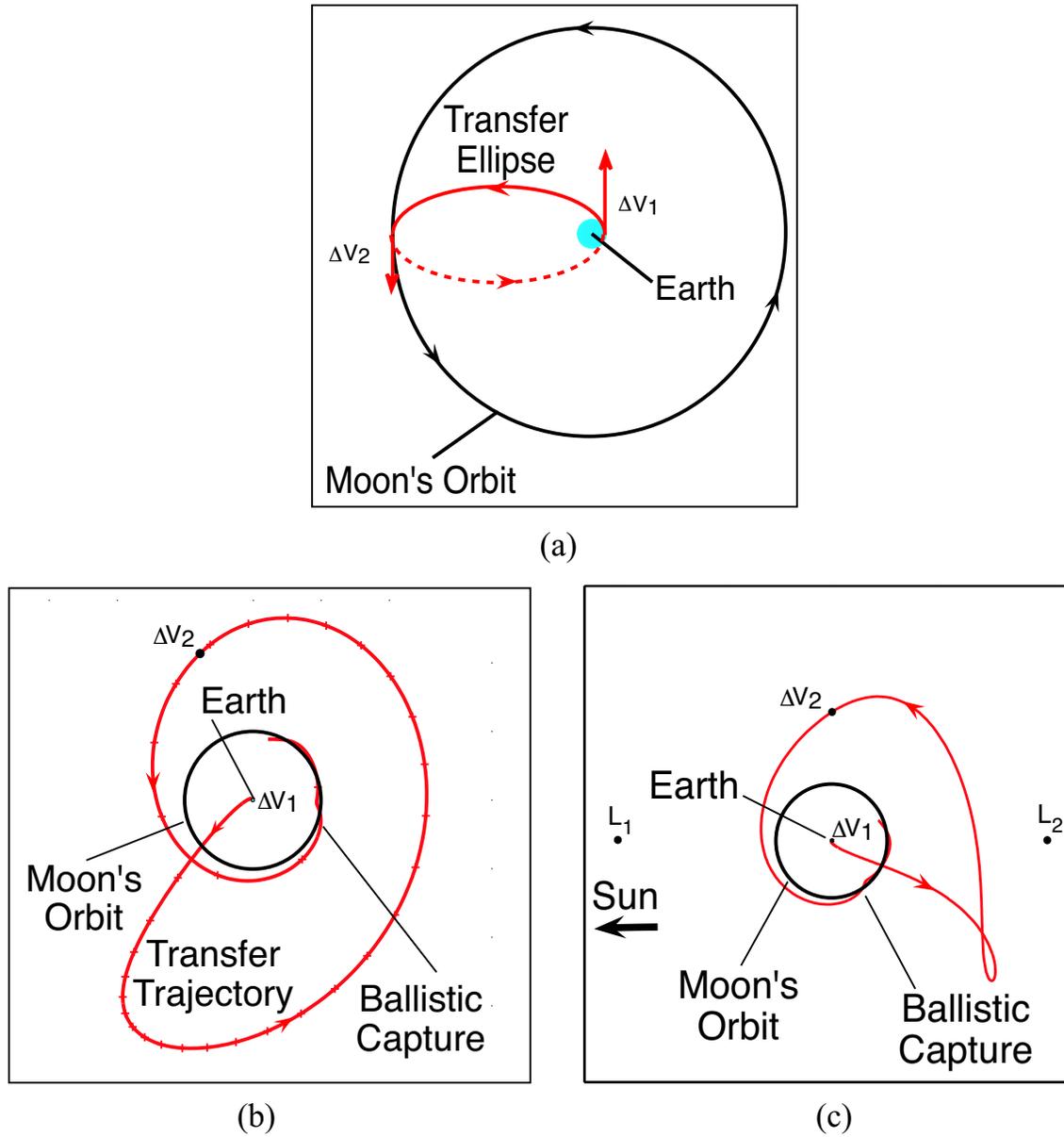
Figure 5.3: (a) Hohmann transfer. (b) Low-energy transfer trajectory in the geocentric inertial frame. (c) Same trajectory in the Sun-Earth rotating frame.

Earth Lagrange points interact with the invariant manifold structure of the Earth-Moon Lagrange points. See Figure 5.4(b). We utilize the sensitivity of the "twisting" of trajectories near the invariant manifold tubes in the Lagrange point region to find a fuel efficient transfer from the Sun-Earth system to the Earth-Moon system. The invariant manifold tubes of the Earth-Moon system provide the dynamical channels in phase space that enable ballistic captures of the spacecraft by the Moon.
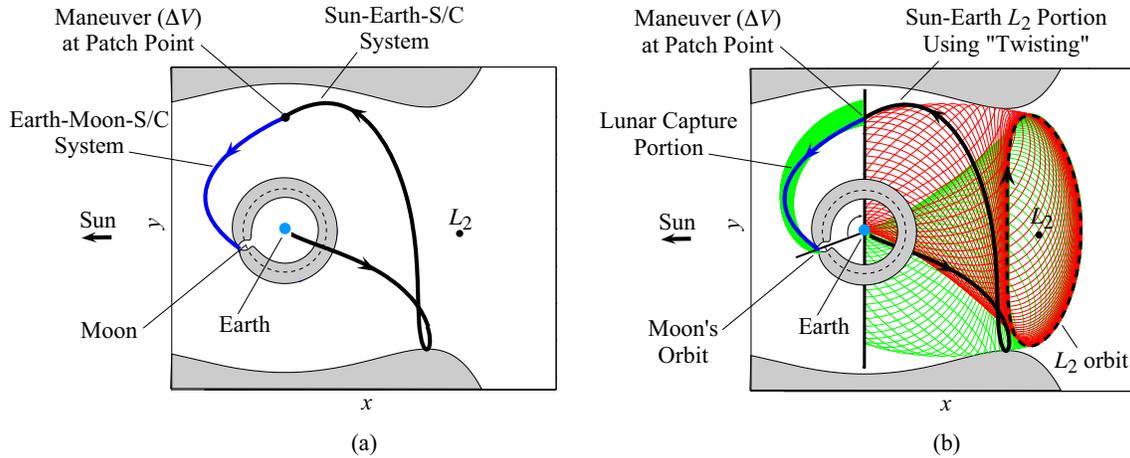
Figure 5.4: (a) Two legs of a Hiten-like trajectory in the Sun-Earth rotating frame. (b) The interaction of invariant manifold tubes of the Sun-Earth and the Earth-Moon systems permits a fuel efficient Earth-to-Moon transfer with the perturbation of the Sun.

The final Earth-to-Moon trajectory is integrated in the bicircular four-body model described in §5.2, where both the Moon and the Earth are assumed to move in circular orbits about the Earth and the Sun, respectively, in the ecliptic, and the spacecraft is an infinitesimal mass point.

The success of this approach depends greatly on the configuration of the specific four bodies of interest. In order for low-energy transfers to take place, the invariant manifold structures of the two three-body systems must intersect within a reasonable time. Otherwise, the transfer may require an impractically long time of flight. For the Sun-Earth-Moon-spacecraft case, this is not a problem. The overlap of these invariant manifold structures provide the low-energy transfers between the Earth and the Moon.

**Construction of Earth-to-Moon Transfer.** The construction is done mainly in the Sun-Earth rotating frame using the Poincaré section $\Gamma$ (along a line of constant $x$-position passing through the Earth). This Poincaré section helps to glue the Sun-Earth Lagrange point portion of the trajectory with the lunar ballistic capture portion.

The basic strategy is to find an initial condition (position and velocity) for a spacecraft on the Poincaré section such that when integrating forward, the spacecraft will be guided by the $L_2$ Earth-Moon manifold and get ballistically captured by the Moon; when integrating backward, the spacecraft will hug the Sun-Earth manifolds and return to Earth.

We utilize two important properties of the Lagrange point dynamics of the three-body problem. The stable manifold tube is key in targeting a capture orbit for the Earth-Moon portion of the design. The twisting of orbits in the equilibrium region is key in finding a fuel efficient transfer for the Sun-Earth Lagrange point portion of the trajectory.

**Lunar Ballistic Capture Portion.** Recall that by targeting the region enclosed by the stable manifold tube of the $L_2$ Lyapunov orbit in the Earth-Moon system, we can construct an orbit which will get ballistically captured by the Moon. When we transform this Poincaré cut of the stable manifold of an Earth-Moon $L_2$ Lyapunov orbit into the Poincaré section of the Sun-Earth system, we obtain a closed curve. A point interior to this curve will approach the Moon when integrated forward. See Figure 5.5. Assuming the Sun is a negligible perturbation to the Earth-Moon-spacecraft three-body dynamics during this leg of the trajectory, any spacecraft with initial conditions within this closed curve will be ballistically captured by the Moon. "Ballistic capture by the Moon" means an orbit which under natural dynamics gets within the sphere of influence of the Moon (approx. 60,000 km) and performs at least one revolution around the Moon. In such a state, a slight $\Delta V$ will result in a stable capture (closing off the necks at $L_1$ and $L_2$).



(a)                                      (b)
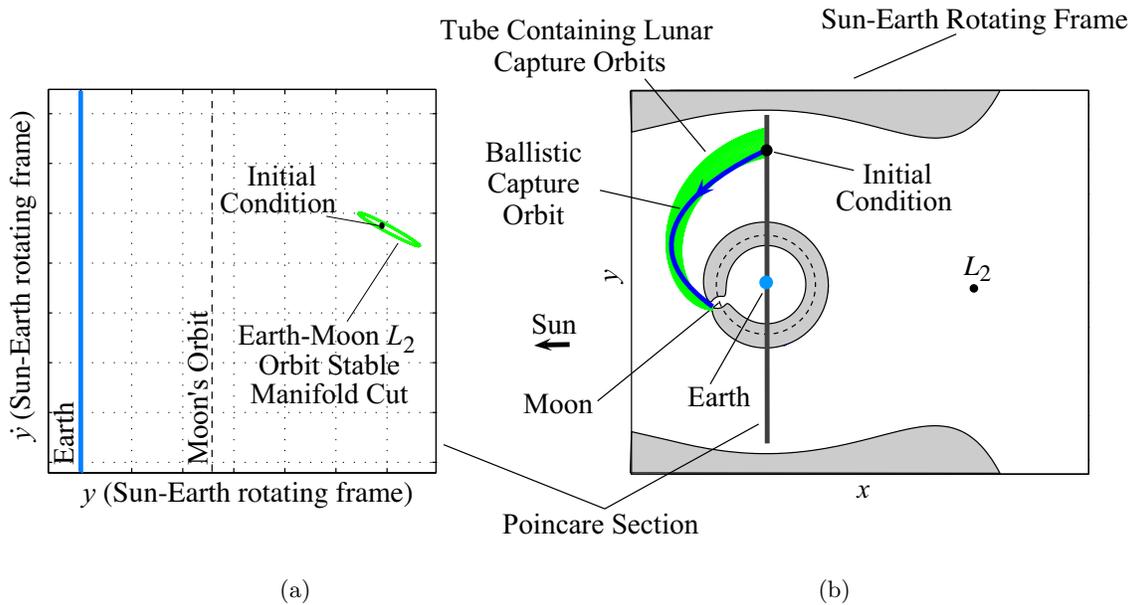
Figure 5.5: (a) The stable manifold cut of an Earth-Moon $L_2$ orbit in the Poincaré section of the Sun-Earth system. (b) A point interior to this cut, with the correct phasing of the Moon, will reach the Moon's ballistic capture region when integrated forward.

