KAM Stability of The Kepler Problem with a General Relativistic Correction Term

By

Majed Sofiani

Submitted to the graduate degree program in Department of Mathematics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

Beat Schaad, Chairperson

Committee members

Atanas Stefanov, Member

Milena Stanislavova, Member

Date defended: May 10, 2017

The Thesis Committee for Majed Sofiani certifies
that this is the approved version of the following thesis :

KAM Stability of The Kepler Problem with a General Relativistic Correction Term

____

Beat Schaad, Chairperson

Date approved: May 10, 2017

Abstract

In this work, we will be investigating a specific Hamiltonian system, namely, the Kepler problem with a correction term $\frac{\delta}{r^3}$ added to the potential energy. Our objective is to show that the system is stable in the sense of the KAM theorem. In the first sections, we introduce essential concepts and tools that will be used in the process of understanding and showing how the KAM theorem works with our system. These concepts and tools are: Hamiltonian formalism, canonical transformations, the Hamilton-Jacobi equation and Action-Angle variables. In the last section, we state the KAM theorem and, based on the results we obtain from previous sections, we can conclude that the system is in fact stable in the sense of the KAM theorem. An informal statement of the KAM theorem is that if the unperturbed Hamiltonian system H_0, expressed in the action variable J, is non-degenerate, then under sufficiently small perturbation εH_1 we have that

$$H(J,\Phi) = H_0(J) + \varepsilon H_1(J,\Phi)$$

for $\varepsilon << 1$, most of the quasiperiodic orbits persist under the small perturbation εH_1. The system under which the perturbation will be acting is the following

$$H_0(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\sin^2\theta}p_\phi^2) - \frac{\gamma}{r} \pm \frac{\delta}{r^3},$$

for $\gamma > 0$.

Acknowledgements

I would like to express my sincere gratitude and appreciation to Prof. Beat Schaad for his continuous guidance, support, and assistance during these past years. This work would not have been completed without his help. It has been a great pleasure working under his supervision.

Special thanks go to my friend Haider Alhazmi, Department of Physics at University of Kansas, for his valuable and insightful comments during our discussions on some ideas related to this work.

I am also indebted to my parents, brothers, and sisters for always being there for me. Thank you for your words of encouragement.

Contents

Chapter 1

Introduction

Celestial mechanics is one of the centuries-old subjects that people have been trying to understand [**1**]. Philosophers, historians, and even religious scholars were always curious to discover the sky. Later, Johannes Kepler (1571-1630) devised three magnificent laws known as Kepler's laws.

- All planets move about the Sun in elliptical orbits, having the Sun as one of the foci.
- A radius vector joining any planet to the Sun sweeps out equal areas in equal periods of time.
- The squares of the sidereal periods (of revolution) of the planets are directly proportional to the cubes of their mean distances from the Sun.

However, Kepler did not explain or prove why the motion of planets obeys these laws. In 1687, Newton stated his laws in *Philosophiae naturalis Principia Mathematica* ("Mathematical Principles of Natural Philosophy"). These laws are known as Newton's laws.

- An object will remain at rest or in uniform motion in a straight line unless acted upon by an external force.
- The vector sum of the forces F on an object is equal to the mass m of that object multiplied by the acceleration a of the object: $\mathbf{F} = m\mathbf{a}$.
- For every external force that acts on an object there is a force of equal magnitude but opposite direction which acts back on the object which exerted that external force.

Newton's laws lasted for a long time until 1905 when Albert Einstein presented his theory of relativity. Unlike what Newton discovered, Einstein considered mass, time, and space as quantities that are dependent on the observer. In other words, two observers within two different frames of reference have different observations. However, Newton's laws still work within a single frame of reference and define the trajectories of the motion in a good manner [**2**], good enough to take us to the moon and back.

About fifty years later, Andrey Kolmogorov, a Russian mathematician, announced a theorem regarding the stability of dynamical systems. Mathematicians realized that this theorem, if true, can be applied to planetary motion to prove the stability of our solar system. Although Kolmogorov did not spell out a rigorous proof of his theory, he outlined the proof-techniques in 1954. Several years later, Jurgen Moser, a German-American mathematician was able to prove a low dimensional version of Kolmogorov theorem. In 1963, Kolmogorov's student, Vladimir Arnold, provided a proof of a valid version of the theorem in all dimensions. Since then, this astonishing result was known as Kolmogorov-Arnold-Moser theory (KAM theory)[**3**].

In this work, we will study in detail the two-body problem subject to the inverse-square force with a quartic correction term. Namely, we will study the motion of the system , define and construct the action-angle variables of the system with this correction term, and see how the KAM theorem can be applied by testing the assumptions of the theorem. To do that, we need to introduce, briefly, some concepts of calculus of variations, classical mechanics, and Hamiltonian formalism.

Formally, the Hamiltonian of the system we are looking at can be written in the following form

$$H_0(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\sin^2\theta}p_\phi^2) - \frac{\gamma}{r} \pm \frac{\delta}{r^3}.$$

For $\gamma > 0$.We show that if the unperturbed Hamiltonian system H_0, expressed in the action variable J, is non-degenerate, then under sufficiently small perturbation such that

$$H(J,\Phi) = H_0(J) + \varepsilon H_1(J,\Phi)$$

for $\varepsilon << 1$, most of the quasiperiodic orbits persist under the small perturbation.

Chapter 2

Motivation

The definition of Schwardschild metric is written in the following form

$$c^2 d\tau^2 = c^2(1 - \frac{2\mu}{r})dt^2 - (1 - \frac{2\mu}{r})^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{2.0.1}$$

Where

τ is the proper time

$2\mu = \frac{2GM}{c^2}$ is Schwardschild radius, where c is the speed of light.

(r, θ, ϕ) spherical coordinates.

Associated with the equations of motion of a test particle in the Schwardschild space-time are two constants of motion

$$(1 - \frac{2\mu}{r})\frac{dt}{d\tau} = \frac{E}{mc^2}$$
$$r^2\frac{d\theta}{d\tau} = h.$$

Where E and h are the total energy and the specific angular momentum, respectively. Using these two constant of motion quantities to eliminate the dependence of the equations of motion on the two variables t and θ leads to the following equation

$$\dot{r}^2 + \frac{h^2}{r^2}(1 - \frac{2\mu}{r}) - \frac{2\mu c^2}{r} = \frac{E^2}{m^2c^2} - c^2. \tag{2.0.2}$$

We see in the last equation that it is written in the form of a kinetic and potential energy, multiplied by 2, of a particle moving in a central field influenced by the following gravitational effective potential energy

$$U(r) = \frac{h^2}{2r^2} - \frac{\mu h^2}{r^3} - \frac{\mu c^2}{r} \tag{2.0.3}$$

This cubic term, as we will see, will not appear in the equivalent Newtonian mechanics. Therefore, we will investigate the affect that this extra term will have on the system when considering it.

