Heteroclinic Connections between Periodic Orbits and Resonance Transitions in Celestial Mechanics

Wang Sang Koon
Control and Dynamical Systems and JPL
Caltech 107-81, Pasadena, CA 91125
koon@cds.caltech.edu

Martin W. Lo
Navigation and Flight Mechanics
Jet Propulsion Laboratory M/S: 301-142
4800 Oak Grove Drive
Pasadena, CA 91109-8099
mwl@trantor.Jpl.Nasa.Gov

Jerrold E. Marsden
Control and Dynamical Systems
Caltech 107-81, Pasadena, CA 91125
marsden@cds.caltech.edu

Shane D. Ross
Control and Dynamical Systems and JPL
Caltech 107-81, Pasadena, CA 91125
shane@cds.caltech.edu

Abstract

This paper applies dynamical systems techniques to the problem of heteroclinic connections and resonance transitions in the planar circular restricted three-body problem. These related phenomena have been of concern for some time in topics such as the capture of comets and asteroids and with the design of trajectories for space missions such as the *Genesis Discovery Mission*. The main new technical result in this paper is the numerical demonstration of the existence of a heteroclinic connection between pairs of periodic orbits, one around the libration point L_1 and the other around L_2, with the two periodic orbits having the same energy. This result is applied to the resonance transition problem and to the explicit numerical construction of interesting orbits with prescribed itineraries. The point of view developed in this paper is that the invariant manifold structures associated to L_1 and L_2 as well as the aforementioned heteroclinic connection are fundamental tools that can aid in understanding dynamical channels throughout the solar system as well as transport between the "interior" and "exterior" Hill's regions and other resonant phenomena.

1

Contents

1   Introduction.

1.1   Background and a Brief Overview.

The three-body problem is a classic problem of astrodynamics. Attempts at its solution laid the foundation for dynamical systems theory and alerted Poincaré to the existence of chaos within Newtonian mechanics. This paper offers a dynamical system explanation for the phenomenon of temporary capture and resonant transition of Jupiter comets within a three-body context. It also explores the possibility of using the transport mechanism discovered in this study for the design of future space missions. For a general introduction to the three-body problem, see Holmes [1990] and Simó [1999].

Resonant Transition in Comet Orbits.   A number of Jupiter comets such as *Oterma* and *Gehrels 3* make a rapid transition from heliocentric orbits outside the orbit of Jupiter to heliocentric orbits inside the orbit of Jupiter and vice versa. During this transition, the comet is frequently captured temporarily by Jupiter for one to several orbits around Jupiter. The interior heliocentric orbit is typically close to the 3:2 resonance (three revolutions around the Sun in two Jupiter periods) while the exterior heliocentric orbit is near the 2:3 resonance (two revolutions around the Sun in three Jupiter periods).

   An important feature of the dynamics of these comets is that during the transition, the orbit passes close to the libration points L_1 and L_2. As we recall below, the points L_1 and L_2 are two of the five equilibrium points for the restricted three-body problem for the Sun-Jupiter system. Equilibrium points are points at which a particle at rest relative to the Sun-Jupiter rotating frame remains at rest. Amongst the equilibrium points, the points L_1 and L_2 are the ones closest to Jupiter, lying on either side of Jupiter along the Sun-Jupiter line.

The Relevance of Invariant Manifolds.   Belbruno and Marsden [1997] attempted to develop a theoretical understanding of the comet transitions using the "fuzzy boundary" concept, which they viewed as "a higher-dimensional analogue of L_1 and L_2." On the other hand, Lo and Ross [1997] began the use of dynamical systems theory to explain this same phenomenon. They used the planar circular restricted three-body problem (***PCR3BP***) as the underlying model with which to begin the investigation. They noticed that the orbits of *Oterma* and *Gehrels 3* (in the Sun-Jupiter rotating frame) follow closely the plots of the invariant manifolds of L_1 and L_2, as in Figure 1.1.

   Having noticed this, Lo and Ross [1997] suggested that one might use invariant manifold theory to study these transitional orbits. The present paper builds on the insights of these works and offers a dynamical system explanation for this phenomenon of temporary capture and resonance transition of Jupiter comets. A key ingredient in our work is the existence of a new heteroclinic connection between periodic orbits around L_1 and L_2 with the same Jacobi constant (a multiple of the Hamiltonian for the PCR3BP) and the dynamical consequences of such an orbit.

The Planar Circular Restricted Three-Body Problem.   The comets of interest (such as *Oterma* and *Gehrels 3*) are mostly heliocentric, but the perturbations of their motion away from Keplerian ellipses are dominated by Jupiter's gravitation. Moreover, their motion is very nearly in Jupiter's orbital plane, and Jupiter's small eccentricity (0.0483) plays little role during the fast resonance transition (which is less than or equal to one Jupiter period in duration). The PCR3BP is therefore an adequate starting model for illuminating the essence of the resonance transition process. However, for a more refined study, especially
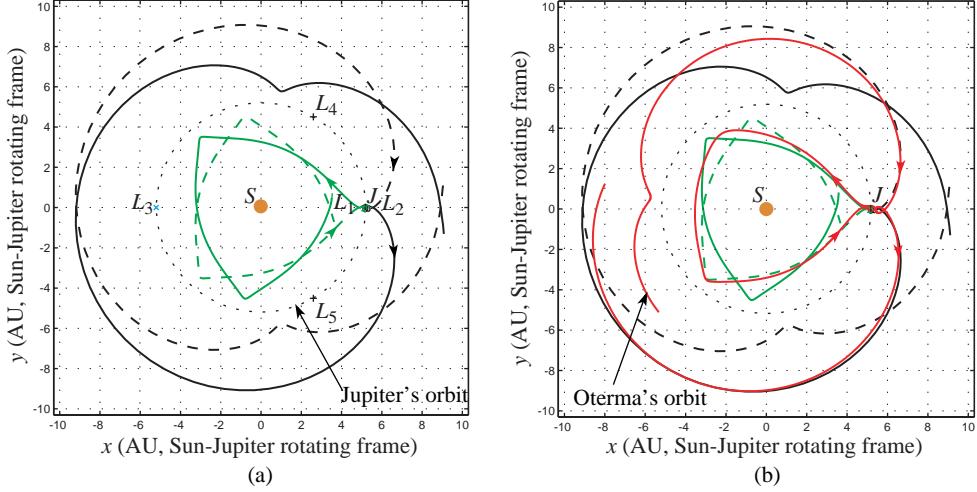
Figure 1.1: (a) Stable (dashed curves) and unstable (solid curves) manifolds of L_1 and L_2 projected to position space in the Sun-Jupiter rotating frame. The L_1 manifolds are green, while the L_2 manifolds are black. (b) The orbit of comet *Oterma* (AD 1915–1980) in the Sun-Jupiter barycentered rotating frame (red) follows closely the invariant manifolds of L_1 and L_2. Distances are in Astronomical Units (AU).

for the cases where the comets have high inclination and are not dominated solely by Jupiter, other models are needed. For additional details, see §6.

Framework of the Paper. The point of view developed in this paper is based on the premise that the invariant manifold structures associated with L_1 and L_2 periodic orbits and the heteroclinic connections are fundamental tools that will further the understanding of the natural transport of material throughout the solar system.

In tackling this problem, we have drawn upon some work of the Barcelona group on the PCR3BP, in particular, Llibre, Martinez and Simó [1985], hereafter denoted LMS. We have also drawn heavily on works of Moser, Conley and McGehee on the same subject. Specific citations are given later.

1.2 Heteroclinic Connections and their Consequences.

Heteroclinic Connection. One of the main new technical results of this paper is the numerical demonstration of a *heteroclinic connection* between a pair of periodic orbits, one around the libration point L_1 and the other around L_2. This heteroclinic connection augments the homoclinic orbits associated with the L_1 and L_2 periodic orbits, which were previously known to exist. By linking these heteroclinic connections and homoclinic orbits, we have found the *dynamical chains* which form the backbone for temporary capture and rapid resonance transition of Jupiter comets. See Figure 1.2.

Existence of Transition Orbits. We have proved the existence of a large class of interesting orbits near a chain which a comet can follow in its rapid transition between the inside and outside of Jupiter's orbit via a Jupiter encounter. The collection of these orbits is called a **dynamical channel**. We also use this term when collections of such chains for separate three body systems, roughly speaking, overlap and are put end to end. We

4
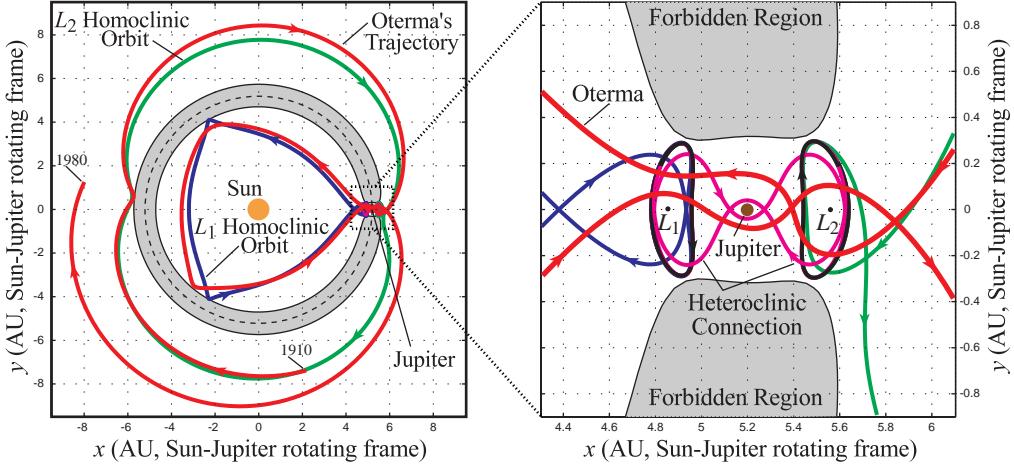
Figure 1.2: A dynamical channel (homoclinic-heteroclinic chain) corresponding to the Jupiter comet *Oterma*. The periodic orbits about $L_1$ and L_2 are black. Their homoclinic orbits are blue and green. The heteroclinic connection between them is magenta. The actual orbit of *Oterma* (AD 1910–1980) is shown in red overlaying the chain. Distances are in Astronomical Units (AU).

can individually label the orbits in a chain with an itinerary giving their past and future whereabouts, making their classification and manipulation possible.

Numerical Construction of Orbits. We not only prove the existence of orbits with prescribed itineraries, but develop a systematic procedure for their numerical construction. This is an important part of the program; it turns a general existence theory into a practical technique for constructing orbits.

Applications to Space Mission Design. The systematic procedures developed here could be used to design spacecraft orbits which explore a large region of space in the vicinity of the Earth (and near Earth's orbit) using low-fuel controls. Behavior related to the dynamical channels has already been observed in the trajectory for NASA's *Genesis Discovery Mission*, which exhibits near-heteroclinic motion between L_1 and L_2 in the Sun-Earth system (Lo, Williams, *et al.* [1998]). Having a better understanding of the underlying homoclinic-heteroclinic structures should allow us to construct and control spacecraft trajectories with desired characteristics (e.g., transfer between L_1 and L_2 orbits, explore the region interior to Earth's orbit and then return to Earth's vicinity).

To give a specific illustration, these techniques can be used to construct a "Petit Grand Tour" of the moons of Jupiter. We can design an orbit which follows a prescribed itinerary in its visit to the many moons (e.g., one orbit around Ganymede, four around Europa, etc.). See Figure 1.3, where we show a preliminary example.

1.3 A Few Key Features of the Three Body Problem.

The Planar Circular Restricted Three-Body Problem. The equations of motion for the PCR3BP will be recalled below, but here we recall a few key features. Two of the bodies, which we call generically the **Sun** and **Jupiter**, have a total mass that is normalized to one. Their masses are denoted, as usual, by $m_S = 1 - \mu$ and $m_J = \mu$ respectively (see Figure 1.4). These bodies rotate in the plane counterclockwise about their common center
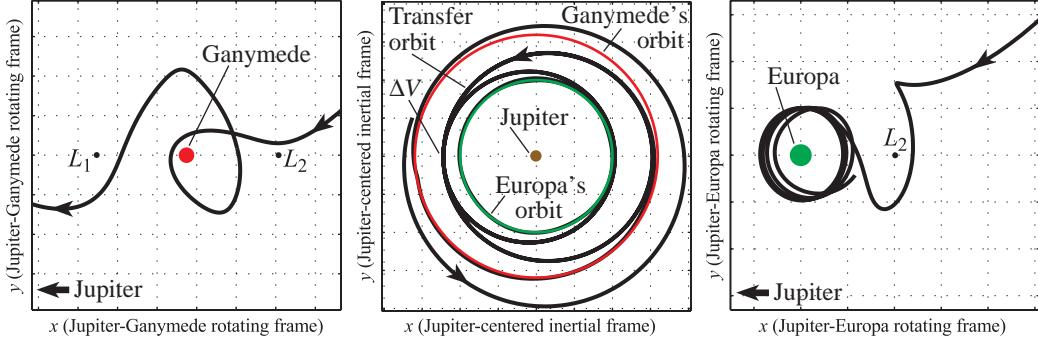
Figure 1.3: The "Petit Grand Tour" space mission concept for the Jovian moons. In our example, we show an orbit coming into the Jupiter system and (a) performing one loop around Ganymede (shown in the Jupiter-Ganymede rotating frame), (b) transferring from Ganymede to Europa using a single impulsive maneuver (shown in the Jupiter-centered inertial frame), and (c) getting captured by Europa (shown in the Jupiter-Europa rotating frame).

of mass and with the angular velocity normalized to one. The third body, which we call the **comet** or the **spacecraft**, has mass zero and is free to move in the plane.

Choose a rotating coordinate system so that the origin is at the center of mass and the Sun (S) and Jupiter (J) are fixed at $(-\mu, 0)$ and $(1-\mu, 0)$ respectively. Then the equations of motion of the comet are an autonomous Hamiltonian system of differential equations with two degrees of freedom. The system has a first integral called the **Jacobi integral** (also called the Jacobi constant), which is a multiple of the Hamiltonian. Following the conventions of the literature, we shall take

$$\text{Jacobi Constant} = -2 \, \times \, \text{Hamiltonian.}$$

Equilibrium Points and Hill's Regions. The system has three unstable collinear equilibrium points on the Sun-Jupiter line, called L_1, L_2 and L_3, whose eigenvalues include one real and one imaginary pair. The level surfaces of the Jacobi constant (which are also energy surfaces) are invariant three-dimensional manifolds. Our main concern here is the behavior of the orbits whose Jacobi constant is just below that of L_2. Recall that the **Hill's region** is the projection of this region defined by the Jacobi integral onto position space. For this case, the Hill's region contains a "neck" about L_1 and L_2, as shown in Figure 1.5(a). Thus, orbits with a Jacobi constant just below that of L_2 are energetically permitted to make a transit through the neck region from the **interior region** (inside Jupiter's orbit) to the **exterior region** (outside Jupiter's orbit) passing through the **Jupiter region**. Part of the methodology we develop is usefully described in terms of an analogy used in Conley [1969]. While this analogy cannot replace the detailed study of the orbit structure of the PCR3BP, it does provide a helpful mental picture. Consider three bowls connected by two troughs so that, when inverted, they look like three mountains with two passes between them. The three bowls correspond to the interior, Jupiter, and exterior regions. The troughs correspond to the L_1 and L_2 equilibrium regions.

The equations of motion of the PCR3BP can be viewed as those describing the motion of a point mass sliding without friction on this "triple bowl." Since the kinetic energy is positive, fixing the value of the Hamiltonian function corresponds to limiting the height to which the mass can go. Our problem corresponds to the case where the mass can go high enough to get from one bowl to the other two with just a little room to spare in the trough.
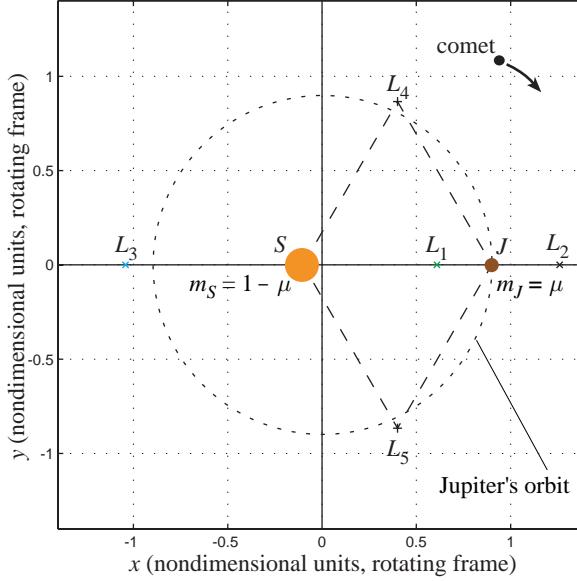
Figure 1.4: Equilibrium points of the planar circular restricted three-body problem as viewed, not in any inertial frame, but in the rotating frame, where the Sun and Jupiter are at fixed positions along the x-axis.

The Flow near the Lagrange Points L_1 and L_2. Having fixed on an appropriate energy level surface, we first study the behavior of orbits near the equilibrium points (see Figure 1.5) which, in the example above, correspond to the saddle points in the troughs connecting the bowls. In §2, we collect the major results on the flow near the equilibrium points L_1 and L_2 from Conley [1968, 1969] and McGehee [1969], both to set notation and for the convenience of the reader. This local study is performed using the linearized system of the PCR3BP. With the aid of a theorem of Moser, all the qualitative results of this linearized system carry over to the full nonlinear equations.

Pieces of stable and unstable manifolds of periodic orbits about L_1 and L_2, made up of asymptotic orbits, separate two types of motion: transit orbits and non-transit orbits. These manifolds play a gate-keeping role for resonance transition. Orbits inside the tubes of these manifolds transit from one region to another. Those outside the tubes bounce back. This observation will be used later in the numerical construction of orbits in §4.

1.4 Outline of the Paper and Summary of the Results.

Transit Orbits. The main result of §2 is that besides the existence of an unstable *periodic* solution called a *Lyapunov* orbit near each equilibrium point, there are also **transit, asymptotic** and **non-transit** solutions. The latter orbits are defined according to whether they make a transit from one region to the other, wind to or from the periodic solution, or come out of one region and pass near the critical point only to fall back into the same region. See Figure 1.5(b).

Homoclinic Orbits and Heteroclinic Connections. In §3 and §4, we make use of the local classification of orbits from §2 to define global classes of orbits in terms of their ultimate behavior with respect to the equilibrium points. As dynamical systems theory suggests, to understand the global dynamics of the flow, one should examine structures

7
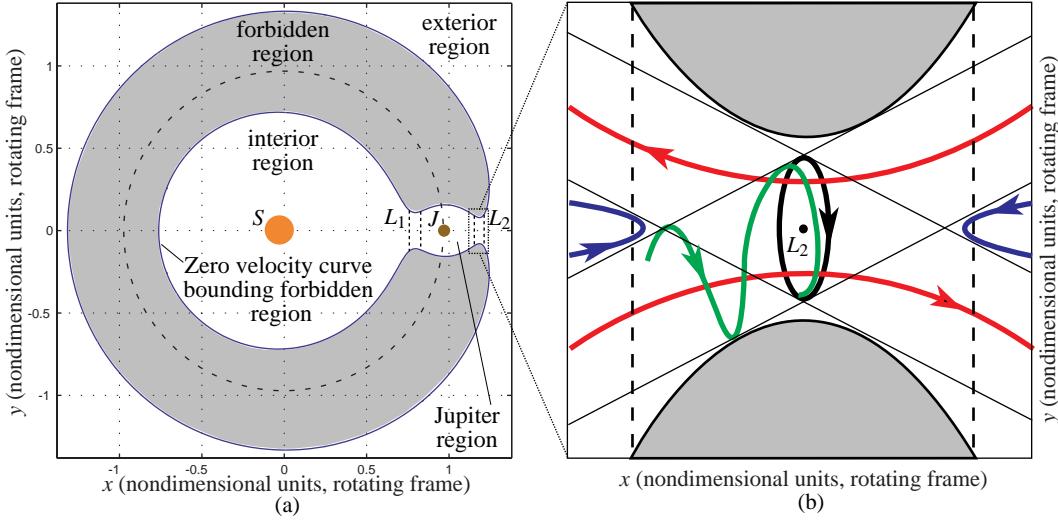
Figure 1.5: (a) Hill's region (schematic, the region in white), which contains a "neck" about L_1 and L_2. (b) The flow in the region near L_2, showing a periodic orbit (black ellipse), a typical asymptotic orbit (green), two transit orbits (red) and two non-transit orbits (blue). A similar figure holds for the region around L_1.

such as homoclinic orbits and heteroclinic connections (see, for example, Moser [1973]).

In this vein, we recall in §3 some results in McGehee [1969], which proved the existence of *homoclinic orbits* in both the interior and exterior regions, which are doubly asymptotic to L_1 and L_2 Lyapunov orbits, respectively.

Then we use semi-analytical methods to show the existence of *heteroclinic connections* in the Jupiter region which asymptotically connect the L_1 and L_2 Lyapunov orbits. Moreover, we also show that with appropriate Jacobi constants, there exist *chains of transversal homoclinic and heteroclinic orbits* (see Figure 1.2). These chains will be used in §4 to organize the distinctively different types of global motions. We use a semi-analytical method by combining symbolic and numerical techniques, which is guided by careful analytical, geometrical and dynamical aspects of the problem.

Global Orbit Structure of the PCR3BP. In §4, we use the chains of homoclinic and heteroclinic orbits to construct a suitable Poincaré map in the neighborhood of the chain which allows 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. We prove a theorem which gives the global orbit structure in the neighborhood of a chain. In simplified form, the theorem essentially says:

For any admissible bi-infinite sequence $(\ldots, u_{-1}; u_0, u_1, u_2, \ldots)$ *of symbols* $\{S, J, X\}$ *where* S, J, *and* X *stand for the interior (Sun), Jupiter, and exterior regions respectively, there corresponds an orbit near the chain whose past and future whereabouts with respect to these three regions match those of the given sequence.*

For example, given the bi-infinite sequence, or itinerary, $(\ldots, S; J, X, J, \ldots)$, there exists an orbit starting in the Jupiter region which came from the interior region and is going to the exterior region and returning to the Jupiter region.

We can then classify the orbits which correspond to qualitatively different varieties of global motions. For example, "oscillating" orbits are (roughly) those which cross from one

8

region to the others infinitely many times; "capture" orbits are those which cross for some amount of time but eventually stay in one region; and asymptotic orbits are those which eventually wind onto the periodic solution. Orbits which exhibit none of these behaviors stay in one region for all time and are called non-transit.

We not only prove the existence of orbits with prescribed itineraries, but develop a systematic procedure for their numerical construction. By following successive intersections of stable and unstable invariant manifolds of L_1 and L_2 Lyapunov orbits with a Poincaré section, we can generate regions of orbits with itineraries of arbitrary length.

Resonance Transition.   In §5, we focus on a limited case of the fast dynamical channel transport mechanism developed in previous sections; the case of transition between resonances. In particular, we study how the invariant manifolds and their heteroclinic intersections connect the mean motion resonances of the interior and exterior regions (e.g., the 3:2 and 2:3 Jupiter resonances) via the Jupiter region.

