GEOMETRIC DERIVATION OF THE DELAUNAY VARIABLES AND GEOMETRIC PHASES*

DONG EUI CHANG[1] and JERROLD E. MARSDEN[2]
[1]*Mechanical and Environmental Engineering, University of California, Santa Barbara, CA 93106, U.S.A., e-mail: dchang@engineering.ucsb.edu*
[2]*Control and Dynamical Systems 107-81, California Institute of Technology, Pasadena, CA 91125, U.S.A., e-mail: marsden@cds.caltech.edu*

(Received: 20 December 2001; revised: 27 September 2002; accepted: 2 November 2002)

Dedicated to Klaus Kirchgässner on the occasion of his 70th birthday

Abstract. We derive the classical Delaunay variables by finding a suitable symmetry action of the three torus $\mathbb{T}^3$ on the phase space of the Kepler problem, computing its associated momentum map and using the geometry associated with this structure. A central feature in this derivation is the identification of the mean anomaly as the angle variable for a symplectic S^1 action on the union of the non-degenerate elliptic Kepler orbits. This approach is geometrically more natural than traditional ones such as directly solving Hamilton–Jacobi equations, or employing the Lagrange bracket. As an application of the new derivation, we give a singularity free treatment of the averaged J_2-dynamics (the effect of the bulge of the Earth) in the Cartesian coordinates by making use of the fact that the averaged J_2-Hamiltonian is a collective Hamiltonian of the $\mathbb{T}^3$ momentum map. We also use this geometric structure to identify the drifts in satellite orbits due to the J_2 effect as geometric phases.

Key words: Kepler vector field, derivation of variables, orbits dynamics and phases

1. Introduction

The purpose of this paper is to prove the following well-known theorem from a viewpoint of geometric mechanics and then to make use of this technique to study the averaged Hamiltonian in the J_2 problem. As a by-product, we also obtain a new interpretation and derivation of known drifts in satellite orbits due to the bulge of the earth as geometric phases.

We now state the theorem about Delaunay variables informally; a more complete statement is given in Proposition 3.2.

THEOREM 1.1. *In the Kepler problem, there is a natural action of the three torus, $\mathbb{T}^3$, on Σ_{elliptic}, the union of nondegenerate Keplerian elliptic orbits, such that the variables of $\mathbb{T}^3$ and its momentum map constitute the Delaunay variables on the set of regular points of the momentum map.*

*This research was partially supported by the Humboldt Foundation and a Max-Planck Research Award and the California Institute of Technology.

The Delaunay variables were first introduced in Delaunay (1860) and have been frequently used as canonical variables in celestial mechanics. In particular, they are action–angle variables. There are two common derivations of these variables; one is to directly solve the Hamilton–Jacobi equations (Born, 1927; Kovalevsky, 1967) and the other is to use the Lagrange brackets (Brouwer and Clemence, 1961; Abraham and Marsden, 1978). The former, being rather brute-force, does not give very much geometric insight. The latter employs the fact that Hamiltonian flows are canonical, and it simplifies computations by evaluating the Lagrange bracket at perigee. Nevertheless, this derivation still lacks the geometric insight that one would like.

In this paper, we give a new derivation of the Delaunay variables; this derivation is straightforward from the geometric mechanics point of view. We derive the coordinates for the Delaunay variables as follows: we first find a $\mathbb{T}^3$ symmetry in the Kepler problem and compute its momentum map. Second, we show that the set of the regular points of the momentum map is a trivial $\mathbb{T}^3$ principal bundle over the set of regular values of the momentum map. The variables of $\mathbb{T}^3$ and its momentum map constitute the Delaunay variables in the set of regular points of the momentum map. Finally, we show that the Delaunay variables so derived are, in fact, action–angle variables (and thus are canonical variables) by a straightforward computation.

The main point in this derivation lies in the new interpretation of the mean anomaly as a symplectic S^1 action on the union of the nondegenerate elliptic Kepler orbits. We compare this new derivation with the general method of finding action–angle variables in Born (1927) and Arnold (1991). In the final section, we apply this new derivation to the study of the averaged J_2 dynamics, that is, the perturbed Kepler motion due to the bulge of the Earth. In particular, this geometric set-up of the problem allows us to interpret and derive the well-known and important phase drifts in satellite orbits as geometric phases.

2. Three Anomalies: Mean, True, and Eccentric

In the Kepler problem, there are three well-known anomalies: the mean, true, and eccentric anomalies. They are measured from the perigee of a given elliptic orbit, so they are not well-defined in a neighborhood of circular orbits. In this section, we will give the three anomalies a new interpretation in terms of the S^1 actions on the union of nondegenerate elliptic orbits. This interpretation is interesting because it has no singularity problems. We also show that the S^1 action corresponding to the mean anomaly is symplectic.

The Kepler vector field. Let $\mathbb{R}^3_0 = \mathbb{R}^3\backslash\{(0, 0, 0)\}$, the configuration manifold. We use $(\mathbf{q}, \mathbf{p})$ as coordinates for the tangent space $T\mathbb{R}^3_0 = (\mathbb{R}^3\backslash\{(0, 0, 0)\}) \times \mathbb{R}^3$, which is identified with the cotangent bundle using the Euclidean inner product.

The manifold $T\mathbb{R}^3_0$ is equipped with the canonical symplectic structure which we write as

$$\Omega = \sum_{i=1}^{3} \mathrm{d}q^i \wedge \mathrm{d}p_i.$$

Since the canonical symplectic form is nondegenerate, it induces the map $\Omega^\sharp$: $T^*T\mathbb{R}^3_0 \to TT\mathbb{R}^3_0$ defined by $\Omega(\Omega^\sharp\alpha, v) = \alpha(v)$ for all $\alpha \in T^*T\mathbb{R}^3_0$ and $v \in TT\mathbb{R}^3_0$. The map $\Omega^\sharp$ can be written in matrix form as

$$\Omega^\sharp = \begin{pmatrix} O_3 & I_3 \\ -I_3 & O_3 \end{pmatrix},$$

where O_3 is the 3×3 zero matrix and I_3 is the 3×3 identity matrix. The Hamiltonian vector field X_F of a Hamiltonian F is defined by $X_F = \Omega^\sharp\, \mathbf{d}F$. The Kepler Hamiltonian H on $T\mathbb{R}^3_0$ (for a satellite whose mass we can take to be unity) is defined by

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\|\mathbf{p}\|^2 - \frac{\mu}{\|\mathbf{q}\|}, \tag{2.1}$$

where μ is a constant (the mass of the primary times the gravitational constant) and $\|\cdot\|$ is the usual Euclidean norm on $\mathbb{R}^3$. The *Kepler vector field X_H* is, by definition, the Hamiltonian vector field of the Hamiltonian H and is given explicitly by

$$X_H(\mathbf{q}, \mathbf{p}) = \left(\mathbf{p}, -\frac{\mu}{\|\mathbf{q}\|^3}\mathbf{q}\right).$$

Let φ_t be the flow of the Kepler vector field X_H; we also write $\varphi_t(\mathbf{q}, \mathbf{p}) = \varphi(t, (\mathbf{q}, \mathbf{p}))$. Besides H, the Kepler motion has the following constants of motion:

$$\mathbf{L}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p}, \tag{2.2}$$

$$\mathbf{A}(\mathbf{q}, \mathbf{p}) = \mathbf{p} \times (\mathbf{q} \times \mathbf{p}) - \mu\frac{\mathbf{q}}{\|\mathbf{q}\|}, \tag{2.3}$$

where $\mathbf{L}$ is the angular momentum and $\mathbf{A}$ is the Laplace (also known as the Runge–Lenz) vector. They satisfy the following (readily verified) relations:

$$\mathbf{L} \cdot \mathbf{A} = 0, \tag{2.4}$$

$$\|\mathbf{A}\|^2 = \mu^2 + 2H\|\mathbf{L}\|^2. \tag{2.5}$$

The set Σ_{elliptic} is defined to be the union of the nondegenerate elliptic Kepler orbits. It is given in terms of the Hamiltonian and the angular momentum by

$$\Sigma_{\text{elliptic}} = \{(\mathbf{q}, \mathbf{p}) \in T\mathbb{R}^3_0 \mid H(\mathbf{q}, \mathbf{p}) < 0, \mathbf{L}(\mathbf{q}, \mathbf{p}) \neq \mathbf{0}\}, \tag{2.6}$$

which is well-known; see for instance, Cushman and Bates (1997) or Chang et al. (2002) for the proof. The flow φ_t induces a flow on the space Σ_{elliptic} and of course is defined for all $t \in \mathbb{R}$.

S^1 Actions and anomalies. There are three anomalies frequently used in celestial mechanics; the mean, true and eccentric anomalies. We will give a new interpretation of them as S^1 group actions[1] on the set Σ_{elliptic}. The S^1 actions related to the three anomalies come from time reparameterizations of the elliptic Kepler flow φ_t restricted to Σ_{elliptic}. To avoid the repetition of similar arguments, we first construct a general setting. Let $T : \Sigma_{\text{elliptic}} \to \mathbb{R}$ be the period of the Kepler flow, which is well-known to be given by

$$T(\mathbf{q}, \mathbf{p}) = \frac{2\pi\mu}{(-2H(\mathbf{q}, \mathbf{p}))^{3/2}}. \tag{2.7}$$

Suppose that a *positive* smooth function $F : \Sigma_{\text{elliptic}} \to \mathbb{R}$ satisfies

$$2\pi = \int_0^{T(\mathbf{q},\mathbf{p})} F \circ \varphi(s, (\mathbf{q}, \mathbf{p}))\, \mathrm{d}s; \tag{2.8}$$

specific F's satisfying this condition will be chosen shortly. Define $h : \mathbb{R} \times \Sigma_{\text{elliptic}} \to \mathbb{R} \times \Sigma_{\text{elliptic}}$ by

$$h(t, (\mathbf{q}, \mathbf{p})) = (\theta(t, (\mathbf{q}, \mathbf{p})), (\mathbf{q}, \mathbf{p})),$$

where

$$\theta(t, (\mathbf{q}, \mathbf{p})) = \int_0^t F \circ \varphi(s, (\mathbf{q}, \mathbf{p}))\, \mathrm{d}s. \tag{2.9}$$

One can verify that h is a global diffeomorphism. Let $\mathrm{pr} : \mathbb{R} \times \Sigma_{\text{elliptic}} \to S^1 \times \Sigma_{\text{elliptic}}$ be the covering map given by

$$\mathrm{pr}(\theta, (\mathbf{q}, \mathbf{p})) = (\mathrm{e}^{\mathrm{i}\theta}, (\mathbf{q}, \mathbf{p})).$$