Another way one could encounter such a form is by considering the gravitational potential energy of a not spherically symmetric mass distribution ρ, say with unit mass, traveling around a massive star with mass M.

$$U(\vec{R}) = \int_D \frac{\rho(\vec{r}')}{|\vec{R} - \vec{r}'|} d\vec{r}' \tag{2.0.4}$$

Where ρ is the mass distribution function and $\vec{R}$ is the separation vector from the position of M. Here $\vec{r'}$ is taken to be small relative to $\vec{R}$. We can write

$$\begin{aligned}
|\vec{R} - \vec{r'}| &= \sqrt{|R^2 - 2\vec{R}\vec{r}' + r'^2|} \\
&= R\sqrt{|1 - 2\frac{\hat{\vec{R}} \cdot \vec{r}'}{R} + (\frac{r'}{R})^2|} \\
&= R\sqrt{|1+S|}
\end{aligned}$$

Where $S = (\frac{r'}{R})^2 - \frac{2\hat{\vec{R}} \cdot \vec{r}'}{R}$ and $\hat{\vec{R}} = \frac{\vec{R}}{R}$.

$$\begin{aligned}
\frac{1}{\sqrt{|1+S|}} &= 1 - \frac{1}{2}S + \frac{3}{8}S^2 + O(s^3) \\
&= 1 + \frac{\hat{\vec{R}} \cdot \vec{r}'}{R} - \frac{1}{2}\frac{r'^2}{R^2} + \frac{3}{2}\frac{(\hat{\vec{R}} \cdot \vec{r}')^2}{R^2} + O(\frac{r'}{R})^3 \\
&= 1 + \frac{\hat{\vec{R}} \cdot \vec{r}'}{R} - \frac{1}{2R^2}(r'^2 - 3(\hat{\vec{R}} \cdot \vec{r}')^2) + O(\frac{r'}{R})^3
\end{aligned}$$

Hence,

$$U(R) \approx \int_D \frac{\rho(\vec{r}')}{R} + \frac{(\hat{\vec{R}} \cdot \vec{r}')\rho(\vec{r}')}{R^2} - \frac{\rho(\vec{r}')(r'^2 - 3(\hat{\vec{R}} \cdot \vec{r}')^2)}{2R^3} d\vec{r}' \tag{2.0.5}$$

These three terms are called Monopole, Dipole, and Quadrupole, respectively. Note that the dipole term, in the case of gravitational potential, vanishes.

Chapter 3

Variational Principles and Euler-Lagrange Equation

3.1 Lagrange Function

Lagrange function or Lagrangian $L = L(t, q(t), \dot{q}(t))$ is a function that describes the dynamics of the system, where $q(t)$ is the position of the particle and $\dot{q}(t)$ is the velocity $v(t)$. We define the Lagrangian of a system as follows:

$$L = T - U,$$

where $T = \frac{mv^2}{2}$ and $U = U(q)$ are **kinetic energy** and **potential energy**, respectively. Associated with such a system is the **action integral**

$$I(q) = \int_{t_0}^{t_1} L(t, \dot{q}(t), q(t))dt \tag{3.1.1}$$

with boundary conditions $q(t_0) = q_0$ and $q(t_1) = q_1$. Moreover, the motion of the system is described by a stationary point of the above functional.

Definition 3.1. For the functional (3.1.1), the equation

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \tag{3.1.2}$$

is called the **Euler-Lagrange equation**.

Definition 3.2. A function $q(t)$ is a **stationary point** of the functional (3.1.1) if the first derivative of (3.1.1) vanishes at $q(t)$.

Theorem 3.1. *Let $I : C^2[t_0, t_1] \to \mathbb{R}$ be a functional defined as* (3.1.1) *where L has continuous partial derivatives of second order with respect to the variables it depends on. Let $G = \{q \in C^2[t_0, t_1] : q(t_0) = q_0 \, and \, q(t_1) = q_1\}$, where q_0 and q_1 are fixed real numbers. Then I is stationary at $q \in G$ if and only if the Euler-Lagrange equation is satisfied along q.[**4**]*

Proof. Assume I is stationary at function $q(t)$. Define $\hat{q} = q + \varepsilon\eta$ such that $\eta(t) \in \{\gamma(t) \in C^2[t_0,t_1] : \gamma(t_0) = \gamma(t_1) = 0\}$. Now, consider

$$I(\hat{q}) = \int_{t_0}^{t_1} L(t,\hat{q},\dot{\hat{q}})dt$$

as a function of ε and apply $\frac{\partial}{\partial\varepsilon}$. By the definition of a stationary point, we have that $I'(\varepsilon)\big|_{\varepsilon=0} = 0$.

$$I'(\varepsilon) = \int_{t_0}^{t_1} [\frac{\partial L(t,\hat{q},\dot{\hat{q}})}{\partial q}\eta + \frac{\partial L(t,\hat{q},\dot{\hat{q}})}{\partial \dot{q}}\dot{\eta}]dt.$$

At $\varepsilon = 0$, we have:

$$I'(0) = \int_{t_0}^{t_1} [\frac{\partial L(t,q,\dot{q})}{\partial q}\eta + \frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\dot{\eta}]dt = 0. \tag{3.1.3}$$

Using integration by parts for the second term in (3.1.3), we obtain:

$$\int_{t_0}^{t_1} \eta\,[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}]dt = 0.$$

Since the last equation holds for all η such that $\eta(t) = \{\eta(t) \in C^2[t_0,t_1] : \eta(t_0) = \eta(t_1) = 0\}$, then we conclude that

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0. \tag{3.1.4}$$

Equation (3.1.4) is what is so called Eluer-Lagrange equation.

Conversely, the functional I is stationary at q means that the derivative of I vanishes at q, that is we want to show that,

$$I'(\varepsilon)\big|_{\varepsilon=0} = 0$$

$$\begin{aligned} I'(0) &= \int_{t_0}^{t_1} [\frac{\partial L(t,q,\dot{q})}{\partial q}\eta + \frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\dot{\eta}]dt \\ &= \int_{t_0}^{t_1} [\frac{d}{dt}\frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\eta + \frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\dot{\eta}]dt. \end{aligned}$$

Using, by assumption, the Euler-Lagrange equation. Again, using integration by parts in the first term, we get

$$I'(0) = \int_{t_0}^{t_1} [-\frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\dot{\eta} + \frac{\partial L(t,q,\dot{q})}{\partial \dot{q}}\dot{\eta}]dt = 0.$$

This completes the proof.

□

Theorem 3.2. *Newton's equation of dynamics coincide with stationary points of the functional I.[5]*

Proof. Consider an n-dimensional system. Newton's equations of dynamics read

$$m_k\ddot{q}_k + \frac{\partial U}{\partial q_k} = 0, \qquad (3.1.5)$$

for $1 \leq k \leq n$. In Theorem 3.1 we showed that a stationary point of any functional must satisfy the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0,$$

where $L = \sum_{k=1}^{n} \frac{m_k \dot{q}_k^2}{2} - U(q_k)$. Obviously, applying the Euler-Lagrange equation to L leads to (3.1.5). □

3.2 Hamilton Function

Consider a mechanical system that has n degrees of freedom. As we know, we can write the Lagrangian function of this system as the following:

$$L = T - U,$$

where $T = \sum_{k=1}^{n} \frac{m_k v_k^2}{2}$, $U = U(\mathbf{q})$ and $\mathbf{q} = (q_1, q_2, q_3, ...q_n)$.