By numerical exploration of the heteroclinic connection between the interior and exterior resonances, we obtain a better picture of the resonance transition of actual Jupiter comets. As our example, we explain the sense in which Jupiter comet *Oterma* transitions between the 3:2 and 2:3 resonances. We discover much about the mixed phase space structure, especially the mean motion resonance structure, of the PCR3BP.

Conclusion and Future Work.   In the conclusion, we make several additional remarks as well as point out some possible directions for future work, such as extensions to three dimensions, many body problems, merging with optimal control, and the transport and distribution of asteroids, comets and Kuiper-belt objects in the solar system.

2   The Flow near the Libration Points L_1 and L_2.

In this section we study the behavior of orbits near the two libration points[1] L_1 and L_2 and particularly those orbits whose Jacobi constant C is just below that of the critical point L_2, that is, $C < C_2$. The Hill's region corresponding to such values of the Jacobi constant contains a "neck" about each libration point; thus, in the case of the Lagrange point L_1 between the two primary masses S and J, orbits on the integral surface can make a transit (through the neck) from the vicinity of one mass point to the other. The aim here is to describe how orbits in the "neck" look. A similar study can be done for the other libration point L_2. Correspondingly, in this section, we shall use L to denote either L_1 or L_2. We will also adopt the convention of using script letters to refer to regions on the energy surface and italicized letters for that same region's projection onto position space. For instance, the equilibrium region $\mathcal{R}$ on the energy surface (the "neck" for either L_1 or L_2) has the position space projection R.

To obtain a good idea of the orbit structure in the "neck" region $\mathcal{R}$, it is sufficient to discuss the equations of motion linearized near the critical point. Indeed, 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.*

2.1   The Planar Circular Restricted Three-Body Problem.

We begin by recalling the equations for the planar circular restricted three-body problem (PCR3BP). See, for example, Abraham and Marsden [1978] or Meyer and Hall [1992] for

---

[1]These points were discovered by Euler before Lagrange discovered the Lagrange points, L_4 and L_5, but it is common to call L_1 and L_2 the Lagrange points despite being historically inaccurate.

more information. As mentioned previously, the two main bodies are called generically the Sun and Jupiter, and have masses denoted $m_S = 1 - \mu$ and $m_J = \mu$. They rotate in the plane in circles counterclockwise about their common center of mass and with angular velocity normalized as one. The third body, which we call the comet or the spacecraft, has mass zero and is free to move in the plane. Choose a rotating coordinate system so that the origin is at the center of mass and the Sun and Jupiter are fixed on the x-axis at $(-\mu, 0)$ and $(1 - \mu, 0)$ respectively (see Figure 1.4). Let (x, y) be the position of the comet in the plane (so these are the position coordinates relative to the positions of the Sun and Jupiter, not relative to an inertial frame).

Methods of Derivation. There are several ways to derive and model the Hamiltonian structure for this system, as discussed at length in the above references. For example, as in Whittaker's book, Abraham and Marsden [1978] use time dependent canonical transformation theory to transform the problem from an inertial frame to a rotating frame. This reference also discusses the Delaunay and the Poincaré models. A simpler technique is to use covariance of the Lagrangian formulation and use the Lagrangian directly in a moving frame (see Marsden and Ratiu [1999]). This method directly gives the equations in Lagrangian form and the associated Hamiltonian form is given by the Legendre transformation.

The Planar Circular Restricted Three-Body Problem Model (PCR3BP). After going through the aforementioned procedure, one finds that the new Hamiltonian function is given by

$$H = \frac{(p_x + y)^2 + (p_y - x)^2}{2} - \frac{x^2 + y^2}{2} - \frac{1 - \mu}{r_1} - \frac{\mu}{r_2} - \frac{\mu(1 - \mu)}{2}, \tag{2.1}$$

where

$$r_1 = \sqrt{(x + \mu)^2 + y^2} \quad \text{and} \quad r_2 = \sqrt{(x - 1 + \mu)^2 + y^2}.$$

The relationship between the momenta and the velocities are a result of either the Legendre transformation (if one is taking a Lagrangian view) or of Hamilton's equations:

$$\dot{x} = \frac{\partial H}{\partial p_x} = p_x + y; \qquad \dot{y} = \frac{\partial H}{\partial p_y} = p_y - x. \tag{2.2}$$

The remaining dynamical equations are

$$\dot{p}_x = -\frac{\partial H}{\partial x} = p_y - x + \Omega_x; \qquad \dot{p}_y = -\frac{\partial H}{\partial y} = -p_x - y + \Omega_y, \tag{2.3}$$

where

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

and where Ω_x, Ω_y are the partial derivatives of Ω with respect to the variables x, y.

On the Lagrangian side we write the equations in terms of the velocities; that is, we make the transformation: $\dot{x} = p_x + y, \dot{y} = p_y - x$, where $\dot{x}, \dot{y}$ correspond to the velocity in the rotating coordinate system. Then the equations can be rewritten in second order form as

$$\ddot{x} - 2\dot{y} = \Omega_x, \qquad \ddot{y} + 2\dot{x} = \Omega_y. \tag{2.4}$$

10

This form of the equations of motion has been studied in detail in Szebehely [1967] and may be more familiar to the astronomy and astrodynamics communities. Equations (2.4) are called the equations of the planar circular restricted three-body problem (PCR3BP). They have a first integral called the **_Jacobi integral_**, which is given by

$$C(x, y, \dot{x}, \dot{y}) = -(\dot{x}^2 + \dot{y}^2) + 2\Omega(x, y) = -2E(x, y, \dot{x}, \dot{y}). \tag{2.5}$$

We shall use E when we regard the Hamiltonian (which is not the kinetic plus potential energy) as a function of the positions and velocities and H when we regard it as a function of the positions and momenta.

Equilibrium Points. The system (2.4) has five equilibrium points, three collinear ones on the x-axis, called L_1, L_2, L_3 and two equilateral points called L_4, L_5 (see Figure 1.4). These equilibrium points are critical points of the (effective potential) function Ω. The value of the Jacobi integral at the point L_i will be denoted by C_i.

2.2   Linearization near the Collinear Equilibria.

Studying the linearization of the dynamics near the equilibria is of course an essential ingredient for understanding the more complete nonlinear dynamics.

To find the linearized equations around the collinear Lagrange point L with coordinates $(k, 0)$, we need the quadratic terms of the Hamiltonian H in equation (2.1) as expanded about $(k, 0)$. After making a coordinate change with $(k, 0)$ as the origin, these quadratic terms form the Hamiltonian function for the linearized equations, which we shall call H_l

$$H_l = \frac{1}{2}\{(p_x + y)^2 + (p_y - x)^2 - ax^2 + by^2\}, \tag{2.6}$$

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

$$\rho = \mu|k - 1 + \mu|^{-3} + (1 - \mu)|k + \mu|^{-3}.$$

A short computation gives the linearized equations in the form

$$\dot{x} = \frac{\partial H_l}{\partial p_x} = p_x + y, \qquad \dot{p}_x = -\frac{\partial H_l}{\partial x} = p_y - x + ax,$$

$$\dot{y} = \frac{\partial H_l}{\partial p_y} = p_y - x, \qquad \dot{p}_y = -\frac{\partial H_l}{\partial y} = -p_x - y - by, \tag{2.7}$$

To make the computations easier and to give the variables simpler geometric meaning, let us introduce the transformation: $v_x = p_x + y, v_y = p_y - x$, where v_x, v_y correspond to velocity in the rotating coordinate system. The transformed equations are then given by

$$\dot{x} = v_x, \qquad \dot{v}_x = 2v_y + ax,$$

$$\dot{y} = v_y, \qquad \dot{v}_y = -2v_x - by, \tag{2.8}$$

which is the linearization of the equations (2.4) around the equilibrium point.

The integral H_l of (2.6) now appears as

$$E_l = \frac{1}{2}(v_x^2 + v_y^2 - ax^2 + by^2), \tag{2.9}$$

which corresponds to the energy integral (E of (2.5)) of the restricted problem. Notice that the zero-surface of the integral E_l corresponds to the Jacobi integral surface which passes

through the libration point. We shall therefore study solutions of equations (2.8) on the surface $E_l = \mathcal{E} > 0$ which corresponds to the case where the Hill's region contains a neck about the libration point.

We remark that this derivation is good for any of the three collinear libration points, though the value of ρ will not be the same for each point. With a mass ratio like that of Jupiter to the Sun, where $\mu = 0.0009537$, the values of a and b are approximately

$$a = 9.892, b = 3.446 \quad \text{for} \quad L_1 \quad \text{and} \quad a = 8.246, b = 2.623 \quad \text{for} \quad L_2,$$

respectively.

2.3 The Geometry of Solutions near the Libration Point.

Now we analyze the linearized equations (2.8). It is straightforward to find that the eigenvalues of this linear system have the form $\pm\lambda$ and $\pm i\nu$, where λ and ν 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 σ and τ are constants. 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

$$\dot{\xi} = \lambda\xi, \qquad \dot{\zeta_1} = \nu\zeta_2,$$
$$\dot{\eta} = -\lambda\eta, \qquad \dot{\zeta_2} = -\nu\zeta_1, \tag{2.10}$$

and the energy function (2.9) becomes

$$E_l = \lambda\xi\eta + \frac{\nu}{2}(\zeta_1^2 + \zeta_2^2). \tag{2.11}$$

Solutions of the equations (2.10) can be conveniently written as

$$\xi(t) = \xi^0 e^{\lambda t}, \qquad \eta(t) = \eta^0 e^{-\lambda t},$$
$$\zeta(t) = \zeta_1(t) + i\zeta_2(t) = \zeta^0 e^{-i\nu t}, \tag{2.12}$$

where the constants ξ^0, η^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 (2.11); namely, the functions $\eta\xi$ and $|\zeta|^2 = \zeta_1^2 + \zeta_2^2$ are both constant along solutions.

The Phase Space. For positive $\mathcal{E}$ and c, the region $\mathcal{R}$, which is determined by

$$E_l = \mathcal{E}, \quad \text{and} \quad |\eta - \xi| \le c, \tag{2.13}$$

is homeomorphic to the product of a two-sphere and an interval; namely, for each fixed value of $\eta - \xi$ between $-c$ and c, we see that the equation $E_l = \mathcal{E}$ determines the two-sphere

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

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 2.1). 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.

The Flow in $\mathcal{R}$. To analyze the flow in $\mathcal{R}$ one simply considers the projections on the (η, ξ)-plane and ζ-plane, respectively. In the first case we see the standard picture of an unstable critical point, and in the second, of a center. Figure 2.1 schematically illustrates the flow in the (η, ξ)-plane. The coordinate axes have been tilted by $45°$ in order to correspond to the direction of the flow in later figures. With regard to the first projection we see that $\mathcal{R}$ itself projects to a set bounded on two sides by the hyperbola $\eta\xi = \mathcal{E}/\lambda$ (corresponding to $|\zeta|^2 = 0$, see (2.11)) and on two other sides by the line segments $\eta - \xi = \pm c$, which correspond to the bounding spheres.

Since $\eta\xi$ is an integral of the equations in $\mathcal{R}$, the projections of orbits in the (η, ξ)-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 (2.12) shows that the orbits move as indicated by the arrows in Figure 2.1.
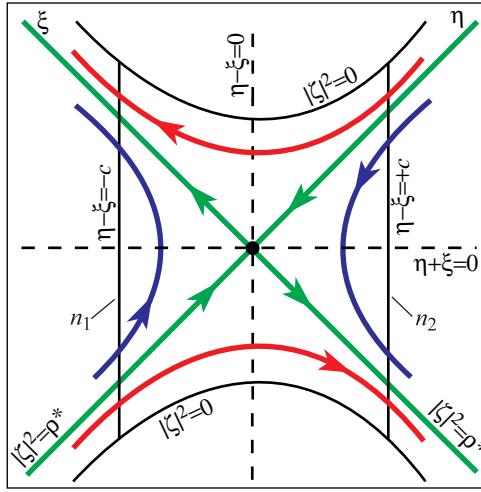
Figure 2.1: The projection onto the (η, ξ)-plane of orbits near the equilibrium point (note, axes tilted $45°$). Shown are the periodic orbit (black dot at the center), the asymptotic orbits (green), two transit orbits (red) and two non-transit orbits (blue).

To interpret Figure 2.1 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 (2.11) that $|\zeta|^2 = \frac{2}{\nu}(\mathcal{E} - \lambda\eta\xi)$. Of course, for points on the bounding hyperbolic segments ($\eta\xi = \mathcal{E}/\lambda$), 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 $\xi = \eta = 0$ corresponds to a ***periodic*** orbit in $\mathcal{R}$ (the Lyapunov orbit). See black dot at center of Figure 2.1.

2. The four half open segments on the axes, $\eta\xi = 0$ (or equivalently $|\zeta|^2 = \rho^*$ where $\rho^* = 2\mathcal{E}/\nu$), correspond to four cylinders of orbits asymptotic to this periodic solution either as time increases ($\xi = 0$) or as time decreases ($\eta = 0$). These are called ***asymptotic*** orbits. See the four green orbits of Figure 2.1.

3. The hyperbolic segments determined by $\eta\xi = \text{constant} > 0$ (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. See the two red orbits of Figure 2.1.

4. Finally the hyperbolic segments determined by $\eta\xi = \text{constant} < 0$ ($|\zeta|^2 > \rho^*$) 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 blue orbits of Figure 2.1.

McGehee Representation. 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 2.2(b).
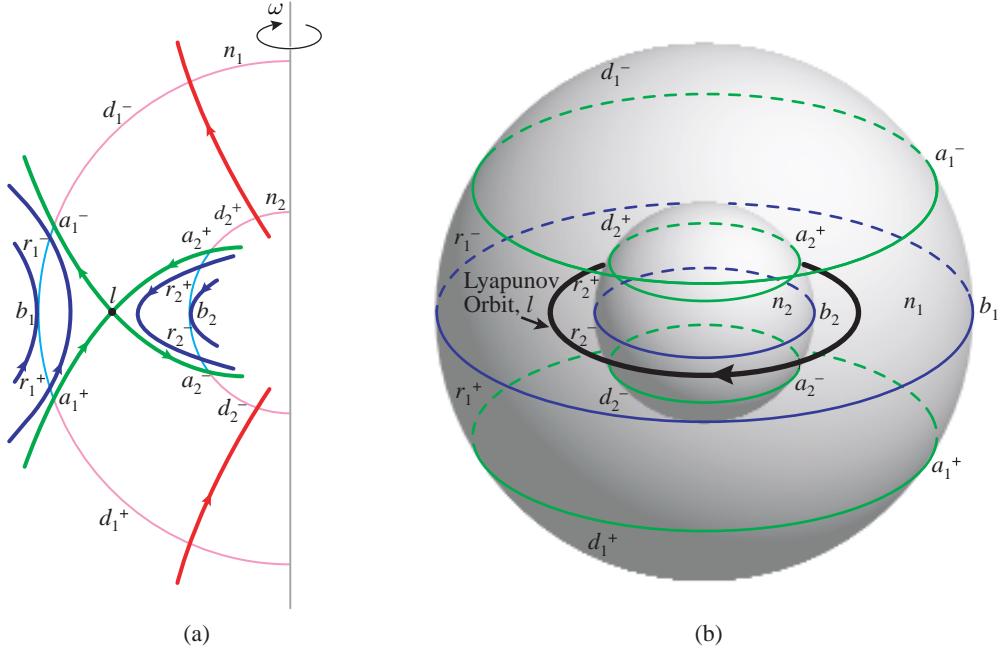
Figure 2.2: (a) The cross-section of the flow in the $\mathcal{R}$ region of the energy surface. (b) The McGehee representation of the flow in the region $\mathcal{R}$.

Figure 2.2(a) is a cross-section of $\mathcal{R}$. Notice that this cross-section is qualitatively the same as the illustration in Figure 2.1. The full picture (Figure 2.2(b)) is obtained by rotating this cross-section, about the indicated axis ω. The following classifications of orbits correspond to the previous four categories:

1. There is an unstable *periodic* orbit l in the region $\mathcal{R}$ corresponding to the point q.

14

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. Then 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^+, 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.

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 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 §4.

2.4 The Flow Mappings in the Equilibrium Region of the Energy Surface.

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 $r^\pm$:

$$\psi_1 : d_1^+ \to d_2^-, \qquad \psi_2 : d_2^+ \to d_1^-, \qquad (2.14)$$
$$\psi_3 : r_1^+ \to r_1^-, \qquad \psi_4 : r_2^+ \to r_2^-. \qquad (2.15)$$

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 (2.12) that

$$\frac{d}{dt} \ \text{arg} \ \zeta = -\nu, \qquad (2.16)$$

we see that the change in the argument of ζ for each of these mappings ψ_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.

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 2.2(b).

In §4, we shall need a simple geometric consequence of the above observation on spiraling stated in terms of "abutting arcs" in the domain, or range of ψ_i. Namely, an arc lying in
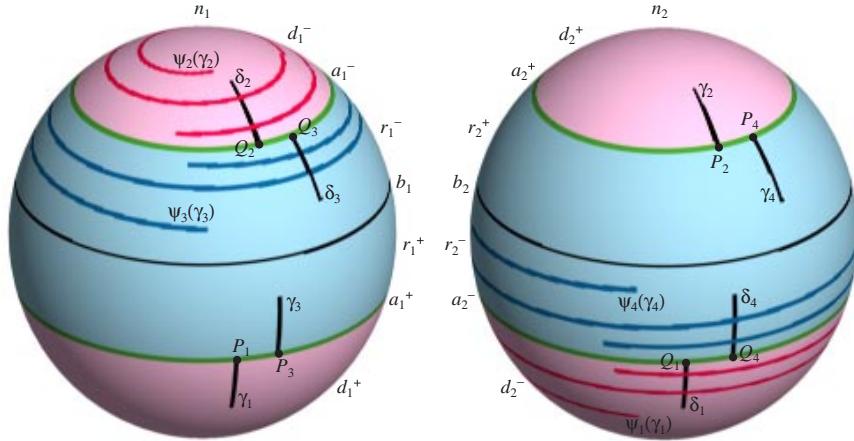
Figure 2.3: Spiraling of the images of arcs γ_i.

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 2.3. For example, let γ_1 be an abutting arc in the domain d_1^+ of ψ_1 with one end point P_1 in a_1^+. Let δ_1 be another abutting arc in the range d_2^- of ψ_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 δ_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 ψ_1; namely the image of γ_1 spirals infinitely many times around and down to a_2^- in the range.

Similarly, let γ_i be an abutting arc in the domain of ψ_i with one end point P_i in a_2^+, a_1^+, a_2^+ for $i = 2, 3, 4$, respectively. Let δ_i be another abutting arc in the range of ψ_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 δ_i an infinite number of times in any neighborhood of the point of abutment Q_i.

2.5   Orbits in the Equilibrium Region of Position Space.

After studying the orbit structure in the equilibrium region $\mathcal{R}$ and its projection on the (η, ξ)-plane, we now examine briefly the appearance of orbits in position space, that is, in the (x, y)-plane.

Recall from §2.3 that the ξ and η 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

16

the main results of Conley [1968].

Recall from §2.3 that the eigenvalues of the linear system (2.8) 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

$$v(t) = (x(t), y(t), \dot{x}(t), \dot{y}(t)) = \alpha_1 e^{\lambda t} u_1 + \alpha_2 e^{-\lambda t} u_2 + 2\text{Re}(\beta e^{i\nu t} w_1), \qquad (2.17)$$

where α_1, α_2 are real and $\beta = \beta_1 + i\beta_2$ is complex. Notice that (2.17), while slightly more complicated, is essentially the same as (2.12).

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

$$x(t) = \alpha_1 e^{\lambda t} + \alpha_2 e^{-\lambda t} + 2(\beta_1 \cos \nu t - \beta_2 \sin \nu t). \qquad (2.18)$$

Thus, if $t \to +\infty$, then $x(t)$ is dominated by its α_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 2.4. The same statement holds if $t \to -\infty$ and α_2 replaces α_1. Different combinations of the signs of α_1 and α_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 projects onto the (x, y)-plane as an ellipse with major axis of length $2\tau\sqrt{\mathcal{E}/\kappa}$ in the direction of the y-axis, and minor axis of length $2\sqrt{\mathcal{E}/\kappa}$ in the direction of the x-axis. The orientation of the orbit is clockwise. Here κ $(= -a + b\tau^2 + \nu^2 + \nu^2\tau^2)$ is a constant. See Figure 2.4. Note that the size of the ellipse goes to zero with $\mathcal{E}$.

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 xy-plane centering around $\bar{u}_2$ and bounded by the lines

$$y = \sigma x \pm 2\sqrt{\mathcal{E}(\sigma^2 + \tau^2)/\kappa}. \qquad (2.19)$$

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

$$y = -\sigma x \pm 2\sqrt{\mathcal{E}(\sigma^2 + \tau^2)/\kappa}. \qquad (2.20)$$

Notice that the width of the strips goes to zero with $\mathcal{E}$.

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 2.4. Since a detailed study will draw us too far afield, we simply state some of the main observations.

In Figure 2.4, S_1 and S_2 are the two strips mentioned above. Outside of each strip $S_i, i = 1, 2$, the sign of α_i is independent of the direction of the velocity. These signs 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 α's are negative, while

17
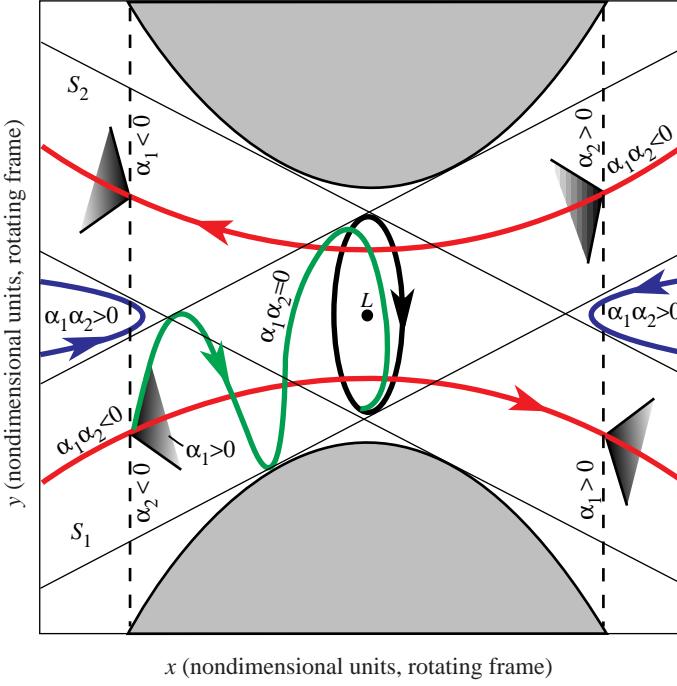
Figure 2.4: The flow in the equilibrium region R of position space. Shown are the periodic orbit (black ellipse), a typical asymptotic orbit (green), two transit orbits (red), and two non-transit orbits (blue).

in the right-most central components both α'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 α_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 2.4. 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 α_1 is positive. The sign of α_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 1.5 and 2.4): the periodic Lyapunov orbits, asymptotic orbits, transit orbits, and, finally, non-transit orbits.