Define an S^1 action, $\Phi : S^1 \times \Sigma_{\text{elliptic}} \to \Sigma_{\text{elliptic}}$ by $\Phi \circ \mathrm{pr} = \varphi \circ h^{-1}$; that is,

$$\Phi(\mathrm{e}^{\mathrm{i}\theta}, (\mathbf{q}, \mathbf{p})) = \varphi \circ h^{-1}(\theta, (\mathbf{q}, \mathbf{p})). \tag{2.10}$$

One can check that Φ is indeed a well-defined smooth group action and that this action is free and proper. The infinitesimal generator of this action is given by

$$Y(\mathbf{q}, \mathbf{p}) := \left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=0} \Phi(\mathrm{e}^{\mathrm{i}\theta}, (\mathbf{q}, \mathbf{p})) = \frac{1}{F(\mathbf{q}, \mathbf{p})} X_H(\mathbf{q}, \mathbf{p}). \tag{2.11}$$

[1] An action of a Lie group G on a manifold M is a smooth mapping $\Phi : G \times M \to M$ such that (i) for all $x \in M$, $\Phi(e, x) = x$ and (ii) for every $g, h \in G$, $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$ for all $x \in M$, where e is the identity element of G. An action is called *free* if, for each $x \in M$, $g \mapsto \Phi(g, x)$ is one-to-one. An action is called *proper* if and only if $\tilde{\Phi} : G \times M \to M \times M$ defined by $\tilde{\Phi}(g, x) = (x, \Phi(g, x))$ is a proper mapping, that is, if $K \subset M \times M$ is compact, then $\tilde{\Phi}^{-1}(K)$ is compact. See, for example, Chapter 9 of Marsden and Ratiu (1999) for more on Lie group actions.

Choices for the three anomalies. Consider the following three positive smooth functions on Σ_{elliptic}:

$$F_1(\mathbf{q}, \mathbf{p}) = \frac{2\pi}{T(\mathbf{q}, \mathbf{p})}, \tag{2.12}$$

$$F_2(\mathbf{q}, \mathbf{p}) = \frac{\|\mathbf{L}(\mathbf{q}, \mathbf{p})\|}{\|\mathbf{q}\|^2}, \tag{2.13}$$

$$F_3(\mathbf{q}, \mathbf{p}) = \frac{\sqrt{-2H(\mathbf{q}, \mathbf{p})}}{\|\mathbf{q}\|}. \tag{2.14}$$

The choice of them comes from the following well-known elementary facts on the Kepler motion. The function F_1 is just the average of the time over one period. The function F_2 comes from the polar expression of the angular momentum, that is, $\|\mathbf{L}\| = r^2\dot{f}$ where r is the polar distance and f is the polar angle or the true anomaly in the orbit plane where the origin is at one of the foci of the elliptic orbit. The function F_3 comes from the trigonometric relation between the true anomaly and the eccentric anomaly, which can be found in pp. 21–22 of Vinti (1998). One can readily check that the three functions F_i, $i = 1, 2, 3$ satisfy (2.8) and that $\theta(t, (\mathbf{q}, \mathbf{p}))$ given by (2.9) with the choices F_i with $i = 1, 2, 3$, corresponds to *the mean anomaly*, *the true anomaly and the eccentric anomaly*, respectively, in the case that $(\mathbf{q}, \mathbf{p})$ is the perigee of a given noncircular elliptic Kepler orbit.

The symplecticity of the mean anomaly action. Let Φ_{F_i} be the action determined by the choice of F_i with $i = 1, 2, 3$. The infinitesimal generator Y_i of Φ_{F_i} is computed, via (2.11), as follows:

$$Y_i = \frac{1}{F_i} X_H.$$

One can check that

$$\mathbf{i}_{Y_1}\Omega = \mathbf{d}I_1, \tag{2.15}$$

where

$$I_1(\mathbf{q}, \mathbf{p}) = \frac{\mu}{\sqrt{-2H(\mathbf{q}, \mathbf{p})}}, \tag{2.16}$$

and where $\mathbf{i}_{Y_1}\Omega$ is the interior product of Y_1 and Ω, which is defined by $\mathbf{i}_{Y_1}\Omega = \Omega(Y_1, \cdot)$ (see Abraham and Marsden, 1978 for more on the interior product). However, one can check that for $i = 2, 3$, the exterior derivative is nonzero:

$$\mathbf{d}(\mathbf{i}_{Y_i}\Omega) \neq 0. \tag{2.17}$$

Since each action Φ_{F_i} can be regarded as the periodic flow (of period 2π) of the vector field Y_i, Proposition 5.4.2 in Marsden and Ratiu (1999) applied to (2.15) and

(2.17) implies that Φ_{F_1} is a symplectic action but Φ_{F_2} and Φ_{F_3} are not. Furthermore, (2.15) implies that the momentum map of the (mean anomaly) action Φ_{F_1} is I_1 in (2.16) (see, e.g. Chapters 11 and 12 of Marsden and Ratiu, 1999 for the definition of the momentum map).

Nontriviality of $\Sigma_{\mathrm{elliptic}}$ *as a principal* S^1 *bundle.* Let S^1 act on $\Sigma_{\mathrm{elliptic}}$ by the Kepler flow as in (2.10). Since this action is free and proper, the set $\Sigma_{\mathrm{elliptic}}$ becomes a principal S^1 bundle. We will show that this bundle is not a trivial bundle, that is, $\Sigma_{\mathrm{elliptic}} \neq S^1 \times (\Sigma_{\mathrm{elliptic}}/S^1)$. This implies that $\Sigma_{\mathrm{elliptic}}$ is not diffeomorphic to $S^1 \times M$ for some 5-dimensional manifold M, on which the Kepler vector field is written as

$$\dot{\theta} = 1, \qquad \dot{m} = 0$$

for $(\theta, m) \in S^1 \times M$. This is one of the reasons why the method of measuring the true anomaly from the perigee in the Kepler elements breaks down at circular orbits. Although one may set a new rule of measuring the true anomaly locally near a circular orbit, this method will not extend globally to cover all of the orbits in $\Sigma_{\mathrm{elliptic}}$.

Let us first identify the base space $\Sigma_{\mathrm{elliptic}}/S^1$. Define the set

$$D = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^2 \times \mathbb{R}^3 \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0, \mathbf{x} \neq 0, \|\mathbf{y}\| < \mu\}.$$

The set D is diffeomorphic to $(0, \infty) \times T_{<\mu} S^2$ where $T_{<\mu} S^2$ is the set of tangent vectors on S^2 of length less than μ. Consider the map $\pi : \Sigma_{\mathrm{elliptic}} \to \mathbb{R}^3 \times \mathbb{R}^3$ defined by

$$\pi(\mathbf{q}, \mathbf{p}) = (\mathbf{L}(\mathbf{q}, \mathbf{p}), \mathbf{A}(\mathbf{q}, \mathbf{p})), \tag{2.18}$$

where $\mathbf{L}$ and $\mathbf{A}$ are defined in (2.2) and (2.3). The following proposition is from Cushman and Bates (1997) (see also Chang et al., 2002 for a short proof):

PROPOSITION 2.1. *The set* $\Sigma_{\mathrm{elliptic}}$ *is a principal* S^1 *bundle over* D *where the bundle projection* $\pi : \Sigma_{\mathrm{elliptic}} \to D$ *is given by* (2.18).

We will show the nontriviality of $\Sigma_{\mathrm{elliptic}}$ by finding a nontrivial subbundle.

LEMMA 2.2. *The subbundle* $\pi|_{\pi^{-1}(S^2 \times 0)} : \pi^{-1}(S^2 \times 0) \to S^2$ *of the bundle* $\pi : \Sigma_{\mathrm{elliptic}} \to D$ *is not a trivial bundle.*

Proof. The set $\pi^{-1}(S^2 \times 0)$ is the union of circular orbits with unit angular momentum vectors. One can see

$$\pi^{-1}(S^2 \times 0) = \{(\mathbf{q}, \mathbf{p}) \in T\mathbb{R}^3_0 \mid \langle \mathbf{q}, \mathbf{p} \rangle = 0, \|\mathbf{q}\| = 1/\mu, \|\mathbf{p}\| = \mu\}$$

which is diffeomorphic, by a linear map $(\mathbf{x}, \mathbf{y}) \mapsto (\mu\mathbf{x}, \mathbf{y}/\mu)$, to the unit tangent bundle $T_1 S^2$ of S^2, which is diffeomorphic to the 2-dimensional real projective space $\mathbb{RP}^2$. It is known that the projective space $\mathbb{RP}^2$ cannot be a trivial principal

S^1 bundle over S^2 (see, for example, Naber, 1997). Hence, the subbundle, $\pi^{-1}(S^2 \times 0) \to S^2$ is not trivial. □

PROPOSITION 2.3. *The principal S^1 bundle Σ_{elliptic} is not a trivial bundle.*

We remark that although one cannot cover the whole bundle Σ_{elliptic} with one local trivialization, it is possible to cover it with two local trivializations. First, notice that $S^2 \times 0$ is a strong deformation retract of D by the homotopy $F : [0, 1] \times D \to D$ defined by $F(t, (\mathbf{x}, \mathbf{y})) = ((1 - t)\mathbf{x} + t\mathbf{x}/\|\mathbf{x}\|, (1 - t)\mathbf{y})$, and that S^2 can be covered by two local charts. Then, using the homotopy lifting property, one can construct two local trivializations to cover the whole bundle Σ_{elliptic}. See Naber (1997) for details.

3. Geometric Derivation of the Delaunay Variables

The objective of this section is to derive the Delaunay variables from the viewpoint of geometric mechanics.

The traditional definition. The Delaunay variables, $(l, g, h, L, G, \tilde{H})$, are classically defined in terms of the Kepler elements as follows:

$$l = n(t - \tau), \qquad L = \sqrt{\mu a}, \tag{3.1}$$

$$g = \omega, \qquad G = \sqrt{\mu a(1 - e^2)}, \tag{3.2}$$

$$h = \Omega, \qquad \tilde{H} = \cos i \sqrt{\mu a(1 - e^2)}, \tag{3.3}$$

where n is the mean motion, a is the semimajor axis, e is the eccentricity, i is the inclination, ω is the argument of the perigee, Ω is the longitude of the ascending node, τ is the time when the satellite passes through the perigee, and $\sim$ was put on H in (3.3) to distinguish it from the Kepler Hamiltonian H defined in (2.1) (see Chapter 9 of Vinti, 1998 e.g. for more details concerning the Delaunay variables).