Now, we introduce a very useful mathematical tool that provides the link between Lagrange function and Hamilton function. For simplicity, consider a function f with one degree of freedom such that $f : \mathbb{R} \to \mathbb{R}$ and $f'' > 0$. Now, for some fixed number p such that $\max\{px - f(x)\}$ exists, define a new function $G(p) = px$

as shown in figure (3.1). We map p to the point x_0 that solves the following equation

$$\max\{px - f(x)\} = 0.$$

In other words,

$$\frac{d}{dx}\left[px - f(x)\right]_{x=x_0} = 0 \Rightarrow p = f'(x_0).$$

Now, we can define a new function

$$g(p) = px_0(p) - f(x_0(p))$$

which is a transformation from $(f(x), x)$ to $(g(p), p)$. We call this transformation **Legendre transformation**.

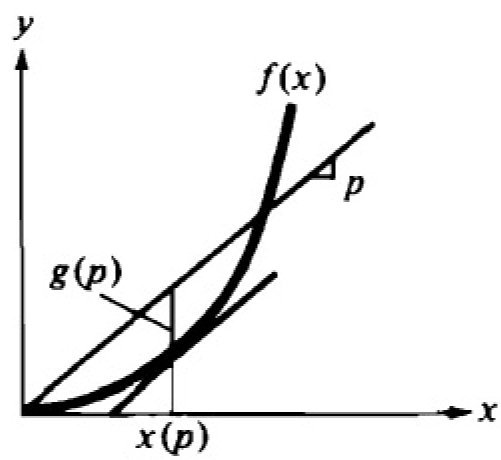

Figure 3.1: Legendre transformation [**5**]

Using Legendre transformation to transform the Lagrangian $L = L(\mathbf{q}, \mathbf{v})$ to some function $H = H(\mathbf{q}, \mathbf{p})$ we write

$$\sum_{k=1}^{n} p_k v_k - L(t, \mathbf{q}, \mathbf{v}) = H(t, \mathbf{q}, \mathbf{p}), \tag{3.2.1}$$

where $p_k = \frac{\partial L}{\partial v_k} = m_k v_k$ is the momentum.

Now we want to write an explicit formula for $H(t,\mathbf{q},\mathbf{p})$,

$$\begin{aligned} H(\mathbf{q},\mathbf{p}) &= \sum_{k=1}^{n} p_k v_k - \frac{m_k v_k^2}{2} + U(\mathbf{q}) \\ &= \frac{m_k v_k^2}{2} + U(\mathbf{q}) \\ &= \frac{p_k^2}{2m_k} + U(\mathbf{q}). \end{aligned}$$

The above equation is called **Hamilton Function** or **Hamiltonian**. Notice that the Hamiltonian $H(\mathbf{q},\mathbf{p})$ is the total energy of the system. Since we showed that Hamilton function can be obtained from the Lagrangian system, we should be able to get the corresponding equations to Euler-Lagrange equations. We start with the total derivative of the Lagrangian:

$$\frac{dL(t,q_k,\dot{q}_k)}{dt} = \sum_{k=1}^{n} \frac{\partial L}{\partial q_k}\dot{q}_k + \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k + \frac{\partial L}{\partial t} = \sum_{k=1}^{n} \frac{\partial L}{\partial q_k}\dot{q}_k + p_k\ddot{q}_k + \frac{\partial L}{\partial t} = \sum_{k=1}^{n} \frac{\partial L}{\partial q_k}\dot{q}_k + \frac{d}{dt}(p_k\dot{q}_k) - \dot{p}_k\dot{q}_k + \frac{\partial L}{\partial t},$$

so we have

$$\frac{d}{dt}(p_k\dot{q}_k - L) = \dot{p}_k\dot{q}_k - \frac{\partial L}{\partial q_k}\dot{q}_k - \frac{\partial L}{\partial t}.$$

Using Legendre transformation (3.2.1) we obtain:

$$\frac{dH}{dt} = \dot{p}\dot{q} - \frac{\partial L}{\partial q}\dot{q} - \frac{\partial L}{\partial t}. \tag{3.2.2}$$

However, since $H = H(t,q_k,p_k)$,

$$\frac{d}{dt}H = \frac{\partial H}{\partial q_k}\dot{q}_k + \frac{\partial H}{\partial p_k}\dot{p}_k + \frac{\partial H}{\partial t}. \tag{3.2.3}$$

Comparing (3.2.2) with (3.2.3) leads to:

$$\frac{\partial H}{\partial p_k} = \dot{q}_k \tag{3.2.4}$$

$$\frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k} = -\dot{p}_k \tag{3.2.5}$$

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{3.2.6}$$

Using the Euler-Lagrange equation in (3.2.5). Equations (3.2.4) and (3.2.5) are the equations corresponding to Euler-Lagrange (3.1.4).

So far, we showed that by means of Legendre transformation, one can convert a Lagrangian system of n second-order differential equations (3.1.4) to a system of $2n$ first-order differential equations (3.2.4) and (3.2.5). The $2n$ equations corresponding to the Euler-Lagrange equations are called **Hamilton's Equations**.

According the above work, we should mention that the equations of motion (Hamilton's equations) of a Hamiltonian system that depends on $2n$ coordinates, say, $\mathbf{x} = (q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$ can be written in the following form

$$\dot{\mathbf{x}} = \mathbf{I} \cdot \nabla H(\mathbf{x}) \tag{3.2.7}$$

where $\nabla = (\partial q_1, \ldots, \partial q_n, \partial p_1, \ldots, \partial p_n)$, $\mathbf{I} = \begin{pmatrix} 0 & I_{n\times n} \\ -I_{n\times n} & 0 \end{pmatrix}$, and I is the identity matrix.

Chapter 4

Canonical Transformations

4.1 Equations of Canonical Transformation

The main purpose of finding a transformation is to obtain a simpler form of the system we are working on. Then, we can solve the simple one and use the transformation equation to get a solution to our original problem or system. In this section, we will find a sufficient condition of what is called a canonical transformation.