3   Existence of Homoclinic Orbits and Heteroclinic Connections.

As mentioned earlier, near the equilibrium point L (i.e., L_1 or L_2), there exists a family of unstable periodic orbits called Lyapunov orbits. For appropriate values of the Jacobi

constant, the energy surface contains exactly one of these periodic solutions around each Lagrange 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 orbits and heteroclinic connections to these L_1 and L_2 Lyapunov orbits.

The local structure of orbits near the libration points gives periodic orbits (the Lyapunov orbits), pieces of the stable and unstable manifolds of these periodic orbits and transit and non-transit orbits. In this section, we explore how these local structures are connected globally. Our goal is to show how homoclinic orbits in the interior region are connected to the homoclinic orbits in the exterior region by a heteroclinic cycle in the Jupiter region. The union of these three structures is called a ***chain***.

The story is completed only in §4 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.

In more detail, this section discusses the following topics.

1. In §3.1 and §3.2, we shall first discuss some of the results in Conley [1968] and McGehee [1969], which have proven the existence of *homoclinic* orbits in both the interior and exterior regions. These are the orbits which are both forward and backward asymptotic to the 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.

2. We shall discuss in §3.3 the main results in LMS [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, LMS [1985] spent their major effort in proving analytically that the intersection is indeed transversal under appropriate conditions, at least in the interior region. We shall summarize their results.

   However, it should be clear from the start that both Theorem 3.3 and 3.4 have been cited only for guidance on how to construct the transversal homoclinic orbits numerically. In §3.4 we shall use the semi-analytical methods developed by the Barcelona group in Gómez, Jorba, Masdemont and Simó [1993] to show numerically the existence of transversal homoclinic orbits in both the interior and exterior regions.

3. In §3.5 we shall use similar semi-analytical methods to show numerically the existence of transversal *heteroclinic* connections in the Jupiter region which connect asymptotically the L_1 and L_2 Lyapunov orbits. A heteroclinic orbit is an orbit lying in the intersection of the stable manifold of one periodic orbit and the unstable manifold of another periodic orbit. Since the PCR3BP is a Hamiltonian system with two degree of freedom, its energy manifold is three dimensional. From the work of Conley, it was known that both the stable and unstable manifolds of the Lyapunov orbits 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.

   Also, in dynamical systems theory, a heteroclinic orbit generally does not provide a mechanism for a part of the phase space to eventually return near to where it

started. But two (and more) heteroclinic orbits forming a cycle may provide this mechanism and generate extremely complicated dynamics. This is indeed the case for the PCR3BP.

4. In §3.6, we shall numerically show that , within an appropriate range of Jacobi constant, there exist chains of two homoclinic orbits and a symmetric heteroclinic cycle, as in Figure 1.2. The existence of these chains will be used in §4 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.

3.1   The Flow Mappings in the Interior and Exterior Regions of the Energy Surface.

Energy Surface and Hill's Region.   We consider equations (2.4) on the energy surface given by setting the Jacobi integral (2.5) equal to a constant. Let $\mathcal{M}$ be that energy surface, i.e.,

$$\mathcal{M}(\mu, C) = \{(x, y, \dot{x}, \dot{y}) \mid C(x, y, \dot{x}, \dot{y}) = \text{constant}\} \tag{3.1}$$

The projection of this surface onto position space is called a Hill's region

$$M(\mu, C) = \{(x, y) \mid \Omega(x, y) \geq C/2\}. \tag{3.2}$$

The boundary of $M(\mu, C)$ is the zero velocity curve. The comet can move only within this region in the (x, y)-plane. For a given μ there are five basic configurations for the Hill's region, the first four of which are shown in Figure 3.1.

Case 5 is where the comet is free to move in the entire plane. In this paper, our main interest is in case 3; but for comparison we shall occasionally bring up case 2 which is the main focus of LMS [1985]. The shaded region is where the motion is forbidden. The small oval region on the right is the *Jupiter region*. The large near circular region on the left is the *interior region* surrounding the Sun. The region which lies outside the shaded forbidden region is the *exterior region* surrounding the Sun (and Jupiter).

The values of C which separate these five cases will be denoted $C_i, i = 1, 2, 3, 4$ which are the values corresponding to the equilibrium points. These values can be easily calculated for small μ and their graphs are shown in Figure 3.2. For case 3, the Jacobi constant lies between C_2 and C_3 which are the Jacobi constants of the libration points L_2 and L_3 respectively. In this case, the Hill's region contains a neck around both L_1 and L_2 and the comet can transit from the interior region to the exterior region and vice versa.

Orbit Segments Winding around a Solid Torus.   In McGehee [1969], the energy surface is broken up further into regions bounded by invariant tori. These invariant tori project onto the darkly shaded annuli shown for case 3 in Figure 3.3.

These annuli separate the Hill's region into sections corresponding to the invariant regions in the energy surface. It is interesting to note that for all of these cases 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 the Sun to Jupiter. Similarly, the two masses are 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 below show that all orbits leaving the vicinity of one of the unstable periodic orbits proceed around the annulus T_1 or T_2 before returning to that vicinity. The direction of procession is the same for all orbits,
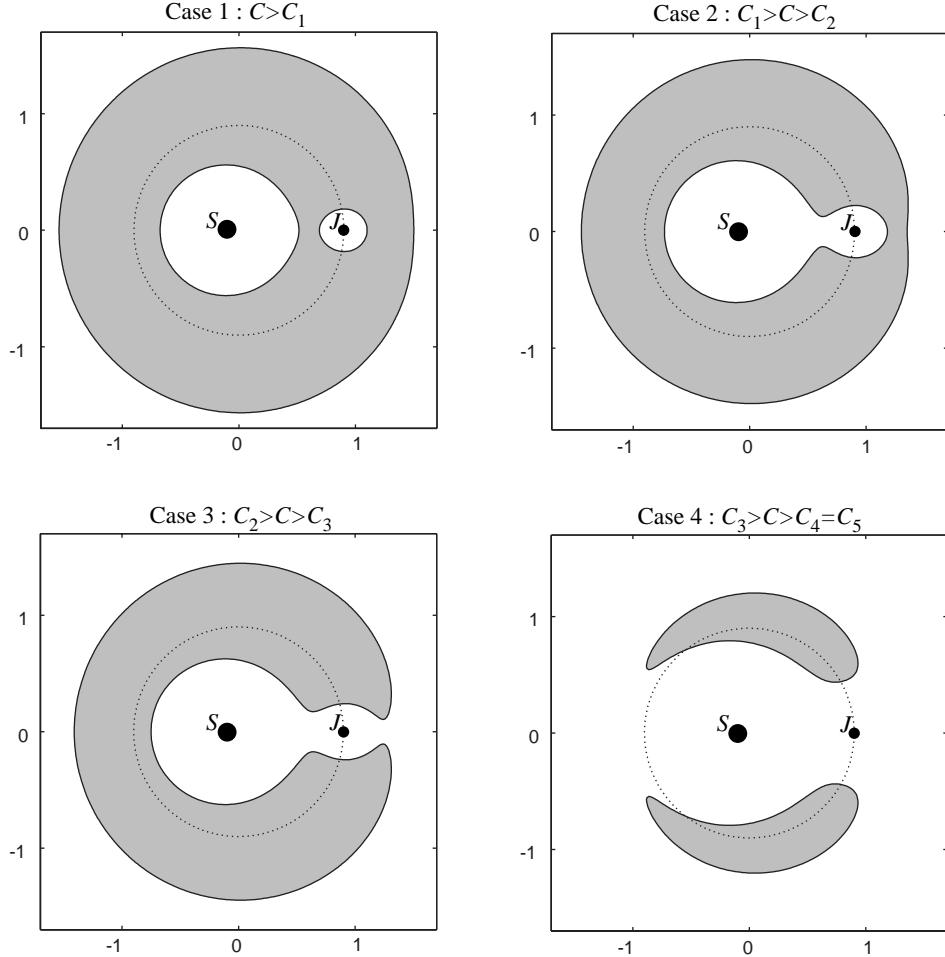
Figure 3.1: Four basic configurations of the Hill's region.

counterclockwise in the interior region and clockwise in the exterior region. In §2, we have studied the regions near the unstable periodic orbits to obtain a qualitative picture of the asymptotic orbits. We shall combine this picture of asymptotic orbits with the fact that orbits in the tori wind around in one direction *to construct homoclinic orbits* in both the interior and exterior regions. See Figure 3.3(b).

Theorems of McGehee. To precisely state the theorems, we must first divide up the Hill's region and the energy surface. We know that for small μ the two equilibrium points occur at a distance $\tilde{\mu}$ on either side of Jupiter with

$$\tilde{\mu} = \frac{2\mu^{1/3}}{3}.$$

We isolate these points by drawing vertical lines on each side of them, i.e., lines at $(1 - \mu \pm c_1\tilde{\mu}, 0)$ and $(1 - \mu \pm b_1\tilde{\mu}, 0)$, where $b_1 < 1 < c_1$. This divides the Hill's region into five sets as shown in Figure 3.4.

Let S and J be the regions that contain the Sun and Jupiter; let region R_1 and region
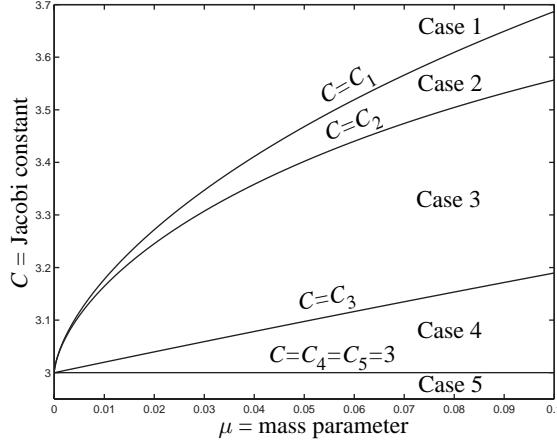
Figure 3.2: The partition of the (μ, C)-plane into five types of Hill's regions.

R_2 be those parts that contain the two equilibrium points L_1 and L_2, respectively; and let X be the region that lies exterior to the orbit of Jupiter. We also divide the energy surface $\mathcal{M}$ into sets projecting onto the regions shown in Figure 3.4. As before, we keep the same name: 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 leads to the assertion that *one can choose the division described above so that we simultaneously have sufficient control of the flow in both regions $\mathcal{S}$ and $\mathcal{R}_1$ to construct a homoclinic orbit. Theorem 3.2 makes the same assertion for regions $\mathcal{X}$ and $\mathcal{R}_2$.*

The analysis of regions $\mathcal{R}_1$ and $\mathcal{R}_2$ is of a local nature. In fact, we limit ourselves to those values of the Jacobi constant for which the linearized equations about the equilibrium point give us the qualitative picture of the flow. The flow for the linearized equations was already analyzed in some detail in §2.

We know that for b_1 and c_1 close to 1, i.e., for the region $\mathcal{R}$ close to the periodic orbit, the flow in $\mathcal{R}$ (which stands for both $\mathcal{R}_1$ and $\mathcal{R}_2$) is that shown in Figure 2.2. But we also know that we cannot make c_1 arbitrarily large without disturbing this qualitative picture for $\mathcal{R}$. 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.

Theorem 3.1 *There exist constants b_1 and c_1 and an open set O_1 in the (μ, C)-plane (see Figure 3.5) containing the graph of $C = C_1(\mu)$ for small $\mu > 0$ such that, for $(\mu, C) \in O_1$:*

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

2. *For $C < C_1(\mu)$, the flow in $\mathcal{R}_1(\mu, C)$ is qualitatively the same as the flow for the linearized equations. (See Figure 2.2)*

3. *If we let $\mathcal{T}_1$ be that submanifold of $\mathcal{M}$ co-bounded by the invariant torus and n_1 (see Figure 3.5), then there exists a function*

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

such that :

22
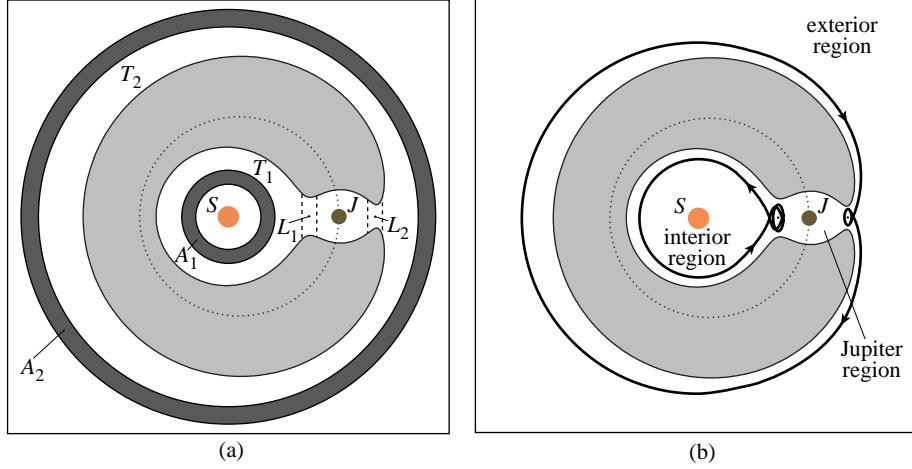
Figure 3.3: (a) The projection of invariant tori (darkly shaded) on position space for case 3. (b) Homoclinic orbits in the interior and exterior regions.
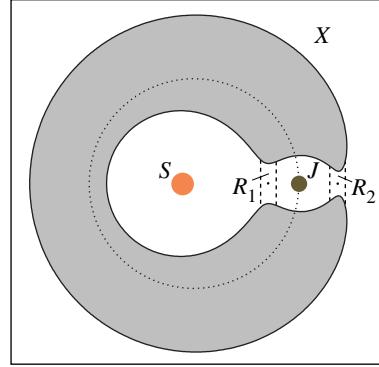
Figure 3.4: Division of Hill's region into five sets.

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

(b) θ is strictly increasing along orbits.

Theorem 3.2 *There exist constants b_1 and c_1 and an open set O_2 in the (μ, C)-plane containing the graph of $C = C_2(\mu)$ for small $\mu > 0$ such that, for $(\mu, C) \in O_2$:*

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

2. *For $C < C_2(\mu)$, the flow in $\mathcal{R}_2(\mu, C)$ is qualitatively the same as the flow for the linearized equations. (See Figure 2.2)*

3. *If we let $\mathcal{T}_2$ be that submanifold of $\mathcal{M}$ co-bounded by the invariant torus and n_2, then there exists a function*

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

such that :
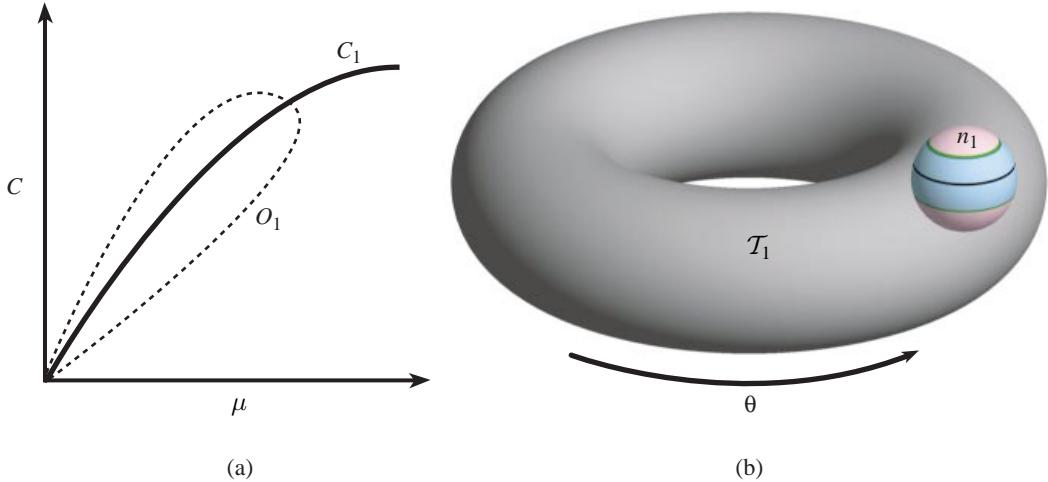
Figure 3.5: (a) Open set O_1 in (μ, C)-plane. (b) The invariant torus.

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

(b) θ is strictly increasing along orbits.

3.2 The Existence of Orbits Homoclinic to the Lyapunov Orbit.

Part 3 of the above theorems gives us the following properties for the flow in $\mathcal{T}$ where $\mathcal{T}$ stands for either $\mathcal{T}_1$ or $\mathcal{T}_2$. The increase in θ along an orbit segment in $\mathcal{T}$ with endpoints in the bounding sphere n is close to a non-zero integer multiple of 2π. The increase in θ 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 2π. Furthermore, the increase of θ along any orbit segment remaining for an arbitrarily long time in $\mathcal{T}$ 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.

A Dichotomy. We assert that *either a transverse 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 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 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. See Figure 3.6. Since $\mathcal{T}_1$ is compact and our flow, which is Hamiltonian, preserves a nondegenerate area element, we can conclude that some orbit which crosses $\mathcal{R}_1$ (and the bounding sphere n_1) and so enters $\mathcal{T}_1$ must also leave $\mathcal{T}_1$ and recross $\mathcal{R}_1$ (and n_1) the other way. See Figure 3.6. 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 that in $\mathcal{R}_1$, the spherical caps d_1^- and d_1^+ are where the flow crosses n_1.
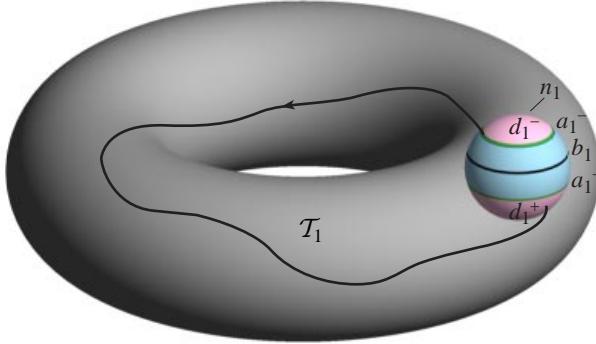
Figure 3.6: The existence of orbits homoclinic to the Lyapunov orbit.

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 γ 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 γ 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 γ' of γ which is carried by the flow to d_1^+. Let r be the first point of γ not in γ', 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.

Now, without loss of generality, we can suppose that we have found an orbit segment with one endpoint, called α, in a_1^- and the other in d^+. We now choose for γ 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 β as an endpoint becomes asymptotic at the other end. If the first possibility holds, we would have a map of d^- to the interior of d^+, contradicting area preservation of Hamiltonian flow. Thus we have proven that either transversal homoclinic orbits exist or total degeneracy occurs for the interior region. The same proof also works for the exterior region.

3.3 The Existence of Transversal Homoclinic Orbits in the Interior Region.

Conley [1968] and McGehee [1969] did not settle the issue of when one has transversality of the homoclinic orbit families for the PCR3BP. Subsequently, LMS [1985] devoted their major effort to show that under appropriate conditions, the invariant manifolds of the L_1 Lyapunov orbits do meet transversally. In this subsection, we shall summarize their analytical results. Moreover, in §3.4 we shall also use the tools of Gómez, Jorba, Masdemont and Simó [1993] to explore numerically the existence of transversal homoclinic orbits in both the interior and

25

exterior regions.

To state the major analytical results of LMS [1985], we first need to set up some notation. As mentioned earlier, near L_1 and for values of $C_1 > C > C_2$ (case 2) there is a family of unstable Lyapunov orbits. When C approaches C_1 from below, 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. In a similar way the L_1 Lyapunov orbit has two-dimensional invariant manifolds $W^s_{L_1,\text{p.o.}}, W^u_{L_1,\text{p.o.}}$, locally diffeomorphic to cylinders. We recall that a homoclinic orbit related to an equilibrium point L or to a periodic orbit $\bar{L}$ is an orbit which tends to L (or $\bar{L}$) as $t \to \pm\infty$. Therefore, it is on the stable and unstable invariant manifolds of the related object (L or $\bar{L}$). A homoclinic orbit is called transversal 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 $\mathcal{M}(\mu, C)$ at the same point.

Notice that equations (2.4) have the following symmetry

$$s : (x, y, \dot{x}, \dot{y}, t) \to (x, -y, -\dot{x}, \dot{y}, -t). \tag{3.3}$$

Therefore, if we know the unstable manifold of L_1 or 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 to find the transversal homoclinic orbits.

Analytical Results for L_1 Lyapunov Orbit in Interior Region. Using the basic framework developed in McGehee [1969], LMS [1985] were able to prove the following two analytical results. Together these two theorems imply that for sufficiently small μ and for an appropriate range of $\Delta C = C_1 - C$, 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 region $\mathcal{S}$ intersect transversally.

Theorem 3.3 *For μ sufficiently small, the branch $W^{u,\mathcal{S}}_{L_1}$ of $W^u_{L_1}$ in the interior region $\mathcal{S}$ has a projection on position space (see Figure 3.7(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, α 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:

$$\mu_k = \frac{1}{N^3 k^3}(1 + o(1)), \tag{3.4}$$

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.

Theorem 3.4 *For μ and $\Delta C = C_1 - C$ 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 region $\mathcal{S}$ of the energy surface intersects the plane $y = 0$ for $x < 0$ in a curve diffeomorphic to a circle (see Figure 3.7(b)).*

In particular, for points in the (μ, C) plane such that there is a μ_k of Theorem 3.3 for which

$$\Delta C > L\mu_k^{4/3}(\mu - \mu_k)^2 \tag{3.5}$$

holds (where L is a constant), there exist symmetric transversal (1,1)-homoclinic orbits.
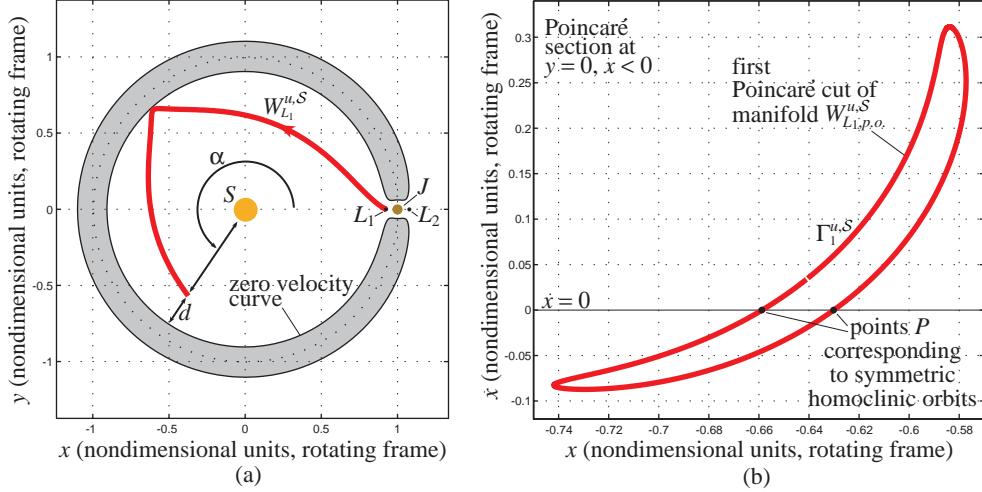
Figure 3.7: (a) Projection of the interior branch of the manifold $W_{L_1}^u$ on the position space. (b) First intersection (Poincaré "cut") $\Gamma_1^{u,\mathcal{S}}$ of the interior branch of $W_{L_1,\text{p.o.}}^u$ with the plane $y = 0$ in the region $x < 0$.