**Twisting of Orbits and Sun-Earth Lagrange Point Portion.** Since the twisting of orbits in the equilibrium region is key in finding the Sun-Earth Lagrange point portion of the design, we would like to review this property briefly. From chapter 2, we learn that orbits twist in the equilibrium region following roughly the Lyapunov orbit. The amount of twist of an orbit depends sensitively on its distance from the manifold tube. The closer to the manifold tube an orbit begins on its approach to the equilibrium region, the more it will be twisted when it exits the equilibrium region. Hence, with small change in the initial condition (such as a small change in velocity at a fixed point), we can change the destination of an orbit dramatically. In fact, we can use this sensitivity to target the spacecraft back to a 200 km Earth parking orbit.

Look at the Poincaré section $\Gamma$ in Figure 5.6(a). Notice that how a minute line strip
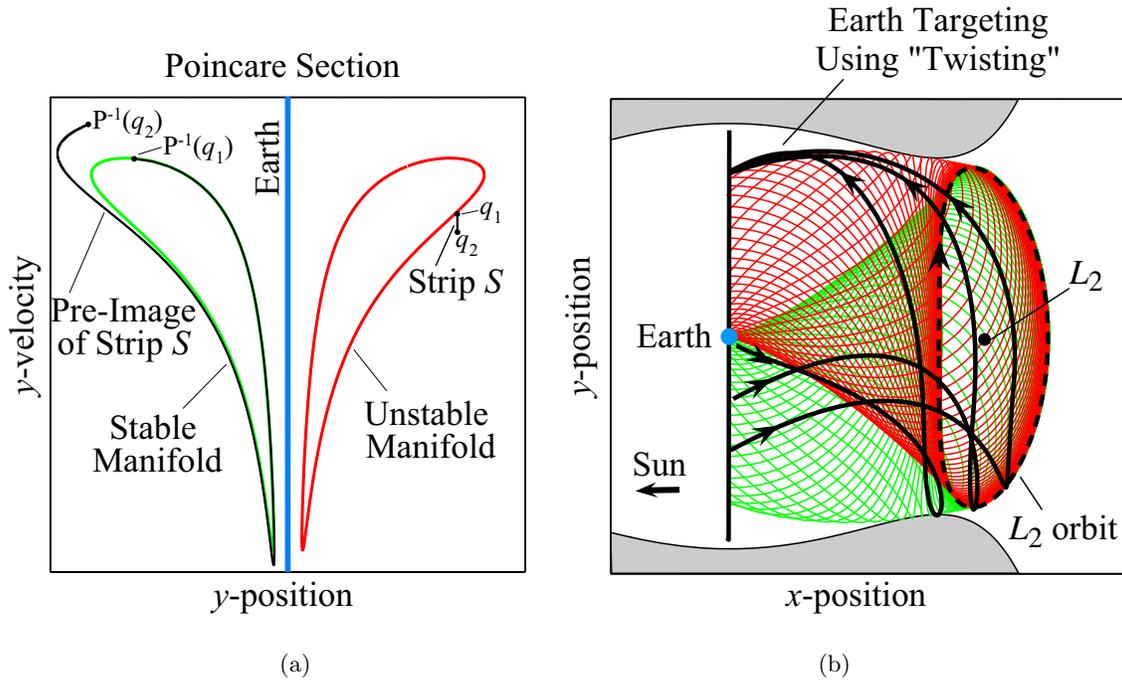


Figure 5.6: (a) Line strip $q_2q_1$ outside of unstable manifold cut gets stretched into a long strip $P^{-1}(q_2)P^{-1}(q_1)$ that wraps around stable manifold cut. (b) With infinitesimal changes in velocity, any point near lower tube cross section can be targeted (integrating backward).

$q_2q_1$ of orbits just outside of the unstable manifold cut, when integrated backward, gets stretched into a long strip $P^{-1}(q_2)P^{-1}(q_1)$ of orbits that wraps around the whole stable manifold cut. Recall that points on $q_2q_1$ represent orbits which have the same position but slightly different velocity. But their pre-image $P^{-1}(q_2)P^{-1}(q_1)$ can reach any position

on the lower line where the stable manifold tube intersects (see Figure 5.6(b)).

Pick an energy in the temporary capture range of the Sun-Earth system which has $L_2$ orbit manifolds that come near a 200 km altitude Earth parking orbit. Compute the Poincaré section $\Gamma$ (see Figure 5.6(a)). The curve on the right is the Poincaré cut of the unstable manifold of the Lyapunov orbit around the Sun-Earth $L_2$. Picking an appropriate initial condition just outside this curve, we can backward integrate to produce a trajectory coming back to the Earth parking orbit.

**Connecting the Two Portions.**  We can vary the phase of the Moon until the Earth-Moon $L_2$ manifold cut intersects the Sun-Earth $L_2$ manifold cut, as illustrated in Figures 5.7(a) and (b). In the region which is in the interior of the Earth-Moon $L_2$ manifold curve but in the exterior of the Sun-Earth $L_2$ manifold curve, an orbit will get ballistically captured by the Moon when integrated forward; when integrated backward, the orbit will hug the unstable manifold back to the Sun-Earth $L_2$ equilibrium region with a twist, and then hug the stable manifold back towards the position of the Earth parking orbit. See Figures 5.7(c) and (d).

With only a slight modification (a small mid-course $\Delta V$ of 34 m/s at the patch point), this procedure produces a genuine solution integrated in the bicircular four-body problem. Since the capture at the Moon is natural (zero $\Delta V$), the amount of on-board fuel necessary is lowered by about 20% compared to a traditional Hohmann transfer (the Hohmann transfer value is taken from Belbruno and Miller [1993]).
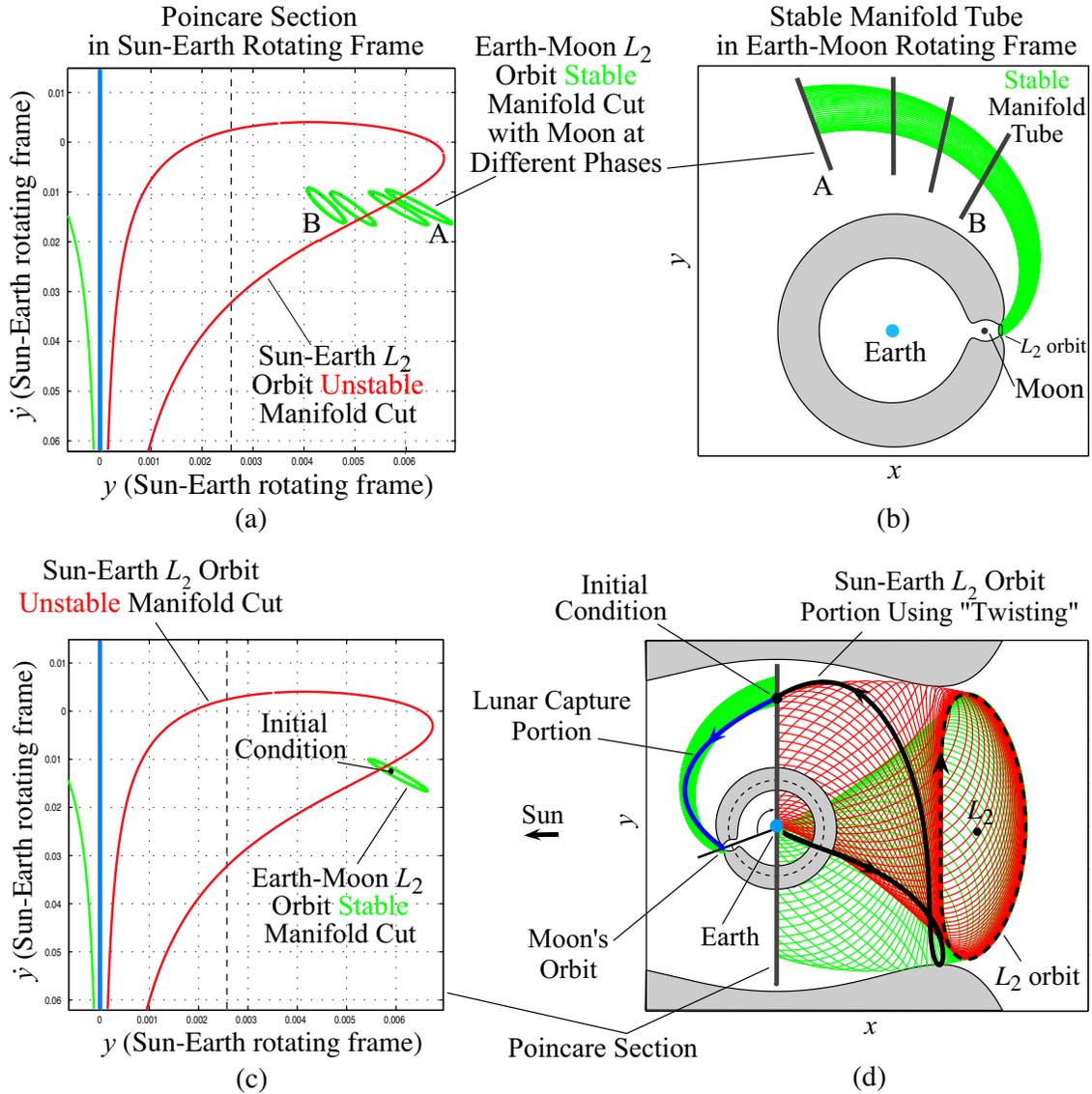
**Figure 5.7:** (a) and (b) Vary the phase of the Moon until Earth-Moon $L_2$ manifold cut intersects Sun-Earth $L_2$ manifold cut. (c) Pick a point in the interior of the Earth-Moon $L_2$ manifold curve but in the exterior of the Sun-Earth $L_2$ manifold curve. (d) An orbit will get ballistically captured by the Moon when integrated foreward; when integrated backward, orbit will hug the invariant manifolds back to the Earth.