The geometric mechanics approach. In this section, we will define a symmetry action of the group $\mathbb{T}^3 = S^1 \times S^1 \times S^1$ (the three torus) and compute its momentum map, from which the Delaunay variables will arise (see, e.g. Chapters 11 and 12 of Marsden and Ratiu, 1999 for the definition of the momentum map). This derivation seems to us to be geometrically clearer than the traditional ones such as the use of the Hamilton–Jacobi equation in Born (1927) and Kovalevsky (1967), and the one using the Lagrange bracket in Brouwer and Clemence (1961) and Marsden and Ratiu (1999). In this paper $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ denotes the standard basis for $\mathbb{R}^3$.

The first S^1 action on Σ_{elliptic} is the mean anomaly action Φ_{F_1} in Section 2. We denote it by Φ_1 and compute it as follows: for $e^{i\theta} \in S^1$ and $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$,

$$\Phi_1(e^{i\theta}, (\mathbf{q}, \mathbf{p})) = \varphi\left(\frac{\theta}{2\pi} T(\mathbf{q}, \mathbf{p}), (\mathbf{q}, \mathbf{p})\right) \tag{3.4}$$

which comes from (2.10) and the constancy of $T(\mathbf{q}(t), \mathbf{p}(t))$. The momentum map was derived in Section 2 to be given as follows:

$$I_1(\mathbf{q}, \mathbf{p}) = \frac{\mu}{\sqrt{-2H(\mathbf{q}, \mathbf{p})}}. \tag{3.5}$$

The second S^1 action on Σ_{elliptic} is the rotation around the angular momentum $\mathbf{L}$, that is, for $e^{i\theta} \in S^1$ and $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$

$$\Phi_2(e^{i\theta}, (\mathbf{q}, \mathbf{p})) = \text{the rotation of } (\mathbf{q}, \mathbf{p}) \text{ around } \mathbf{L}(\mathbf{q}, \mathbf{p}) \text{ by angle } \theta \tag{3.6}$$

which can be concretely written as

$$\Phi_2(e^{i\theta}, (\mathbf{q}, \mathbf{p})) = \mathbf{R}_{\mathbf{L}(\mathbf{q},\mathbf{p})}\mathbf{R}_z(\theta_2)\mathbf{R}^{-1}_{\mathbf{L}(\mathbf{q},\mathbf{p})} \cdot (\mathbf{q}, \mathbf{p}), \tag{3.7}$$

where

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.8}$$

and

$$\mathbf{R}_{\mathbf{L}} = \begin{cases} \left(\dfrac{\mathbf{j} \times \mathbf{L}}{\|\mathbf{j} \times \mathbf{L}\|} \quad \dfrac{\mathbf{L} \times (\mathbf{j} \times \mathbf{L})}{\|\mathbf{L} \times (\mathbf{j} \times \mathbf{L})\|} \dfrac{\mathbf{L}}{\|\mathbf{L}\|} \right) & \text{if } \mathbf{L} \text{ is not parallel to } \mathbf{j}, \\ \left(\dfrac{\mathbf{k} \times \mathbf{L}}{\|\mathbf{k} \times \mathbf{L}\|} \quad \dfrac{\mathbf{L} \times (\mathbf{k} \times \mathbf{L})}{\|\mathbf{L} \times (\mathbf{k} \times \mathbf{L})\|} \dfrac{\mathbf{L}}{\|\mathbf{L}\|} \right) & \text{if } \mathbf{L} \text{ is not parallel to } \mathbf{k}, \end{cases}$$

where $\cdot$ in (3.7) is the diagonal action on each $\mathbb{R}^3$ component. The corresponding momentum map I_2 is computed as

$$I_2(\mathbf{q}, \mathbf{p}) = \|\mathbf{L}(\mathbf{q}, \mathbf{p})\|, \tag{3.9}$$

that is, I_2 satisfies the following:

$$\Omega^\sharp \, \mathbf{d}I_2(\mathbf{q}, \mathbf{p}) = \left.\frac{d}{dt}\right|_{t=0} \Phi_2(e^{it}, (\mathbf{q}, \mathbf{p})) = (\mathbf{e}_{\mathbf{L}} \times \mathbf{q}, \mathbf{e}_{\mathbf{L}} \times \mathbf{p}),$$

where $\mathbf{e}_{\mathbf{L}} := \mathbf{L}(\mathbf{q}, \mathbf{p})/\|\mathbf{L}(\mathbf{q}, \mathbf{p})\|$.

The third S^1 action on Σ_{elliptic} is by the rotation around the z-axis, that is, for $e^{i\theta} \in S^1$ and $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$,

$$\Phi_3(e^{i\theta}, (\mathbf{q}, \mathbf{p})) = \mathbf{R}_z(\theta) \cdot (\mathbf{q}, \mathbf{p}). \tag{3.10}$$

The corresponding momentum map I_3 is given by

$$I_3(\mathbf{q}, \mathbf{p}) = \mathbf{L}(\mathbf{q}, \mathbf{p}) \cdot \mathbf{k} \tag{3.11}$$

that is the z-component of the angular momentum $\mathbf{L}(\mathbf{q}, \mathbf{p})$. Notice that the three angles θ_i's denoting the three S^1 group actions are an extension of the three classical Delaunay variables l, g, h and that the three momentum maps I_i's are the same as the three Delaunay variables, $L, G, \tilde{H}$.

We remark that the momentum maps I_2 and I_3 can be easily derived from the momentum map corresponding to the SO(3) action on Σ_{elliptic}, where the SO(3) momentum map is the angular momentum vector $\mathbf{L}$ (see Chapter 12.2 of Marsden and Ratiu, 1999 for the derivation). Notice that the momentum maps I_2 and I_3 are projections of $\mathbf{L}$ to the rotation axes, $\mathbf{e_L}$ and $\mathbf{k}$.

Commutativity of the actions. A reason why we choose these particular three S^1 actions is that they commute with one another, that is,

$$\Phi_j(\mathrm{e}^{\mathrm{i}\theta_j}, \Phi_k(\mathrm{e}^{\mathrm{i}\theta_k}, (\mathbf{q}, \mathbf{p}))) = \Phi_k(\mathrm{e}^{\mathrm{i}\theta_k}, \Phi_j(\mathrm{e}^{\mathrm{i}\theta_j}, (\mathbf{q}, \mathbf{p})))$$

for $j, k = 1, 2, 3$. Each S^1 action, Φ_i can be regarded as a (periodic) flow of the Hamiltonian vector field X_{I_i} for $i = 1, 2, 3$. Let X_{I_j} be the infinitesimal generators corresponding the action Φ_j for $j = 1, 2, 3$, that is, $X_{I_j}(\mathbf{q}, \mathbf{p}) = \mathrm{d}/\mathrm{d}t|_{t=0}\Phi_j(\mathrm{e}^{\mathrm{i}t}, (\mathbf{q}, \mathbf{p}))$. We want to show $[X_{I_j}, X_{I_k}] = 0$ for $j, k = 1, 2, 3$. One can directly show this. Alternatively, One first computes

$$\{I_i, I_j\} = 0 \tag{3.12}$$

for $1 \leqslant i, j \leqslant 3$ where $\{\, ,\, \}$ is the canonical Poisson bracket on $T\mathbb{R}^3 = \mathbb{R}^3 \times \mathbb{R}^3$. This involutivity implies

$$[X_{I_i}, X_{I_j}] = -X_{\{I_i, I_j\}} = 0$$

which, in turn, implies that the three S^1 actions, Φ_1, Φ_2 and Φ_3, commute with one another. We can define the $\mathbb{T}^3 = S^1 \times S^1 \times S^1$ group action on Σ_{elliptic} by

$$\Phi((\mathrm{e}^{\mathrm{i}\theta_1}, \mathrm{e}^{\mathrm{i}\theta_2}, \mathrm{e}^{\mathrm{i}\theta_3}), (\mathbf{q}, \mathbf{p})) = \Phi_1(\mathrm{e}^{\mathrm{i}\theta_1}, \Phi_2(\mathrm{e}^{\mathrm{i}\theta_2}, \Phi_3(\mathrm{e}^{\mathrm{i}\theta_3}, (\mathbf{q}, \mathbf{p})))) \tag{3.13}$$

for $(\mathrm{e}^{\mathrm{i}\theta_1}, \mathrm{e}^{\mathrm{i}\theta_2}, \mathrm{e}^{\mathrm{i}\theta_3}) \in \mathbb{T}^3$ and $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$.

Remark. We now give some additional intuition and strategy behind the commutativity of the three S^1 actions. This will also help explain why one would choose the particular three S^1 actions defined above. First, the choice of Φ_1 is a natural one because it is the (time-reparametrized) Kepler flow. Since the Kepler dynamics have a rotational symmetry, the (time-reparametrized) Kepler flow Φ_1 commutes with rotations. Hence, one can choose Φ_3, the rotation around the z-axis, for an another S^1 action without loss of generality. Another rotation with a *fixed* axis other than the z-axis, will not commute with Φ_3, so we need to find a *dynamical rotation* whose rotation axis gets rotated by Φ_3. Consider an arbitrary point $(\mathbf{q}, \mathbf{p})$ in Σ_{elliptic} and the plane $\pi_{\mathbf{q}\times\mathbf{p}}$ spanned by $\mathbf{q}$ and $\mathbf{p}$ with the normal vector in the direction of $\mathbf{q} \times \mathbf{p}$. Let $\mathbf{R}$ be a rotation about the z-axis. Then, the rotation of $(\mathbf{q}, \mathbf{p})$ by $\mathbf{R}$ followed by the rotation of $(\mathbf{Rq}, \mathbf{Rp})$ in the plane $\mathbf{R}(\pi_{\mathbf{q}\times\mathbf{p}})$ by an angle θ is the same as the rotation of $(\mathbf{q}, \mathbf{p})$ in the plane $\pi_{\mathbf{q}\times\mathbf{p}}$ through the angle θ followed by the rotation $\mathbf{R}$. The rotation of $(\mathbf{q}, \mathbf{p})$ in the plane $\pi_{\mathbf{q}\times\mathbf{p}}$ is exactly the rotation about the angular momentum vector $\mathbf{q} \times \mathbf{p}$. This leads one to the choice we made above for the S^1 action Φ_2. Hence, Φ_2 and Φ_3 commute. Of course, Φ_2 and Φ_1 commute because of the rotational symmetry of the Kepler dynamics.