For details of the proofs, see LMS [1985]. We would like to make a few comments about these results which are pertinent to the main thrust of our paper.

1. The main objective of both theorems is to study the transversality of the invariant manifolds for the L_1 Lyapunov orbit on the energy surface whose Jacobi constant C is slightly less than $C_1(\mu)$ as one varies μ and C. The main step is to obtain an expression for the first intersection $\Gamma_1^{u,\mathcal{S}}$ of the unstable manifold $W_{L_1,\text{p.o.}}^{u,\mathcal{S}}$ with the plane $y = 0$ in the region $x < 0$. While formulas were provided in LMS [1985] for this closed curve as a function of μ and ΔC 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 3.3, the set of values of μ for which we have a symmetric (1,1)-homoclinic orbit associated to L_1 is discrete and is given by equation (3.4). Then for any other value of μ 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 ΔC is too small, $\Gamma_1^{u,\mathcal{S}}$ does not cut the x-axis and hence (by symmetry) $\Gamma_1^{s,\mathcal{S}}$ of the stable manifold $W_{L_1,\text{p.o.}}^{s,\mathcal{S}}$ 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 μ, if we increase ΔC, we hope that $\Gamma_1^{u,\mathcal{S}}$ of the unstable manifold will become large. Therefore we can look for some value of ΔC 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 reversibility of the PCR3BP, $\Gamma_1^{s,\mathcal{S}}$ 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 3.7(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 3.4 say that the above phenomenon occurs if $\Delta C > L\mu_k^{4/3}(\mu - \mu_k)^2$ holds.

2. Using the results of Theorem 3.4, LMS [1985] was able to draw the mesh of homoclinic tangencies for the $(\mu, \Delta C)$-plane. The numbers in Figure 3.8 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 μ and ΔC. For us, the key point of the theorems is that for the wide

range of μ which exist in the solar system, the invariant manifolds of the L_1 Lyapunov orbit intersect transversally for sufficiently large ΔC.
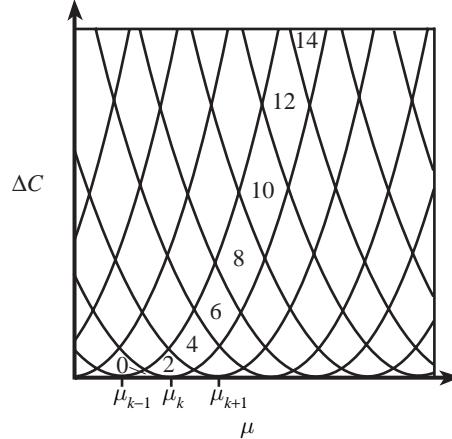

Figure 3.8: Partition of the $(\mu, \Delta C)$-plane according to the number of symmetric (1,1)-homoclinic points found in the first intersection of $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ with the plane $y = 0, x < 0$.

3. The heart of the proofs of these two theorems is to obtain expressions for $W^{u,\mathcal{S}}_{L_1}$ as a function of μ and for $W^{u,\mathcal{S}}_{L_1,\text{p.o.}}$ as a function of μ and ΔC. By using the basic framework of McGehee [1969], LMS [1985] divided the annulus T_1 in the interior region S into two parts: a small neighborhood H near R_1 and the rest of the region 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 PCR3BP in the limit when μ approaches zero. In the rest of the region 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, LMS [1985] were able to obtain these analytical results.

Summary. Conley [1968] and McGehee [1969] have proved the existence of homoclinic orbits for both the interior and exterior region, and LMS [1985] have shown analytically the existence of transversal symmetric (1,1)-homoclinic orbits in the interior region under appropriate conditions. For our problem, we need to find transversal homoclinic orbits in both interior and exterior regions as well as transversal heteroclinic cycles for the L_1 and L_2 Lyapunov orbits. We shall perform some numerical explorations using the tools developed by the Barcelona group. For more details on finding invariant manifolds numerically, see Gómez, Jorba, Masdemont and Simó [1993] and references therein.

3.4 The Existence of Transversal Homoclinic Orbits in the Exterior Region.

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 region. Though there are no analytical results proving the existence of transversal homoclinic orbits in the $\mathcal{X}$ region, 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 $\mathcal{X}$ region flows clockwise in the position space. We refer to this exterior branch of the manifold as $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$. See Figure 3.9(a). This two-dimensional manifold "tube" $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ first intersects the plane $y = 0$ on the part of T_2 which is opposite to L_2 with respect to the Sun (i.e., $x < 0$). The intersection, as one would expect geometrically, is a curve diffeomorphic to a circle. We call this intersection the first "cut" of $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ with $y = 0$. See Figure 3.9(b). Note that in order to define the first cut we
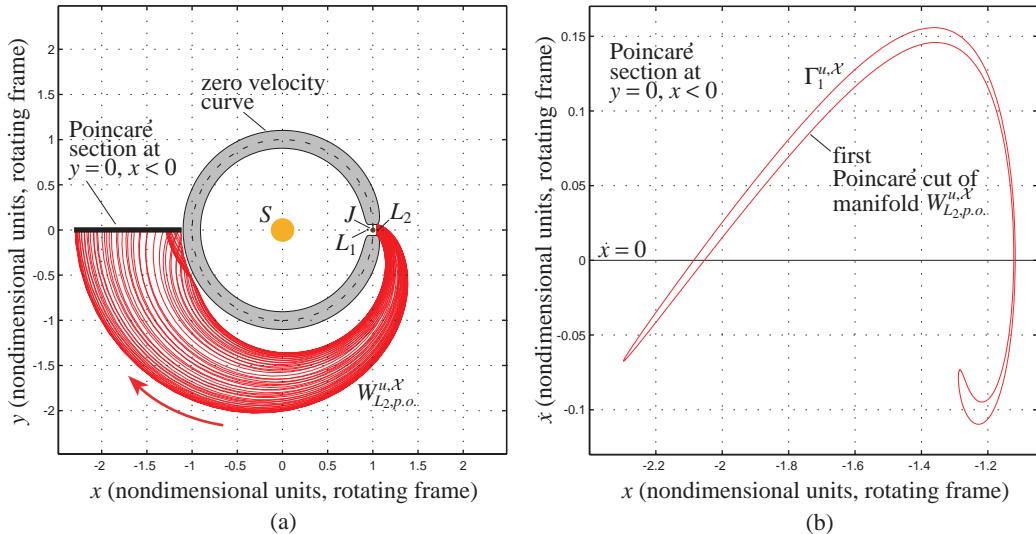

Figure 3.9: (a) The position space projection of the unstable manifold "tube" $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ until the first intersection with the Poincaré section at $y = 0, x < 0$. (b) The first Poincaré cut $\Gamma^{u,\mathcal{X}}_1$ of the manifold $W^{u,\mathcal{X}}_{L_2,\text{p.o.}}$ on the plane $y = 0, x < 0$.

exclude a neighborhood of n_2 in the $\mathcal{X}$ region. Some arcs of this curve produce successive intersections without leaving the $\mathcal{X}$ region. 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 the exterior region branch of $W^s_{L_2,\text{p.o.}}$.

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. Other intersections (for larger q and p) may exist, but we will restrict ourselves for now to the first. Suppose that the unstable manifold intersection $\Gamma^{u,\mathcal{X}}_q$ is a closed curve γ 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 reversibility of the PCR3BP, 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 γ intersects the $\dot{x} = 0$ line of the $(x, \dot{x})$-plane. Points P along the curve γ which intersect the $\dot{x} = 0$ line are (q, q)-homoclinic points, corresponding to (symmetric) orbits homoclinic to the Lyapunov orbit. If the curve γ 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 γ 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 appear.

Consider Figure 3.9(b), where we used the values $\mu = 0.0009537$ and $\Delta C = C_2 - C = .01$ to compute the unstable Poincaré cut. If we also plotted the stable cut $\Gamma_1^{s,\mathcal{X}}$, which is the mirror image of unstable cut $\Gamma_1^{u,\mathcal{X}}$, we would find several points of intersection. In Figure 3.10(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 $\mathcal{X}$ region. The transversal symmetric $(1,1)$-homoclinic orbit corresponding to the left $\dot{x} = 0$ intersection is shown in Figure 3.10(b).
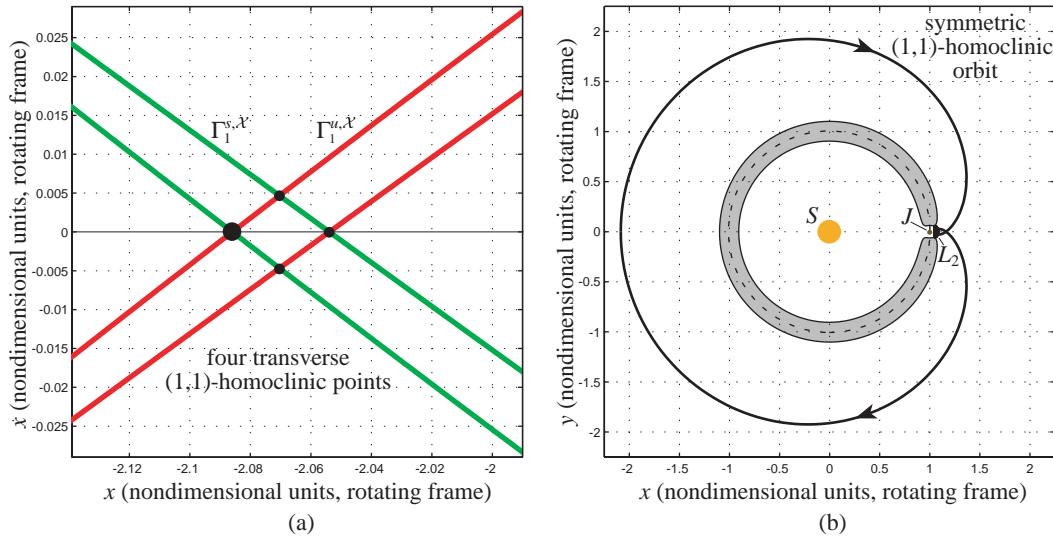
Figure 3.10: (a) A group of four transverse $(1,1)$-homoclinic points. (b) The symmetric $(1,1)$-homoclinic orbit corresponding to the left $\dot{x} = 0$ $(1,1)$-homoclinic point (the large black dot in (a)).

We also notice two off-axis intersections in Figure 3.10(a), completing the local transversal intersection of two closed loops in the $(x, \dot{x})$-plane. As these two intersections occur near the line $\dot{x} = 0$, they will be nearly symmetric. A more pronounced case of nonsymmetry occurs for the other group of intersection points centered near $x = -1.15$, for which we have the nonsymmetric $(1,1)$-homoclinic orbit given in Figure 3.11.
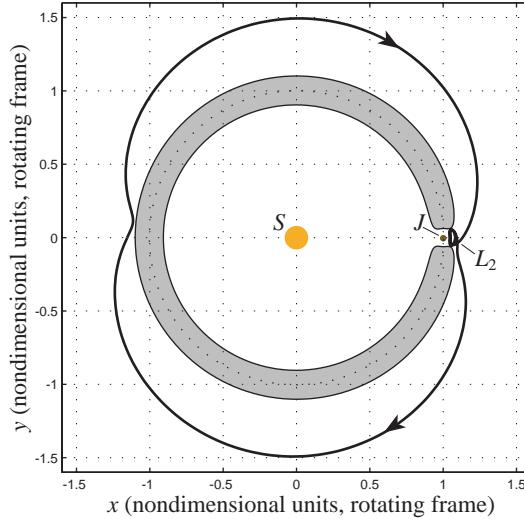
Figure 3.11: A nonsymmetric $(1, 1)$-homoclinic point.

A similar procedure can numerically produce homoclinic orbits in the interior region as well as in the Jupiter region. We can even look at cuts beyond the first. See Figure 3.12(a).

For example, in Figure 3.12(b) we show an interior region $(1, 3)$-homoclinic orbit (note, also $(2, 2)$ and $(3, 1)$, using $\bar{q} + \bar{p} = q + p$) associated to an L_1 Lyapunov orbit for $\mu = .1, \Delta C = C_1 - C = 0.0743$.
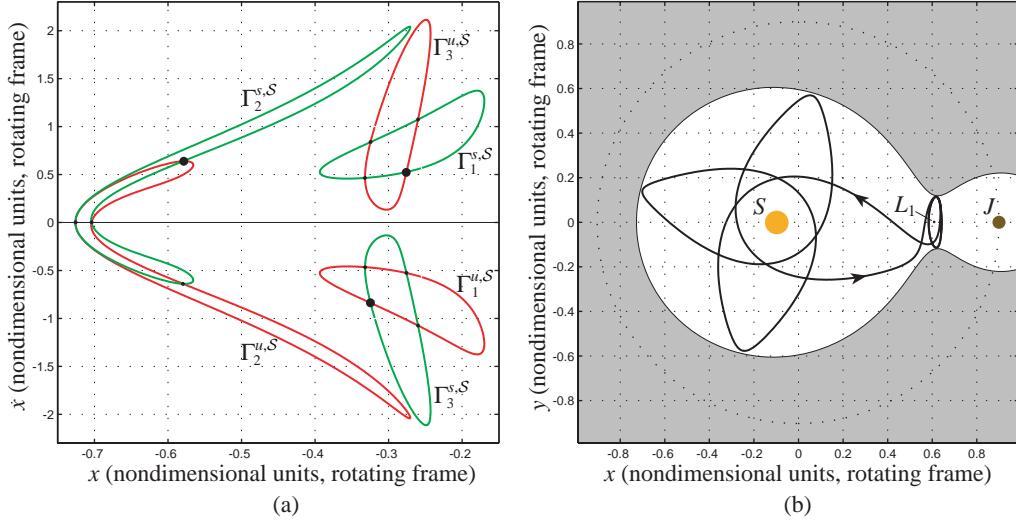
Figure 3.12: (a) The first three Poincaré cuts of the unstable ($W_{L_1,\text{p.o.}}^{u,\mathcal{S}}$) and stable ($W_{L_1,\text{p.o.}}^{s,\mathcal{S}}$) manifolds with the plane $y = 0$. (b) A nonsymmetric $(1,3)$-homoclinic orbit in the interior region (corresponding to the large dot in (a)).

3.5 The 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 $\mathcal{J}$ region. 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 ($C_2 > C > C_3$, see Figure 3.1), 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 $\mathcal{J}$ region be denoted $W_{L_1,\text{p.o.}}^{u,\mathcal{J}}$. On the same energy surface (same C value) there is an L_2 Lyapunov orbit, whose stable manifold in the $\mathcal{J}$ region we shall similarly denote $W_{L_2,\text{p.o.}}^{s,\mathcal{J}}$. The projection of the two-dimensional manifold tubes onto the position space is shown in Figure 3.13(a).

To find intersections between these two tubes, we cut the flow by the plane $x = 1 - \mu$. See Figure 3.13(b).

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

Numerical experiments show that the L_1 Lyapunov orbit unstable manifold $W_{L_1,\text{p.o.}}^{u,\mathcal{J}}$ does not coincide with the L_2 Lyapunov orbit stable manifold $W_{L_2,\text{p.o.}}^{s,\mathcal{J}}$. Moreover, for a wide range of μ and C values (where $C_2 > C > C_3$), numerical explorations show that they do intersect transversally. While it is true that for certain values of μ and C, there are tangencies between the stable and unstable manifold, we will not deal with this interesting case in this study. Hence, from now on, we will concentrate our numerical explorations only on the cases where the stable and unstable manifold intersect transversally.
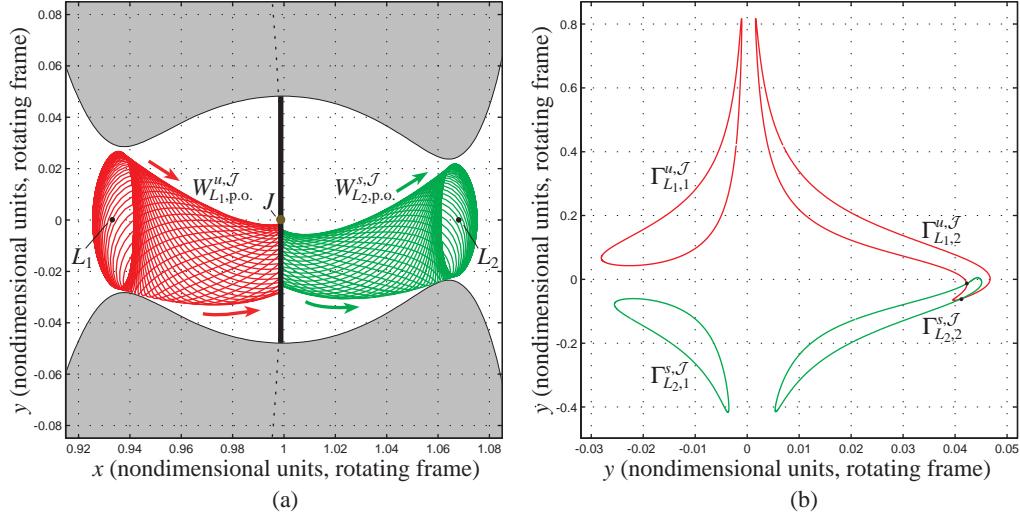
Figure 3.13: (a) The projection of invariant manifolds $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$ and $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$ in the region J of the position space. (b) The first two Poincaré cuts of the invariant manifolds with the plane $x = 1 - \mu$.

Now, suppose that $\Gamma^{u,\mathcal{J}}_{L_1,q}$ and $\Gamma^{s,\mathcal{J}}_{L_2,p}$ 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., $\Gamma^{u,\mathcal{J}}_{L_1,q} \cap \Gamma^{s,\mathcal{J}}_{L_2,p}$) 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. Our objective is to obtain the first intersection point (or group of points) of the curve $\Gamma^{u,\mathcal{J}}_{L_1,q}$ with the curve $\Gamma^{s,\mathcal{J}}_{L_2,p}$ 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 ($\mu = 0.0009537$), we took $C = 3.037$ and proceeded numerically to obtain the intersections of the invariant manifolds $W^{u,\mathcal{J}}_{L_1,\text{p.o.}}$ and $W^{s,\mathcal{J}}_{L_2,\text{p.o.}}$ with the plane $x = 1 - \mu$. In Figure 3.13(b) we show the curves $\Gamma^{u,\mathcal{J}}_{L_1,q}$ for $q = 1, 2$ and $\Gamma^{s,\mathcal{J}}_{L_2,p}$ for $p = 1, 2$. Notice that $\Gamma^{u,\mathcal{J}}_{L_1,2}$ and $\Gamma^{s,\mathcal{J}}_{L_2,2}$ intersect in two points (the black dots in Figure 3.13(b) near $y = 0.042$). Thus, the minimum q and p for a heteroclinic point to appear for this particular value of μ, C is $q = 2$ and $p = 2$. The $(2,2)$-heteroclinic points can each be forward and backward integrated to produce heteroclinic trajectories going from the L_1 Lyapunov orbit to the L_2 Lyapunov orbit. We show one of the heteroclinic orbits in Figure 3.14. Notice that the number of revolutions around Jupiter is given by $(q + p - 1)/2$. The reverse trajectory, going from the L_2 Lyapunov orbit to the L_1 Lyapunov orbit, is easily given by the symmetry s (3.3). It would be the mirror image (about the x-axis) of the trajectory in Figure 3.14, with the direction arrows reversed. These two *heteroclinic connections* together form a symmetric *heteroclinic cycle*.
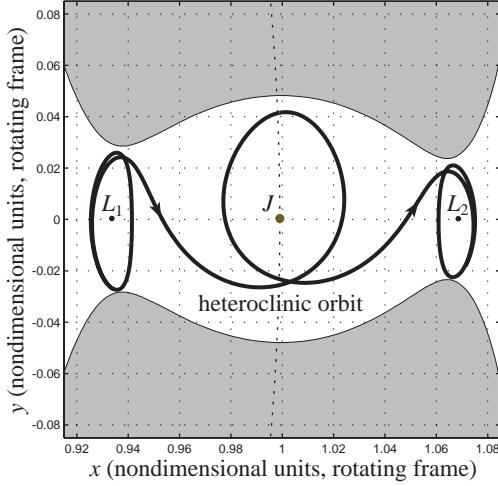
Figure 3.14: The existence of a transversal $(2, 2)$-heteroclinic orbit in the J region.

3.6 The Existence of Chains of Homoclinic Orbits and Heteroclinic Cycles.

We have used a combination of analytical and numerical techniques to show 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 Jacobi constant 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 regions.

As an example, we again choose the Sun-Jupiter system ($\mu = 0.0009537$), but now a Jacobi constant value similar to that of comet *Oterma* during its Jupiter encounters ($C = 3.03$). Using the described methodologies, we obtain an interior region orbit homoclinic to the L_1 Lyapunov orbit, an exterior region orbit homoclinic to the L_2 Lyapunov orbit, and a heteroclinic cycle connecting the L_1 and L_2 Lyapunov orbits. The union of these orbits is a *homoclinic-heteroclinic chain*. See Figure 1.2. The existence of homoclinic-heteroclinic chains has important consequences, which will be expanded upon further in §4.

4 Global Orbit Structure.

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 shall use the chain of two homoclinic orbits and one symmetric heteroclinic cycle (previously generated in §3) 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 shall find an invariant set for this map near some transversal homoclinic and heteroclinic points along the chain where "Smale horseshoe"-like dynamics exist. We shall 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 regions. Not only shall we prove the existence of the invariant set, but we shall also numerically approximate it, gaining further insight into the complex global dynamics associated with the chains.

Here is additional detail about how we shall proceed:

1. In §4.1, we shall construct a Poincaré map P transversal to the flow whose domain U consists of four different squares $U_i, i = 1, 2, 3, 4$, located in different regions of phase space in the neighborhood of the chain. See Figures 4.1 and 4.2.
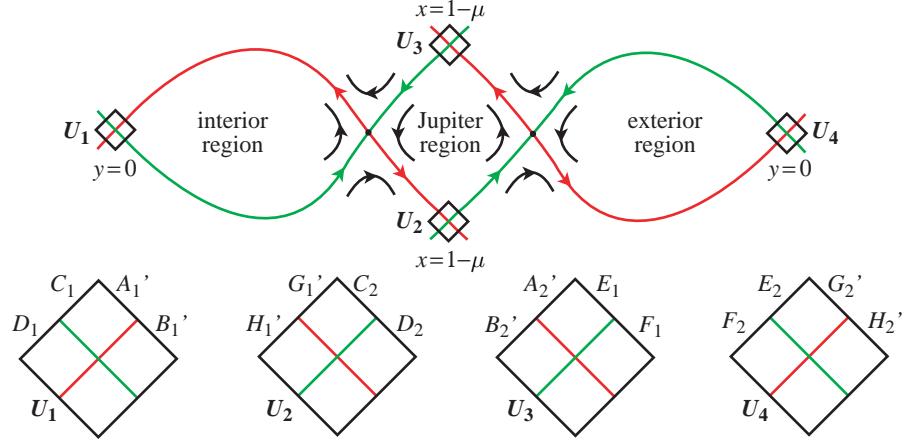

Figure 4.1: The construction of a suitable Poincaré map.

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 region, 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 region 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 ($P^n(q), n = 0, \pm1, \pm2, \dots$) 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.