# Chapter 6

# Statistical Theory of Interior-Exterior Transition and Collision Probabilities for Minor Bodies in the Solar System

## 6.1   Introduction

Several Jupiter-family comets, such as P/Oterma, P/Gehrels 3, and P/Helin-Roman-Crockett, make a transition from heliocentric orbits inside the orbit of Jupiter to heliocentric orbits outside the orbit of Jupiter and vice versa (Carusi, Kresák, Pozzi and Valsecchi [1985]; Koon, Lo, Marsden and Ross [2001b]). During this transition, the comet can be captured temporarily by Jupiter for one to several orbits around Jupiter (Tancredi, Lindgren and Rickman [1990] and Howell, Marchand and Lo [2000]). The Tisserand parameters of these objects, termed the *quasi-Hildas* (hereafter QHs) by Kresák [1979], are slightly in excess of 3. The possible pre-capture orbital history of D/Shoemaker-Levy 9 (henceforth referred to as SL9) also places it within this group (Benner and McKinnon [1995]).

An important feature of the motion of these comets is that during the phase right before and after their encounter with Jupiter, their orbits pass close to the libration points $L_1$ and $L_2$ of the sun-Jupiter system. This has been pointed out by many authors, including Tancredi, Lindgren and Rickman [1990] Valsecchi [1992], and Belbruno and Marsden [1997]. Hence objects with low velocity relative to these points (i.e., orbits with aphelion near $L_2$ or perihelion near $L_1$) are most likely to be captured (Kary and Dones [1996]).

During the short time just before an encounter with Jupiter, the most important orbital perurbations are due to Jupiter alone, as suggested by the passages of comets by $L_1$ and

$L_2$. $N$-body effects of Saturn and the other large planets surely play a significant role over significantly longer times, but we concentrate here on the time right before a comet's encounter with Jupiter. To simplify the analysis, we use the most rudimentary dynamical model, namely, the circular, planar restricted three-body model (PCR3BP), to determine the basic phase space structure which causes the dynamical behavior of the QH comets. Furthermore, since the PCR3BP is an adequate starting model for many other systems, results can be applied to other phenomena in the solar system, such as the near-Earth asteroid (NEA) problem, wherein one considers the motion of an asteroid on an energy surface in the sun-Earth system where libration point dynamics are important.

Lo and Ross [1997] suggested that studying the $L_1$ and $L_2$ invariant manifold structures would be a good starting point for understanding the capture and transition of these comets. Koon, Lo, Marsden and Ross [2000] studied the stable and unstable invariant manifolds associated to $L_1$ and $L_2$ periodic orbits. They took the view that these manifolds, which are topologically tubes within an energy surface, are phase space conduits transporting material to and from Jupiter and between the interior and exterior of Jupiter's orbit.

In the present paper, we wish to extend the results of Koon, Lo, Marsden and Ross [2000] to obtain *statistical* results. In particular, we wish to address two basic questions about QHs and NEAs: How likely is a QH collision with Jupiter or a NEA collison with Earth? How likely is a P/Oterma-like interior-exterior resonance transition? With this work, we put SL9, NEA impacts, and interior-exterior transitions into the broader context of generic motion in the restricted three-body problem.

The paper is broken up into two sections. In section 6.2, we discuss some phenomena of the QH comets, namely interior-exterior and collisions with Jupiter. In section 6.3, we frame the above questions as a transport problem, viewing the PCR3BP as the underlying dynamical system. We also summarize the results and suggest future directions.

## 6.2 The Quasi-Hilda Group of Comets

The QH group of comets is a small group of strongly Jupiter-interacting comets having a Tisserand parameter slightly above 3, characterized by repeated and long-lasting temporary captures (Benner and McKinnon [1995]). As authors have noted, the capture process frequently moves bodies from orbits outside Jupiter's orbit to inside Jupiter's orbit, pass-

ing by $L_1$ and $L_2$ in the process of approaching or departing from Jupiter's vicinity (e.g., Kary and Dones [1996]). We will refer to this type of transition as an *interior-exterior* transition.

**Interior-Exterior Transition.** In Figure 6.1(a), we show the interior-exterior transition of QH P/Oterma in a sun-centered inertial frame. The interior orbit is in an exact 3:2 mean motion resonance with Jupiter[1] while the exterior orbit is near the 2:3 resonance with Jupiter. In Figure 6.1(b), we show a homoclinic-heteroclinic chain of orbits in the PCR3BP as seen in the rotating frame. This is a set of orbits on the intersection of $L_1$ and $L_2$ stable and unstable manifolds with energies equal to that of P/Oterma. The homoclinic-heteroclinic chain is believed to form the backbone for temporary capture and interior-exterior transition of QHs, as can be seen when the orbit of P/Oterma in the rotating frame is overlayed as in Figure 6.1(c) (Koon, Lo, Marsden and Ross [2000]).
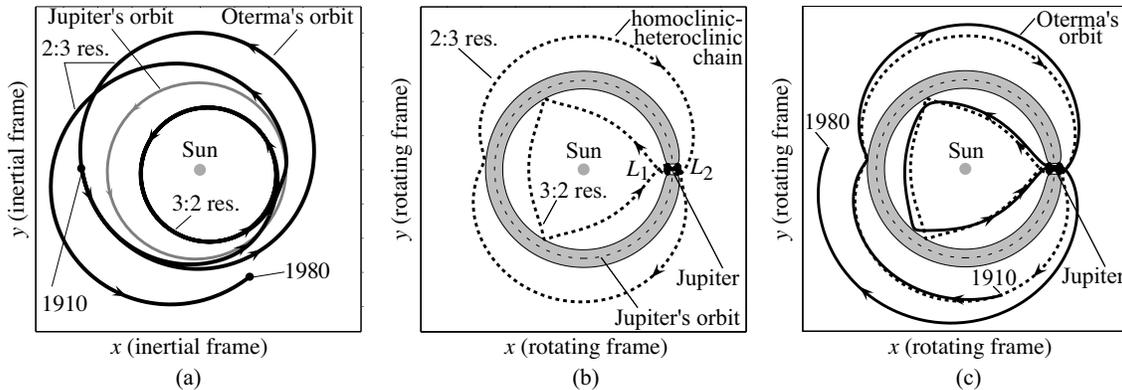


Figure 6.1: (a) Orbit of quasi-Hilda comet P/Oterma in sun-centered inertial frame during time interval AD 1910–1980 (ecliptic projection). (b) A homoclinic-heteroclinic chain for the energy of P/Oterma in the circular, planar restricted three-body problem, as seen in the rotating frame with the sun and Jupiter fixed. (c) The orbit of P/Oterma, transformed into the rotating frame, overlaying the chain.

**Collision with Jupiter.** At the time of its discovery, SL9 was only 0.3 AU from Jupiter and broken up into several fragments due to tidal disruption on an earlier approach within the planet's Roche limit (Marsden [1993]). Integrations indicated that it would collide with the planet (Chodas and Yeomans [1993]), which it subsequently did in July 1994.

---

[1]By *exact*, we mean that P/Oterma orbits the sun three times while Jupiter orbits the sun twice, as seen in an inertial frame.

**Likely Pre-Collision Heliocentric Orbit of SL9.** Pre-collision integrations of individual SL9 fragmemts (Benner and McKinnon [1995]) suggest that the SL9 progenitor approached Jupiter by passing by $L_1$ or $L_2$ from a short-period heliocentric orbit between either Jupiter and Mars or between Jupiter and Saturn (Figure 6.2(a)). The distribution of heliocentric $a$ and $e$ determined from these fragment integrations are shown in Figure 6.2(b). The pre-collision fragments have Tisserand parameters of about $T = 3.02 \pm 0.01$. From this value and the similarity of the pre-collision orbits to the known QHs, Benner and McKinnon [1995] suggest a QH origin for SL9.



(a)　　　　　　　　　　　　　(b)
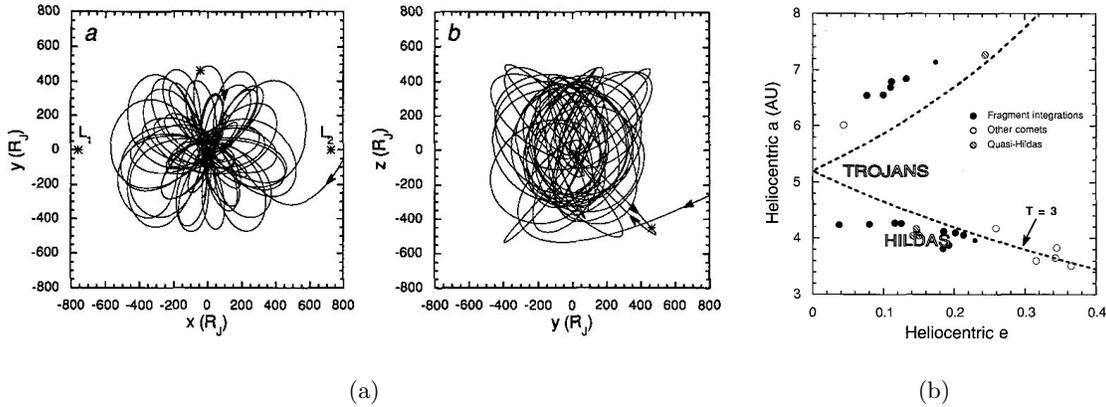
Figure 6.2: (a) A typical SL9 trajectory showing the passage past a libration point and subsequent capture. The sun is to the right. (Reproduced from Benner and McKinnon [1995]. According to their terminology, their $L_2$ is our $L_1$, and vice versa.) (b) Heliocentric $a$ and $e$ of possible SL9 progenitor orbits, based on fragment integrations. The positions of selected comets and two major outer belt asteroid groups, the Trojans and the Hildas, are shown. The dashed curves are for Tisserand parameter $T = 3$ (for zero inclination); orbits above the upper curve and below the lower curve have $T > 3$ and are generally not Jupiter-crossing, while those between the two curves ($T < 3$) are Jupiter-crossing. (Reproduced from Benner and McKinnon [1995].)