The momentum map of the torus action. The corresponding $\mathbb{T}^3$ momentum map $\mathbf{J} : \Sigma_{\text{elliptic}} \to \mathbb{R}^3$ is given by

$$\mathbf{J} = (I_1, I_2, I_3) \tag{3.14}$$

with I_i's in (3.5), (3.9) and (3.11). The image of $\mathbf{J}$ is

$$\operatorname{Im} \mathbf{J} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid |x_3| \leqslant x_2 \leqslant x_1\}.$$

LEMMA 3.1. *The set B of the regular values of the map* $\mathbf{J}$ *is given by*

$$B = \{(I_1, I_2, I_3) \in \mathbb{R}^3 \mid |I_3| < I_2 < I_1\} \tag{3.15}$$

and its inverse image $\Sigma_B = \mathbf{J}^{-1}(B)$ is given by

$$\Sigma_B = \{(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}} \mid \mathbf{L}(\mathbf{q}, \mathbf{p}) \times \mathbf{k} \neq 0,\ \mathbf{A}(\mathbf{q}, \mathbf{p}) \neq 0\}, \tag{3.16}$$

where $\mathbf{L}$ *and* $\mathbf{A}$ *are defined in* (2.2) *and* (2.3). *The set Σ_B is the union of nondegenerate, noncircular and nonequatorial Kepler orbits.*

Proof. The second statement is implied by (2.5). To prove the first statement, notice that the rank of $\mathbf{dJ} = (\mathbf{d}I_1, \mathbf{d}I_2, \mathbf{d}I_3)$ is the same as the rank of $\{X_{I_1}, X_{I_2}, X_{I_3}\}$ because Ω is nondegenerate and $X_{I_i} = \Omega^\sharp\, \mathbf{d}I_i$ for $i = 1, 2, 3$. Consider the set

$$A_c := \quad \mathbf{J}^{-1}(\{(I_1, I_2, I_3) \in \mathbb{R}^3 \mid |I_3| \leqslant I_2 = I_1\}).$$

Since $\mathbb{T}^3$ is an Abelian group, the momentum map $\mathbf{J}$ is $\mathbb{T}^3$ invariant. Hence, the actions Φ_1 and Φ_2 leave A_c invariant and coincide with each other on A_c. It follows that $X_{I_1} = X_{I_2}$ on A_c. In the same manner, the following holds:

$$X_{I_2} = X_{I_3} \quad \text{on } A_{\text{eq}} := \mathbf{J}^{-1}(\{(I_1, I_2, I_3) \in \mathbb{R}^3 \mid |I_3| = I_2 \leqslant I_1\}).$$

We now show that B is indeed the set of regular values by showing that the points in Σ_B are regular points of $\mathbf{J}$. We first compute the infinitesimal generators corresponding the three actions as follows:

$$X_{I_1}(\mathbf{q}, \mathbf{p}) = \frac{T(\mathbf{q}, \mathbf{p})}{2\pi} \left(\mathbf{p},\ -\mu \frac{\mathbf{q}}{\|\mathbf{q}\|^3} \right), \tag{3.17}$$

$$X_{I_2}(\mathbf{q}, \mathbf{p}) = (\mathbf{e_L} \times \mathbf{q},\ \mathbf{e_L} \times \mathbf{p}), \tag{3.18}$$

$$X_{I_3}(\mathbf{q}, \mathbf{p}) = (\mathbf{k} \times \mathbf{q}, \mathbf{k} \times \mathbf{p}), \tag{3.19}$$

where, as above, $\mathbf{e_L} := \mathbf{L}(\mathbf{q}, \mathbf{p})/\|\mathbf{L}(\mathbf{q}, \mathbf{p})\|$. First notice that none of $X_{I_i}(\mathbf{q}, \mathbf{p})$ vanish on Σ_B. We will show that $\{X_{I_1}, X_{I_2}, X_{I_3}\}$ is linearly independent on Σ_B.

Fix an arbitrary $(\mathbf{q}, \mathbf{p}) \in \Sigma_B$. Notice that both components of each of $X_{I_1}(\mathbf{q}, \mathbf{p})$ and $X_{I_2}(\mathbf{q}, \mathbf{p})$ are orthogonal to $\mathbf{L}(\mathbf{q}, \mathbf{p})$. We claim that both components of $X_{I_3}(\mathbf{q}, \mathbf{p})$ cannot be simultaneously orthogonal to $\mathbf{L}(\mathbf{q}, \mathbf{p})$. To see this, suppose that

$$\mathbf{L}(\mathbf{q}, \mathbf{p}) \cdot (\mathbf{k} \times \mathbf{q}) = \mathbf{0} \quad \text{and} \quad \mathbf{L}(\mathbf{q}, \mathbf{p}) \cdot (\mathbf{k} \times \mathbf{p}) = \mathbf{0}.$$

It follows that both $\mathbf{q}$ and $\mathbf{p}$ are orthogonal to the vector $\mathbf{k} \times \mathbf{L}(\mathbf{q}, \mathbf{p})$. Then $\mathbf{q}$ and $\mathbf{p}$ are parallel because both of them are perpendicular to the two linearly independent vectors $\{\mathbf{L}(\mathbf{q}, \mathbf{p}),\ \mathbf{k} \times \mathbf{L}(\mathbf{q}, \mathbf{p})\}$. It follows that $\mathbf{L}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p} = \mathbf{0}$, which is impossible for $(\mathbf{q}, \mathbf{p}) \in \Sigma_B$. Hence $X_{I_3}(\mathbf{q}, \mathbf{p})$ is not spanned by the set $\{X_{I_1}(\mathbf{q}, \mathbf{p}),\ X_{I_2}(\mathbf{q}, \mathbf{p})\}$.

We next show that the two vectors $X_{I_1}(\mathbf{q}, \mathbf{p})$ and $X_{I_2}(\mathbf{q}, \mathbf{p})$ are linearly independent. Suppose that it is not. Then there is $c \neq 0$ such that $X_{I_1}(\mathbf{q}, \mathbf{p}) = cX_{I_2}(\mathbf{q}, \mathbf{p})$, which implies

$$\mathbf{p} = c(\mathbf{e}_{\mathrm{L}} \times \mathbf{q}), \qquad -\frac{\mu}{\|\mathbf{q}\|^3}\mathbf{q} = c(\mathbf{e}_{\mathrm{L}} \times \mathbf{p}).$$

This implies $\mathbf{A}(\mathbf{q}, \mathbf{p}) = 0$, which is not the case for $(\mathbf{q}, \mathbf{p}) \in \Sigma_B$. It follows that the two vectors $X_{I_1}(\mathbf{q}, \mathbf{p})$ and $X_{I_2}(\mathbf{q}, \mathbf{p})$ are linearly independent. Therefore, the three vectors $X_{I_1}(\mathbf{q}, \mathbf{p})$, $X_{I_2}(\mathbf{q}, \mathbf{p})$ and $X_{I_3}(\mathbf{q}, \mathbf{p})$ are linearly independent for all $(\mathbf{q}, \mathbf{p}) \in \Sigma_B$. This completes the proof that B is the set of regular values of $\mathbf{J}$. We now prove the third statement. Notice that the set A_c is the union of the circular orbits and the set A_{eq} is the union of equatorial elliptic orbits. Hence, Σ_B is the union of nondegenerate, noncircular and nonequatorial Kepler orbits. □

The canonical nature of the Delaunay variables. We now show that (Σ_B, Ω) is symplectomorphic to $(\mathbb{T}^3 \times B, \mathrm{d}\theta_i \wedge \mathrm{d}I_i)$. This implies that the Delaunay variables are canonical variables. Notice that $\mathbb{T}^3$ acts on Σ_B freely and properly. Define a map $s = (\mathbf{q}_s, \mathbf{p}_s) : B \to \Sigma_B$ by

$$\mathbf{q}_s = \frac{\|\mathbf{l}\|^2\mathbf{a}}{(\mu + \|\mathbf{a}\|)\|\mathbf{a}\|}, \qquad \mathbf{p}_s = \frac{\mathbf{l} \times \mathbf{q}_s}{\|\mathbf{q}_s\|^2}, \tag{3.20}$$

where $\mathbf{l}, \mathbf{a} : B \to \mathbb{R}^3$ are defined by

$$\begin{aligned}
\mathbf{l}(x_1, x_2, x_3) &= \left(\sqrt{x_2^2 - x_3^2},\ 0,\ x_3\right), \\
\mathbf{a}(x_1, x_2, x_3) &= \frac{\mu\sqrt{x_1^2 - x_2^2}}{x_1 x_2}\left(x_3,\ 0,\ -\sqrt{x_2^2 - x_3^2}\right).
\end{aligned}$$

One can check that $\mathbf{J} \circ s = \mathrm{Id}_B$. Notice that $s = (\mathbf{q}_s, \mathbf{p}_s)$ is the perigee of the Kepler orbit with the angular momentum vector $\mathbf{l}$ and the Laplace vector $\mathbf{a}$. Define a map $\phi : \mathbb{T}^3 \times B \to \Sigma_B$ by

$$\phi(g, \mathbf{x}) = \Phi_g \circ s(\mathbf{x})$$

for

$$(g, \mathbf{x}) := ((\theta_1, \theta_2, \theta_3), (x_1, x_2, x_3)) \in \mathbb{T}^3 \times B,$$

where Φ_g is the $\mathbb{T}^3$ action defined in (3.13). Since $\mathbf{J}$ is $\mathbb{T}^3$ invariant, it follows from $\mathbf{J} \circ s = \mathrm{Id}_B$ that:

$$\mathbf{J} \circ \phi(g, \mathbf{x}) = \mathbf{x}$$

which with the free action of $\mathbb{T}^3$ implies the injectivity of ϕ.

We now show the surjectivity of ϕ. Take any $(\mathbf{q}, \mathbf{p}) \in \Sigma_B$. Let C be the elliptic Kepler orbit containing $(\mathbf{q}, \mathbf{p})$. There is $g_1 = (\theta_1, 0, 0)$, namely, along the Kepler flow, such that $\Phi_{g_1}(\mathbf{q}, \mathbf{p})$ is at the perigee of the Kepler orbit $C = \Phi_{g_1}(C)$ (see Figure 3.1). By simple geometry, one can see that there is $g_2 = (0, \theta_2, 0)$, namely, a rotation about the angular momentum of the orbit C such that the z-component of the position variable of the point $\Phi_{g_2 g_1}(\mathbf{q}, \mathbf{p})$ is the smallest of all the z-components of points in the orbit $\Phi_{g_2 g_1}(C)$. In other words, the orbit $g_2 g_1(C)$ has its perigee in the lowest place in the configuration space $\mathbb{R}^3$. Then, there is $g_3 = (0, 0, \theta_3)$, a rotation about the z-axis, such that $\Phi_{g_3 g_2 g_1}(\mathbf{q}, \mathbf{p})$ is contained in the set

$$\Pi := \{((x, 0, z), (0, v_y, 0)) \in T\mathbb{R}^3 \mid z < 0,\ v_y > 0\}. \tag{3.21}$$

Notice that Π is isotropic, that is, $\Omega(v, w) = 0$ for all $v, w \in T\Pi = \Pi \times \Pi$. Since $s(B) \subset \Pi$, it follows:

$$s^*\Omega = 0. \tag{3.22}$$

Let $g = g_3 g_2 g_1 = (\theta_1, \theta_2, \theta_3) \in \mathbb{T}^3$ and $\mathbf{x} = \mathbf{J}(g(\mathbf{q}, \mathbf{p})) = \mathbf{J}(\mathbf{q}, \mathbf{p}) \in B$. One can check that $s(\mathbf{x}) = \Phi_g(\mathbf{q}, \mathbf{p})$. It follows that $\phi(g^{-1}, \mathbf{x}) = (\mathbf{q}, \mathbf{p})$. This proves the subjectivity of ϕ.