2. 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 §4.2 and §4.3, we shall extend the symbolic dynamics results of LMS [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., $(\dots, u_{i_{-1}}; u_{i_0}, u_{i_1}, u_{i_2}, \dots)$ 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.

3. In §4.4, we shall state the main theorem of this section 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

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

there exist initial conditions near the transversal homoclinic and heteroclinic points 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 $2, 4$. 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}. See Figures 4.1 and 4.2. 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 shall provide a sketch of the proof in §4.4 and §4.6 both for completeness and for the convenience of the reader. For more details, one can consult Moser [1973], LMS [1985] and Wiggins [1988, 1993].

4. In §4.5 we numerically construct sets of orbits with prescribed itineraries. By successive application of the Poincaré map P to a transversal plane in the neighborhood of a chain, we can generate regions of orbits with itineraries of any size.

4.1 Construction of a Suitable Poincaré Map.

In §3, we have shown that with an appropriate Jacobi constant, 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 regions and a symmetric $(1, 1)$-transversal heteroclinic cycle in the Jupiter region. 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 ϵ_1 and ϵ_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 $A_1, B_1, C_1, D_1, E_1, F_1, G_1$ and H_1 as the set of points

of $d_{1,1}^-, r_{1,1}^-, r_{1,1}^+, d_{1,1}^+, d_{1,2}^+, r_{1,2}^+, r_{1,2}^-$ and $d_{1,2}^-$, respectively, such that $||\zeta|^2 - \rho^*| < \epsilon$. These sets correspond to thin strips on the bounding sphere centered on the asymptotic sets $a_{1,1}^-, a_{1,1}^+, a_{1,2}^+$ and $a_{1,2}^-$, respectively. Similarly, we can define corresponding strips for the bounding spheres $n_{2,1}$ and $n_{2,2}$ of the equilibrium region $\mathcal{R}_2$. See Figure 4.2.
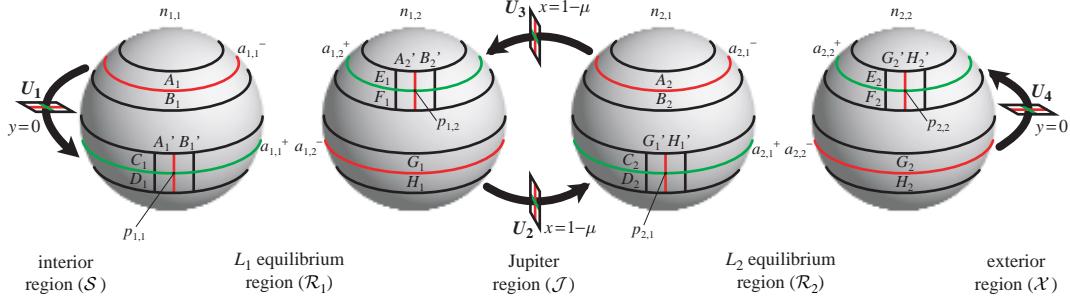
Figure 4.2: The strips near the asymptotic sets on the spheres $n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}$.

If ϵ_1 and ϵ_2 are small enough, the flow is transversal to the surfaces just defined. Recall from §2.4 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 $i = 1, 2, 3, 4$. 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) region. 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 4.2 provided ϵ_1 and ϵ_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 region. 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 mappings sending A_2, B_2, G_1, H_1 into 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 4.2.

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 4.1 and 4.2. 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 $P : U \to U$ 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 shall 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 shall consider the invariant set of points, Λ, which remain in U under all forward

and backward iterations by P. Thus Λ is defined as

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

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.

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

Domain of the Poincaré Map P. Instead of studying the first return map $\bar{f}$ (induced by 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. See Figure 4.3.
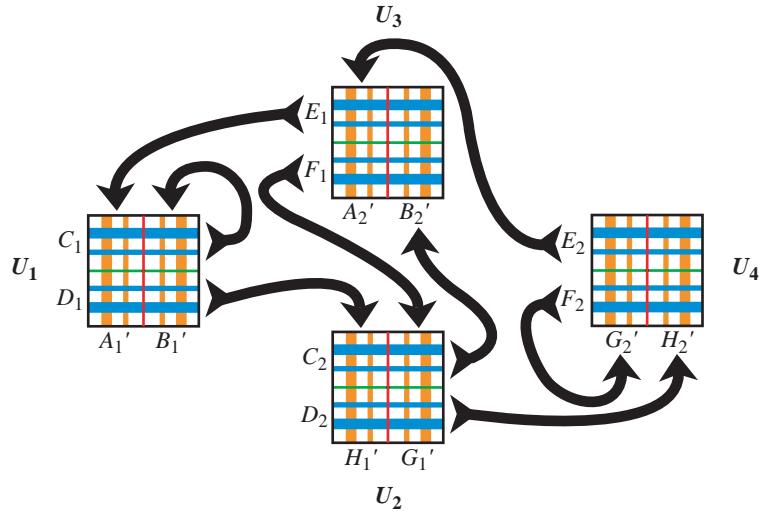
Figure 4.3: The families of horizontal strips (blue) and their images (orange) under P.

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) region, 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 shall return to this point at the end of §4.3.

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. See Figure 4.4.

Horizontal and Vertical Strips. For the standard textbook example, the first return map $\bar{f}$ (induced by f) on the square Q qualitatively looks like a Smale horseshoe map.
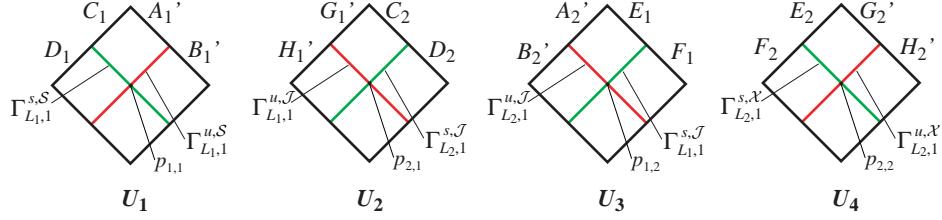
Figure 4.4: The domain $U = U_1 \cup U_2 \cup U_3 \cup U_4$ of the Poincaré map P.

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, 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.
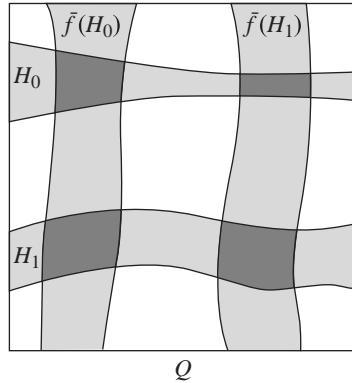
Figure 4.5: Generalization of the notion of horizontal and vertical rectangles for the Conley-Moser conditions.

For the Poincaré map P constructed in this section, 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 shall discuss first the issue of each family having an infinite number of strips.

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 $(D_1 \cap B_1')$. See Figure 4.6.

Recall the observation in §2.4 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 $(D_1 \cap B_1')$ under P is a strip contained in H_1' of arbitrarily long length, cutting U_2 an infinite number of times
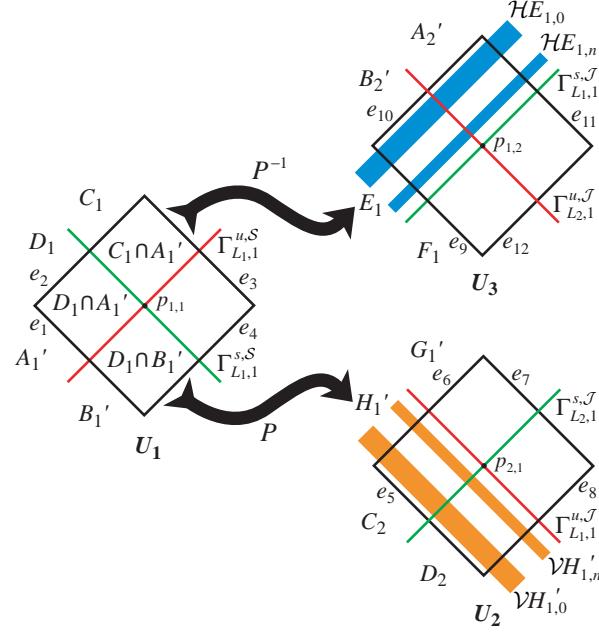
Figure 4.6: The topological squares and the images of some rectangles. We show schematically only two strips although there is an infinite number.

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. 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

$$\mathcal{V}H_{1,\infty}' = \lim_{n \to \infty} \mathcal{V}H_{1,n}'.$$

Clearly, $\mathcal{V}H_{1,\infty}'$ is simply the vertical curve $\Gamma_{L_1,1}^{u,\mathcal{J}}$ which is on the Jupiter region 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 4.6. 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

$$B_2'(\{V_n^{32}\}), H_2'(\{V_n^{42}\}), A_1'(\{V_n^{13}\}), G_1'(\{V_n^{23}\}), A_2'(\{V_n^{34}\}), G_2'(\{V_n^{44}\}).$$

40

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

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. 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.

Now we shall 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 ϵ_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 ϵ_1, the orbit which pierces V_n^{21} has wound around L_1 for $(n + r_{\min})$ times. In the rest of §4, we shall assume that we have adjusted the widths (the ϵ_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.

4.2   The Generalized Conley-Moser Conditions

For the standard textbook example (introduced in §4.1), 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 $0 < \nu_h < 1$.

We shall call Condition 1 the ***strip condition***. 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 §4.1, 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 shall state below the theorem that the Poincaré map P of the PCR3BP satisfies the generalized Conley- Moser conditions but shall leave its proof to §4.5.

Theorem 4.1 *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.

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 where 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. See Figure 4.6.

Even though the proof will be deferred to §4.6, we shall use this result to prove the main theorem on the global orbit structure of the PCR3BP in §4.3 and §4.4

4.3 Symbolic Dynamics.

In §4.1 and §4.2, 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 D into a horseshoe and intersects it with the square, one has an infinite Cantor set of trapped points p in the invariant set Λ_h. Here,

$$\Lambda_k = \cap_{n=-\infty}^{\infty} h^n(D),$$

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

Recall that p can be defined by

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

where s_i denotes one of the elements in $S = \{0, 1\}$ 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 Σ^2) can be attached to every point p in the invariant set Λ_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

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

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 Λ_h is to draw the regions that remain trapped for one forward and one backward iteration in the square D. 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. See Figure 4.7. 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 $\{0, 1\}$ shown above.

For the Poincaré map P, we can use a similar technique to visualize the invariant set Λ and its associated set of bi-infinite sequences. Instead of one square D, 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 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. Recall from §4.1 that for $\{V_n^{ji}\}$, the index ji indicates that the family is in the square U_j and it came from the

...0,0; ...1,0;   ...1,1; ...0,1;

;1,0...

;1,1...

...1,0;1,1...

;0,1...

;0,0...

D
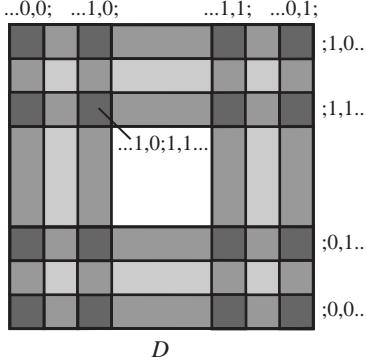
Figure 4.7: The invariant set Λ_h of the horseshoe map h.

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. See Figure 4.8.

For simplicity of illustration, we draw Figure 4.8 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" $Q_{m;n}^{3;12}$ as an example. Since $Q_{m;n}^{3;12}$ 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" $(Q_{m;n}^{3;12})$:

1. they are currently in U_1 and will go to U_2 and on their way their projection will wind around L_1 for $(n+r_{\min})$ revolutions where $r_{\min}$ is the minimum number of revolutions discussed earlier in §4.1;

2. they came from U_3 and their 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 Λ 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 at the beginning of §4). It is used to keep track 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 member of each vertical or horizontal family ($\{V_n^{ji}\}$ or $\{H_n^{ij}\}$). As mentioned at the end of §4.1, this set of integers also corresponds to the number of revolutions that the projection of an orbit winds around either L_1 and L_2.
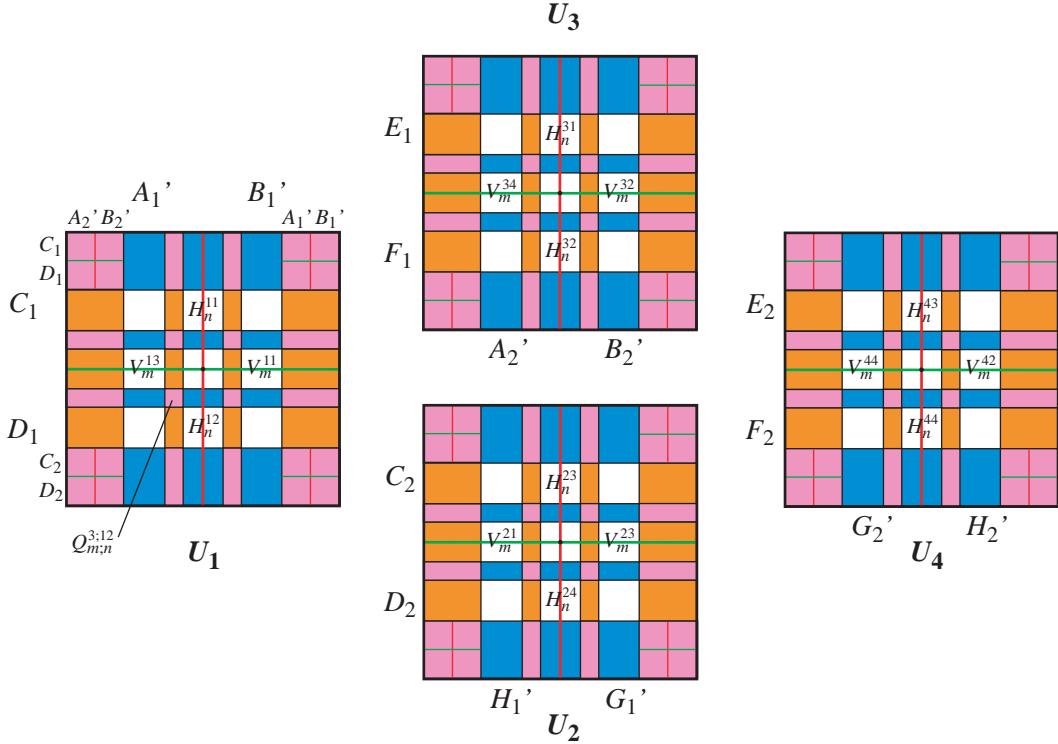
Figure 4.8: The invariant set Λ of the Poincaré map P.

Singular Poincaré Map. Now we shall discuss briefly the issue of the singular Poincaré map and how it relates to certain modifications of the space of symbol sequences Σ. 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 Σ 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** α from now on. Moreover, the shift map σ on Σ 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 the Figure 4.8, it should be clear that H^{12}_∞ (or H^{11}_∞) is simply the horizontal curve $\Gamma^{s,\mathcal{S}}_{L_1,1}$ which is on the interior (Sun) region 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 β 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 γ 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 δ, we have either a homoclinic or a heteroclinic orbit.

4.4 Global Orbit Structure.

Now we are ready to put together all the results in §4.2 and §4.3 and to state the main theorem of §4 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 regions and a symmetric $(1,1)$-heteroclinic cycle in the Jupiter region. Now we shall 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 region together with two homoclinic orbits, one of which is a (q_1, p_1) orbit in the interior region and the other is a (q_3, p_3) orbit in the exterior region.

Theorem 4.2 *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*

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

   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 revolution 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 region) is a constant which depends on the region and the given chain of homoclinic orbits and heteroclinic cycle. For the Jupiter region, the number is $(q_2 + p_2 - 1)/2$. For the interior and exterior regions, 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*

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

   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*

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

   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 shall provide a sketch of the proof here, which makes use of the major results in §4.2 and §4.3. 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 §4.2, we shall defer their proofs to §4.6 so that we can discuss first the implications of this theorem.

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 region, 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 region 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 §4.3 on symbolic dynamics shows that for any bi-infinite sequence of type α, $\alpha = (u, r)$, 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 Additional Comments on the Implications of the Theorem.   Type α orbits include "oscillating," "capture" and "non-transit" orbits. Recall that oscillating orbits are orbits which cross from one region to the other infinitely many times, capture orbits are orbits which cross sometime but eventually stay in one region, and non-transit orbits always stay in the same region. Type β and type γ orbits are asymptotic orbits which wind to one of the Lyapunov orbits. Type δ 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 regions.

4.5 Numerical Construction of Orbits with Prescribed Itineraries

Throughout this paper, we have been developing a framework for understanding transport in the PCR3BP. Fundamental to our approach has been the homoclinic-heteroclinic chain, those objects which are the union of two homoclinic orbits and a symmetric heteroclinic cycle. Early in our investigations, we noticed the similarity between observations of actual comet orbits like *Oterma* and homoclinic-heteroclinic chains of the same energy. See Figure 4.9. Noting this similarity, we deduced that the same dynamics governing the motion of the comets was at work in the chains. By exploring and cataloging the phase space objects related to the chain, we gain insight into the dynamics of the temporary capture and resonance transition of actual comets.
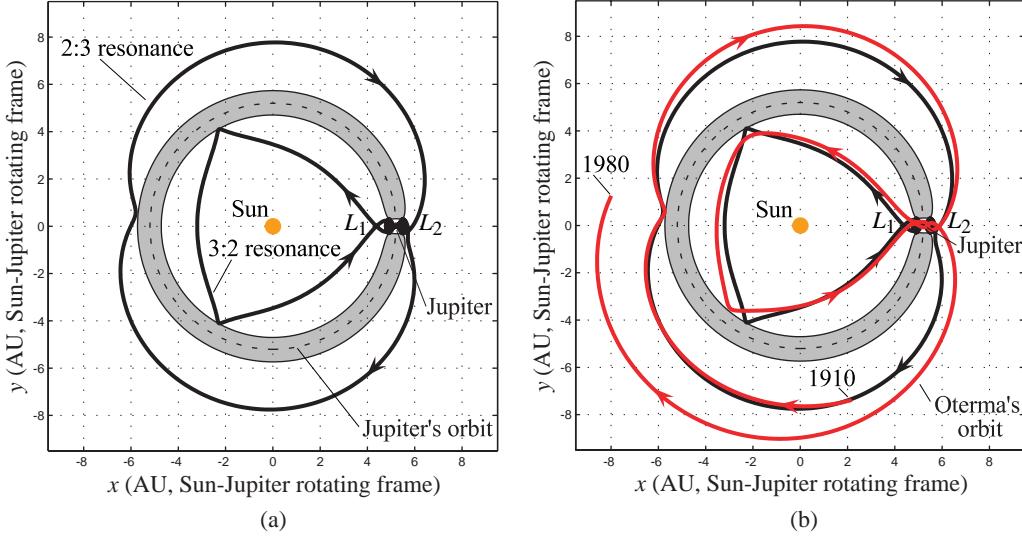

Figure 4.9: (a) The homoclinic-heteroclinic chain corresponding to the Jupiter comet *Oterma*. (b) The actual orbit of *Oterma* (AD 1910–1980) overlaying the chain.

In this section, we make this observation more concrete by exploring the complex orbit structure in the neighborhood of a chain. What we have found is an invariant set of orbits, to each of which we can attach an itinerary (e.g., $(\ldots, X, J, S, J, \ldots)$ in the informal notation) describing the future and past history of the orbit for all time. Furthermore, Theorem 4.1 shows us that all permissible itineraries exist in the neighborhood of a chain.

The invariant set is a theoretical construct, and though useful for guiding our understanding and classification of the dynamics, its infinite nature renders it powerless to provide us with usable trajectories. Computational and numerical methods must be brought to bear which iteratively approximate the invariant set.

Numerical Construction of Orbits with Prescribed Itineraries. The description of the construction of the invariant set in §4.3 involved successive iterations of the Poincaré map P. Finite areas of finite central block itineraries evolved under successive application of the map P into a "cloud of points," the invariant set Λ of points with bi-infinite itineraries. If we truncate the construction of the invariant set at some finite number of iterations of P, we will find regions of phase space which have a certain finite itinerary. Orbits in such regions will be robust. More specifically, the essential feature of the orbit, its itinerary, will be robust because all the nearby orbits in phase space have the same finite itinerary. Thus,

by truncating our construction of the invariant set Λ at some finite number of applications of P, we can generate a set of robust orbits with different finite itineraries. The sets of orbits with different itineraries are easily visualizable on our chosen Poincaré section as areas in which all the orbits have the same finite itinerary. We will also no longer be limited to a small neighborhood of a chain, but can obtain more global results.

Example Itinerary: (X, J, S, J, X). In what follows, we shall illustrate the numerical construction of sets of orbits with prescribed itineraries. We shall of course be limited to an itinerary of finite size, a central block. However, using our simple procedure, the size of this central block can be arbitrarily large. We shall use the less formal sequence notation using the symbols $\{S, J, X\}$ to denote the location of the orbit in the interior (Sun), Jupiter, or exterior regions, respectively.

As our example, we shall construct an orbit with the central block (X, J, S, J, X) which roughly corresponds to the behavior of comet *Oterma* (AD 1910–1980) with respect to the Sun-Jupiter system. This central block denotes an orbit which went from the exterior region into the interior (Sun) region via the Jupiter region, and will then return to the exterior region via the Jupiter region.

We seek regions of phase space which have the sequences $(\dots, X, J, S, J, X, \dots)$ with the central block (X, J, S, J, X). We shall therefore systematically seek regions on a suitably chosen Poincaré section which correspond to this central block. We shall take $C = 3.038$ (just below C_2 in case 3) as our Jacobi constant. We choose this Jacobi constant because, though it differs from *Oterma*'s ($C = 3.03$), it makes the visualization easier and preserves the dynamics of *Oterma*'s transition. Moreover, in order to link the present numerical construction with the earlier theoretical framework and terminology, we shall adopt the following convention. The U_1 and U_4 (Poincaré) sections will be the planes $(y = 0, x < 0)$ in the interior region, and $(y = 0, x < -1)$ in the exterior region, respectively. The U_2 and U_3 sections will be the planes $(x = 1 - \mu, y < 0)$ and $(x = 1 - \mu, y > 0)$ in the Jupiter region, respectively.
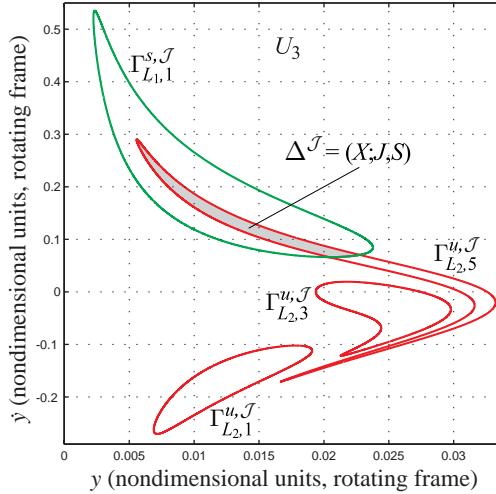

Figure 4.10: The first few transversal cuts of the L_1 (stable) and L_2 (unstable) Lyapunov orbit manifolds on the U_3 section in the Jupiter region. Notice the intersection region, in which all orbits have the central block itinerary $(X; J, S)$.