Twice as many fragments came from the outer asteroid belt as compared to the inner transjovian region. However, Benner and McKinnon [1995] do not conclude that SL9 originated from the outer asteroid belt. Instead, they say that "the chaos in SL9's orbit is so strong...that what is being seen is a statistical scrambling of all possible trajectories for an object as loosely bound as SL9." The bias toward an asteroid origin is a measure of the relative ease of capture (or escape) toward $L_1$ versus $L_2$, a known result (Heppenheimer and Porco [1977]). The statistical likelihood of a pre-collision interior orbit depends on

the relative populations of interacting comets interior or exterior to Jupiter. If there are roughly equal populations, then a pre-collision interior origin is favored.

## 6.3    Transport in the Planar Circular Restricted Three-Body Problem

When the dynamics are chaotic, statistical methods may be appropriate (Wiggins [1992]). By following ensembles of phase space trajectories, we can determine transition probabilities concerning how likely particles are to move from one region to another.

Following Wiggins [1992], suppose we study the motion on a manifold $\mathcal{M}$. Further, suppose $\mathcal{M}$ is partitioned into disjoint regions

$$R_i, i = 1, \dots, N_R,$$

such that

$$\mathcal{M} = \bigcup_{i=1}^{N_R} R_i.$$

At $t = 0$, region $R_i$ is *uniformly covered* with species $S_i$. Thus, species type of a point indicates the region in which it was located initially.

The statement of the transport problem is then as follows:

*Describe the distribution of species $S_i, i = 1, \dots, N_R$, throughout the regions $R_j, j = 1, \dots, N_R$, for any time $t > 0$.*
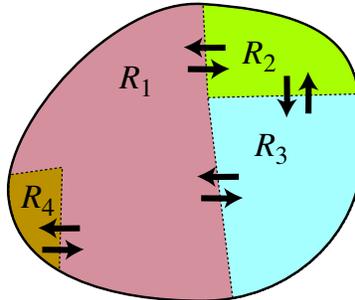


Figure 6.3: The manifold $\mathcal{M}$ is partitioned into the regions $R_i, i = 1, \dots, N_R$. If points are distributed uniformly over $\mathcal{M}$ at $t = 0$, we want to compute the movement of points between these regions for all times $t > 0$.

Some quantities we would like to compute are: $T_{i,j}(t)$, the amount of species $S_i$ contained in region $R_j$, and $F_{i,j}(t) = \frac{dT_{i,j}}{dt}(t)$, the flux of species $S_i$ into region $R_j$ (see Figure 6.3). For some problems, the probability of transport between two regions or the probability of an event occurring (e.g., collision), may be more relevant.

**Planar Circular Restricted Three-Body Problem.** Here we only review the material concerning the PCR3BP which has relevance toward our discussion of transport. See details in Szebehely [1967] and Koon, Lo, Marsden and Ross [2001b].

Consider motion in the standard rotating coordinate system as shown in Figure 6.4 with the origin at the center of mass, and the sun and Jupiter fixed on the $x$-axis at the points $(-\mu, 0)$ and $(1 - \mu, 0)$, respectively. Let $(x, y)$ be the position of the comet in the plane, then the equations of motion in this rotating frame are:

$$\ddot{x} - 2\dot{y} = -\bar{U}_x,$$
$$\ddot{y} + 2\dot{x} = -\bar{U}_y,$$

where

$$\bar{U}(x, y) = -\frac{1}{2}(x^2 + y^2) - \frac{1 - \mu}{r_1} - \frac{\mu}{r_2},$$

is the effective potential and the subscripts denote its partial derivatives and $r_1, r_2$ are the distances from the comet to the sun and the Jupiter, respectively.

These equations are autonomous and can be put into Hamiltonian form. They have an energy integral

$$E = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \bar{U}(x, y).$$

which is related to the Jacobi integral $C$ by $C = -2E$. The Jacobi integral can be expressed approximately in terms of the comet's semimajor axis, $a$, and eccentricity, $e$, in a form known as the Tisserand parameter, $T$, i.e., $C = T + \mathcal{O}(\mu)$, where

$$T = \frac{1}{a} + 2\sqrt{a(1 - e^2)}.$$

The energy manifolds,

$$\mathcal{M}(\mu, e) = \{(x, y, \dot{x}, \dot{y}) \mid E(x, y, \dot{x}, \dot{y}) = e\},$$

where $e$ is a constant are three-dimensional surfaces foliating the four-dimensional phase space. For fixed $\mu$ and $e$, the Hill's region is the projection of the energy manifold onto the position space

$$M(\mu, e) = \{(x, y) \mid \bar{U}(x, y) \leq e\},$$

and is the region in the $xy$-plane where the comet is energetically permitted to move The forbidden region is the region which is not accessible for the given energy. See Figure 6.4(b).
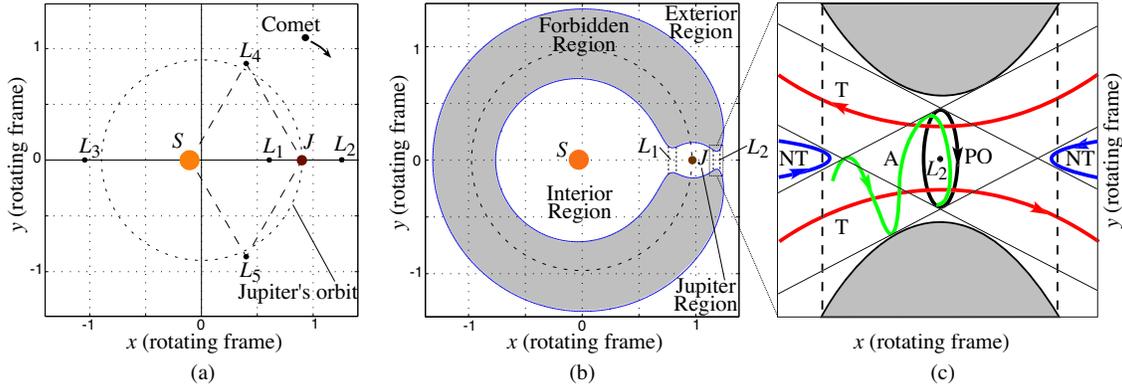


Figure 6.4: (a) The rotating frame showing the libration points, in particular $L_1$ and $L_2$, of the planar, circular restricted three-body problem. (b) The energetically forbidden region is the gray "C." The Hill's region, $M(\mu, e)$ (region in white), contains a bottleneck about $L_1$ and $L_2$. (c) The flow in the region near $L_2$, showing a periodic orbit around $L_2$ (labeled PO), a typical asymptotic orbit winding onto the periodic orbit (A), two transit orbits (T) and two non-transit orbits (NT). A similar figure holds for the region around $L_1$.

Eigenvalues of the linearized equations at $L_1$ and $L_2$ have one real and one imaginary pair, having a saddle $\times$ center structure. Our main concern is the behavior of orbits whose energy is just above that of $L_2$, for which the Hill's region is a connected region with an *interior* region (inside Jupiter's orbit), *exterior* region (outside Jupiter's orbit), and a *Jupiter* region (bubble surrounding Jupiter). We will use the terminology interior, exterior, and Jupiter regions to mean regions in the Hill's region and the corresponding regions of the energy surface, $\mathcal{M}(\mu, e)$. Thus, we have a useful partition for our problem for which we can compute transport properties. These regions are connected by bottlenecks about $L_1$ and $L_2$ and the comet can pass between the regions only through these bottlenecks. Inside each bottleneck, adjacent regions, e.g., the interior and Jupiter regions, share a common

boundary in the energy surface. This common boundary is known as the transition state and has been used previously in astrodynamical transport calculations (Jaffé, Ross, Lo, Marsden, Farrelly and Uzer [2002]). For our analysis of transport, we must focus on the bottlenecks.

In each bottleneck (one around $L_1$ and one around $L_2$), there exist 4 types of orbits, as given in Conley [1968] and illustrated in Figure 6.4(c): (1) an unstable *periodic* Lyapunov orbit; (2) four cylinders of *asymptotic* orbits that wind onto or off this period orbit, which form pieces of stable and unstable manifolds; (3) *transit* orbits which the comet must use to make a transition from one region to the other; and (4) *nontransit* orbits where the comet bounces back to its original region.
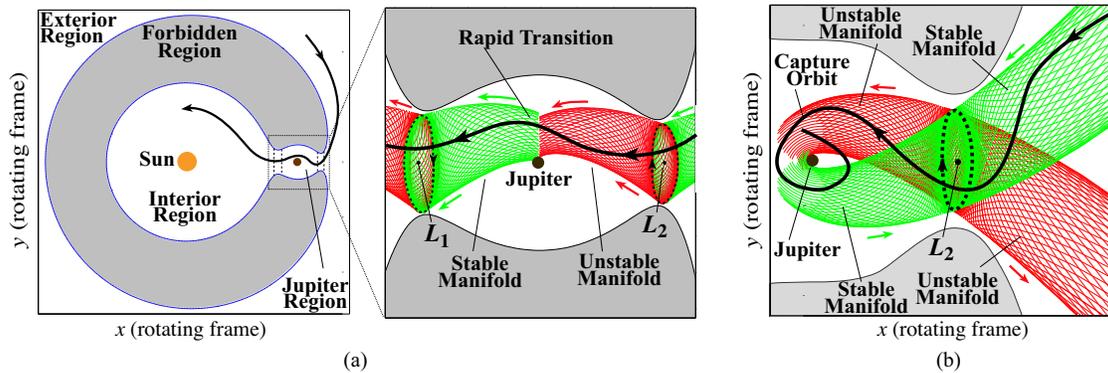


Figure 6.5: (a) An example of an interior-exterior transit orbit. This on goes from outside to inside Jupiter's orbit, passing by Jupiter. The tubes containing transit orbits—bounded by the cylindrical stable (lightly shaded) and unstable (darkly shaded) manifolds—intersect such that a transition is possible. (b) An orbit beginning inside the stable manifold tube in the exterior region is temporarily captured by Jupiter. When the tubes intersect the surface of Jupiter, a collision is possible.