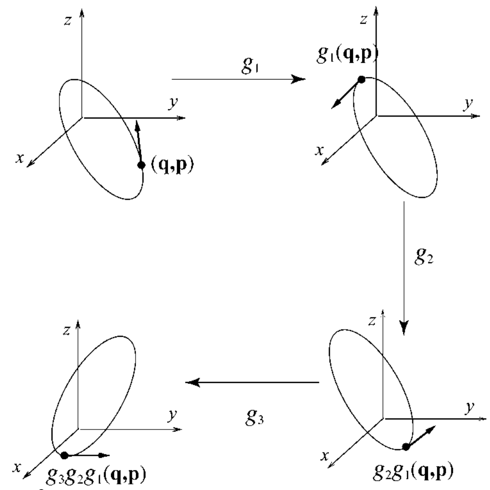

Figure 3.1. Illustration of the $\mathbb{T}^3$-action where g_1 is by Kepler flow, g_2 is a rotation about the angular momentum, and g_3 is a rotation about the z-axis.

We will show that ϕ is symplectic, which will also imply (by the inverse function theorem) that ϕ is a diffeomorphism. We will use $(g, \mathbf{x})$ as coordinates for $\mathbb{T}^3 \times B$ in the following computation to avoid any possible confusion. Then

$$\phi^*\Omega\,(\partial_{\theta_i}, \partial_{\theta_j}) = \Omega\,(T\phi(\partial_{\theta_i}), T\phi(\partial_{\theta_j})) = \Omega\,(X_{I_i}, X_{I_j}) = \{I_i,\ I_j\} = 0$$

and

$$\begin{aligned}\phi^*\Omega\,(\partial_{\theta_i}, \partial_{x_j}) &= \Omega\,(T\phi(\partial_{\theta_i}), T\phi(\partial_{x_j})) = \Omega\,(X_{I_i}, T\phi(\partial_{x_j})) = \mathrm{d}I_i\,(T\phi(\partial_{x_j}))\\ &= \mathrm{d}(I_i \circ \phi)(\partial_{x_j}) = \mathrm{d}x_i\,(\partial_{x_j}) = \delta_{ij}.\end{aligned}$$

Also

$$\phi^*\Omega\,(\partial_{x_i}, \partial_{x_j}) = s^*\Phi_g^*\Omega\,(\partial_{x_i}, \partial_{x_j}) = s^*\Omega\,(\partial_{x_i}, \partial_{x_j}) = 0,$$

where the $\mathbb{T}^3$ action Φ_g being symplectic was used in the second equality, and (3.22) was used in the third equality. Hence, we have shown that

$$\phi^*\Omega = \sum_{i=1}^{3} \mathrm{d}\theta_i \wedge \mathrm{d}I_i \tag{3.23}$$

for $((\theta_1, \theta_2, \theta_3), (I_1, I_2, I_3)) \in \mathbb{T}^3 \times B$. This proves that ϕ is symplectic. In other words, the Delaunay variables are canonical variables. Equation (3.23) can be written in the traditional Delaunay variables as

$$\phi^*\Omega = \mathrm{d}l \wedge \mathrm{d}L + \mathrm{d}g \wedge \mathrm{d}G + \mathrm{d}h \wedge \mathrm{d}\tilde{H}.$$

In the following proposition, we summarize what we have derived. It is the procedure of derivation of Proposition 3.2 that is the main result of the current paper, rather than the proposition itself.

PROPOSITION 3.2. *There is a* $\mathbb{T}^3$ *action on* Σ_{elliptic} *given by* (3.13) *with the associated momentum map* $\mathbf{J}$ *given by* (3.14), *where* Σ_{elliptic} *is the union of the nondegenerate elliptic Kepler orbits given in* (2.6). *The set* Σ_B *of regular points of* $\mathbf{J}$ *in* (3.16) *is a trivial* $\mathbb{T}^3$ *principal bundle symplectomorphic to* $\mathbb{T}^3 \times B$ *where* B, *defined in* (3.15), *is the set of regular values of* $\mathbf{J}$ *and* $\mathbb{T}^3 \times B$ *is equipped with the canonical symplectic form* $\sum_{i=1}^{3} \mathrm{d}\theta_i \wedge \mathrm{d}I_i$ *with* $((\theta_1, \theta_2, \theta_3), (I_1, I_2, I_3)) \in \mathbb{T}^3 \times B$. *The coordinates* $((\theta_1, \theta_2, \theta_3), (I_1, I_2, I_3))$ *coincide with the traditional Delaunay variables* $(l, g, h, L, G, \tilde{H})$ *defined in* (3.1)–(3.3).

Remark. We now compare our approach with the derivation of the Delaunay variables using Hamilton–Jacobi theory. The derivation by Hamilton–Jacobi theory can be found in Sections 21 and 22 of Born (1927) (see also Chapter 10 of Goldstein, 1980 for an exposition of the method), which is summarized as follows. The rotational symmetry of the Kepler Hamiltonian allows one to use separation of variables in the Hamilton–Jacobi equation, yielding three action variables. This step involves a special integration trick using complex variables due to Sommerfeld. Then, one makes use of the degeneracy of the corresponding angle variables to obtain a new set of angle and action variables so that two of the three angle variables do not change in time. Finally, one seeks the physical meaning of this set of action–angle variables, which requires nontrivial geometric intuition. This Hamilton–Jacobi approach is rather different from that in this paper.

Another approach to constructing the Delaunay variables can be based on the Liouville–Arnold theorem. This approach is sketched in Arnold (1991, p. 117), which refers to Charlier (1927) for details. In this approach one must begin with first integrals in involution. Even though this general machinery guarantees one to get a set of action–angle variables, it lacks geometric insight and it involves some complicated integrations.

4. First-order Averaged J_2-Dynamics and Geometric Phases

We now study the *(first-order) averaged dynamics* of the perturbed Kepler motion with the perturbation due to the bulge of the earth. This is usually called J_2-dynamics because the coefficient of the first biggest perturbation term is referred to as J_2 (see Chapter 15 of Vinti, 1998). Traditionally, Delaunay variables are used for this study. In the current paper, we will derive the explicit expression of the flow equation of the averaged J_2 dynamics using the set up of the previous sections. We will also preform the $S^1 \times S^1$ reduction and study the associated phases and bifurcations. Our study of the $S^1 \times S^1$ reduction is a review and improvement of the work of Cushman (1991) and Coffey et al. (1986). The improvement comes about through the use of topological arguments to show the presence of the Hopf fibration, and by the addition of new results on phases.

In this section, we will use H_0 for the unperturbed or Kepler Hamiltonian in (2.1) *and H for the perturbed Hamiltonian. So one should replace H in the previous sections by H_0 in this section.*

4.1. DERIVATION OF THE FLOW OF THE AVERAGED J_2-DYNAMICS

In order to derive the flow of the averaged J_2-dynamics, we will use Cartesian coordinates, which are superior to other coordinates because the Cartesian coordinates do not have singularities, together with the observation that the averaged Hamiltonian is a collective Hamiltonian of the $\mathbb{T}^3$ momentum map $\mathbf{J}$ in (3.14). Vinti (1998) has an excellent analysis of the J_2-dynamics using the traditional perturbation method and Cushman (1991) employs normal form theory and reduction theory to study the same problem. In particular, the averaged Hamiltonian in this paper is the first-order normal form. For the sake of simplicity, we will omit the phrase *first-order* in the following.

The J_2 and averaged Hamiltonian. The Hamiltonian H for the J_2 problem is given by

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\|\mathbf{p}\|^2 - \frac{\mu}{\|\mathbf{q}\|} + \frac{J_2 \mu R_e^2 (2q_3^2 - q_1^2 - q_2^2)}{2\|\mathbf{q}\|^5}, \tag{4.1}$$

where $J_2 \approx 1082.63 \times 10^{-6}$ and R_e is the radius of the earth. One averages this Hamiltonian over the unperturbed flow, that is, the Kepler flow of H_0, in order to study its secular motion (in fact, this was justified using normal form theory in Cushman, 1991). The averaged Hamiltonian $\bar{H}$ on Σ_{elliptic} is defined by

$$\bar{H}(\mathbf{q}, \mathbf{p}) := \frac{1}{T(\mathbf{q}, \mathbf{p})} \int_0^{T(\mathbf{q}, \mathbf{p})} H \circ \varphi_t^{H_0}(\mathbf{q}, \mathbf{p}) \, \mathrm{d}t$$

for $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$ where $\varphi_t^{H_0}$ is the unperturbed Kepler flow, that is, the Hamiltonian flow of the Kepler Hamiltonian H_0 and T is the period of the Kepler flow defined in (2.7) (again, one must replace H by H_0 in (2.7)). One can then compute the averaged Hamiltonian $\bar{H}$ as follows (see, e.g. Chapter 17 of Vinti, 1998)

$$\bar{H} = H_0 + \frac{J_2 \mu R_e^2 \sqrt{(-2H_0)^3} (\|\mathbf{L}\|^2 - 3L_3^2)}{4\|\mathbf{L}\|^5}, \tag{4.2}$$

where $\mathbf{L}$ is the angular momentum and $L_3 := \mathbf{L} \cdot \mathbf{k}$ is the third component of $\mathbf{L}$.