In Figure 4.10, we show the first few transversal Poincaré cuts of the L_1 and L_2 Lyapunov

orbit manifolds on the U_3 section in the Jupiter region $\mathcal{J}$. Note that the Poincaré map is area preserving owing to the Hamiltonian nature of the flow and the particular choice of Poincaré section. Notice that there is an intersection $\Delta^{\mathcal{J}}$ for $p+q=6$, where $p=1$ is the cut number for the L_1 Lyapunov orbit stable manifold and $q=5$ is the cut number for the L_2 Lyapunov orbit unstable manifold. The interior $\Delta_{L_1,1}^{s,\mathcal{J}}$ of $\Gamma_{L_1,1}^{s,\mathcal{J}}$ (the first cut of the L_1 Lyapunov orbit stable manifold) is connected to the spherical cap $d_{1,2}^+$ of the bounding sphere $n_{1,2}$ by the stable manifold tube $W_{L_1,\text{p.o.}}^{s,\mathcal{J}}$. Hence, $\Delta_{L_1,1}^{s,\mathcal{J}}$ contains all the orbits that will go from the Jupiter region to the interior (Sun) region during their next close approach to the L_1 equilibrium region. Similarly, the interior $\Delta_{L_2,5}^{u,\mathcal{J}}$ of $\Gamma_{L_2,5}^{u,\mathcal{J}}$ (the fifth cut of the L_2 Lyapunov orbit unstable manifold with the $x=1-\mu$ plane, following the convention of §3) is connected to the spherical cap $d_{2,1}^-$ of the bounding sphere $n_{2,1}$ by the unstable manifold tube $W_{L_2,\text{p.o.}}^{u,\mathcal{J}}$. Thus, $\Delta_{L_2,5}^{u,\mathcal{J}}$ contains all the orbits that entered the Jupiter region from the exterior region and have completed two revolutions around Jupiter.

Therefore, the intersection

$$\Delta^{\mathcal{J}} = \Delta_{L_1,1}^{s,\mathcal{J}} \cap \Delta_{L_2,5}^{u,\mathcal{J}}$$

contains all the orbits that have come from the exterior region $\mathcal{X}$ into the Jupiter region $\mathcal{J}$, have gone around Jupiter $2\frac{1}{2}(=(p+q-1)/2)$ times, and will enter the interior region $\mathcal{S}$. The region $\Delta^{\mathcal{J}}$ is the intersection of the image of the spherical cap $d_{2,1}^-$ and the pre-image of the spherical cap $d_{1,2}^+$. Therefore, from the discussion in §2, we know that orbits contained in the intersection $\Delta^{\mathcal{J}}$ are those which entered the L_2 equilibrium region $\mathcal{R}_2$ from the exterior region $\mathcal{X}$ and which will exit the L_1 equilibrium $\mathcal{R}_1$ into the interior region $\mathcal{S}$. The orbits are currently in the Jupiter region $\mathcal{J}$. We can therefore attach the central block label $(X; J, S)$ to the intersection $\Delta^{\mathcal{J}}$.

To determine regions of phase space with additional symbols of our desired central block, we take the $(X; J, S)$ region and evolve it forward under the equations of motion until it intersects the U_1 section in the interior region. In Figure 4.11 we show this Poincaré section. Notice that the $(X, J; S)$ region lies entirely within the interior $\Delta_{L_1,1}^{u,\mathcal{S}}$ of the first interior region cut $\Gamma_{L_1,1}^{u,\mathcal{S}}$ of the L_1 Lyapunov orbit unstable manifold. We also see that a couple of segments of the $(X, J; S)$ region intersect the interior $\Delta_{L_1,1}^{s,\mathcal{S}}$ of the first interior region stable manifold cut $\Gamma_{L_1,1}^{s,\mathcal{S}}$. Any orbit within $\Delta_{L_1,1}^{s,\mathcal{S}}$, and therefore within the stable manifold tube $W_{L_1,\text{p.o.}}^{s,\mathcal{S}}$, will be brought back to the Jupiter region. These intersecting segments $\Delta^{\mathcal{S}}$ therefore carry the label $(X, J; S, J)$ and bring us one symbol closer (J) to our desired central block.

We take the larger of the two intersecting segments and evolve it forward in time until it re-enters the Jupiter region and intersects the U_2 Poincaré section. See Figure 4.12. Notice that the $(X, J, S; J)$ region (the image of the larger segment of $\Delta^{\mathcal{S}}$) lies entirely within the interior $\Delta_{L_1,1}^{u,\mathcal{J}}$ of the first Jupiter region cut $\Gamma_{L_1,1}^{u,\mathcal{J}}$ of the L_1 Lyapunov orbit unstable manifold. This thin filament has a segment intersecting the interior $\Delta_{L_2,5}^{s,\mathcal{J}}$ of the Poincaré cut $\Gamma_{L_2,5}^{s,\mathcal{J}}$ of the L_2 Lyapunov orbit stable manifold. Any orbit in this intersection region Δ will escape from Jupiter into the exterior region. Thus, any orbit in this segment Δ can be labeled with the central block $(X, J, S; J, X)$, which is our desired finite itinerary.

We have forward and backward integrated an initial condition within this region to illustrate the characteristics of an orbit corresponding to the $(X, J, S; J, X)$ region. See Figure 4.13. 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 be generated with this same systematic procedure. Not only do we know such orbits exist, but we have a relatively simple method for producing them.

50
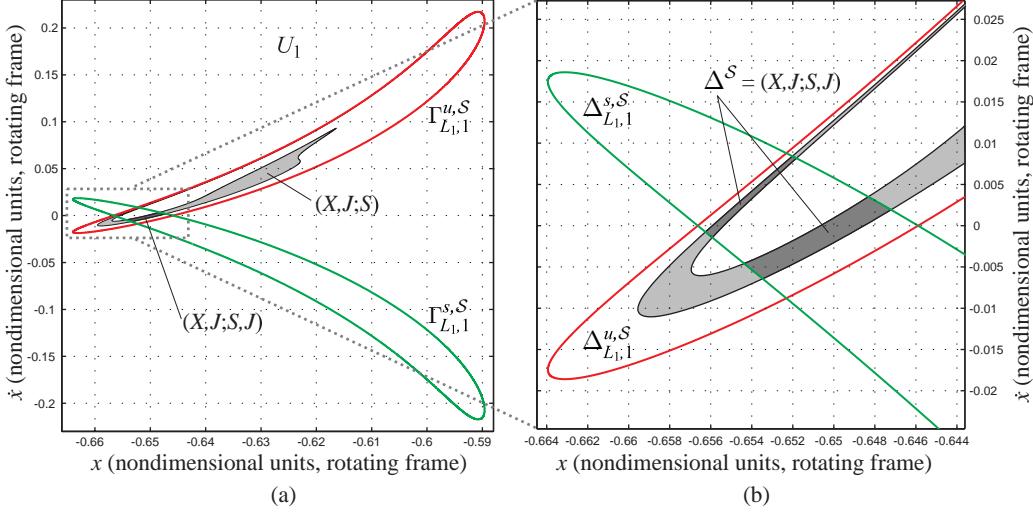
Figure 4.11: (a) Taking the $(X; J, S)$ region of the U_3 Poincaré section (see Figure 4.10), we evolve it until it intersects the U_1 Poincaré section in the interior region (lightly shaded). (b) A close-up of the intersection of the $(X, J; S)$ region with the interior $\Delta_{L_1,1}^{s,\mathcal{S}}$ of $\Gamma_{L_1,1}^{s,\mathcal{S}}$, the first stable manifold cut of the L_1 Lyapunov orbit. Note the regions labeled $(X, J; S, J)$, which will return to the Jupiter region.

4.6 The Poincaré Map Satisfies The Generalized Conley-Moser Conditions

The proof that the Poincaré map P satisfies the generalized Conley-Moser conditions follows the same pattern as the proof given in LMS [1985]. We shall provide a sketch here mainly for the convenience of the reader. For more details, see Moser [1974].

Strip Condition. The fact that the Poincaré map P satisfies the strip condition follows from the lemma below. Since we have a heteroclinic cycle in our case, the proof of this crucial lemma is slightly different from the proof in LMS [1985]. Hence, more detail will be provided here.

Lemma 4.3 *The Poincaré map P maps horizontal strips to vertical strips, i.e.,*

$$P(\mathcal{H}C_{1,n}) = \mathcal{V}B'_{1,n} \quad P(\mathcal{H}D_{1,n}) = \mathcal{V}H'_{1,n} \quad P(\mathcal{H}C_{2,n}) = \mathcal{V}B'_{2,n} \quad P(\mathcal{H}D_{2,n}) = \mathcal{V}H'_{2,n}$$
$$P(\mathcal{H}E_{1,n}) = \mathcal{V}A'_{1,n} \quad P(\mathcal{H}F_{1,n}) = \mathcal{V}G'_{1,n} \quad P(\mathcal{H}E_{2,n}) = \mathcal{V}A'_{2,n} \quad P(\mathcal{H}F_{2,n}) = \mathcal{V}G'_{2,n}$$

for all positive integer n.

Proof. We illustrate the methods for the case $P(\mathcal{H}E_{1,n}) = \mathcal{V}A'_{1,n}$. Since this case involves the heteroclinic cycle, it is typical in our study. The other cases can be proved similarly.

Recall that the equations of the PCR3BP have a symmetry s which we have used earlier to construct the stable manifold out of the unstable manifold. Since the heteroclinic cycle in our chain is a symmetric one, we shall have the following relation:

$$P^{-1} = s^{-1} \circ P \circ s,$$

where s is regarded as the symmetry s restricted to the domain U of the Poincaré map. Note $s = s^{-1}$. In the following, we shall regard all operations on sets as taking place in U.

51
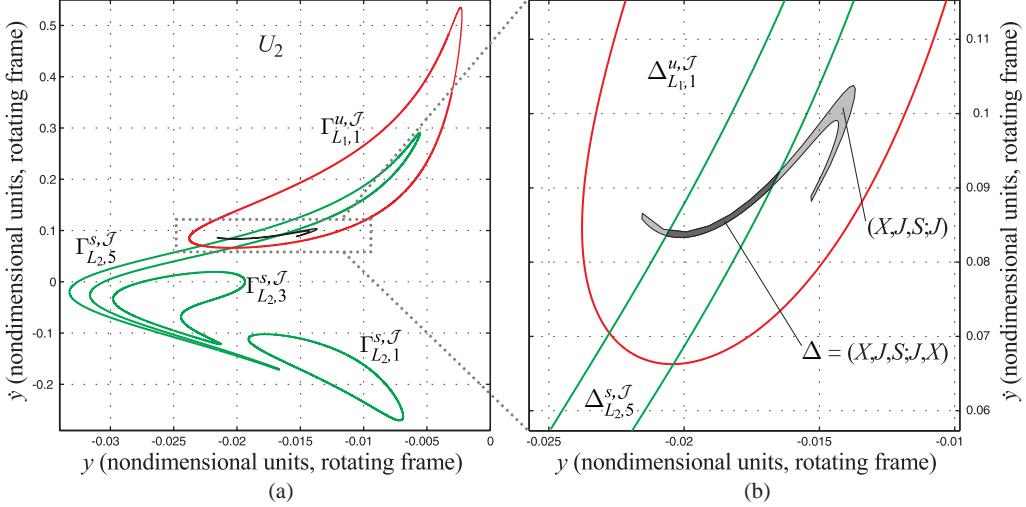
Figure 4.12: (a) Taking the $(X, J; S, J)$ region of the U_1 Poincaré section (see Figure 4.11), we evolve it until it intersects the U_2 Poincaré section ($x = 1 - \mu, y < 0$) in the Jupiter region (lightly shaded). Part (b) shows a close-up of the intersection of the $(X, J, S; J)$ region with the interior $\Delta_{L_2,5}^{s,\mathcal{J}}$ of $\Gamma_{L_2,5}^{s,\mathcal{J}}$, the fifth stable manifold cut of the L_2 Lyapunov orbit. Note the region labeled $(X, J, S; J, X)$, which will return to the exterior region. This region contains orbits with the desired finite itinerary.

Also recall that

$$\mathcal{H}E_{1,n} \subset P^{-1}((C_1 \cap A_1') \cup (D_1 \cap A_1')) = sPs((C_1 \cup D_1) \cap A_1') = sP(D_1 \cap (A_1' \cup B_1')).$$

But $P(D_1 \cap (A_1' \cup B_1'))$ is the family of vertical strips in H_1'. It is equal to $P(D_1) \cap H_1'$. Therefore, we have

$$\mathcal{H}E_{1,n} \subset s(P(D_1) \cap H_1') = sP(D_1) \cap E_1 = P^{-1}s(D_1) \cap E_1 = P^{-1}(A_1') \cap E_1.$$

Applying the Poincaré map on both sides, we obtain

$$P(\mathcal{H}E_{1,n}) \subset A_1' \cap P(E_1) = \cup_{n=0}^{\infty} \mathcal{V}A_{1,n}'. \tag{4.1}$$

Similarly, we have

$$P^{-1}(\mathcal{V}A_{1,n}') = sPs(\mathcal{V}A_{1,n}') = sP(\mathcal{H}D_{1,n}) \subset s(\cup_{n=0}^{\infty} \mathcal{V}H_{1,n}') = \cup_{n=0}^{\infty} \mathcal{H}E_{1,n}.$$

Therefore,

$$\mathcal{V}A_{1,n}' \subset \cup_{n=0}^{\infty} P(\mathcal{H}E_{1,n}). \tag{4.2}$$

Using the relations (4.1) and (4.2) we find that

$$\cup_{n=0}^{\infty} P(\mathcal{H}E_{1,n}) = \cup_{n=0}^{\infty} \mathcal{V}A_{1,n}'.$$

Since the strips of the type $\mathcal{H}E_{1,n}$ or $\mathcal{V}A_{1,n}'$ are pairwise disconnected, each one of the strips $\mathcal{H}E_{1,n}$ must be mapped by P onto one of the strips $\mathcal{V}A_{1,m}'$. It remains to show that $m = n$.

Let γ be a diagonal line in the square $E_1 \cap B_2'$. Clearly, γ intersects $\mathcal{H}E_{1,n}$ for all n. Pick a point γ_n in each intersection $\gamma \cap \mathcal{H}E_{1,n}$. Recall that $\{\mathcal{H}E_{1,n}\}$ is ordered with respect

52
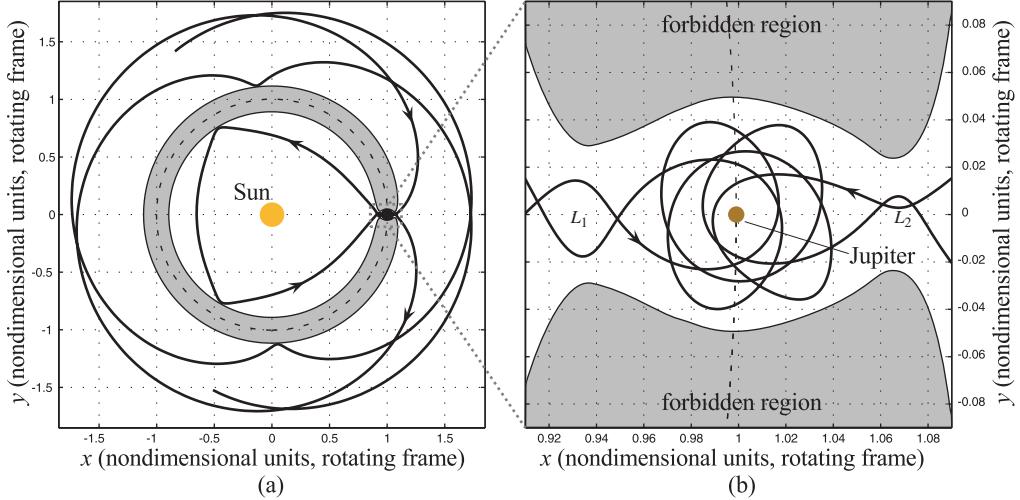
Figure 4.13: (a) An orbit with the itinerary $(\ldots, X, J, S, J, X, \ldots)$ computed using an initial condition inside the Δ region of the U_2 section (See Figure 4.12). (b) A close-up of this orbit in the Jupiter region.

to its distance from the longer edge e_E of the rectangle E_1. Therefore, the set $\{\gamma_n\}$ can be made into an ordered set $(\gamma_0, \gamma_1, \ldots, \gamma_n, \ldots)$ with respect to the ordering by distance from the point γ_n to the edge e_E.

After one iteration by P,

$$\gamma \cap \cup_{n=0}^{\infty} \mathcal{H}E_{1,n} \quad \text{is mapped into} \quad P(\gamma) \cap \cup_{n=0}^{\infty} \mathcal{V}A'_{1,n}.$$

The key observation is that since $P(\gamma)$ spirals inward from the longer edge e_A of the rectangle A'_1, the set $\{P(\gamma_n)\}$ can also be made into an order set $(P(\gamma_0), P(\gamma_1), \ldots, P(\gamma_n), \ldots)$ with respect to the distance from the point $P(\gamma_n)$ to the edge e_A. Recall that

1. every $\mathcal{V}A'_{1,m}$ must contain one and only one $P(\gamma)$ and

2. $\{\mathcal{V}A'_{1,m}\}$ is also ordered with respect to its distance from the longer edge e_A.

It follows from this that $m = n$.

Hyperbolicity Condition. As pointed out earlier, for the standard textbook example (introduced in §4.1), it is well known that if the first return map $\bar{f}$ (induced by f) on the square Q satisfies the Conley-Moser Conditions 1 and 2, then there exists an invariant set $\Lambda_{\bar{f}}$ of Q on which $\bar{f}$ has chaotic dynamics. However, a direct verification of whether $\bar{f}$ satisfies Condition 2 or not is non-trivial. When one thinks of stretching and contraction of maps, it is natural to think of the properties of the derivative of the map $(D\bar{f})$ at different points. Hence, when the map $\bar{f}$ is continuously differentiable, Condition 2 is usually replaced by another equivalent condition (Condition 3) that is based solely on the properties of the derivative of $\bar{f}$. Compared with Condition 2, Condition 3 is easier to check. While we shall state only the Generalized Condition 3 in the following, the standard Condition 3 is exactly the same with a couple of obvious modifications.

Define the unstable sector bundle $\mathcal{S}^u$ (in the tangent bundle TU) over the families of the horizontal strips as follows

$$\mathcal{S}^u_q = \{(v, w) \in T_p U \mid |v| \le \kappa |w|\},$$

53

where $0 < \kappa < 1/2$ and q is a point in a horizontal strip. Similarly, the stable sector bundle $\mathcal{S}^s$ over the families of the vertical strips is defined as

$$\mathcal{S}^s_q = \{(v, w) \in T_p U \mid |w| \leq \kappa|v|\},$$

where q is a point in a vertical strip. Then the Poincaré map P is said to satisfy the **Generalized Condition 3** if the following two conditions are met:

(a) $DP(\mathcal{S}^u_q) \subset \mathcal{S}^u_{P(q)}$ and $|w_1| \geq \kappa^{-1}|w_0|$ where (v_1, w_1) is the image of (v_0, w_0) under DP; i.e., the vertical component of a tangent vector gets amplified at least by κ^{-1} under DP.

(b) Similarly, $DP^{-1}(\mathcal{S}^s_q) \subset \mathcal{S}^s_{P^{-1}(q)}$ and $|v_{-1}| \geq \kappa^{-1}|v_0|$ where (v_{-1}, w_{-1}) is the image (v_0, w_0) under DP^{-1}; i.e., the horizontal component of a tangent vector gets amplified at least by κ^{-1} under DP^{-1}.

Since the Generalized Condition 3 is based solely on the local properties of the derivative of a map, the proof that Generalized Conditions 1 and 3 imply Generalized Condition 2 is essentially the same as the standard proof that the Conditions 1 and 3 imply Condition 2 with some obvious modifications and hence will be skipped. For more details on the standard case, see Moser [1973] and Wiggins [1990].

As for the proof that the Poincaré map P satisfies the Generalized Condition 3, the key observation is that all the stretching and contraction by the map P takes place inside the equilibrium regions $\mathcal{R}_1$ and $\mathcal{R}_2$. Recall that $\mathcal{R}_i$ is bounded by pairs of spheres $n_{i,1}$ and $n_{i,2}$ (for $i = 1, 2$) which contains the domain U of the map P (or more accurately four squares whose union is diffeomorphic to U). See Figures 2.3 and 4.2. Inside these equilibrium regions, the flow is exactly given by the linear equations (see §2.3) in suitable coordinates. This flow satisfies the generalized Condition 3 with a constant κ that can be chosen as large as desired provided that U is sufficiently small.

5 Resonance Transitions

5.1 Introduction.

Our new dynamical mechanism effecting transfer between the interior and exterior regions is the heteroclinic intersection between the L_1 and L_2 Lyapunov orbit manifold tubes in the Jupiter region. As mentioned previously, the orbits interior to these tubes are the transit orbits of each equilibrium region. Therefore, their intersection is a set of orbits which come from one heliocentric region (S or X) and exit to the other (X or S).

This was an unexpected result. It was previously believed that a third degree of freedom was necessary for resonance transition or that "Arnold diffusion" was somehow involved. But as we have seen, only the planar CR3BP is necessary. The dynamics and phase space geometry involved in the heteroclinic connection now give us a language with which to discuss and further explore resonance transition.

The dynamical channels discussed in previous sections are a generic transport mechanism connecting the interior and exterior Hill's regions. We shall now focus on a limited case of this generic transport mechanism; the case of transport between resonances. In particular, we shall study how this homoclinic-heteroclinic transport mechanism connects the mean motion resonances of the interior and exterior regions (e.g., the 3:2 and 2:3 Jupiter resonances) via the Jupiter region.

Using numerical exploration of the heteroclinic connection between the interior and exterior resonances, we shall obtain a deeper understanding of the mean motion resonance

transition of actual Jupiter comets. In particular, we shall try to explain in more precise terms the sense in which *Oterma* transitions between the 3:2 and 2:3 resonances. In the process, we shall discover much about the mixed phase space structure, especially the mean motion resonance structure, of the PCR3BP.

Recall that in §3.6 we constructed a homoclinic-heteroclinic chain $\mathcal{C}$ for the Sun-Jupiter system and with a Jacobi constant value similar to that of comet *Oterma* during its Jupiter encounters ($C=3.03$). See Figures 1.2 and 4.9. This chain is a union of four orbits: an interior region orbit homoclinic to the L_1 Lyapunov orbit, an exterior region orbit homoclinic to the L_2 Lyapunov orbit, and a symmetric heteroclinic cycle (two orbits) connecting the L_1 and L_2 Lyapunov orbits. For simplicity of exposition, we chose this particular chain because both of its homoclinic orbits are of (1,1)-type and were constructed using the first Poincaré cuts of their respective stable and unstable manifolds. Limiting our chain to (1,1)-type meant, for this particular energy regime, that two different resonance connections were possible; 3:2 to 1:2 and 3:2 to 2:3. We chose the 3:2 to 2:3 chain for our exploration.