McGehee [1969] was the first to observe that the asymptotic orbits are pieces of the two-dimensional stable and unstable invariant manifold tubes associated to the Lyapunov orbit and that they form the boundary between transit and nontransit orbits. The transit orbits, passing from one region to another, are those inside the cylindrical manifold tube. The nontransit orbits, which bounce back to their region of origin, are those outside the tube. Most importantly, to transit from outside Jupiter's orbit to inside (or vice versa), or get temporarily captured, a comet *must* be inside a tube of transit orbits, as in Figures 6.5(a) and 6.5(b). The invariant manifold tubes are global objects—they extend far beyond

the vicinity of the bottleneck, partitioning the energy manifold.

**Numerical Computation of Invariant Manifolds.** Key to our analysis is the computation of the invariant manifolds of Lyapunov orbits, thus we include some notes on computation methods. Periodic Lyapunov orbits can be computed using a high order analytic expansion (see Llibre, Martinez and Simó [1985]) or by using continuation methods (Doedel, Paffenroth, Keller, Dichmann, Galan and Vanderbauwhede [2003]). Their stable and unstable manifolds can be approximated as given in Parker and Chua [1989]. The basic idea is to linearize the equations of motion about the periodic orbit and then use the monodromy matrix provided by Floquet theory to generate a linear approximation of the stable manifold associated with the periodic orbit. The linear approximation, in the form of a state vector, is numerically integrated in the nonlinear equations of motion to produce the approximation of the stable manifold. All numerical integrations were performed with a standard seventh-eighth order Runge-Kutta method.

**Interior-Exterior Transition Mechanism.** The heart of the transition mechanism from outside to inside Jupiter's orbit (or vice versa) is the intersection of tubes containing transit orbits. We can see the intersection clearly on a two-dimensional Poincaré surface-of-section in the three-dimensional energy manifold. We take our surface to be $\Sigma_{(\mu,e)} = \{(y, \dot{y})|x = 1 - \mu, \dot{x} < 0\}$, along a vertical line passing through Jupiter's center as in Figure 6.6(a). In Figure 6.6(b), we plot $\dot{y}$ versus $y$ along this line, we see that the tube cross-sections are distorted circles. Upon magnification in Figure 6.6(c), it is clear that the tubes indeed intersect.

Any point within the region bounded by the curve corresponding to the stable tube cut is on an orbit that will go from the Jupiter region into the interior region. Similarly, a point within the unstable tube cut is on an orbit that came from the exterior region into the Jupiter region. A point inside the region bounded by the intersection of both curves (lightly shaded in Figure 6.6(c)) is on an orbit that makes the transition from the exterior region to the interior region, via the Jupiter region.

**Interior-Exterior Transition Probability.** Note that since $p_y = \dot{y} + x$ and $x$ is constant, the $(y, \dot{y})$ plane is a linear displacement of the canonical plane $(y, p_y)$. Furthermore,
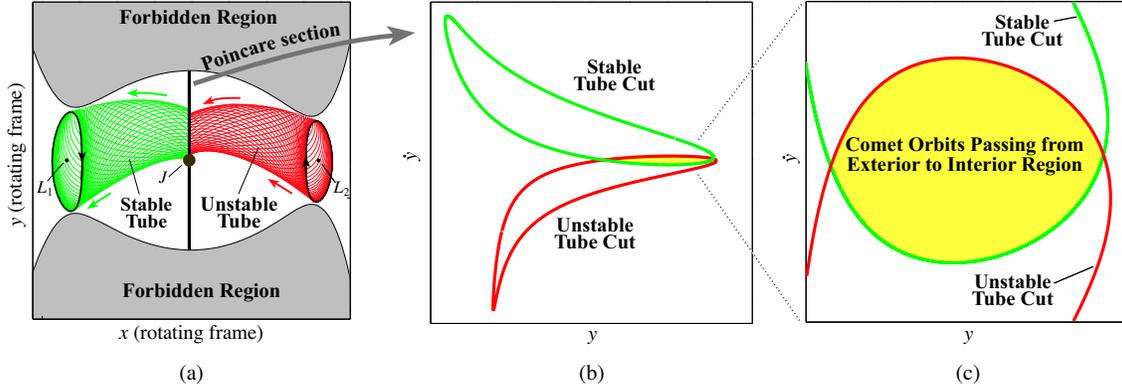
Figure 6.6: (a) We take a Poincaré surface-of-section $\Sigma_{(\mu,e)} = \{(y,\dot{y})|x = 1 - \mu, \dot{x} < 0\}$, along a vertical line through the center of Jupiter ($J$). Both the $L_1$ and $L_2$ periodic orbit invariant manifold tubes intersect $\Sigma_{(\mu,e)}$ transversally. (b) On $\Sigma_{(\mu,e)}$, we see the first unstable tube cut for $L_2$ and first stable tube cut for $L_1$. (c) A small portion of the interior of the tubes intersect—this set in the energy manifold $\mathcal{M}(\mu,e)$ containing the comet orbits which pass from the exterior to the interior region.

the action integral around any closed loop $\Gamma$ on $\Sigma_{(\mu,e)}$,

$$S = \oint_{\Gamma} p \cdot dq = \oint p_y \, dy,$$

is simply the area enclosed by $\Gamma$ on the surface-of-section $\Sigma_{(\mu,e)}$ (Meiss [1992]).

The agreement between a Monte-Carlo simulation and a Markov approximation in an earlier paper (Jaffé, Ross, Lo, Marsden, Farrelly and Uzer [2002]) suggests that for energies slightly above $L_1$ and $L_2$, there are components of the energy surface for which the motion is "well mixed" (cf. Meiss [1992]). Thus, the Markov approximation is a good one. Let $R_1$ be the interior region and $R_2$ be the exterior region. In the Markov approximation, the probability of a particle going from region $R_i$ to $R_j$ is

$$P_{ij} = \frac{\mathcal{F}_{ij}}{A_j}$$

where $A_j$ is the area of the first unstable tube cut on $\Sigma_{(\mu,e)}$, containing transit orbits from $R_j$, and $\mathcal{F}_{ij} = \mathcal{F}_{ji}$ is the area of overlap of the first unstable tube cut from $R_j$ and the first stable tube cut from $R_i$ on $\Sigma_{(\mu,e)}$. This transition probability is exact for one iterate of the Poincaré map; however, it is typically only qualitatively correct for longer times.

In Figure 6.7, we give the results of the calculations of $P_{12}$ and $P_{21}$ for mass paramter

$\mu = 9.537 \times 10^{-4}$ and a variety of energies in the range of QH Jupiter-family comets. This is the probability of a comet to move from the interior to the exterior and vice versa during its first pass through the surface-of-section $\Sigma_{(\mu,e)}$.
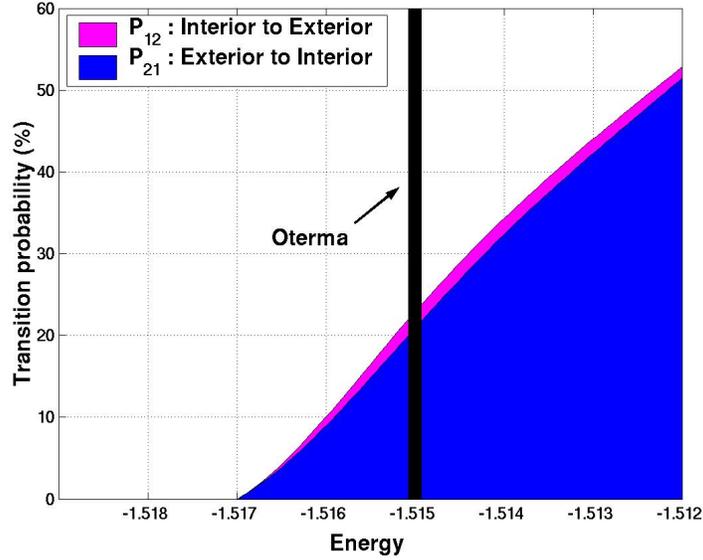


Figure 6.7: **Interior-exterior transition probabilities for quasi-Hilda Jupiter-family comets.** The probability of a comet to move from the interior to the exterior and vice versa during its first pass through the surface-of-section $\Sigma_{(\mu,e)}$ is plotted as a function of energy in the planar, circular restricted three-body problem. The energy value of P/Oterma is shown for comparison. Note that interior to exterior transitions are slightly more probable than the reverse transition.

A few comments regarding this result are due. (1) Notice that there is a lower limit in energy, $E_t \approx -1.517$. For $E \leq E_t$, the tube cuts do not overlap and no direct transition is possible. After more loops around Jupiter, transition may be possible (cf. Koon, Lo, Marsden and Ross [2000]). (2) The probability increases as a function of energy. (3) Quasi-Hilda P/Oterma is located in the region of $\approx 25\%$ probability. (4) Finally, notice that $P_{12} > P_{21}$, which is a result of $A_1 > A_2$, the slight asymmetry we shuld expect for a mass parameter of this value or larger (cf. Simó and Stuchi [2000]).

**Collision Probabilities.** Collision probabilities can be computed for objects coming through the $L_1$ and $L_2$ bottlenecks from the interior and exterior regions, respectively. We augment the procedure for computing interior-exterior transition probalities in the

following way. Instead of computing $\mathcal{F}_{ij}$, we now compute the overlap of the first unstable manifold cut with the *diameter* of the secondary (e.g., Jupiter). Since the surface $\Sigma_{(\mu,e)}$ passes through the center of secondary, any particle located on $\Sigma_{(\mu,e)}$ with $|y| \leq R$ will have collided with the secondary, where $R$ is the radius of secondary in units of the primary-secondary distance. This is illustrated in Figure 6.8.