Consider a Hamiltonian $H(q,p)$. A canonical transformation is a function, say, ϕ that maps (q,p) into (Q,P) and gives a new Hamiltonian $K = K(Q,P)$ such that the structure of the Hamilton's equations is preserved, that is,

$$\frac{\partial K}{\partial P} = \dot{Q} \tag{4.1.1}$$

$$\frac{\partial K}{\partial Q} = -\dot{P}. \tag{4.1.2}$$

Considering $P = P(q,p)$ and $Q = Q(q,p)$ functions of q and p gives the total time derivative as the following:

$$\dot{P}(q,p) = \frac{\partial P}{\partial p}\dot{p} + \frac{\partial P}{\partial q}\dot{q}$$
$$\dot{Q}(q,p) = \frac{\partial Q}{\partial p}\dot{p} + \frac{\partial Q}{\partial q}\dot{q}.$$

So we can write this system of two equations as follows

$$\begin{pmatrix} \dot{P} \\ \dot{Q} \end{pmatrix} = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial Q}{\partial q} \end{pmatrix} \begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial P}{\partial q} \\ \frac{\partial Q}{\partial p} & \frac{\partial Q}{\partial q} \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial p} \\ \frac{\partial H}{\partial q} \end{pmatrix},$$

where I is the identity matrix.

Let $K = H \circ \phi^{-1}$ be the transformed Hamiltonian.

$$\frac{\partial H}{\partial p} = \frac{\partial K}{\partial P}\frac{\partial P}{\partial p} + \frac{\partial K}{\partial Q}\frac{\partial Q}{\partial p}$$
$$\frac{\partial H}{\partial q} = \frac{\partial K}{\partial P}\frac{\partial P}{\partial q} + \frac{\partial K}{\partial Q}\frac{\partial Q}{\partial q}$$

Which read

$$\begin{pmatrix} \frac{\partial H}{\partial p} \\ \\ \frac{\partial H}{\partial q} \end{pmatrix} = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\ \\ \frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q} \end{pmatrix} \begin{pmatrix} \frac{\partial K}{\partial P} \\ \\ \frac{\partial K}{\partial Q} \end{pmatrix}.$$

Therefore, we have

$$\begin{pmatrix} \dot{P} \\ \dot{Q} \end{pmatrix} = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial P}{\partial q} \\ \\ \frac{\partial Q}{\partial p} & \frac{\partial Q}{\partial q} \end{pmatrix} \begin{pmatrix} 0 & -I \\ \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial p} \\ \\ \frac{\partial H}{\partial q} \end{pmatrix} = \left[\begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial P}{\partial q} \\ \\ \frac{\partial Q}{\partial p} & \frac{\partial Q}{\partial q} \end{pmatrix} \begin{pmatrix} 0 & -I \\ \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\ \\ \frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q} \end{pmatrix} \right] \begin{pmatrix} \frac{\partial K}{\partial P} \\ \\ \frac{\partial K}{\partial Q} \end{pmatrix}.$$

On the other hand (4.1.1) and (4.1.2) read as follows

$$\begin{pmatrix} \dot{P} \\ \dot{Q} \end{pmatrix} = \begin{pmatrix} 0 & -I \\ \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial K}{\partial P} \\ \\ \frac{\partial K}{\partial Q} \end{pmatrix}.$$

Therefore, the change of coordinates function ϕ is canonical if and only if the Jacobian matrix of ϕ is **symplectic**.

Definition 4.1. A $2n \times 2n$ matrix J is symplectic if

$$J \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} J^T = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Where I is the $n \times n$ identity matrix.

To conclude, the sufficient condition to have a canonical transformation is having a symplectic Jacobian matrix of the transformation mapping from the old coordinates, say (q, p) phase space, to the new ones, say (Q, P) phase space [7]. In other words, a canonical transformation is a transformation, say, ϕ such that $\phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined by $\phi(q, p) = (Q, P)$ with the property that Hamilton's equations structure is invariant.

Remark 4.1. Let the Jacobian matrix of ϕ be denoted as $\partial\phi$. Then we say that in order to have a canonical transformation this matrix must be symplectic. that is,

$$\partial\phi \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \partial\phi^T = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Moreover, one can use some properties for determinants to find that the determinant of $\partial\phi$.

$$\begin{aligned} &\det \left[\partial\phi \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \partial\phi^T\right] = \det \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \\ &\det[\partial\phi] \det[\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}] \det[\partial\phi^T] = 1 \\ &[\det \partial\phi]^2 = 1 \Rightarrow \det \partial\phi = \pm 1. \end{aligned}$$

Of course, the fact that the determinant of $\partial\phi$ does not vanish is crucial because we need to solve for our original variables q and p. Therefore, having $\det \partial\phi \neq 0$ means that it is a non-singular matrix which means we can locally find (q, p) using $(q, p) = \phi^{-1}(Q, P)$.

Remark 4.2. From remark 4.1 we showed that the determinant of the Jacobian matrix of a canonical transformation is ± 1. This property, actually, can be used to show that a canonical transformation that maps (q, p) to (Q, P) preserves the volume in phase space. This is what *Liouville's theorem* basically says. [9]

4.2 Generating Functions

Now that we discussed what is so called a canonical transformation and how we can obtain a simpler Hamiltonian without loosing some important features of the original Hamiltonian such as the structure of

Hamilton's equations and the volume in phase space, it is time to introduce an important method by which we can find a canonical transformation.

Considering change in the action I from a path to another, as we have seen in chapter (1), gives the Euler-Lagrange equation. However, the action can be written in this form

$$I = \int_{t_0}^{t_1} \sum_{k=1}^{n} p_k \dot{q}_k - H(t, \mathbf{q}, \mathbf{p})\, dt \tag{4.2.1}$$

and we can obtain Hamilton's equations by the same technique we used to obtain the Euler-Lagrange equation in Chapter (1). Now assume we have new variables $(\mathbf{Q}, \mathbf{P})$ that satisfy Hamilton's equations

$$\dot{Q}_k = \frac{\partial K}{\partial P_k} \tag{4.2.2}$$

$$\dot{P}_k = -\frac{\partial K}{\partial Q_k}, \tag{4.2.3}$$

for some Hamiltonian $K(\mathbf{Q}, \mathbf{P})$. Since the new variables satisfy Hamilton's equations, then the following holds

$$\delta \hat{I} = \delta \int_{t_0}^{t_1} \sum_{k=1}^{n} P_k \dot{Q}_k - K(t, \mathbf{Q}, \mathbf{P})\, dt = 0 \tag{4.2.4}$$

$$\delta I = \delta \int_{t_0}^{t_1} \sum_{k=1}^{n} p_k \dot{q}_k - H(t, \mathbf{q}, \mathbf{p})\, dt = 0, \tag{4.2.5}$$

where δI is the independent variation of the q's and the p's and we can write it as

$$\delta \int_{t_0}^{t_1} \sum_{k=1}^{n} p_k \dot{q}_k - H(t, \mathbf{q}, \mathbf{p})\, dt = \int_{t_0}^{t_1} \delta p_k dq_k + p_k d\delta q_k - \frac{\partial H}{\partial q_k} \delta q_k dt - \frac{\partial H}{\partial p_k} \delta p_k dt. \tag{4.2.6}$$