The collective form of the Hamiltonian. Notice that, in agreement with general theory about dual pairs and collectivization – see Marsden and Ratiu (1999) for an elementary discussion – $\bar{H}$ is *collective*, that is, $\bar{H}$ is a function of the momentum map $\mathbf{J}$ as follows:

$$\bar{H} = F \circ \mathbf{J},$$

where

$$\mathbf{J} = (I_1, I_2, I_3) = \left(\frac{\mu}{\sqrt{-2H_0}}, \|\mathbf{L}\|, L_3 \right)$$

is the momentum map in (3.14) and where F is defined by

$$F(x_1, x_2, x_3) = -\frac{\mu^2}{2x_1^2} + \frac{J_2 \mu^4 R_e^2 (x_2^2 - 3x_3^2)}{4x_1^3 x_2^5}.$$

Then the Hamiltonian vector field $X_{\bar{H}}$ is expressed as a linear combination of $\{X_{I_1}, X_{I_2}, X_{I_3}\}$ by Theorem 12.4.2 of Marsden and Ratiu (1999) as follows:

$$X_{\bar{H}} = a_1 X_{I_1} + a_2 X_{I_2} + a_3 X_{I_3},$$

where X_{I_i}'s are in (3.17)–(3.19) and the three functions $a_i := \partial F/\partial x_i \circ \mathbf{J}$ are given by

$$a_1 = \frac{\mu^2}{I_1^3} - \frac{3J_2\mu^4 R_e^2(I_2^2 - 3I_3^2)}{4I_1^4 I_2^5}, \tag{4.3}$$

$$a_2 = \frac{3J_2\mu^4 R_e^2(5I_3^2 - I_2^2)}{4I_1^3 I_2^6}, \tag{4.4}$$

$$a_3 = -\frac{3J_2\mu^4 R_e^2 I_3}{2I_1^3 I_2^5}. \tag{4.5}$$

One can also derive this using elementary properties of the Poisson bracket $\{\,,\,\}$.

Consequences of involutivity. The involution relation (3.12) implies that

$$\{I_i,\, a_j\} = 0 \tag{4.6}$$

with $i = 1, 2, 3$. One can check the Lie bracket of any two of $a_i X_{I_i}$ with $i = 1, 2, 3$ is zero. It follows that the flows $\varphi_t^{a_i X_{I_i}}$ of the vector fields $a_i X_{I_i}$ commute with one another. Hence, the flow $\varphi_t^{\bar{H}}$ is given by

$$\varphi_t^{\bar{H}} = \varphi_t^{a_3 X_{I_3}} \circ \varphi_t^{a_2 X_{I_2}} \circ \varphi_t^{a_1 X_{I_1}}, \tag{4.7}$$

where the order of the composition of the three flows does not matter. We need the following lemma.

LEMMA 4.1. *Suppose two functions f and H are in involution, that is, $\{f, H\} = 0$. Then the flow $\varphi_t^{fX_H}$ of the vector field fX_H is a time reparameterization of the flow $\varphi_t^{X_H}$ of the vector field X_H as follows:*

$$\varphi^{fX_H}(t, (\mathbf{q}, \mathbf{p})) = \varphi^{X_H}(f(\mathbf{q}, \mathbf{p})t, (\mathbf{q}, \mathbf{p})).$$

Proof. Let $\psi(t, (\mathbf{q}, \mathbf{p})) = \varphi^{X_H}(f(\mathbf{q}, \mathbf{p})t, (\mathbf{q}, \mathbf{p}))$. Notice that $f(\mathbf{q}, \mathbf{p})$ is constant along the flow φ^{X_H} by the involution assumption. It is straightforward to show that ψ_t is a one-parameter group. Then one checks that $\mathrm{d}\psi_t/\mathrm{d}t = fX_H$. $\square$

By (4.6), the function a_i is constant along the flow $\varphi_t^{I_j}$ of the Hamiltonian vector fields X_{I_j} for $1 \leqslant i, j \leqslant 3$ where the flows $\varphi_t^{I_1}, \varphi_t^{I_2}, \varphi_t^{I_3}$ are given in (3.4), (3.6), (3.10), respectively, with θ replaced by t. Hence, by Lemma 4.1, the flow $\varphi_t^{\bar{H}}$ in (4.7) is given by

$$\varphi_t^{\bar{H}}(\mathbf{q}, \mathbf{p}) = \varphi_{a_3(\mathbf{q},\mathbf{p})t}^{I_3} \circ \varphi_{a_2(\mathbf{q},\mathbf{p})t}^{I_2} \circ \varphi_{a_1(\mathbf{q},\mathbf{p})t}^{I_1}(\mathbf{q}, \mathbf{p}) \tag{4.8}$$

for $(\mathbf{q}, \mathbf{p}) \in \Sigma_{\text{elliptic}}$. Recall that $\varphi_{a_3(\mathbf{q},\mathbf{p})t}^{I_3}$ is the rotation about the z-axis, that $\varphi_{a_2(\mathbf{q},\mathbf{p})t}^{I_2}$ is the rotation about the angular momentum $\mathbf{L}(\mathbf{q}, \mathbf{p})$, and that $\varphi_{a_1(\mathbf{q},\mathbf{p})t}^{I_1}$ is the time reparameterized Kepler flow.

Secular drifts. One can also check that the three functions a_i's coincide with the three functions c_i in (19.110), (19.123) and (19.126) of Vinti (1998) up to the first-order in J_2. These are the well-known formulas of the secular drift rates. Notice that the flow $\varphi_t^{I_1}$ in (3.4) can be written as

$$\varphi_t^{I_1}(\mathbf{q},\mathbf{p}) = \varphi^{H_0}\left(\frac{I_1(\mathbf{q},\mathbf{p})^3}{\mu^2}t, (\mathbf{q},\mathbf{p})\right)$$

by (2.7) where $\varphi_t^{H_0}$ is the Kepler flow. Hence, $\varphi_{a_1(\mathbf{q},\mathbf{p})t}^{I_1}(\mathbf{q},\mathbf{p})$ can be written as

$$\varphi_{a_1(\mathbf{q},\mathbf{p})t}^{I_1}(\mathbf{q},\mathbf{p}) = \varphi_{(t+\tilde{a}_1(\mathbf{q},\mathbf{p})t)}^{H_0}(\mathbf{q},\mathbf{p}),$$

where

$$\tilde{a}_1 = -\frac{3J_2\mu^2 R_e^2(I_2^2 - 3I_3^2)}{4I_1 I_2^5}. \tag{4.9}$$

We can write the flow of $\varphi^{\bar{H}}$ of the averaged Hamiltonian $\hat{H}$ as follows:

$$\varphi_t^{\bar{H}}(\mathbf{q},\mathbf{p}) = \varphi_{a_3(\mathbf{q},\mathbf{p})t}^{I_3} \circ \varphi_{a_2(\mathbf{q},\mathbf{p})t}^{I_2} \circ \varphi_{\tilde{a}_1(\mathbf{q},\mathbf{p})t}^{H_0} \circ \varphi_t^{H_0}(\mathbf{q},\mathbf{p}), \tag{4.10}$$

where $\varphi_{\tilde{a}_1(\mathbf{q},\mathbf{p})t}^{H_0}$ is the secular drift along the Kepler flow (or, the drift of the anomaly), $\varphi_{a_2(\mathbf{q},\mathbf{p})t}^{I_2}$ is the secular drift by the rotation about the angular momentum $\mathbf{L}(\mathbf{q},\mathbf{p})$ (or, the drift of the argument of the perigee ω), and $\varphi_{a_3(\mathbf{q},\mathbf{p})t}^{I_3}$ is the secular drift by the rotation about the z-axis (or the drift of the longitude of the ascending node Ω).

The critical inclination. Historically, the inclination $i := \cos^{-1}(I_3/I_2)$, which satisfies $a_2(\mathbf{q},\mathbf{p}) = 0$, that is, $5\cos^2 i - 1 = 0$, is called the *critical inclination*. At this inclination, we do not have secular drift of the argument of the perigee ω *on average* (Figure 4.1). Notice that the sign of the drift rate $\tilde{a}_1$ along the Kepler orbit changes its sign at the inclination $i = \cos^{-1}(1/\sqrt{3})$, as can be seen in (4.9). In summary, we derived the flow of the averaged J_2-Hamiltonian in (4.2) and decomposed it into four mutually commuting flows in (4.10) by using the fact that the averaged Hamiltonian is a collective Hamiltonian of the $\mathbb{T}^3$-momentum map $\mathbf{J}$ in (3.14), so that we were able to identify the well-known secular drift terms and

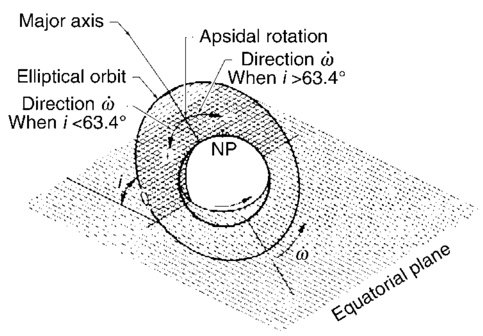

Figure 4.1. Phase drifts in the J_2 problem (the figure is taken from Chobotov, 1996).

the drift rates easily in Cartesian coordinates while coordinates with singularities such as the traditional Delaunay variables were used in the past.

4.2. REDUCTIONS, BIFURCATIONS, AND PHASES

The papers of Cushman (1991) and Coffey et al. (1986) studied the role of symplectic reduction in the problem of J_2-dynamics. We will review it with simpler and more topological arguments and then discuss how geometric phases are involved in the reduction picture. For a numerical study of the J_2-dynamics with Poincaré maps, refer to Broucke (1994).

Symplectic reductions for the J_2 problem. Recall that the two actions Φ_1 and Φ_3 commute and that the averaged Hamiltonian $\bar{H}$ in (4.2) is invariant under the $S^1 \times S^1$ action $\Phi_1 \times \Phi_3$ (the invariance of $\bar{H}$ is proven by showing that $\{\bar{H}, I_1\} = \{\bar{H}, I_3\} = 0$; its invariance under Φ_1 follows from the averaging construction). We will perform the $S^1 \times S^1$ reduction by stages, first by Φ_1 and then by Φ_3 (reduction by stages for commuting group actions was given in Marsden and Weinstein, 1974; more general approaches to reduction by stages are found in Marsden et al., 1998).