**Poincare Section: Tube Intersecting a Planet**
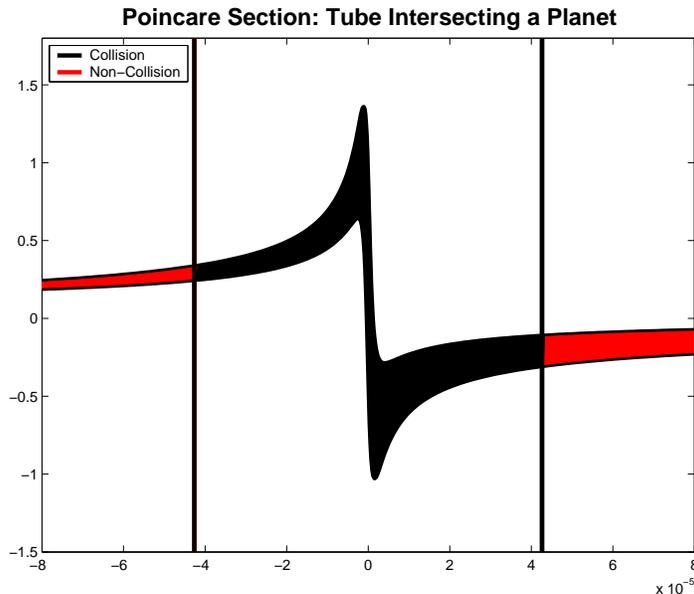
Figure 6.8: The surface-of-section, $\Sigma_{(\mu,e)}$, is shown, with $y$ vs. $\dot{y}$. The area inside the first unstable manifold tube cut with $|y| \leq R$ is shown in in black. These are orbits that collide with the surface of the secondary. The two vertical lines are at $y = \pm R$.

There is a singularity at the center of the secondary, $y = 0$ on $\Sigma_{(\mu,e)}$, so the calculation is actually performed along a nearby parallel surface-of-section, where $x = 1 - \mu \pm c$, with $c$ a small number on the order of the integration tolerance (the '+' sign is for orbits coming from the exterior, and the '−' for orbits coming from the interior).

Collision probabilities for the sun-Jupiter case ($\mu = 9.537 \times 10^{-4}$, $R = 8.982 \times 10^{-5}$) are given in Figure 6.9. We notice the following. (1) The probability is not monotonically increasing as in Figure 6.7. (2) The energy range of possible pre-collision Shoemaker-Levy 9 orbits (from Benner and McKinnon [1995]) lies in the range of highest collision probability, suggesting the utility of this approach. (3) There is an asymmetry in orbits coming from the interior or the exterior, and now there are two lower energy cutoffs, $E_c^1 \approx -1.5173$ and $E_c^2 \approx -1.5165$, below which no collision can occur on the first pass
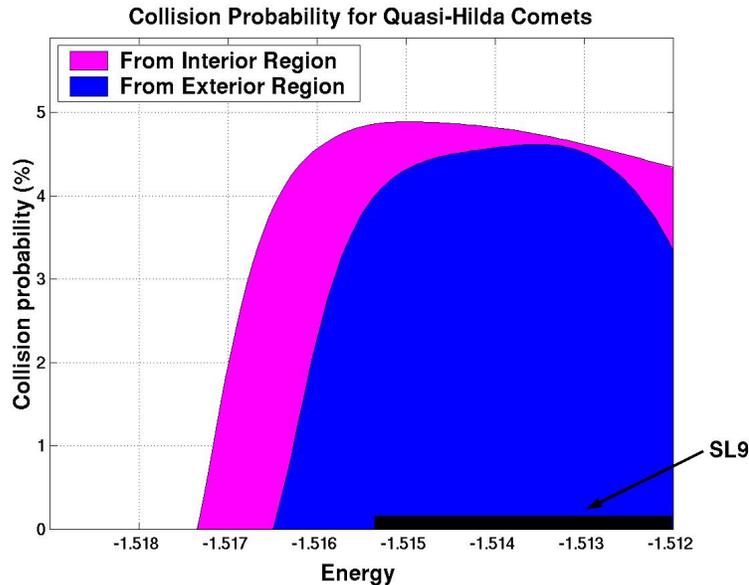
Figure 6.9: **Collision probabilities for quasi-Hilda comets.** The probability of collision for orbits making their first pass through the surface-of-section $\Sigma_{(\mu,e)}$ is plotted as a function of energy. The energy range of possible pre-collision D/Shoemaker-Levy 9 orbits is shown for comparison.

by Jupiter. The asymmetry may be too slight to differentiate an interior origin from an exterior origin for SL9.

As a final computation, we address the NEA collision problem. For a mass parameter corresponding to the sun-Earth-asteroid problem ($\mu = 3.036 \times 10^{-6}$, $R = 4.258 \times 10^{-5}$), we compute the collision probability. The result is shown in Figure 6.10. It is interesting that the collision probabilities are nearly twice those for the quasi-Hilda case, even though Jupiter has a much larger mass and radius than the Earth. The asymmetry in interior/exterior originating orbits is not as pronounced as in Figure 6.9, owing to the smaller value of $\mu$, and $E_c^1 \approx E_c^2 \approx -1.5 - 4.03 \times 10^{-4}$.

## 6.4   Conclusions

We address some questions regarding nonlinear comet and asteroid behavior by applying statistical methods to the planar, circular restricted three-body problem. In particular, we make a Markov assumption regarding the phase space and compute probabilities of interior-exterior transition and collision with the secondary. Theory and observation are
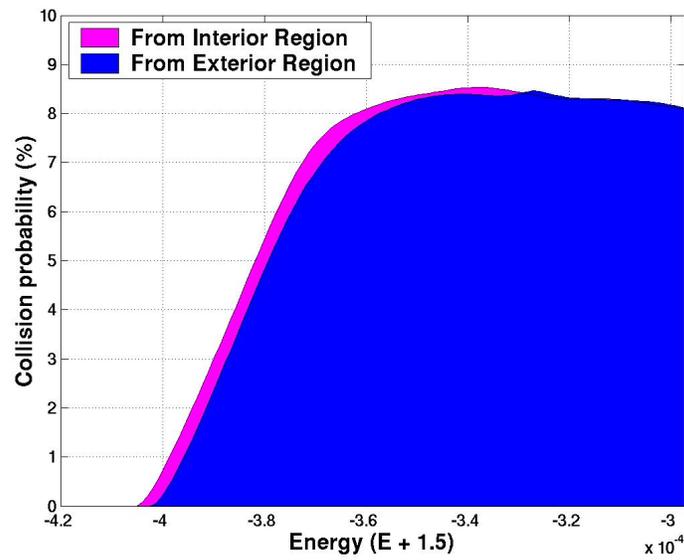
Figure 6.10: **Collision probabilities for near-Earth asteroids.** Note that the collision probabilities are nearly twice those for the quasi-Hilda case in Figure 6.9, even though Jupiter has a much larger mass and radius than the Earth.

seen to agree for the comets P/Oterma and D/Shoemaker-Levy 9.

# Chapter 7

# Conclusion

In this thesis, we have applied dynamical systems techniques to consider the global dynamics of the planar circular restricted three-body problem (PCR3BP). We consider the stable and unstable manifold tubes associated to periodic orbits about the libration points $L_1$ and $L_2$. The periodic orbits considered reside in bottleneck regions of the energy manifold, separating large realms associated with motion about one mass, the other mass, or both masses. The cylinders have the physical property that all motion through the bottlenecks must occur through the interior of these surfaces. The cylinders thus mediate the global transport of test particles between large zones of the energy surface which are separated by the bottlenecks.

One of the main results is the numerical demonstration of the existence of a transversal heteroclinic orbit connecting pairs of periodic orbits: one around $L_1$ and the other around $L_2$, with the two periodic orbits having the same energy. This result, initially reported by Koon, Lo, Marsden and Ross [2000], was rigorously proven using a computer assisted approach by Wilczak and Zgliczyński [2003] and Kirchgraber and Stoffer [2004]. This augments the known homoclinic orbits associated to the $L_1$ and $L_2$ Lyapunov orbits which were proven to exist by McGehee [1969] and Llibre, Martinez and Simó [1985]. By linking these heteroclinic orbits with homoclinic orbits on the same energy surface, we prove a theorem on the global orbit structure of the PCR3BP. This theorem can also be taken as a proof of "horseshoe-like" chaos in the system.

**Spacecraft Trajectory Design.** Spacecraft can utilize the sensitive dynamics in this regime of motion to explore a large region of space near Earth (and near Earth's orbit) using low-fuel controls. Behavior related to the dynamical channels has already been

observed by Lo, Williams, Bollman, Han, Hahn, Bell, Hirst, Corwin, Hong, Howell and Barden [1998] in the trajectory for the *Genesis Discovery Mission*, which exhibits near-heteroclinic motion between $L_1$ and $L_2$ in the Sun-Earth system. See Figure 7.1. With
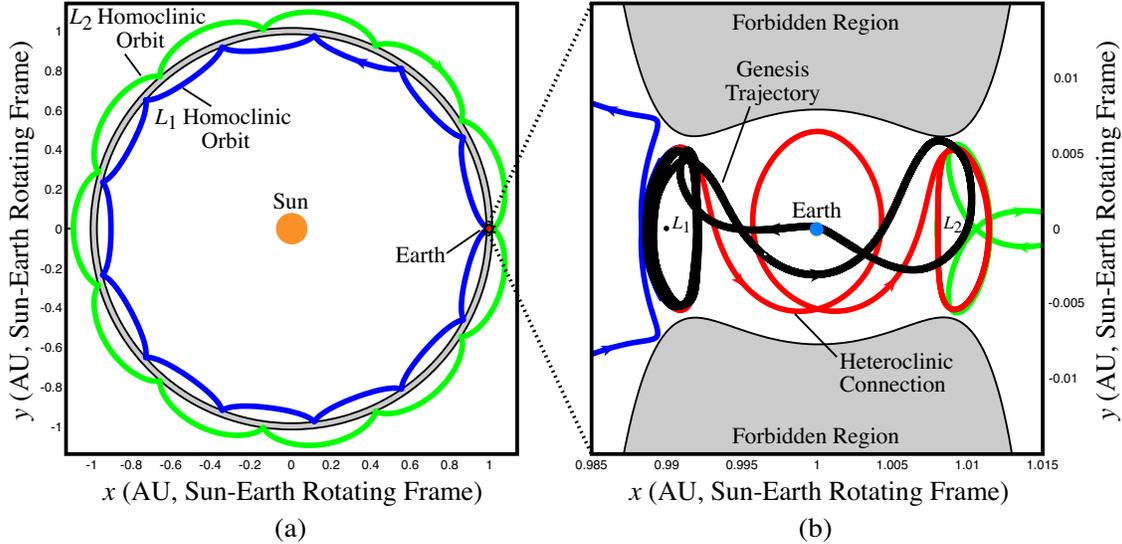


Figure 7.1: (a) A homoclinic-heteroclinic chain on the *Genesis Discovery Mission* trajectory's energy surface. (b) Close-up of the chain in Earth's vicinity. The actual *Genesis Discovery Mission* trajectory is shown in black overlaying the chain, and in particular, the heteroclinic connection from $L_1$ to $L_2$.

a better understanding of the underlying homoclinic-heteroclinic structures we should be able to construct and control spacecraft trajectories with desired exotic characteristics (e.g., transfer between $L_1$ and $L_2$ orbits, explore interior region and then return to Earth's vicinity).