Integrating the second term by parts

$$\int_{t_0}^{t_1} p_k d\delta q_k = p_k \delta q_k |_{t_0}^{t_1} - \int_{t_0}^{t_1} \delta q_k dp_k, \tag{4.2.7}$$

we obtain

$$\delta \int_{t_0}^{t_1} \sum_{k=1}^{n} p_k \dot{q}_k - H(t, \mathbf{q}, \mathbf{p})\, dt = \int_{t_0}^{t_1} \delta p_k [dq_k - \frac{\partial H}{\partial p_k} dt] + p_k \delta q_k |_{t_0}^{t_1} - \int_{t_0}^{t_1} \delta q_k [dp_k - \frac{\partial H}{\partial q_k} dt]. \tag{4.2.8}$$

As we did before, we require that the variation of q_k at the end points vanishes, so we have the variation of the paths starting and ending from fixed points at t_0 and t_1. Finally, by the independent variation of the q's and the p's, we conclude that the only way to obtain $\delta I = 0$ is to make the two integrals vanish, and hence we have Hamilton's equations. Similarly, the equations of motion of the Hamiltonian expressed in terms of $(\mathbf{Q}, \mathbf{P})$ can be derived.

The two actions are variationally equivalent if there exists a function ψ such that

$$\sum_{k=1}^{n} p_k \dot{q}_k - H(t, \mathbf{q}, \mathbf{p}) = \sum_{k=1}^{n} P_k \dot{Q}_k - K(t, \mathbf{Q}, \mathbf{P}) + \frac{d}{dt}\psi(t, \mathbf{q}, \mathbf{p}). \tag{4.2.9}$$

Since

$$\int_{t_0}^{t_1} \frac{d}{dt}\psi(t, \mathbf{q}, \mathbf{p})dt = \psi(t_1, \mathbf{q}(t_1), \mathbf{p}(t_1)) - \psi(t_0, \mathbf{q}(t_0), \mathbf{p}(t_0))$$

is a constant, then the variation of the actions $I(\mathbf{q}, \mathbf{p})$ and $\hat{I}(\mathbf{Q}, \mathbf{P})$ is not affected [1]. Rewriting (4.2.9), by means of $Q_k = Q_k(t, \mathbf{q}, \mathbf{p})$, we can write $\psi(t, \mathbf{q}, \mathbf{p}) = \psi_1(t, \mathbf{q}, \mathbf{Q})$ and

$$\frac{d}{dt}\psi_1(t, \mathbf{q}, \mathbf{Q}) = \sum_{k=1}^{n} p_k \dot{q}_k - \sum_{k=1}^{n} P_k \dot{Q}_k + K(t, \mathbf{Q}, \mathbf{P}) - H(t, \mathbf{q}, \mathbf{p}). \tag{4.2.10}$$

Also, we can write

$$\frac{d}{dt}\psi_1(t, \mathbf{q}, \mathbf{Q}) = \sum_{k=1}^{n} \left[\frac{\partial \psi_1}{\partial q_k}\dot{q}_k + \frac{\partial \psi_1}{\partial Q_k}\dot{Q}_k\right] + \frac{\partial \psi_1}{\partial t}. \tag{4.2.11}$$

Comparing (4.2.10) with (4.2.11)

$$\sum_{k=1}^{n} p_k \dot{q}_k - \sum_{k=1}^{n} P_k \dot{Q}_k + K(t, \mathbf{Q}, \mathbf{P}) - H(t, \mathbf{q}, \mathbf{p}) = \sum_{k=1}^{n} \left[\frac{\partial \psi_1}{\partial q_k}\dot{q}_k + \frac{\partial \psi_1}{\partial Q_k}\dot{Q}_k\right] + \frac{\partial \psi_1}{\partial t},$$

which implies the following set of equations

$$p_k = \frac{\partial \psi_1}{\partial q_k} \tag{4.2.12}$$

[1] Think of it as adding a constant to a function that has a stationary point, this point is invariant under the addition of a constant.

$$P_k = -\frac{\partial \psi_1}{\partial Q_k} \tag{4.2.13}$$

$$H(t,\mathbf{q},\mathbf{p}) + \frac{\partial \psi_1}{\partial t} = K(t,\mathbf{Q},\mathbf{P}). \tag{4.2.14}$$

The function ψ_1 is called the **Generating Function** of the canonical transformation. The associated transformation with ψ_1 is $(Q_k, P_k) = (Q_k(t,\mathbf{q},\mathbf{p}), -\frac{\partial \psi_1}{\partial Q_k}(t,\mathbf{q},\mathbf{Q}))$ and from (4.2.12) we can write $Q_k = Q_k(t,\mathbf{q},\mathbf{p})$ given that $|\frac{\partial^2 \psi_1}{\partial q_n \partial Q_n}| \neq 0$.

Observe that we can use the transformations

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

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

to construct four types of generating functions that give different canonical mappings [**4**]. For example, one can write

$$\psi(t,\mathbf{q},\mathbf{p}) = \psi(t,\mathbf{q},\mathbf{p}(t,\mathbf{q},\mathbf{P})) = -P_k Q_k + \psi_2(t,\mathbf{q},\mathbf{P}), \tag{4.2.15}$$

where $Q_k = Q_k(t,\mathbf{q},\mathbf{p(t,q,P)}) = Q_k(t,\mathbf{q},\mathbf{P})$. Plugging (4.2.15) into (4.2.9) gives,

$$-\sum_{k=1}^{n} \dot{P}_k Q_k - \sum_{k=1}^{n} P_k \dot{Q}_k + \sum_{k=1}^{n} \left[\frac{\partial \psi_2}{\partial q_k}\dot{q}_k + \frac{\partial \psi_2}{\partial P_k}\dot{P}_k\right] + \frac{\partial \psi_2}{\partial t} = \sum_{k=1}^{n} p_k \dot{q}_k - \sum_{k=1}^{n} P_k \dot{Q}_k + K(t,\mathbf{Q},\mathbf{P}) - H(t,\mathbf{q},\mathbf{p}).$$

Last equation implies the following,

$$p_k = \frac{\partial \psi_2}{\partial q_k} \tag{4.2.16}$$

$$Q_k = \frac{\partial \psi_2}{\partial P_k} \tag{4.2.17}$$

$$H(t,\mathbf{q},\mathbf{p}) + \frac{\partial \psi_2}{\partial t} = K(t,\mathbf{Q},\mathbf{P}). \tag{4.2.18}$$

The associated canonical transformation to the generating function $\psi_2(t,\mathbf{q},\mathbf{P})$ is $(Q_k, P_k) = (\frac{\partial \psi_2}{\partial P_k}(t,\mathbf{q},\mathbf{P}), P_k(t,\mathbf{q},\mathbf{p}))$.