Theorem 4.1, or more accurately its simplified version, tells us that in a neighborhood of this particular $\mathcal{C}$, there exists an orbit $\mathcal{O}$ whose symbolic sequence $(\dots, J, X, J, S, J, \dots)$ is periodic and has a central block itinerary (J, X, J, S, J). Because this orbit transitions between the interior and exterior regions (the neighborhood of the 3:2 and 2:3 resonances, in particular), we call this kind of itinerary a resonance transition block. This orbit makes a rapid transition from the exterior region to the interior region and vice versa, passing through the Jupiter region. It will repeat this pattern ad infinitum.

We have commented earlier that while an orbit with this exact itinerary is very fragile, the structure of nearby orbits whose symbolic sequences have a central block like the orbit $\mathcal{O}$, namely (J, X, J, S, J), is quite robust. In fact, we have devised simple procedures to construct sets of orbits with such specific characteristics (as encoded in the central block itinerary) in the previous section.

We will study how this particular chain $\mathcal{C}$ and its nearby dynamical channels connect the 3:2 resonance of the interior region and the 2:3 resonance of the exterior region.

Delaunay Variables. Recall that the PCR3BP is a perturbation of the two-body problem. Hence, outside of a small neighborhood of L_1, the trajectory of a comet in the interior region follows essentially a two-body orbit around the Sun. In the heliocentric inertial frame, the orbit is nearly elliptical. The mean motion resonance of the comet with respect to Jupiter is equal to $a^{-3/2}$ where a is the semi-major axis of this elliptical orbit. Recall that the Sun-Jupiter distance is normalized to be 1 in the PCR3BP. The comet is said to be in p:q resonance with Jupiter if $a^{-3/2} \approx p/q$, where p and q are small integers. In the heliocentric inertial frame, the comet makes roughly p revolutions around the Sun in q Jupiter periods. See Figure 5.1, where we illustrate a numerically constructed orbit $\mathcal{O}'$, which has a central block sequence (J, X, J, S, J). Similar observations also hold for orbits in the exterior region outside of a small neighborhood of L_2.

To study the process of resonance transition, we shall use a set of canonical coordinates, called Delaunay variables, which make the study of the two-body regime of motion particularly simple, and thus simplify the perturbation arguments for the PCR3BP. Tradition holds that the Delaunay variables in the rotating coordinates are denoted $l, \bar{g}, L$, and G. See Figure 5.2. The quantity G is the angular momentum, while L is related to the semi-major axis a, by $L = a^{1/2}$, and hence encodes the mean motion resonance (with respect to Jupiter in the Sun-Jupiter system). Both l and $\bar{g}$ are angular variables defined modulo 2π. The angle $\bar{g}$ is the argument of the perihelion relative to the rotating axis. The angle l is the mean anomaly. It is the ratio of the area swept out by the ray from the Sun to the comet starting from its perihelion passage to the total area. For more detail, see Szebehely [1967], Abraham and Marsden [1978], and Meyer and Hall [1992].
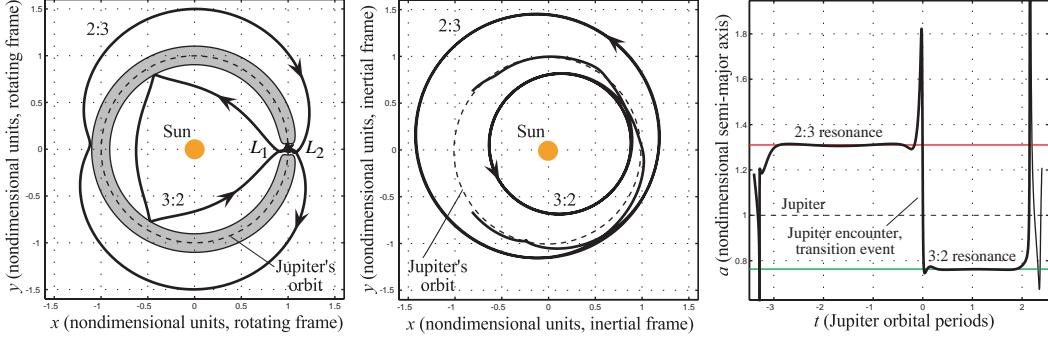
Figure 5.1: (a) The orbit $\mathcal{O}'$, with itinerary (J, X, J, S, J), in the rotating frame. (b) The orbit $\mathcal{O}'$ in the heliocentric inertial frame. (c) Plot of a versus t for the orbit $\mathcal{O}'$. Important mean motion resonances 3:2 and 2:3 are also shown for comparison.
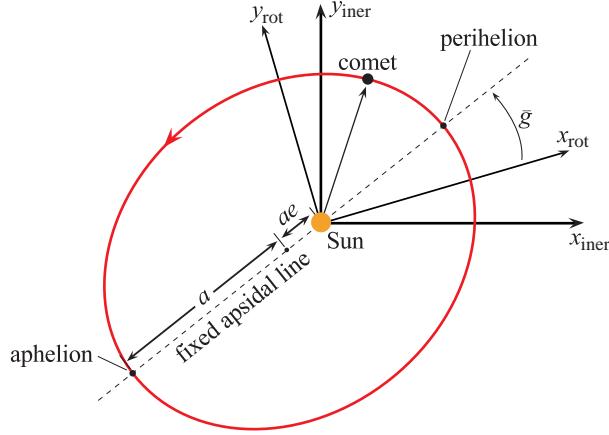

Figure 5.2: Geometry of the Delaunay variables. Elliptical orbits in the fixed (inertial) and rotating frames.

5.2 Interior and Exterior Resonances.

Interior Resonances. Figure 5.3 shows the first Poincaré cuts of the stable and unstable manifolds of an L_1 Lyapunov orbit with the U_1 section $(y = 0, x < 0)$. They have been plotted using Delaunay variables L and $\bar{g}$.

The striking thing is that the first cuts of the stable and unstable manifolds intersect exactly at the region of the 3:2 resonance. Recall that the interior $\Delta_{L_1,1}^{s,\mathcal{S}}$ of $\Gamma_{L_1,1}^{s,\mathcal{S}}$ (the first cut of the stable manifold) is connected to the spherical cap $d_{1,1}^+$ of the bounding sphere $n_{1,1}$ by the stable manifold tube. Hence, $\Delta_{L_1,1}^{s,\mathcal{S}}$ contains all the orbits that will go from the interior (Sun) region to the Jupiter region during the next close approach to the L_1 equilibrium region. Similarly, the interior $\Delta_{L_1,1}^{u,\mathcal{S}}$ of $\Gamma_{L_1,1}^{u,\mathcal{S}}$ (the first cut of the unstable manifold) contains all the orbits that came from the Jupiter region into the interior (Sun) region during their previous close approach to the L_1 equilibrium region. Therefore, their intersection $\Delta^{\mathcal{S}} = \Delta_{L_1,1}^{s,\mathcal{S}} \cap \Delta_{L_1,1}^{u,\mathcal{S}}$ contains all the orbits that have come from the Jupiter region $\mathcal{J}$ into the interior region $\mathcal{S}$, gone around the Sun once (in the rotating frame), and

56
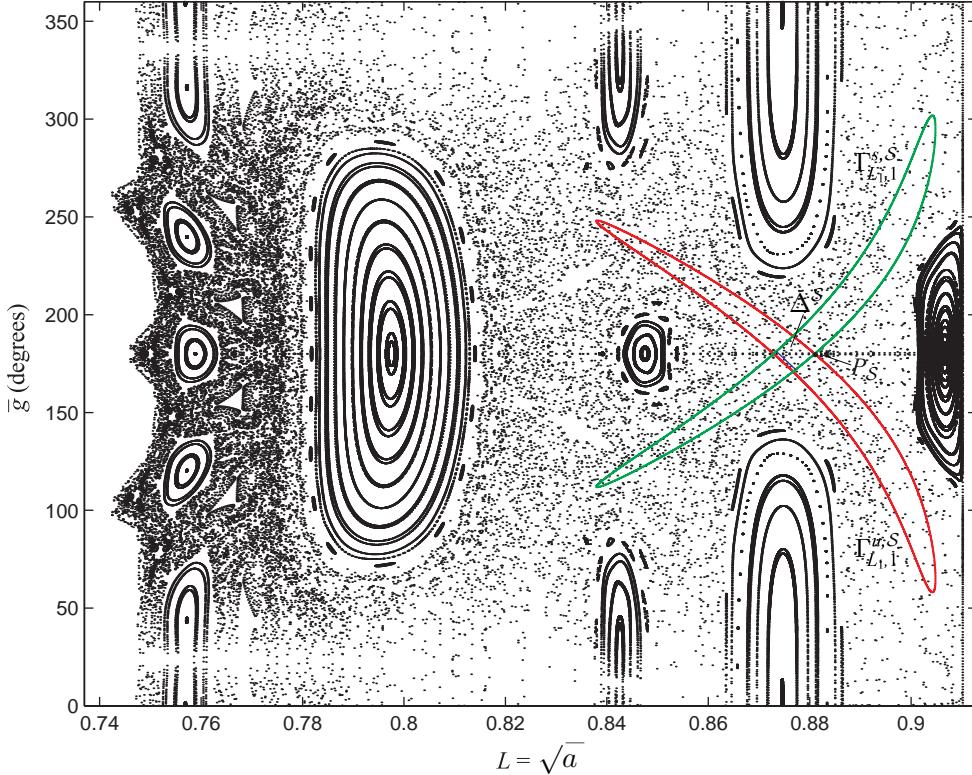
Figure 5.3: The interior region U_1 Poincaré section showing the first cuts of the stable $(\Gamma^{s,\mathcal{S}}_{L_1,1})$ and unstable $(\Gamma^{u,\mathcal{S}}_{L_1,1})$ manifolds of an L_1 Lyapunov orbit. Notice their intersection at the 3:2 resonance. The background points reveal the mixed phase space of stable periodic and quasiperiodic tori "islands" embedded in a bounded chaotic "sea."

will return to the Jupiter region. In the heliocentric inertial frame, these orbits are nearly elliptical outside a neighborhood of L_1. See Figure 5.1. They have a semi-major axis which corresponds to 3:2 resonance by Kepler's law (i.e., $a^{-3/2} = L^{-3} \approx 3/2$). Therefore, any Jupiter comet which has an energy similar to *Oterma*'s and which circles around the Sun once in the interior region must be in 3:2 resonance with Jupiter.

Also note that the point $P_{\mathcal{S}}$, which is on the boundary of $\Delta^{\mathcal{S}}$, is a symmetric (1,1)-homoclinic point which we have used to construct the symmetric (1,1)-homoclinic orbit in Figure 1.2. This also explains the reason for marking it as a homoclinic orbit which corresponds to the 3:2 resonance.

The black background points in Figure 5.3 reveal the character of the interior region phase space for this Jacobi constant surface. They were generated by picking one hundred evenly spaced initial points along the $y = 0, \dot{x} = 0$ line (with the same Jacobi constant $C = 3.03$). These initial points were each integrated for several hundred iterations of the Poincaré map on the U_1 section and then transformed into Delaunay variables.

The background points reveal a mixed phase space of stable periodic and quasiperiodic tori "islands" embedded in a bounded chaotic "sea." The families of stable tori, where a "family" denotes those tori islands which lie along a strip of nearly constant L, correspond to mean motion resonances. The size of the tori island corresponds to the dynamical significance of the resonance. The number of tori islands equals the order of the resonance (e.g.,

3:2 is order 1, 5:3 is order 2). In the center of each island, there is a point corresponding to an exactly periodic, stable, resonant orbit. In between the stable islands of a particular resonance (i.e., along a strip of nearly constant L), there is a saddle point corresponding to an exactly periodic, unstable, resonant orbit. In Figure 5.3, the intersection region $\Delta^{\mathcal{S}}$ is centered on this saddle point for the 3:2 resonance.

A subset of the interior resonance intersection region $\Delta^{\mathcal{S}}$ is connected to exterior resonances through a heteroclinic intersection in the Jupiter region. We have plotted this subset as the small strip inside $\Delta^{\mathcal{S}}$. This subset is part of the dynamical channel which connects the interior and exterior resonances. This is the robust resonance transition mechanism which we have sought. More on the resonance transition will be discussed below.

Exterior Resonances. Similar to Figure 5.3 for the interior region, Figure 5.4 shows the first exterior region Poincaré cuts of the stable and unstable manifolds of an L_2 Lyapunov orbit with the U_4 section on the same Jacobi constant surface ($C = 3.03$). They have been plotted, as before, using the Delaunay variables L and $\bar{g}$.
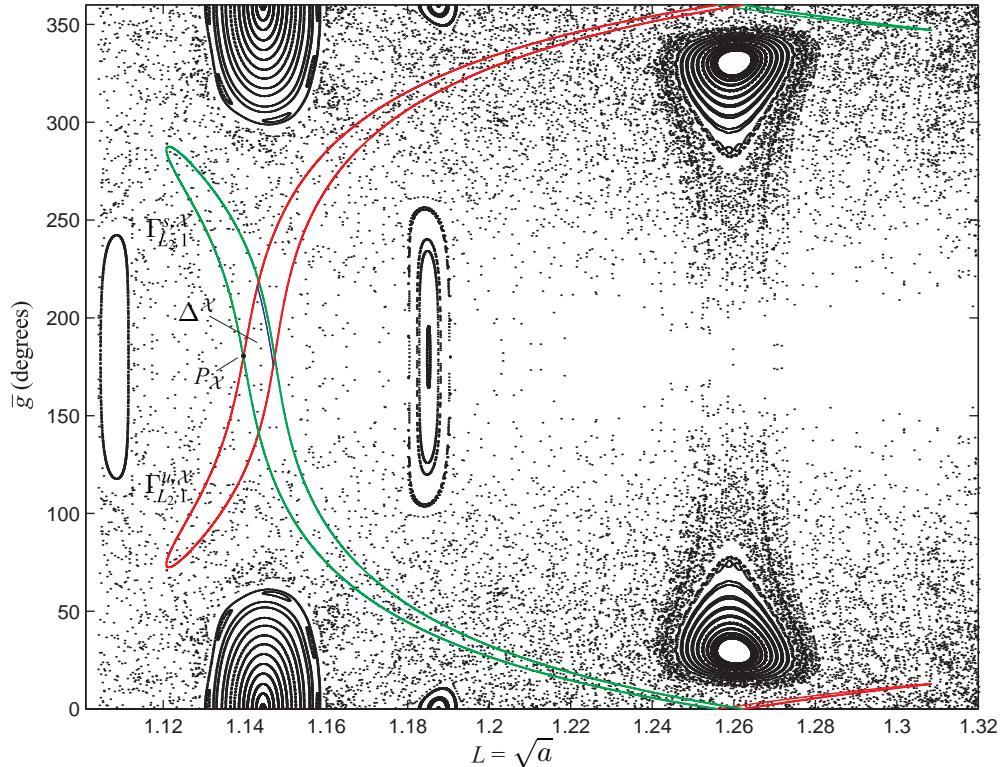

Figure 5.4: The exterior region U_4 Poincaré section showing the first cuts of the stable ($\Gamma^{s,\mathcal{X}}_{L_2,1}$) and unstable ($\Gamma^{u,\mathcal{X}}_{L_2,1}$) manifolds of an L_2 Lyapunov orbit. Notice their intersections at the 2:3 and 1:2 resonances. The background points reveal a mixed phase space similar to that of Figure 5.3.

Notice that the first cuts of the stable and unstable manifolds intersect at two places; one of the intersections is exactly at the region of the 2:3 resonance, the other is at the 1:2 resonance. We would like to point out that $\bar{g}$ is an angle variable modulo 2π and hence the two intersections near $L = 1.26$ should be identified.

Recall that the interior $\Delta^{s,\mathcal{X}}_{L_2,1}$ of $\Gamma^{s,\mathcal{X}}_{L_2,1}$ (the first cut of the stable manifold) is connected

to the spherical cap $d_{2,2}^+$ of the bounding sphere $n_{2,2}$ by the stable manifold tube. Hence, $\Delta_{L_2,1}^{s,\mathcal{X}}$ contains all the orbits that will go from the exterior region to the Jupiter region in the next round. Similarly, the interior $\Delta_{L_2,1}^{u,\mathcal{X}}$ of $\Gamma_{L_2,1}^{u,\mathcal{X}}$ (the first cut of the unstable manifold) contains all the orbits that have come from the Jupiter region into the exterior region in the previous round. Therefore, their intersection

$$\Delta^{\mathcal{X}} = \Delta_{L_2,1}^{s,\mathcal{X}} \cap \Delta_{L_2,1}^{u,\mathcal{X}}$$

contains all the orbits that have come from the Jupiter region $\mathcal{J}$ into the exterior region $\mathcal{X}$, have gone around the Sun once (in the rotating frame), and will return to the Jupiter region. Notice that $\Delta^{\mathcal{X}}$ has two components, one at the 2:3 resonance region and the other at the 1:2 resonance region.

In the heliocentric inertial frame, these orbits are nearly elliptical outside a neighborhood of L_2. They have a semi-major axis which corresponds to either 2:3 or 1:2 resonance by Kepler's law. Therefore, any Jupiter comet which has an energy similar to *Oterma*'s and which circles around the Sun once in the exterior region must be in either 2:3 or 1:2 resonance with Jupiter.

Note that the point $P_{\mathcal{X}}$, which is on the boundary of $\Delta^{\mathcal{X}}$ at the 2:3 resonance region, is a symmetric (1,1)-homoclinic point that we have used to construct the symmetric (1,1)-homoclinic orbit of the exterior region in Figure 1.2. This also explains why we have marked it as a homoclinic orbit which corresponds to the 2:3 resonance.

The background points in Figure 5.4 were generated by a technique similar to those in Figure 5.3. They reveal a similar mixed phase space, but now the resonances are exterior resonances (exterior to the orbit of Jupiter). We see that the exterior resonance intersection region $\Delta^{\mathcal{X}}$ envelops both the 2:3 and the 1:2 unstable resonance points.

A portion of $\Delta^{\mathcal{X}}$ is connected to interior resonances through a heteroclinic intersection in the Jupiter region. In particular, a subset of the 2:3 intersection region of $\Delta^{\mathcal{X}}$ connects to the 3:2 intersection region of $\Delta^{\mathcal{S}}$ via a heteroclinic intersection in the Jupiter region. We have plotted this subset as the small strip inside $\Delta^{\mathcal{X}}$. Note that this strip is the pre-image of the strip in $\Delta^{\mathcal{S}}$ of Figure 5.3. This is the resonance transition dynamical channel shadowed by the Jupiter comet *Oterma* during its recent resonance transition.

5.3   Resonance Transitions.

We have made reference to a heteroclinic intersection connecting the interior $\Delta^{\mathcal{S}}$ and exterior $\Delta^{\mathcal{X}}$ resonance intersection regions. In Figure 5.5, we show the image of $\Delta^{\mathcal{X}}$ (the 2:3 resonance portion) and the pre-image of $\Delta^{\mathcal{S}}$ in the $\mathcal{J}$ region. Their intersection $\Delta^{\mathcal{J}}$ contains all the orbits whose itineraries have the central block $(J, X; J, S, J)$, corresponding to at least one transition between the exterior 2:3 resonance and interior 3:2 resonance. The orbit $\mathcal{O}'$ of Figure 5.1 is such an orbit passing through the region $\mathcal{J}$.

Note the point $P_{\mathcal{J}}$, which lies in the intersection of the boundaries of $\Delta^{\mathcal{X}}$ (the 2:3 resonance portion) and the pre-image of $\Delta^{\mathcal{S}}$. This point $P_{\mathcal{J}}$ corresponds to a heteroclinic connection between the exterior 2:3 and interior 3:2 resonances. In a neighborhood of $P_{\mathcal{J}}$, the dynamical channel connects the 3:2 interior resonance region with the 2:3 exterior resonance region. The periodic orbit $\mathcal{O}$ referred to earlier, which goes from 3:2 to 2:3 and back again ad infinitum, lies in this neighborhood.

The orbit of comet *Oterma* (from 1910 to 1980) also lies in the neighborhood of $P_{\mathcal{J}}$, in the region with itinerary (X, J, S, J, X), as determined from §4.5. *Oterma* does not perform the "exact" exterior to interior homoclinic-heteroclinic resonance transition defined by the sequence (J, X, J, S, J), but as a nearby trajectory (and "nearby" itinerary), it exhibits a similar transient behavior. We note that *Oterma* exhibits only one transition during the
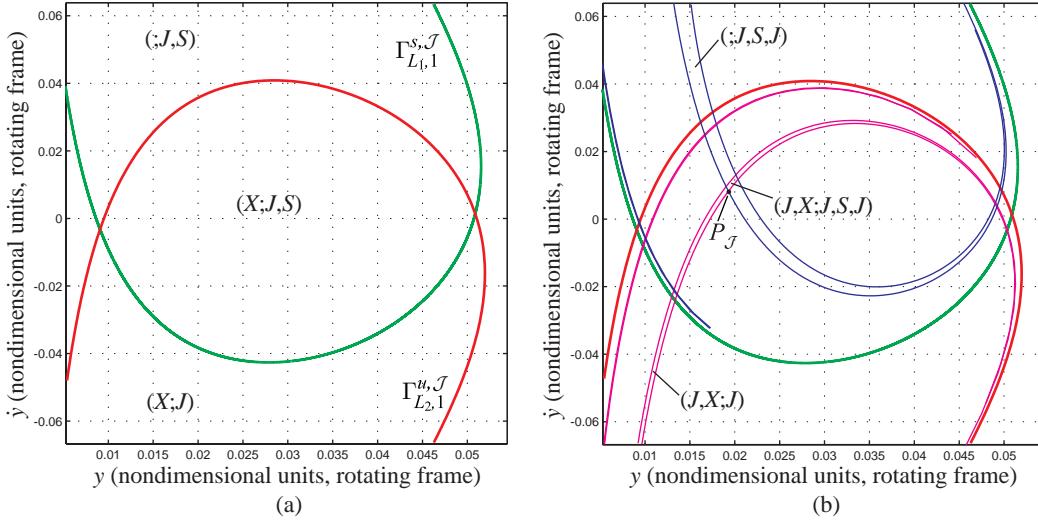
Figure 5.5: The Jupiter region U_3 Poincaré section showing the image of $\Delta^{\mathcal{X}}$ (the 2:3 resonance portion) and the pre-image of $\Delta^{\mathcal{S}}$ (the 3:2 resonance). Notice their intersections, the largest of which is labeled $(J, X; J, S, J)$, corresponding to the itinerary of this group of orbits.

time interval (a few hundred years, centered on the present) for which there is reliable orbit data. It begins in the exterior region close to the 2:3 resonance (i.e., $a^{-3/2} \approx 2/3$), is perturbed by Jupiter into an exactly homoclinic 3:2 resonance (3 revolutions around the Sun in 2 Jupiter periods), and is then nearly symmetrically perturbed into the exterior region, slightly beyond the 2:3 resonance. See Figure 4.9.

It is reasonable to conclude that, within the full three-dimensional model, *Oterma*'s orbit lies within an analogous region of phase space which carries the label (X, J, S, J, X). It is therefore within the L_1 and L_2 manifold tubes, whose complex global dynamics lead to intermittent behavior, including resonance transition.

More study is needed for a thorough understanding of the resonance transition phenomenon. The tools developed in this paper (dynamical channels, symbolic dynamics, etc.) should lay a firm theoretical foundation for any such future studies.