In fact, the channels can be utilized around any planet or moon system. These dynamical channels can be exploited for the systematic construction of low energy spacecraft trajectories, such as the Sun-perturbed Earth-to-Moon example given in Chapter 5 which saves 20% of the fuel compared with an Apollo-like Hohmann transfer.

Greater space mission flexibility could be achieved post-launch owing to the sensitivity of the phase space in these dynamical channels. Miniscule fuel expenditures could lead to dramatically different spacecraft trajectories. One could turn a near-Earth mission into an asteroid rendezvous and return mission *in situ* with an appropriately placed small thrust. Rather than being a hindrance to orbital stability, sensitivity facilitates mission versatility.

**Material Transport Throughout the Solar System.** Small bodies such as asteroids and comets, through their interactions with the planets and the Sun, can exhibit strongly chaotic motion. Nevertheless, the ability to predict the behavior of populations of these small but numerous objects is essential for understanding such problems as the evolution of both short- and long-range comets originating in the Kuiper Belt and the Oort Cloud, respectively, the dynamics of near-Earth asteroids, and zodiacal and circumplanetary dust dynamics. These problems illustrate the fundamental importance of understanding mass transport in the solar system.

In principle, the computation of probabilities rates of mass transport can be accomplished by large numerical simulations in which the orbits of vast numbers of test particles are propagated in time including as many interactions as desirable. However, such calculations are computationally demanding and it may be difficult to extract from them information about key dynamical mechanisms. They do have the considerable advantage, however, that a variety of nongravitational effects can easily be included, even if these destroy the Hamiltonian nature of the problem.

In the work of Chapter 6 and Jaffé, Ross, Lo, Marsden, Farrelly and Uzer [2002], a complementary approach has been developed that can be used provided that the problem is of autonomous Hamiltonian form.

In the interior-exterior transition and collision probability problem studied in Chapter 6, the dynamics was confined to the plane so as to allow the simplest illustration of the method. However, the phase space transport theory is most powerful for multidimensional degree-of-freedom systems for which simulations become more difficult and insight into the dynamical mechanisms is harder to extract. Examples include the evolution of long-range comets and circumplanetary dust escape from nonequatorial "halo" orbits. While we have not considered nongravitational forces our methods allow the inclusion, for example, of interactions of charged dust grains with planetary magnetic fields or the effect of solar radiation pressure.

# Appendix A

# Table of Parameter Values

The following table lists the mass parameters for some important two-body systems in the solar system, some of which are covered in the text. For system $m_1$-$m_2$ (where $m_1 > m_2$), the mass parameter,

$$\mu = \frac{m_2}{m_1 + m_2},$$

is given. In addition the conversion factors to convert to dimensional units (e.g., km, km/s, sec) are also given. The conversion from units of distance, velocity, and time in the unprimed, normalized system to the primed, dimensionalized system is

$$\text{distance} \quad d' = Ld,$$
$$\text{velocity} \quad s' = Vs,$$
$$\text{time} \quad t' = \frac{T}{2\pi}t,$$

where $L$ is the distance between the centers of $m_1$ and $m_2$, $V$ is the orbital velocity of $m_1$, $T$ is the orbital period of $m_1$ and $m_2$.

| System | $\mu$ | $L$ (km) | $V$ (km/s) | $T$ (sec) |
|---|---|---|---|---|
| Sun-Jupiter | $9.537 \times 10^{-4}$ | $7.784 \times 10^8$ | 13.102 | $3.733 \times 10^8$ |
| Sun-(Earth+Moon) | $3.036 \times 10^{-6}$ | $1.496 \times 10^8$ | 29.784 | $3.147 \times 10^7$ |
| Earth-Moon | $1.215 \times 10^{-2}$ | $3.850 \times 10^5$ | 1.025 | $2.361 \times 10^6$ |
| Mars-Phobos | $1.667 \times 10^{-8}$ | $9.380 \times 10^3$ | 2.144 | $2.749 \times 10^4$ |
| Jupiter-Io | $4.704 \times 10^{-5}$ | $4.218 \times 10^5$ | 17.390 | $1.524 \times 10^5$ |
| Jupiter-Europa | $2.528 \times 10^{-5}$ | $6.711 \times 10^5$ | 13.780 | $3.060 \times 10^5$ |
| Jupiter-Ganymede | $7.804 \times 10^{-5}$ | $1.070 \times 10^6$ | 10.909 | $6.165 \times 10^5$ |
| Jupiter-Callisto | $5.667 \times 10^{-5}$ | $1.883 \times 10^6$ | 8.226 | $1.438 \times 10^6$ |
| Saturn-Mimas | $6.723 \times 10^{-8}$ | $1.856 \times 10^5$ | 14.367 | $8.117 \times 10^4$ |
| Saturn-Titan | $2.366 \times 10^{-4}$ | $1.222 \times 10^6$ | 5.588 | $1.374 \times 10^6$ |
| Neptune-Triton | $2.089 \times 10^{-4}$ | $3.548 \times 10^5$ | 4.402 | $5.064 \times 10^5$ |
| Pluto-Charon | $1.097 \times 10^{-1}$ | $1.941 \times 10^4$ | 0.222 | $5.503 \times 10^5$ |

Table A.1: **Table of $m_1$-$m_2$ systems in the solar system.** Source: The first three are the values used in Koon, Lo, Marsden and Ross [2000] and Koon, Lo, Marsden and Ross [2001a]. The others are from the Jet Propulsion Laboratory's solar system dynamics website: `http://ssd.jpl.nasa.gov/`.

# Appendix B

# Moser's Theorem and Local Integrals Near the Collinear Equilibrium Points

Given a time-independent, analytic Hamiltonian system of differential equations with two degree of freedom. Suppose these equations have a non-degenerate equilibrium point with one pair of real and one pair of imaginary eigenvalues, $\pm\lambda$ and $\pm i\nu$. We can assume, without loss of generality, that the phase space coordinates $x_1, x_2, y_1, y_2$ are chosen so that the Hamiltonian function assumes the following form:

$$H(x, y) = \lambda x_1 y_1 + \tfrac{1}{2}\nu(x_2^2 + y_2^2) + O_3(x, y),$$

where $x = (x_1, x_2), y = (y_1, y_2)$ and the symbol $O_n(\cdot, \cdot)$ denotes terms of order $n$ or higher in the variables deplayed.

In particular, the equilibrium point has coordinates $x = y = 0$ and the differential equations are obtained from $H$ as

$$
\begin{aligned}
\dot{x}_1 &= H_{y_1} = \lambda x_1 + O_2(x, y), \\
\dot{y}_1 &= -H_{x_1} = -\lambda y_1 + O_2(x, y), \\
\dot{x}_2 &= H_{y_2} = \nu y_2 + O_2(x, y), \\
\dot{y}_2 &= -H_{y_1} = \nu x_2 + O_2(x, y).
\end{aligned}
$$

The linearized equations are similarly obtained from a Hamiltonian function which consists of the quadratic terms of $H$ or, equivalently, by dropping the terms of order two or higher in the above equations. Solutions of these linearized equations are conveniently

written as

$$x_1(t) = x_1^0 e^{\lambda t}, \qquad x_2(t) = x_2^0(t)e^{-\lambda t},$$

$$z(t) = x_2(t) + iy_2(t) = z^0 e^{-\nu t},$$

where the constants $x_1^0, y_1^0$ and $z^0 = x_2^0 + iy_2^0$ are the initial conditions.

These linearized equations admit integrals in addition to the Hamiltonian function; namely, the functions $x_1 y_1$ and $|z|^2 = x_2^2 + y_2^2$ are both constant along solutions. A special case of a theorem by Moser [1958] states that the full non-linear equations admit "local" integrals analogous to these: thus there are two power series in $x$ and $y$ beginning respectively with quadratic terms $x_1 y_1$ and $x_1^2 + y_2^2$ which converge in some neighborhood of $x = y = 0$ and such that the corresponding functions are constanta along pieces of solutions lying in the domain of convergence.

A special case of Moser's theorem is stated by Conley [1969] in a form suited to the PCR3BP. In this statement $\xi$ and $\eta$ are real variables and $\zeta$ is complex.

**Theorem (Moser).** Let $x = y = 0$ correspond to a critical point as decribed above. Then there exists a (real) analytic, transformation

$$x_1 = \xi + O_2(\xi, \eta, \zeta, \bar{\zeta}), \qquad y_1 = \eta + O_2(\xi, \eta, \zeta, \bar{\zeta}),$$

$$z = x_2 + iy_2 = \zeta + O_2(\xi, \eta, \zeta, \bar{\zeta}),$$

as well as power series $\alpha$ and $\beta$ in the variables $\chi = \xi\eta$ and $|\zeta|^2$ of the form

$$\alpha = \lambda + O_1(\chi, |\zeta|^2),$$

$$\beta = -i\nu + O_1(\chi, |\zeta|^2),$$

such that solutions of the transformed equations are given by

$$\xi(t) = \xi^0 e^{t\alpha}, \qquad \eta(t) = \eta^0 e^{-t\alpha},$$

$$\zeta(t) = \zeta^0 e^{t\beta}, \qquad \bar{\zeta} = \bar{\zeta}^0 e^{-t\beta},$$

where $\xi^0, \eta^0$ and $\zeta^0$ are determined from the initial conditions and $\bar{\zeta}$ is the complex conjugate of $\zeta$. Furthermore, the coefficients of $\alpha$ and $\beta$ are real and complex, respectively,

from which it follows that the functions $\xi\eta = x_1y_1 + O_3(x,y)$ and $|\zeta|^2 = x_2^2 + y_2^2 + O_3(x,y)$ are local integrals, as are $\alpha$ and $\beta$.