The first reduction. Define the *Runge vector* $\mathbf{R}$ on Σ_{elliptic} by

$$\mathbf{R}(\mathbf{q}, \mathbf{p}) = \frac{1}{\sqrt{-2H_0(\mathbf{q}, \mathbf{p})}}\mathbf{A}(\mathbf{q}, \mathbf{p}),$$

where $\mathbf{A}$ is the Laplace vector defined in (2.3). By (2.4) and (2.5), the angular momentum vector $\mathbf{L}$ and the Runge vector $\mathbf{R}$ satisfy

$$\mathbf{L} \cdot \mathbf{R} = 0, \tag{4.11}$$

$$\|\mathbf{L}\|^2 + \|\mathbf{R}\|^2 = I_1^2. \tag{4.12}$$

Define the mappings $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ on Σ_{elliptic} by

$$\mathbf{a} = \tfrac{1}{2}(\mathbf{L} + \mathbf{R}), \qquad \mathbf{b} = \tfrac{1}{2}(\mathbf{L} - \mathbf{R}). \tag{4.13}$$

One can show by a straightforward calculation (see also Cushman and Bates, 1997) that

$$\{a_i, a_j\} = \epsilon_{ijk} a_k, \qquad \{b_i, b_j\} = \epsilon_{ijk} b_k, \qquad \{a_i, b_j\} = 0. \tag{4.14}$$

By (4.11)–(4.14) and the fact (Proposition 2.1) that a nondegenerate elliptic Keplerian orbit is uniquely determined by a pair of $\mathbf{L} \neq 0$ and $\mathbf{R}$ (or, $\mathbf{A}$ with $\|\mathbf{A}\| < \mu$), the first reduced space $I_1^{-1}(\nu_1)/S^1$ is symplectomorphic to the space

$$\begin{aligned} &S^2_{\nu_1/2} \times S^2_{\nu_1/2} \backslash \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \mathbf{a} + \mathbf{b} = 0\} \\ &\quad = \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\mathbf{a}\| = \|\mathbf{b}\| = \nu_1/2,\ \mathbf{a} \neq -\mathbf{b}\} \end{aligned}$$

with the product symplectic structure induced from the Lie–Poisson structure of $(\mathbb{R}^3, \times)$. Notice that the points $(\mathbf{a}, -\mathbf{a})$ correspond to degenerate Keplerian orbits, that is, $\mathbf{L} = 0$. The reduced Hamiltonian $\bar{H}_{\nu_1}$ is given by

$$\bar{H}_{\nu_1} = -\frac{\mu^2}{2\nu_1^2} + \frac{J_2\mu^4 R_{\mathrm{e}}^2(\|\mathbf{a}+\mathbf{b}\|^2 - 3(a_3+b_3)^2)}{4\nu_1^3\|\mathbf{a}+\mathbf{b}\|^5}.$$

For aesthetic purposes, we will add the degenerate Keplerian orbits to $I_1^{-1}(\nu_1)/S^1$ and work with $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$. However, we will also keep track of these additional points.

The second reduction. The S^1 action by Φ_3 on Σ_{elliptic} induces the S^1 action on $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$ defined by

$$\mathrm{e}^{\mathrm{i}\theta} \cdot (\mathbf{a}, \mathbf{b}) = (\mathbf{R}_z(\theta)\mathbf{a},\ \mathbf{R}_z(\theta)\mathbf{b})$$

with $\mathbf{R}_z(\theta)$ in (3.8), which follows from (4.13) and the property $\mathbf{R}(\mathbf{x} \times \mathbf{y}) = \mathbf{R}\mathbf{x} \times \mathbf{R}\mathbf{y}$ for $\mathbf{R} \in \mathrm{SO}(3)$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. The corresponding momentum map I_{3,ν_1} on $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$ is given by

$$I_{3,\nu_1}(\mathbf{a}, \mathbf{b}) = a_3 + b_3,$$

where I_{3,ν_1} can also be understood as the induced function from I_3 in (3.11) via (4.13). The function I_{3,ν_1} has four critical points on $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$, namely the points

$$\{\pm((0, 0, \nu_1/2), (0, 0, \pm\nu_1/2))\},$$

which are also the fixed points of the S^1 action on $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$. The points

$$\{\pm((0, 0, \nu_1/2), (0, 0, \nu_1/2))\}$$

correspond to the equatorial circular Kepler orbits of $I_1 = \nu_1$. The points

$$\{\pm((0, 0, \nu_1/2), (0, 0, -\nu_1/2))\}$$

correspond to the degenerate Kepler orbits of $I_1 = \nu_1$ along the polar axis.

The range of I_{3,ν_1} is given by

$$|I_{3,\nu_1}(\mathbf{a}, \mathbf{b})| \leqslant \nu_1$$

for $(\mathbf{a}, \mathbf{b}) \in S^2_{\nu_1/2} \times S^2_{\nu_1/2}$. The level set $I_{3,\nu_1}^{-1}(\nu_3)$ for $|\nu_3| \leqslant \nu_1$ is given by

$$\begin{aligned} &I_{3,\nu_1}^{-1}(\nu_3) \\ &\quad = \{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\mathbf{a}\| = \|\mathbf{b}\| = \nu_1/2,\ b_3 = \nu_3 - a_3,\ a_3 \in \mathcal{I}\}, \end{aligned} \tag{4.15}$$

where $\mathcal{I}$ is the interval defined by

$$\mathcal{I} = [\max\{-\nu_1/2, \nu_3 - \nu_1/2\},\ \min\{\nu_1/2, \nu_3 + \nu_1/2\}].$$

We will show that $I_{3,\nu_1}^{-1}(\nu_3)$ is diffeomorphic to the three sphere S^3 for $0 < |\nu_3| < \nu_1$. We think of S^3 as follows:

$$\begin{aligned} S^3 &= \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\} \\ &= \{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 = k, |z_2|^2 = 1 - k, k \in [0, 1]\}. \end{aligned} \tag{4.16}$$

Using the stereographic projection, one can construct a diffeomorphism of $I_{3,\nu_1}^{-1}(\nu_3)$ onto S^3 by (4.15) and (4.16). The S^1 action on $I_{3,\nu_1}^{-1}(\nu_3)$ induces the S^1 action on S^3 as follows:

$$e^{i\theta} \cdot (z_1, z_2) = (e^{i\theta} z_1, e^{i\theta} z_2). \tag{4.17}$$

Notice that $S^1 \to S^3 \to S^3/S^1$ is the Hopf fibration with the S^1 action in (4.17). Hence, the bundle $S^1 \to I_{3,\nu_1}^{-1}(\nu_3) \to I_{3,\nu_1}^{-1}(\nu_3)/S^1$ is bundle-isomorphic to the Hopf fibration for $0 < |\nu_3| < \nu_1$.

We emphasize that the degenerate Kepler orbits we added to $I_1^{-1}(\nu_1)/S^1$ are not contained in $I_{3,\nu_1}^{-1}(\nu_3)$ for $0 < |\nu_3| < \nu_1$. They are contained in the zero level set $I_{3,\nu_1}^{-1}(0)$. However, the zero level set $I_{3,\nu_1}^{-1}(0)$ is *not a smooth manifold* because of the critical points

$$\pm((0, 0, \nu_1/2), (0, 0, -\nu_1/2)). \tag{4.18}$$

The zero level set $I_{3,\nu_1}^{-1}(0)$ is *homeomorphic* to a quotient space $S^3/\sim$ where the equivalence relation identifies $S^1 \times 0$ with a point and $0 \times S^1$ with another point.

The reduced space $I_{3,\nu_1}^{-1}(0)/S^1$ is *homeomorphic* to S^2 but has *singular pinch points*. This is consistent with the case of the singular reduction (the critical points in (4.18) are the fixed points of the S^1 action). We will not study this singular case in the current paper. See Cushman (1991) and references therein for more details on this case.

Lastly, notice that

$$\begin{aligned} I_{3,\nu_1}^{-1}(\nu_1) &= ((0, 0, \nu_1/2), (0, 0, \nu_1/2)), \\ I_{3,\nu_1}^{-1}(-\nu_1) &= -((0, 0, \nu_1/2), (0, 0, \nu_1/2)). \end{aligned}$$

This completes our study of the level sets of I_{3,ν_1}.

We construct the Hopf fibration $S^1 \to I_{3,\nu_1}^{-1}(\nu_3) \to I_{3,\nu_1}^{-1}(\nu_3)/S^1$ *in coordinates* for $0 < |\nu_3| < \nu_1$. Coffey et al. (1986) suggested the following projection:

$$\mathbf{w} = (w_1, w_2, w_3) : I_{3,\nu_1}^{-1}(\nu_3) \subset S^2_{\nu_1/2} \times S^2_{\nu_1/2} \to I_3^{-1}(\nu_3)/S^1 \simeq S^2 \subset \mathbb{R}^3,$$

where

$$\begin{aligned} w_1 &= (\mathbf{L} \times \mathbf{R}) \cdot \mathbf{k} = 2(a_2 b_1 - a_1 b_2), \\ w_2 &= \|\mathbf{L}\|(\mathbf{R} \cdot \mathbf{k}) = \|\mathbf{a} + \mathbf{b}\|(a_3 - b_3), \\ w_3 &= \tfrac{1}{2}(\|\mathbf{L} \times \mathbf{k}\|^2 - \|\mathbf{R}\|^2) = \tfrac{1}{2}((a_1 + b_1)^2 + (a_2 + b_2)^2 - \|\mathbf{a} - \mathbf{b}\|^2). \end{aligned}$$

The map $\mathbf{w}$ satisfies

$$w_1^2 + w_2^2 + w_3^2 = \left(\frac{\nu_1^2 - \nu_3^2}{2}\right)^2$$

for $(\mathbf{a}, \mathbf{b}) \in I_{3,\nu_1}^{-1}(\nu_3)$. One can check

$$I_{3,\nu_1}^{-1}(\nu_3)/S^1 = \{\mathbf{w} \in \mathbb{R}^3 \mid \|\mathbf{w}\| = (\nu_1^2 - \nu_3^2)/2\}. \tag{4.19}$$

As a map on $S^2_{\nu_1/2} \times S^2_{\nu_1/2}$, $\mathbf{w} = (w_1, w_2, w_3)$ satisfy

$$\{w_i, w_j\} = 2\|\mathbf{a} + \mathbf{b}\|\epsilon_{ijk} w_k = 2(\nu_1^2 + w_3 - \|\mathbf{w}\|)^{1/2}\epsilon_{ijk} w_k. \tag{4.20}$$

It follows that the *reduced symplectic structure* on $I_{3,\nu_1}^{-1}(\nu_3)/S^1$ can be regarded as the one induced from the Poisson structure of $\mathbb{R}^3$ in (4.20). Thus, the symplectic structure on $I_{3,\nu_1}^{-1}(\nu_3)/S^1$ is the restriction of that in (4.20) to $I_{3,\nu_1}^{-1}(\nu_3)/S^1$, which is given by

$$\{w_i, w_j\} = 2\big(w_3 + \tfrac{1}{2}(\nu_1^2 + \nu_3^2)\big)^{1/2}\epsilon_{ijk} w_k.$$

The *reduced Hamiltonian* $\bar{H}_{\nu_1,\nu_3}$ on $I_{3,\nu_1}^{-1}(\nu_3)/S^1$ is given by

$$\bar{H}_{\nu_1,\nu_3} = -\frac{\mu^2}{2\nu_1^2} + \frac{J_2\mu^4 R_e^2(w_3 + (\nu_1^2 - 5\nu_3^2)/2)}{4\nu_1^3(w_3 + (\nu_1^2 + \nu_3^2)/2)^{5/2}}.$$