Similarly, we can choose

$$\psi = q_k p_k + \psi_3(t,\mathbf{Q},\mathbf{p}),$$

which leads to

$$P_k = -\frac{\partial \psi_3}{\partial Q_k}$$

$$q_k = -\frac{\partial \psi_3}{\partial p_k}$$

$$H(t,\mathbf{q},\mathbf{p}) + \frac{\partial \psi_3}{\partial t} = K(t,\mathbf{Q},\mathbf{P}).$$

Finally, we can write

$$\psi_4 = q_k p_k - Q_k P_k + \psi_4(t,\mathbf{p},\mathbf{P}).$$

We end up with the following equations

$$q_k = \frac{\partial \psi_4}{\partial p_k}$$

$$Q_k = -\frac{\partial \psi_4}{\partial P_k}$$

$$H(t,\mathbf{q},\mathbf{p}) + \frac{\partial \psi_4}{\partial t} = K(t,\mathbf{Q},\mathbf{P}).$$

Chapter 5

Hamilton-Jacobi equation and Action-Angle Variables

5.1 Deriving Hamilton-Jacobi equation From Hamilton's Principle

In Chapter (1), we defined the action (1) as

$$I(q) = \int_{t_0}^{t_1} L(t,q,\dot{q})\,dt,$$

such that $q(t_0) = q_0$, $q(t_1) = q_1$. In fact, the minimizing function q of (1) is the actual motion or path. Let $q = q(t)$ be the actual motion. In other words, functional I has its minimum value at $q(t)$. Let $\hat{q}(t) = q(t) + \varepsilon\eta(t)$, where $\eta(t)$ is an arbitrary function of t. The change in I as we replace q by $\hat{q}$ is

$$I(\hat{q}) - I(q) = \int_{t_0}^{t_1} L(t,\hat{q},\dot{\hat{q}})\,dt - \int_{t_0}^{t_1} L(t,q,\dot{q})\,dt.$$

Using Taylor's theorem to expand $L(t,\hat{q},\dot{\hat{q}})$, we get,

$$L(t,\hat{q},\dot{\hat{q}}) = L(t,q,\dot{q}) + (\eta\frac{\partial L}{\partial q} + \dot{\eta}\frac{\partial L}{\partial \dot{q}})\varepsilon + o(\varepsilon^2).$$

We can write the difference as,

$$\begin{aligned} I(\hat{q}) - I(q) &= \int_{t_0}^{t_1} (\eta\frac{\partial L}{\partial q} + \dot{\eta}\frac{\partial L}{\partial \dot{q}})\varepsilon\,dt + 0(\varepsilon^2) \\ &= \varepsilon\delta I(q,\eta) + o(\varepsilon^2), \end{aligned}$$

where the term $\delta I(q,\eta)$ is called the **first variation** of I. By the assumption that q is the actual motion, we get

$$\delta I(q,\eta) = \int_{t_0}^{t_1} (\eta\frac{\partial L}{\partial q} + \dot{\eta}\frac{\partial L}{\partial \dot{q}}), dt = 0.$$

Integrate the second term by parts leads to

$$\frac{\partial L}{\partial \dot{q}}\eta\Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\eta\, dt + \int_{t_0}^{t_1} \frac{\partial L}{\partial q}\delta q\, dt = 0$$
$$\frac{\partial L}{\partial \dot{q}}\eta\Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} [\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}]\,\eta\, dt = 0.$$

Recall that if we assume that $\hat{q}(t) = q(t) + \varepsilon\eta(t)$ be such that $\eta(t_0) = \eta(t_1) = 0$, we obtain Euler-Lagrange equation.

Instead of considering paths starting and ending at common points as we have done in Chapter (1), consider paths with common starting point $q(t_0) = q_0$ but different ending points at fixed time t_1,i.e., $q(t_1) \neq \hat{q}(t_1)$ which means $\eta(t_1) \neq 0$. From the first variation formula above, we have

$$\delta I = \frac{\partial L}{\partial \dot{q}}\eta(t_1) + \int_{t_0}^{t_1} [\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}]\,\eta\, dt.$$

Since the actual motion satisfies E-L equation, then

$$\delta I = \frac{\partial L}{\partial \dot{q}}\eta = p\eta. \tag{5.1.1}$$

Since we started with an arbitrary function η, we can take it to be a function which is small everywhere in the interval (t_0, t_1) and hence it is considered to be a variation of q. We write $\eta = \delta q$ and from (5.1.1) we get

$$\frac{\partial I}{\partial q} = p. \tag{5.1.2}$$

Similarly, we can consider paths starting from t_0 such that $q(t_0) = \hat{q}(t_0)$ ($\eta(t_0) = 0$) and ending at fixed q_1 with various $t_1 = t$. Then, from (1) we know $\frac{dI}{dt} = L$. However, the action, in this way of considering it, is a function of q and t. Hence,

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \frac{\partial I}{\partial q}\dot{q} = \frac{\partial I}{\partial t} + p\dot{q}.$$

Compare,

$$\frac{\partial I(q,t)}{\partial t} + p\dot{q} = L$$
$$\frac{\partial I(q,t)}{\partial t} + H(q, \frac{\partial I(q,t)}{\partial q}) = 0,$$

using Legendre transformation and (5.1.2) in the last step.

5.2 Hamilton-Jacobi Equation

From the previous chapter, we showed how one can have a canonical transformation by constructing a generating function that gives a map from our old phase space to the new one. However, it can be naturally asked that how we can derive such a generating function. In this section we will discuss this problem and try to find a way by which we will be able at least to reduce the difficulties in solving the Hamiltonian. Of course, this is the primary objective of transferring our system from a phase space to another. In Section 4, we have proven that generating functions of first type $\psi_1(t, \mathbf{q}, \mathbf{Q})$ give (4.2.14). Now we want to get the simplest possible transformed Hamiltonian $K(t, \mathbf{Q}, \mathbf{P})$ which is $K = 0$. Therefore, we will try to solve for ψ_1 in the following equation

$$H(t, \mathbf{q}, \mathbf{p}) + \frac{\partial \psi_1(t, q, Q)}{\partial t} = 0, \tag{5.2.1}$$

so that the equations of motion of the transformed Hamiltonian K are

$$\dot{Q} = \frac{\partial K}{\partial P_i} = 0 \tag{5.2.2}$$

$$\dot{P} = -\frac{\partial K}{\partial Q_i} = 0, \tag{5.2.3}$$

which means, $Q_i = \alpha_i$, $P_i = \beta_i$ Where α_i and β_i are constants for all $i = 1, 2...n$. Using (4.2.12) we can write (5.2.1) as

$$H(t, q_1, q_2, ..., q_n, \frac{\partial \psi_1(t, \mathbf{q})}{\partial q_1}, \frac{\partial \psi_1(t, \mathbf{q})}{\partial q_2}, ..., \frac{\partial \psi_1(t, \mathbf{q})}{\partial q_n}) + \frac{\partial \psi_1(t, \mathbf{q})}{\partial t} = 0. \tag{5.2.4}$$

The first order partial differential equation (5.2.4) with $n+1$ variables is called the **Hamilton-Jacobi Equation**. Now, instead of solving a system of $2n$ ordinary differential equations, we try to find a complete solution of the partial differential equation (5.2.4).