Finally, the transformation of the Hamiltonian function has the form

$$K(\xi,\eta,\zeta,\bar{\zeta}) = H(x,y) = \lambda\xi\eta + \tfrac{1}{2}|\zeta|^2 + O_2(\chi,|\zeta|^2),$$

and in particular depends only on the variables $\chi = \xi\eta$ and $|\zeta|^2$.

# Bibliography

Abraham, R. and J. E. Marsden [1978] *Foundations of Mechanics*. Addison-Wesley, New York, 2nd edn.

Belbruno, E. and B. Marsden [1997] Resonance hopping in comets. *The Astronomical Journal* **113**, 1433–1444.

Belbruno, E. A. and J. K. Miller [1993] Sun-perturbed Earth-to-Moon transfers with ballistic capture. *Journal of Guidance, Control and Dynamics* **16**, 770–775.

Benner, L. A. M. and W. B. McKinnon [1995] On the orbital evolution and origin of comet Shoemaker-Levy 9. *Icarus* **118**, 155–168.

Carusi, A., L. Kresák, E. Pozzi and G. B. Valsecchi [1985] *Long Term Evolution of Short Period Comets*. Adam Hilger, Bristol, UK.

Chodas, P. W. and D. K. Yeomans [1993] The upcoming collision of comet Shoemaker-Levy 9 with Jupiter. *Bull. Am. Astron. Soc.* **25**.

Conley, C. C. [1968] Low energy transit orbits in the restricted three-body problem. *SIAM J. Appl. Math.* **16**, 732–746.

Conley, C. C. [1969] On the ultimate behavior of orbits with respect to an unstable critical point. I. Oscillating, asymptotic, and capture orbits. *J. Differential Equations* **5**, 136–158.

Devaney, R. L. [1981] Singularities in classical mechanical systems. In A. Katok, ed., *Ergodic Theory and Dynamical Systems*, pages 211–333. Birkhäuer, Basel.

Doedel, E. J., R. C. Paffenroth, H. B. Keller, D. J. Dichmann, J. Galan and A. Vanderbauwhede [2003] Continuation of periodic solutions in conservative systems with application to the 3-body problem. *Int. J. Bifurcation and Chaos* To appear.

Euler, L. [1767] De motu rectilineo trium corporum se mutuo attrahentium. *Novi commentarii academiae scientarum Petropolitanae* **11**, 144–151. In Oeuvres, Seria Secunda tome XXV Commentationes Astronomicae (page 286).

Gómez, G., A. Jorba, J. Masdemont and C. Simó [1991] Study refinement of semi-analytical halo orbit theory, Final Report, ESOC Contract No.:8625/89/D/MD(SC).

Gómez, G., W. S. Koon, M. W. Lo, J. E. Marsden, J. Masdemont and S. D. Ross [2001] Invariant manifolds, the spatial three-body problem and space mission design (Paper No. AAS 01-301). In D. B. Spencer, C. C. Seybold, A. K. Misra and R. J. Lisowski, eds., *Advances in the Astronautical Sciences*, vol. 109, pages 3–22. Univelt, San Diego.

Gómez, G., J. Masdemont and C. Simó [1993] Study of the transfer from the earth to a halo orbit around the equilibrium point $l_1$. *Cel. Mech. Dyn. Astro.* **56**, 541–562.

Heppenheimer, T. A. and C. Porco [1977] New contributions to the problem of capture. *Icarus* **30**, 385–401.

Holmes, P. [1990] Poincaré, celestial mechanics, dynamical-systems theory and "chaos". *Physics Reports* **193**, 137–163.

Howell, K., B. Barden and M. Lo [1997] Application of dynamical systems theory to trajectory design for a libration point mission. *The Journal of the Astronautical Sciences* **45**, 161–178.

Howell, K. C., B. G. Marchand and M. W. Lo [2000] Temporary satellite capture of short period Jupiter family comets from the perspective of dynamical systems. In *AAS/AIAA Space Flight Mechanics Meeting*. Clearwater, Florida. Paper No. AAS 00-155.

Jaffé, C., S. D. Ross, M. W. Lo, J. E. Marsden, D. Farrelly and T. Uzer [2002] Theory of asteroid escape rates. *Physical Review Letters* **89**, 011101–1.

Kary, D. M. and L. Dones [1996] Capture statistics of short-period comets: implications for comet D/Shoemaker-Levy 9. *Icarus* **121**, 207–224.

Kirchgraber, U. and D. Stoffer [2004] Possible chaotic motion of comets in the Sun-Jupiter system—a computer-assisted approach based on shadowing. *Nonlinearity* **17**, 281–300.

Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [1999] Constructing a low energy transfer between Jovian moons. In *Celestial Mechanics : an international conference on celestial mechanics*. Evanston, IL.

Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [2000] Heteroclinic connections between periodic orbits and resonance transitions in celestial mechanics. *Chaos* **10**, 427–469.

Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [2001a] Low energy transfer to the Moon. *Celest. Mech. Dyn. Astron.* **81**, 63–73.

Koon, W. S., M. W. Lo, J. E. Marsden and S. D. Ross [2001b] Resonance and capture of Jupiter comets. *Celest. Mech. Dyn. Astron.* **81**, 27–38.

Kresák, L. [1979] Dynamical interrelations among comets and asteroids. In T. Gehrels, ed., *Asteroids*, pages 289–309. Univ. of Arizona Press, Tucson.

Lagrange, J.-L. [1772] Essai sur le problème des trois corps. Prix de l'Académie Royale des Sciences de Paris, tome IX, in vol. 6 of *Oeuvres de Lagrange* (Gauthier-Villars, Paris, 1873), 272–282.

Llibre, J., R. Martinez and C. Simó [1985] Transversality of the invariant manifolds associated to the Lyapunov family of periodic orbits near L2 in the restricted three-body problem. *J. Diff. Eqns.* **58**, 104–156.

Lo, M., 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 and B. Barden [1998] Genesis mission design. Paper No. AIAA 98-4468.

Lo, M. W. and S. D. Ross [1997] Surfing the interplanetary tides. *Bull. Am. Astron. Soc.* **29**, 1021.

Marchal, C. [1990] *The Three-Body Problem*. Elsevier, Amsterdam.

Marsden, B. G. [1993] *IAU Circ.* **5726**.

Marsden, J. E. and T. S. Ratiu [1999] *Introduction to Mechanics and Symmetry*, vol. 17 of *Texts in Applied Mathematics*. Springer-Verlag.

McGehee, R. [1969] *Some homoclinic orbits for the restricted three-body problem*. Ph.D. thesis, University of Wisconsin, Madison.

Meiss, J. D. [1992] Symplectic maps, variational principles, and transport. *Rev. Mod. Phys.* **64**, 795–848.

Meyer, K. R. and R. Hall [1992] *Hamiltonian Mechanics and the N-Body Problem*. Texts in Applied Mathematics Science. Springer-Verlag, Berlin.

Moser, J. [1958] On the generalization of a theorem of Liapunov. *Comm. Pure Appl. Math.* **11**, 257–271.

Moser, J. [1973] *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics*. Princeton University Press.

Parker, T. S. and L. O. Chua [1989] *Practical Numerical Algorithms for Chaotic Systems*. Springer-Verlag, New York.

Poincaré, H. [1890] Sur la problème des trois corps et les équations de la dynamique. *Acta Math.* **13**, 1–271.

Poincaré, H. [1892-1899] *Les Méthodes Nouvelles de la Mécanique Céleste (New Methods in Celestial Mechanics)*, vol. 1-3. Gauthier-Villars, Paris. Reprinted by Dover, New York, 1957.

Ross, S. D., W. S. Koon, M. W. Lo and J. E. Marsden [2003] Design of a multi-moon orbiter. In *13th AAS/AIAA Space Flight Mechanics Meeting*. Ponce, Puerto Rico. Paper No. AAS 03-143.

Ross, S. D., W. S. Koon, M. W. Lo and J. E. Marsden [2004] Application of dynamical systems theory to a very low energy transfer. In *14th AAS/AIAA Space Flight Mechanics Meeting*. Maui, Hawaii. Paper No. AAS 04-289.

Simó, C. [1999] Dynamical systems methods for space missions on a vicinity of collinear libration points. In C. Simó, ed., *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, vol. 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 223–241. Kluwer Acad. Publ., Dordrecht.

Simó, C., G. Gómez, A. Jorba and J. Masdemont [1995] The bicircular model near the triangular libration points. In A. E. Roy and B. A. Steves, eds., *From Newton to Chaos*, pages 343–370. Plenum Press, New York.

Simó, C. and T. J. Stuchi [2000] Central stable/unstable manifolds and the destruction of KAM tori in the planar Hill problem. *Physica D* **140**, 1–32.

Szebehely, V. [1967] *Theory of Orbits: The Restricted Problem of Three Bodies*. Academic, New York.

Tancredi, G., M. Lindgren and H. Rickman [1990] Temporary satellite capture and orbital evolution of comet P/Helin-Roman-Crockett. *Astronomy and Astrophysics* **239**, 375–380.

Valsecchi, G. B. [1992] Close encounters, planetary masses, and the evolution of cometary orbits. In J. A. Fernández and H. Rickman, eds., *Periodic Comets*, pages 143–157. Univ. de la Republica, Montevideo, Uruguay.

Whittaker, E. T. [1927] *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies: With an Introduction to the Problem of Three Bodies*. Cambridge University Press, Cambridge.

Wiggins, S. [1990] *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, vol. 2 of *Texts in Applied Mathematics Science*. Springer-Verlag, Berlin.

Wiggins, S. [1992] *Chaotic Transport in Dynamical Systems*, vol. 2 of *Interdisciplinary Appl. Math.*. Springer, Berlin-Heidelberg-New York.

Wiggins, S. [1993] *Global Dynamics, Phase Space Transport, Orbits Homoclinic to Resonances, and Applications*. Fields Institute Monographs. American Mathematical Society.

Wilczak, S. and P. Zgliczyński [2003] Heteroclinic connections between periodic orbits in planar restricted circular three body problem - a computer assisted proof. *Commun. Math. Phys.* **234**, 37–75.