6 Conclusions and Future Work.

In this paper, we have applied dynamical systems techniques to the problem of heteroclinic connections and resonance transitions in the planar circular restricted three-body problem (PCR3BP). One of the main results in this paper is the semi-analytical discovery of a heteroclinic connection between L_1 and L_2 periodic (Lyapunov) orbits having the same energy (Jacobi constant). 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 LMS [1985]. By linking these heteroclinic connections with homoclinic orbits on the same Jacobi constant surface, we have found dynamical channels that provide a fast transport mechanism between the interior and exterior Hill's regions. This rapid transport mechanism, which occurs with only two degrees of freedom, is a dynamical systems phenomenon not to be confused with Arnold diffusion.

The channels provide a starting point for understanding the transport mechanisms con-

necting mean motion resonances, and in particular, those mechanisms which link interior and exterior resonances (e.g., the 3:2 and 2:3 Jupiter resonances) via the Jupiter capture region. By comparing observations of the orbits of Jupiter comets like *Oterma* with the dynamical channels discovered herein, we conclude that the comets are guided by these dynamical channels. See Figures 1.2 and 4.9.

Moreover, these dynamical channels could be exploited by spacecraft to explore a large region of space near Earth (and near Earth's orbit) using low-fuel controls. In fact, the channels can be utilized around any planet or moon system. Behavior related to the dynamical channels has already been observed by Lo, Williams, *et al.* [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 6.1. With 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).
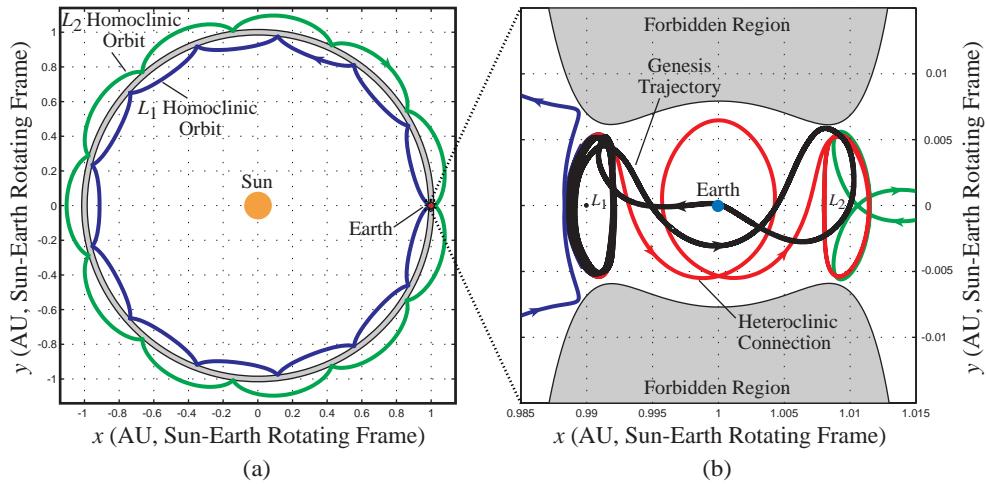

Figure 6.1: (a) A homoclinic-heteroclinic chain on the *Genesis Discovery Mission* trajectory's energy surface. (b) Close-up of the chain in Earth's vicinity. The actual *Genesis Discovery Mission* trajectory is shown in black overlaying the chain, and in particular, the heteroclinic connection from L_1 to L_2.

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.

Extension to Three Dimensions. The natural extension of our work is to apply the same methodology to the three-dimensional CR3BP. We will seek homoclinic and heteroclinic orbits associated with three-dimensional periodic "halo" and quasi-periodic "quasi-halo" and Lissajous orbits about L_1 and L_2. Their union would be three-dimensional homoclinic-heteroclinic chains around which the symbolic dynamics could be used to track a variety of exotic orbits.

The three-dimensional chains would provide an initial template for the construction of actual spacecraft trajectories. By presenting a more complete portrait of the phase space geometry near L_1 and L_2, the three-dimensional channels will be of enormous benefit in the de-

sign and control of constellations of spacecraft in these regions. The homoclinic-heteroclinic structures suggest natural low-fuel paths for deployment of constellation spacecraft to and from Earth. They will aid in the design of control schemes necessary for space missions such as NASA's Terrestrial Planet Finder (TPF) which must maintain precise coordinated pointing and relative separation of the formation flying spacecraft.

The three-dimensional dynamical channels may also provide a more complete understanding of phase space transport mechanisms. In particular, they may elucidate the resonance transition process for Jupiter comets which have large excursions out of Jupiter's orbital plane.

Coupling of Two Three-Body Systems. To obtain a better grasp of the dynamics governing transport between adjacent planets (or moons), we could apply our methodology to the coupled PCR3BP. The coupled PCR3BP considers two nested co-planar three-body systems, such as for two adjacent giant planets competing for control of the same comet (e.g., Sun-Jupiter-comet and Sun-Saturn-comet). When close to the orbit of one of the planets, the comet's motion is dominated by the corresponding planet's three-body dynamics. Between the two planets, the comet's motion is mostly heliocentric, but is precariously poised between two competing three-body dynamics. In this region, heteroclinic orbits connecting Lyapunov orbits of the two different three-body systems may exist, leading to complicated transfer dynamics between the two adjacent planets. See Figure 1.3.

This transfer dynamics, which may be realized in actual comet behavior, could be exploited for free transfers of spacecraft between adjacent moons in the Jovian and Saturnian systems (Lo and Ross [1998]). For instance, one could conduct a "Petit Grand Tour" of the Jovian moon system, an example of which is shown in Figure 1.3. By systematically seeking heteroclinic connections between libration point orbits of adjacent moons, one could design trajectories which transfer from the vicinity of one moon to another using fuel-minimizing controlled thrusts.

Merging Optimal Control and Stabilization with Dynamical Systems Theory. The construction of exotic spacecraft orbits using homoclinic-heteroclinic dynamical channels requires optimal thruster controls to navigate these dynamically sensitive regions of phase space. Using optimal, fuel minimizing impulsive and continuous thrust, is the most efficient and natural way to take advantage of the delicate dynamics.

Lawden [1963] developed Primer Vector Theory, the first successful application of optimal control theory to minimize fuel consumption for trajectories with impulsive thrusts in the two-body problem. The extension of Primer Vector Theory to continuous low-thrust control for the restricted three-body problem is a current area of active research. Our work on this problem indicates that developing optimal control theory within the dynamical systems framework shows promise for producing a numerical solution in the three-body context.

In our ongoing effort to use the methods of optimal controls to study the orbit transfer problem for certain JPL space missions, we are exploring the "direct" method for solving the optimal control problem. In the direct method, the optimal control problem can be first approximated by a discrete optimization problem using a collocation or multiple shooting discretization scheme. Then the resulting optimization problem is solved numerically with a sophisticated sequential quadratic programming (SQP) technique. While the numerical algorithm of the direct method is quite robust for certain types of two-body problems, we do not expect that application to the three-body regime will be completely straightforward. It would also be interesting to explore the ways in which optimal control in the presence of mechanics (as in, for example, Koon and Marsden [1997]) is useful in this problem.

As usual, for any numerical algorithm, a good initial guess is vital, especially if the problem is very sensitive numerically. Dynamical systems theory can provide geometrical

insight into the structure of the problem and even good approximate solutions. For example, in finding low-thrust optimal transfers to L_1 halo orbits in the Sun-Earth system, it is important to know that the invariant manifolds of the halo orbits extend to the vicinity of the Earth and any trajectory on these manifolds can be used as a super-highway for free rides to and from the halo orbits. See Figure 6.2.
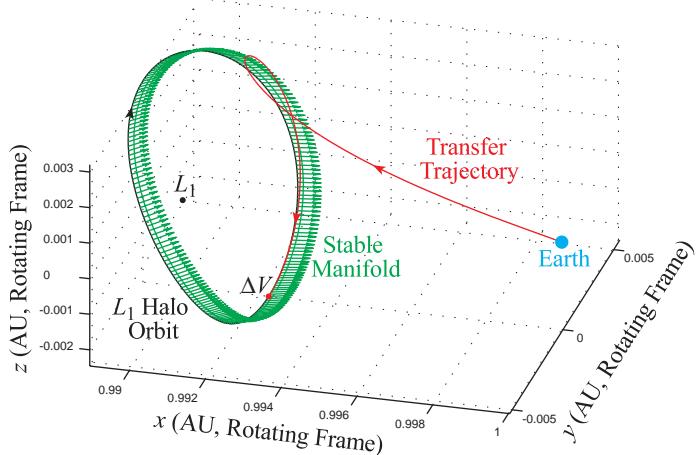

Figure 6.2: A transfer trajectory from low Earth orbit to an L_1 halo orbit. This trajectory was constructed using the stable manifold of the halo orbit. The arrows attached to the halo orbit point in the direction of the stable manifold.

Clearly, this theoretical insight and its derivative numerical tools can aid in the construction of superior initial guesses that lead to a convergent solution.

A deeper understanding of the dynamical structure of the restricted three-body problem, including the ideas we have contributed in this paper, may suggest alternative formulations of the optimizing scheme which are based more on the geometry of the phase space. Instead of "numerically groping in the dark", algorithms could be developed with the natural dynamics built in, thereby yielding better convergence properties.

In addition to the optimal control problem of getting to a halo orbit, there are well known techniques for stabilizing the dynamics once one gets there. Some of these techniques are related to the general theory of stabilizing dynamics near saddle points and homoclinic or heteroclinic orbits, as in Bloch and Marsden [1989]. In addition, it would be of interest to explore the use of other stabilization techniques that make use of the mechanical structure for problems of this sort, as in Bloch, Leonard and Marsden [1997].

Symplectic Integrators. The use of symplectic integrators for the long time integrations of the solar system is well known through the work of Tremaine, Wisdom and others. In many problems in which the dynamics is delicate or where there are delicate controls, care is needed with integration algorithms. The area of integration algorithms for mechanical systems continues to develop and be implemented; see for example, Wendlandt and Marsden [1997], Kane, Marsden and Ortiz [1999], and Kane, Marsden, Ortiz and West [1999] and references therein. These techniques are very effective for both conservative mechanical systems as well as systems with forcing, such as controlled systems. It would be of interest to explore these numerical methods in the context of space mission design and other orbital mechanics problems.

Pattern Evocation. The resonant structures that one sees in the rotating frames of interest in the present paper appear similar to what one sees in the phenomenon of *pattern evocation* (see Marsden and Scheurle [1995], Marsden, Scheurle and Wendlandt [1996]) when rotationally symmetric systems are viewed from the point of view of an appropriate rotating frame. Of course for the restricted three body problem there is a simple and natural choice of a rotating frame. However, for the full three body problem or other situations, the general theory still suggests that appropriate rotating frames can be found relative to which simple resonant phenomena would be evoked. It would be of interest to explore this link further.

Four or More Body Problems. While the planar CR3BP model provides an adequate explanation for a class of Jupiter comets whose Jacobi constant is close to (and less than) C_2 and whose motion is close to the plane of Jupiter's orbit, it fails to explain resonance transition phenomena for high inclination Jupiter comets and comets not dominated solely by Jupiter. For this second class of comets, other effects such as out-of-plane motion and perturbation by other giant planets, most notably Saturn, are quite strong and need to be considered. Though the Jupiter comets exhibit their transitions on relatively short time-scales (tens to hundreds of years), rare terrestrial planet encounters (with Earth and Mars) also need to be considered. In short, the study of this second class of comets require the complete storehouse of tools needed in the study of the near-Earth asteroids, regarded by many as the most challenging topic in celestial mechanics.

However, since the mean motion resonances (mostly with Jupiter) and their associated transport mechanisms still play the dominant role in solar system material transport, this paper can be seen as laying a firm foundation for any future studies in this direction. We may need to consider other more complicated models like the full three-dimensional CR3BP and the coupled PCR3BP as mentioned above. As Lo and Ross [1997] suggested, further exploration of the phase space structure as revealed by the homoclinic-heteroclinic structures and their association with mean motion resonances may provide deeper conceptual insight into the evolution and structure of the asteroid belt (interior to Jupiter) and the Kuiper Belt (exterior to Neptune), plus the transport between these two belts and the terrestrial planet region. See Figure 6.3(a).

Potential Earth-impacting asteroids may utilize the dynamical channels as a pathway to Earth from nearby, seemingly harmless heliocentric orbits which are in resonance with the Earth. The same dynamics which allows us to construct libration point space missions such as the *Genesis Discovery Mission*, which is on a natural Earth collision orbit, is also the dynamics that could bring unexpected Earth impactors. This phenomena has been observed recently in the impact of comet *Shoemaker-Levy 9* with Jupiter, which was in 2:3 resonance with Jupiter (one of the resonances dynamically connected to the Jupiter region) just before impact.

Zodiacal Dust Cloud. Numerical simulations of the orbital evolution of asteroidal dust particles show that the Earth is embedded in a circumsolar ring of asteroidal dust known as the zodiacal dust cloud (Dermott *et al.* [1994]). Both simulations and observations reveal that the zodiacal dust cloud has structure. When viewed in the Sun-Earth rotating frame, there are several high density clumps ($\sim$10% greater than the background) which are mostly evenly distributed throughout the Earth's orbit. The simulations of Dermott *et al.* [1994] considered the gravitational effects of the actual solar system and non-gravitational forces: radiation pressure, Poynting-Robertson light drag, and solar wind drag. The dust particles are believed to spiral in towards the Sun from the asteroid belt, becoming trapped temporarily in exterior mean motion resonances with the Earth. They are then scattered by close encounters with the Earth leading to further spiraling towards, and eventual collision with, the Sun.
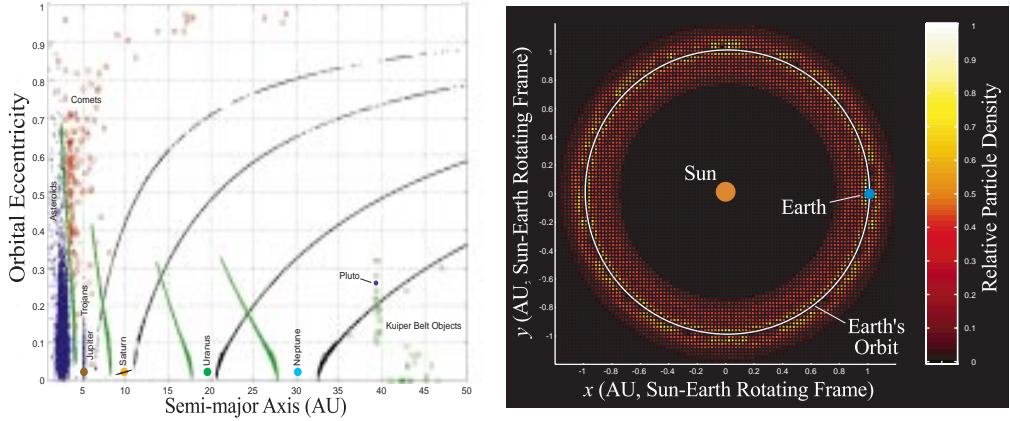
Figure 6.3: (a) Dynamical channels in the solar system. We plot the (local) semi-major axis versus the orbital eccentricity. We show the L_1 (grey) and L_2 (black) manifolds for each of the giant outer planets. Notice the intersections between manifolds of adjacent planets, which leads to chaotic transport. Also shown are the asteroids (dots), comets (circles), and Kuiper Belt objects (lighter circles). (b) The zodiacal dust ring around the Earth's orbit, as modeled by Earth's L_1 and L_2 stable and unstable manifolds. We show the Sun-Earth rotating frame. Notice the "clumps" in Earth's orbit.

We suspect that the gross morphology of the ring is given by a simpler CR3BP model involving the homoclinic and heteroclinic structures (the dynamical channels) associated with L_1 and L_2 (Lo and Ross [1997]). See Figure 6.3(b).

The drag forces do not destroy the dynamical channel structure, but instead seem to lead to convergence onto the structure for particles spiraling in from the inner asteroid belt. Once trapped in a channel, the dynamics naturally lead to transport (via an Earth encounter) into the interior region, where drag forces dominate once more.

As with the Earth, the structure of any extrasolar terrestrial planet's zodiacal dust ring is probably dominated by the three-body dynamics. As the particular features of the ring structure (i.e., width of ring, number of high density clumps) are characteristic of the particular mass ratio of the planet to the star, one could use the structure observed in an extrasolar zodiacal dust ring to determine the mass of the planet, assuming the mass of the star could be determined using spectroscopic methods. The Terrestrial Planet Finder mission could use such a scheme to detect terrestrial planets embedded in the zodiacal dust rings of nearby stars.

A New Paradigm for a New Millennium. A century has passed since Poincaré introduced dynamical systems theory to study the restricted three-body problem. Yet this system still enchants us with its rich structure and dazzling spectrum of behaviors despite its deceptively simple formulation. With the fundamental dynamical systems tools developed herein, we stand poised to appreciate and utilize this rich structure in ways Poincaré could only imagine.

65

References

Abraham, R. and J.E. Marsden [1978] *Foundations of Mechanics.* Second Edition, Addison-Wesley.

Belbruno, E and B. Marsden [1997] Resonance hopping in comets, *The Astronomical Journal* **113(4)**, 1433-1444.

Belbruno, E and J. Miller [1993] Sun-perturbed earth-to-moon transfers with ballistic capture, *Journal of Guidance, Control, and Dynamics* **16**, 770-775.

Bloch, A.M., P.S. Krishnaprasad, J.E. Marsden, and G. Sánchez de Alvarez [1992] Stabilization of rigid body dynamics by internal and external torques, *Automatica* **28**, 745–756.

Bloch, A.M. and J.E. Marsden [1989] Controlling homoclinic orbits, *Theoretical and Computational Fluid Mechanics* **1**, 179–190.

Bloch, A.M., N. Leonard and J.E. Marsden [1997] Stabilization of mechanical systems using controlled Lagrangians, *Proc CDC* **36**, 2356–2361.

Conley, C. [1963] On some new long periodic solutions of the plane restricted three body problem. *Comm. Pure Appl. Math.* **16**, 449–467.

Conley, C. [1968] Low energy transit orbits in the restricted three-body problem. *SIAM J. Appl. Math.* **16**, 732–746.

Dermott, S. *et al.* [1994] A circumsolar ring of asteroidal dust in resonant lock with the Earth, *Nature* **369**, 719-723.

Devaney, R. L. [1981] Singularities in classical mechanical systems, *Ergodic Theory and Dynamical Systems* (A. Katok, ed.), 211-333, Birkhäuer, Basel.

Enright, P. J. and B. A. Conway [1992] Discrete approximations to optimal trajectories using direct transcription and nonlinear programming, *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 4, 1992, pp. 994-1002.

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)*, Barcelona, April, 1991.

Herman, A. L. and B.A. Conway [1996] Direct optimization using collocation based on high-order Gauss-Lobatto quadrature rules, *Journal of Guidance, Control, and Dynamics*, **19**, 592–599.

Hiday-Johnston, L. and Howell, K. [1994] Transfers Between Libration-Point Orbits in the Elliptic Restricted Problem, *Celestial Mechanics and Dynamical Astronomy* **58**, 317–337.

Holmes, P. [1990] Poincareé, celestial mechanics, dynamical systems theory and chaos. *Physics Reports* **193**, 3 (1990), 137-163.

Howell, C., 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(2)**, April 1997, 161-178.

Kane, C, J.E. Marsden, and M. Ortiz [1999] Symplectic energy momentum integrators (*J. Math. Phys., to appear*).

Kane, C, J.E. Marsden, M. Ortiz and M. West [1999] Variational Integrators and the Newmark Algorithm for Conservative and Dissipative Mechanical Systems (*preprint*).

Koon, W.S. and J.E. Marsden [1997] Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim.* **35**, 901-929.

Lawden, D.F. [1963] *Optimal Trajectories for Space Navigation.* Butterworth, London.

Llibre, J., R. Martinez and C. Simó [1985] Transversality of the invariant manifolds associated to the Lyapunov family of perodic orbits near L2 in the restricted three-body problem, *Journal of Differential Equations* **58**, 104-156.

Lo, M. and S. Ross [1997] SURFing the Solar System: Invariant Manifolds and the Dynamics of the Solar System *JPL IOM 312/97*, 2-4.

Lo, M. and S. Ross [1998] Low energy interplanetary transfers using invariant manifolds of L1, L2 and halo orbits, *AAS/AIAA* Space Flight Mechanics Meeting, Monterey, California, 9-11 February 1998.

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, B. Barden, and R. Wilson [1998] Genesis Mission Design, *Paper No. AIAA 98-4468*.

Marsden, J.E. [1992], *Lectures on Mechanics* London Mathematical Society Lecture note series, **174**, Cambridge University Press.

Marsden, J.E., O.M. O'Reilly, F.J. Wicklin, and B.W. Zombro [1991] Symmetry, stability, geometric phases, and mechanical integrators, *Nonlinear Science Today* **1**, 4–11, **1**, 14–21.

Marsden, J.E. and T.S. Ratiu [1999] *Introduction to Mechanics and Symmetry.* Texts in Applied Mathematics, **17**, Springer-Verlag, 1994. Second Edition, 1999.

McCord, C.K., K.R. Meyer and Q. Wang [1998] *The Integral Manifolds of the Three Body Problem*, Memoirs of the Am. Math. Soc. **628**.

McGehee, R. P. [1969] Some homoclinic orbits for the restricted three-body problem, Ph.D. thesis, University of Wisconsin.

McGehee, R. P. [1973] Parabolic orbits in the three-body problem. Dynamical systems *Proc. Sympos., Univ. Bahia, Salvador, 1971* 249–254. Academic Press, New York.

McGehee, R. P. [1974] Triple collision in the collinear three-body problem. *Invent. Math.* **27**, 191–227.

Easton, R. W. and McGehee, R. [1979] Homoclinic phenomena for orbits doubly asymptotic to an invariant three-sphere. *Indiana Univ. Math. J.* **28**, 211–240.

Marsden, J.E. and J. Scheurle [1995] Pattern evocation and geometric phases in mechanical systems with symmetry, *Dyn. and Stab. of Systems.* **10**, 315–338.

Marsden, J.E., J. Scheurle and J. Wendlandt [1996] Visualization of orbits and pattern evocation for the double spherical pendulum, *ICIAM 95: Mathematical Research, Academie Verlag, Ed. by K. Kirchgässner, O. Mahrenholtz and R. Mennicken* **87**, 213–232.

Meyer, K.R. and R. Hall [1992] *Hamiltonian Mechanics and the n-body Problem.* Springer-Verlag, Applied Mathematical Sciences.

Moser, J. [1973] *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics.* Princeton University Press.

Simó, C. [1990] Dynamical systems methods for space missions on a vicinity of collinear libration points, *Hamiltonian Systems with Three or More Degrees of Freedom* (ed. C. Simó), 223-241.

Szebehely, V. [1967] *Theory of Orbits.* Academic Press, New York/London.

Wendlandt, J.M. and J.E. Marsden [1997] Mechanical integrators derived from a discrete variational principle, *Physica D* **106**, 223–246.

Wiggins, S. [1990] *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* Texts Appl. Math. Sci. **2**, Springer-Verlag.

Wiggins, S. [1993] Global dynamics, phase space transport, orbits homoclinic to resonances, and applications, *Fields Institute Monographs, American Mathematical Society*, 1993.