Definition 5.1. A solution $\psi = \psi(t, \mathbf{q}, \alpha)$, where $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$ are constants, is called a **complete solution** or **complete integral** if ψ has continuous second derivatives with respect to the coordinates q_i, α_i, and t, and the matrix $\left[\frac{\partial^2\psi}{\partial\mathbf{q}\partial\alpha_1}, \frac{\partial^2\psi}{\partial\mathbf{q}\partial\alpha_2}, ..., \frac{\partial^2\psi}{\partial\mathbf{q}\partial\alpha_n}\right]$ is nonsingular.

The idea of using Hamilton-Jacobi equation is that we find its complete solution which is a function $\psi = \psi(t, q_1, \ldots, q_n, \alpha_1, \ldots, \alpha_n)$. Then, we use ψ_1 as a generating function depending on the old coordinates and the new coordinates $(\alpha_1, \alpha_2, ..., \alpha_n)$, from section 4.2,

$$-\frac{\partial\psi}{\partial\alpha_i} = \beta_i,$$

where $\beta_1, \ldots, \beta_n)$ are the new momenta. Now we have n algebraic equations from which we can solve for every q_i to get $q_i = q_i(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)$. Then, using

$$\frac{\partial\psi}{\partial q_i} = p_i$$

we can write $p_i = p_i(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)$. Then, we have found the coordinates q_i and momenta p_i as functions of time and constants α_i, β_i.

Before we end this section, we should discuss a special but important case. Suppose we have a Hamiltonian that is independent explicitly of time. We can immediately notice that a complete solution of Hamilton-Jacobi equation has the form

$$\psi(t, \mathbf{q}, \alpha) = W(\mathbf{q}, \alpha) - f(\alpha)t,$$

where ψ is a generating function of type 1 defined as $\psi = \psi_2(t, \mathbf{q}, \alpha)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$. The Hamilton-Jacobi equation can be written as

$$H(q_1, \ldots, q_n, \frac{\partial W}{\partial q_1}, \ldots, \frac{\partial W}{\partial q_n}) = f(\alpha).$$

Now, we seek to find a function $W(q_1, \ldots, q_n, \alpha_1, \ldots, \alpha_n)$ such that, using transformation equations

$$\frac{\partial\psi}{\partial q_i} = \frac{\partial W}{\partial q_i} = p_i$$

$$-\frac{\partial W}{\partial \alpha_i} = \beta_i$$

H is transformed into a function $f(\alpha_1, \ldots, \alpha_n)$. Now, equations of motion for f are

$$-\frac{\partial f}{\partial \alpha_i} = \dot{\beta}_i$$

$$\frac{\partial f}{\partial \beta_i} = 0 = \dot{\alpha}_i.$$

Integrating,

$$\alpha_i = constant$$

$$\beta_i = -\frac{\partial f}{\partial \alpha_i} t + c_i = \omega_i t + c_i,$$

where $\omega_i = \omega_i(\alpha_1, \ldots, \alpha_n)$ and c_i are constants for all $i = 1, 2 \ldots, n$

5.3 Analysis of Kepler Problem

In this section, we will look in detail to the motion in a central field. More specifically, let us consider a central field with potential energy $U(r) = \frac{-\gamma}{r}$ whrer $\gamma > 0$. Next remark shows an important result regarding the motion of any system in a central field.

Remark 5.1. In any central field, angular momentum is conserved (constant of motion).

Proof. First, angular momentum A is defined as $A = \vec{r} \times \dot{\vec{r}}$. We want to show that $\frac{dA}{dt} = 0$.

$$\frac{dA}{dt} = \dot{\vec{r}} \times \dot{\vec{r}} + \vec{r} \times \ddot{\vec{r}}.$$

Clearly, first term vanishes and then we are left with the second term. Using Newton's second law

$$\ddot{\vec{r}} = -\nabla U(r),$$

where $U(r)$ is the potential energy and it depends only on the distance between the particles. Now we have

$$\frac{dA}{dt} = \vec{r} \times -\nabla U(r)$$

which vanishes according to Newton's third law. □

This, in fact, shows that the direction of the angular momentum is constant which implies that the motion lies on the perpendicular plane (r,ϕ) to A. Moreover, remark 5.1 can be generalized to every system which is not effected by external forces.

Remark 5.2. In a central field, the total energy of a system that is explicitly independent of time is conserved.

Proof. Starting with the kinetic energy of the system $T = \frac{1}{2}\dot{r}^2$,

$$\dot{T} = \dot{r}\cdot\ddot{r} = \dot{r}\cdot -\nabla U(r) = -\dot{U}(r)$$

which implies that $T+U = E = constant$. □

We now can write the Lagrangian as

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + \frac{\gamma}{r}.$$

Using consistency of the angular momentum gives

$$A = (r,0,0)\times(m\dot{r}, mr\dot{\phi}, 0) = (0,0,mr^2\dot{\phi}) \in \mathbb{R}^3. \tag{5.3.1}$$

We can also write the total energy as follows

$$E = \frac{1}{2}m(\dot{r}^2 + \frac{A^2}{m^2r^2}) - \frac{\gamma}{r}. \tag{5.3.2}$$

Therefore, we have reduced the problem to a one degree of freedom system by using the fact that the angular moment is constant of motion. Now, from equation (5.3.2) we can see that total energy E is not exactly the sum of the kinetic and potential energy, we have an extra term which produces a new potential called the effective potential energy. We denote it by $U_e(r)$.

$$U_e(r) = -\frac{\gamma}{r} + \frac{A^2}{2mr^2}.$$

Solving (5.3.2) for $\dot{r}$, we obtain

$$\dot{r} = \frac{dr}{dt} = \sqrt{\frac{2}{m}(E + \frac{\gamma}{r}) - \frac{A^2}{m^2 r^2}}, \tag{5.3.3}$$

which can be expressed as

$$t = \int_{r_0}^{r} \frac{ds}{\sqrt{\frac{2}{m}(E + \frac{\gamma}{s}) - \frac{A^2}{m^2 s^2}}}. \tag{5.3.4}$$

The constant angular momentum A gives a relation between $d\phi$ and dt that can be used to replace dt in the last equation. We, as a result, obtain $dt = \frac{mr^2}{A} d\phi$, then

$$d\phi = \frac{A}{mr^2 \sqrt{\frac{2}{m}(E + \frac{\gamma}{r}) - \frac{A^2}{m^2 r^2}}} dr$$

$$\phi(r) = \int_{r_0}^{r} \frac{A}{ms^2 \sqrt{\frac{2}{m}(E + \frac{\gamma}{s}) - \frac{A^2}{m^2 s^2}}} ds. \tag{5.3.5}$$

Equations (5.3.4) and (5.3.5) give relations between r with t and r with ϕ, respectively. These two equations provide a complete solution of the problem. The first one gives r as an implicit function of time and the second one gives the equation of the path.