The *reduced Hamiltonian vector field*, $\dot{\mathbf{w}} = X_{\bar{H}_{\nu_1,\nu_3}}$ on $I_{3,\nu_1}^{-1}(\nu_3)/S^1$ is given by

$$\begin{aligned}
\dot{w}_1 &= \frac{3J_2\mu^4 R_e^2(\nu_1^2 - 9\nu_3^2 + 2w_3)w_2}{\nu_1^3(\nu_1^2 + \nu_3^2 + 2w_3)^3},\\
\dot{w}_2 &= -\frac{3J_2\mu^4 R_e^2(\nu_1^2 - 9\nu_3^2 + 2w_3)w_1}{\nu_1^3(\nu_1^2 + \nu_3^2 + 2w_3)^3}, \qquad \dot{w}_3 = 0.
\end{aligned}$$

Then, the flow of $X_{\bar{H}_{\nu_1,\nu_3}}$ with the initial condition $\mathbf{w}(0) = (w_{10}, w_{20}, w_{30})$ is easily read off as follows:

$$\mathbf{w}(t) = (r_0 \cos(\alpha t), r_0 \sin(\alpha t), w_{30}), \tag{4.21}$$

where

$$r_0 = \sqrt{(w_{10})^2 + (w_{20})^2} = \sqrt{\left(\frac{\nu_1^2 - \nu_3^2}{2}\right)^2 - (w_{30})^2},$$

and

$$\alpha = -\frac{3J_2\mu^4 R_e^2(\nu_1^2 - 9\nu_3^2 + 2w_{30})}{\nu_1^3(\nu_1^2 + \nu_3^2 + 2w_{30})^3}.$$

A direct computation shows that this angular frequency, α, coincides with a_2 in (4.4), the drift rate of the argument of the perigee, ω. This is an expected result because the flows φ^{I_1} and φ^{I_3} are reduced out in (4.8) by the $S^1 \times S^1$-reduction by $\Phi_1 \times \Phi_3$. Thus, the reduced flow is the same as the projection of $\varphi^{I_2}_{a_2 t}$ in (4.8). This explains why α coincides with a_2.

Bifurcations in the reduction picture. The equilibria of the reduced Hamiltonian dynamics on $I^{-1}_{3,\nu_1}(\nu_3)/S^1$ for $0 < |\nu_3| < \nu_1$ are easily computed from $X_{\bar{H}_{\nu_1,\nu_3}}$ as follows:

$$\begin{cases} \left\{\left(0, 0, \pm\frac{\nu_1^2-\nu_3^2}{2}\right)\right\} & \text{if } \nu_1/\sqrt{5} < |\nu_3| < \nu_1, \\ \left\{\left(0, 0, \pm\frac{\nu_1^2-\nu_3^2}{2}\right)\right\} \cup \left\{\mathbf{w} \in \mathbb{R}^3 \;\middle|\; \|\mathbf{w}\| = \frac{\nu_1^2-\nu_3^2}{2},\, w_3 = \frac{9\nu_3^2-\nu_1^2}{2}\right\} & \\ & \text{if } 0 < |\nu_3| < \nu_1/\sqrt{5}. \end{cases}$$

Notice that a bifurcation happens at the *critical value*

$$\nu_1^2 - 5\nu_3^2 = 0. \tag{4.22}$$

For $0 < |\nu_3| < \nu_1/\sqrt{5}$, a ring of new equilibria appears (Figure 4.2).

This ring of equilibria at $w_3 = (9\nu_3^2 - \nu_1^2)/2$, corresponds to the *critical inclination i* satisfying

$$\cos i = \frac{L_3}{\|\mathbf{L}\|} = \frac{1}{\sqrt{5}}$$

because of (4.22), (4.12), and the definition of w_3. This ring of equilibria is due to the averaging performed over the (unperturbed) Kepler flow. To break this *fictitious* degeneracy, one needs to study with the higher order normal form of the J_2-Hamiltonian in (4.1) (see Cushman, 1991 for more details). However, one should keep in mind that the reduction of the symplectic spaces performed in the current paper is independent of the Hamiltonians involved.

Geometric phases in the second reduction. In the Hopf fibration, $S^1 \to I^{-1}_{3,\nu_1}(\nu_3) \to I^{-1}_{3,\nu_1}(\nu_3)/S^1$, we want to compute the amount of the angle

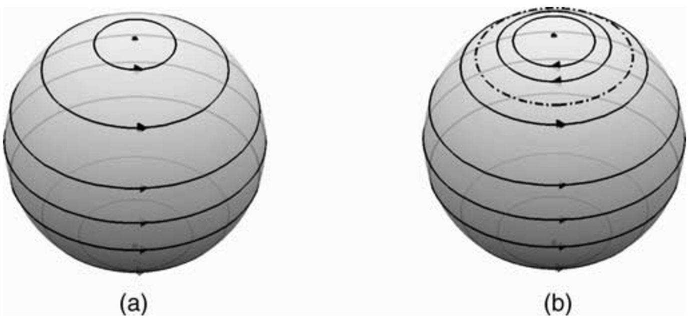

Figure 4.2. Flows on the reduced sphere. (a) $\nu_1/\sqrt{5} < |\nu_3| < \nu_1$; (b) $0 < |\nu_3| < \nu_1/\sqrt{5}$.

displacement in the fiber during one period of the reduced flow in (4.21). The period Δt is given by

$$\Delta t = \frac{2\pi}{\alpha}.$$

Notice that the fiber variables corresponds to Ω, the longitude of the ascending node, in a local trivialization. The quantity a_3 in (4.5) can be expressed in terms of the w_3 component of the reduced flow as well as ν_1 and ν_3 as follows:

$$a_3(w_3; \nu_1, \nu_3) = -\frac{3J_2\mu^4 R_e^2 \nu_3}{2\nu_1^3(w_3 + (\nu_1^2 + \nu_3^3)/2)^{5/2}}.$$

Hence, a_3 is constant along the flow on the level set $I_{3,\nu_1}^{-1}(\nu_3)$, so the phase displacement, $\Delta\Omega$, during the time interval Δt is given by

$$\Delta\Omega = a_3\, \Delta t. \tag{4.23}$$

Thus, one can also write

$$\Delta\Omega = \frac{4\pi\nu_3\sqrt{w_3 + (\nu_1^2 + \nu_3^2)/2}}{w_3 + (\nu_1^2 - 9\nu_3^2)/2}. \tag{4.24}$$

Since the reduced integral curves shown in Figure 4.2 lie in level sets of w_3, one can replace w_3 with its initial value. Interestingly, this formula for the phase depends only on the basic data in the geometry of the reduction construction (ν_1, ν_3 and initial data w_{30} on the reduced space that picks out a particular reduced trajectory) and not on the Hamiltonian; specifically, it is written in a way so that J_2 does not appear. In *this sense* it is *geometric*. Figure 4.3 illustrates, in a schematic way, the phase displacement. Of course at this point one could bring in the general

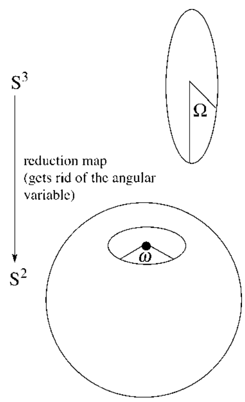

Figure 4.3. Schematic of the bundle whose associated geometric phase gives the angular drift $\Delta\Omega$ given by (4.24) during one period Δt of the reduced dynamics.

machinery of geometric and dynamic phases and reconstruction to bear on the problem for additional geometric insight, as in Marsden et al. (1990, 2000), but we shall not pursue the matter further in the present paper.

References

Abraham, R. and Marsden, J. E.: 1978, *Foundations of Mechanics*, 2nd edn, Addison-Wesley, Reading, MA.

Arnold, V. I.: 1991, *Dynamical Systems*, Springer, New York, Vol. III.

Born, M.: 1927, *The Mechanics of the atom*, G. Bells and Sons, Ltd, London.

Broucke, R. A.: 1994, Numerical integration of periodic orbits in the main problem of artificial satellite theory, Celestial Mechanics and Dynamical Astronomy **58**, 99–123.

Brouwer, D. and Clemence, G. M.: 1961, *Methods of Celestial Mechanics*, Academic Press, New York.

Chang, D. E., Chichka, D. F. and Marsden, J. E.: 2002, 'Lyapunov-based transfer between elliptic Keplerian orbits', *Discrete Contin. Dyn. Syst. B* **2**, 57–67.

Charlier, C. L.: 1927, *Die Mechanik des Himmels*, *Bd.I*, *II*, 2nd edn, Walter de Gruyter, Berlin.

Chobotov, V. A. (ed.): 1996, *Orbital Mechanics*, 2nd edn, AIAA Education Series, American Institute of Aeronautics and Astronautics, Inc., New York.

Coffey, S. L., Depit, A. and Miller, B. R.: 1986, 'The critical inclination in artificial satellite theory', *Cel. Mech.*, **39**, 365–406.

Cushman, R.: 1991, In: K. Jones et al. (eds), A Survey of Normalization Techniques Applied to Perturbed Keplerian Systems, *Dynamics Reported*, Springer, New York, Vol. 1.

Cushman, R. H. and Bates, L. M.: 1997, *Global Aspects of Classical Integrable Systems*, Birkhüser, Boston.

Delaunay, C.: 1860, Théorie du mouvement de la lune, Mem. **28** (1860); **29** (1867), *Acad. Sci. France*, Paris.

Goldstein, H.: 1980, *Classical Mechanics*, 2nd edn, Addison-Wesley, Reading, MA.

Kovalevsky, J.: 1967, *Introduction to Celestial Mechanics*, Springer, New York.

Marsden, J. E., Misiolek, G., Perlmutter, M. and Ratiu, T.: 1998, Symplectic reduction for semidirect products and central extensions, *Diff. Geom. Appl.* **9**, 173–212.

Marsden, J. E., Montgomery, R. and Ratiu, T. S.: 1990, *Reduction, Symmetry and Phases in Mechanics*, Vol. 436 of *Memoirs of the AMS*, American Mathematical Society, Providence, RI.

Marsden, J. E. and Ratiu, T. S.: 1999, *Introduction to Mechanics and Symmetry*, 2nd edn, Springer, New York.

Marsden, J. E., Ratiu, T. S. and Scheurle, J.: 2000, Reduction theory and the Lagrange–Routh equations, *J. Math. Phys.* **41**, 3379–3429.

Marsden, J. E. and Weinstein, A.: 1974, Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5**, 121–130.

Naber, G. L.: 1997, *Topology, Geometry, and Gauge Fields: Foundations*, Springer, New York.

Vinti, J. P.: 1998, *Orbital and Celestial Mechanics*, AIAA, Virginia.