As shown in figure (5.1), having a negative total energy gives a bounded or finite radial motion. Therefore, we can rewrite (5.3.5) as

$$\Delta\phi(r) = 2(\phi(r_+) - \phi(r_-)) = 2 \int_{r_-}^{r_+} \frac{A}{mr^2 \sqrt{\frac{2}{m}(E + \frac{\gamma}{r}) - \frac{A^2}{m^2 r^2}}} dr, \tag{5.3.6}$$

where r_+ and r_- are the maximum and minimum values of r, respectively. In other words, the path of the particle is bounded by the two circles $r = r_-$ and $r = r_+$. By evaluating integral (5.3.6) and checking that $\Delta\phi = 2\pi$, using complex integration (appendix .1), we conclude that after one complete period, the radius makes one complete revolution and returns back to the starting point. In fact, we can generalize that for any potential energy, that is, when $\Delta\phi = 2\pi \frac{m}{n}$, for some integers m and n, then the path is closed.

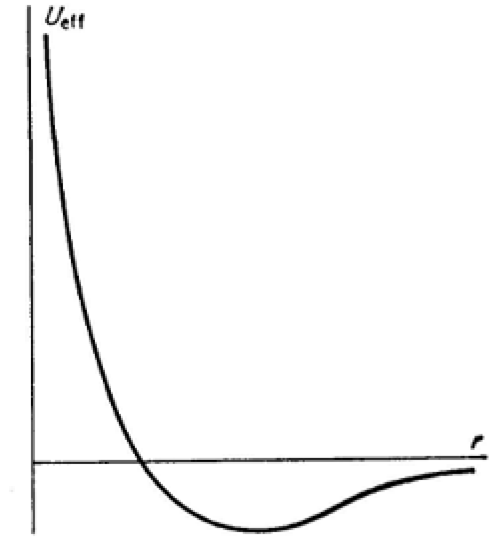

Figure 5.1: Effective potential [**9**]

Before we end this section we will find an explicit formula for r_- and r_+ in terms of some geometric parameters. This gives some nice relations between the physical quantities and the geometry of the orbit. We first start with the ordinary differential equation (5.3.3),

$$\dot{r} = \frac{dr}{dt} = \sqrt{\frac{2}{m}(E + \frac{\gamma}{r}) - \frac{A^2}{m^2 r^2}}.$$

Since r_- and r_+ are the minimum and the maximum of r, respectively, we can conclude that the right hand side of the above equation is zero at $r = r_-$ and $r = r_+$. Notice that this does not mean that there is no motion. It just means that r attains its minimum and maximum at these points. We can say that the minimum and the maximum solve the following equation

$$\frac{2}{m}(E + \frac{\gamma}{r}) - \frac{A^2}{m^2 r^2} = 0. \tag{5.3.7}$$

From which this follows

$$r^2 + \frac{\gamma}{E} r - \frac{A^2}{2mE} = 0. \tag{5.3.8}$$

Quadratic formula leads to the following form

$$r = -\frac{\gamma}{2E}\left(1 \pm \sqrt{1 + \frac{2A^2E}{m\gamma^2}}\right). \tag{5.3.9}$$

According to Landau and Lifshitz [**9**], we have

$$e = \sqrt{1 + \frac{2A^2E}{m\gamma^2}}, \tag{5.3.10}$$

$$F = \frac{A^2}{m\gamma}, \tag{5.3.11}$$

where e and F are the eccentricity and the semi-latus rectum, respectively. Now, using (5.3.10) to rewrite (5.3.9)

$$r = -\frac{\gamma}{2E}(1 \pm e). \tag{5.3.12}$$

We can immediately see that having the eccentricity $0 < e < 1$ implies

$$r_- = -\frac{\gamma}{2E}(1 - e),$$

$$r_+ = -\frac{\gamma}{2E}(1 + e).$$

Notice that the above two equations for the minimum and the maximum have some physical quantities and a geometric parameter. In order to turn them into pure geometric expressions, we should notice that

$$r_- + r_+ = 2a,$$

where a is semi-major axis of the ellipse. Therefore, we obtain the following

$$r_- = a(1 - e)$$

$$r_+ = a(1 + e).$$

Now, with the help of the well-known solution of the integral (5.3.5)

$$\phi(r) = \cos^{-1}\frac{\frac{A^2-m\gamma r}{rA}}{\sqrt{2mE+\frac{m^2\gamma^2}{A^2}}},$$

we can use (5.3.10) and (5.3.11) to rewrite the above equation in the following form

$$r(\phi) = \frac{F}{1+e\cos\phi}. \tag{5.3.13}$$

5.4 Action-Angle Variables

In physics, systems with periodic motion and autonomous Hamiltonian are of special importance. First, we would like to explain what we mean by periodic motion. We have two types of periodic motions that depend on whether the velocity changes sign over the complete period of the motion or not. **Libration motion** occurs when the velocity $\dot{q}$ changes sign and **rotation motion** is when velocity always keeps the same sign. Note that, in phase space, when motion is described by circle, ellipse, or any closed curve, then the motion is a libration and if the motion is described as a periodic function of coordinates $\mathbf{q}$ then the motion is rotation.

In Section (5.2), we discussed the case where $H = H(q, p)$ and we showed how one can construct a generating function that produces a canonical transformation such that the transformed Hamiltonian depends only on the new momenta. Now, for the same kind of Hamiltonian $H = H(q, p)$ we introduce a new variable J defined as

$$J_i = \frac{1}{2\pi}\oint p_i\, dq_i \tag{5.4.1}$$

Where the integration is taken over a complete period of each q_i.

The idea of action-angle variables, for one degree of freedom, is that for a closed bounded motion one can see that the phase space trajectories are closed curves. This means that the motion is periodic. Hamilton-Jacobi equation, as we have already seen, can be written as

$$H(q, \frac{\partial W(q,I)}{\partial q}) = \beta = \hat{H}(I). \tag{5.4.2}$$

Here, we take the constant of motion to be the transformed Hamiltonian. Where W is a generating function of type 2, $W = W(q,I)$, compare with (4.2.15).

$$\frac{\partial W(q,I)}{\partial q} = p,$$

$$\frac{\partial W(q,I)}{\partial I} = \Phi.$$

Now, for a fixed β, we look for (q_i, p_i) such that (5.4.2) holds. From the previous second transformation equation,

$$\frac{d\Phi}{dq} = \frac{\partial}{\partial I}\frac{\partial W}{\partial q} = \frac{\partial}{\partial I}p.$$

We require that Φ changes form 0 to 2π in one cycle around the curve, say ξ, that satisfies $H(q_i, p_i) = \beta$. Consequently,

$$d\Phi = \frac{\partial}{\partial I}p\,dq.$$

Which implies,

$$2\pi = \oint_{\xi} d\Phi = \frac{\partial}{\partial I}\oint_{\xi} p\,dq.$$

Hence, the previous equation is satisfied if

$$I = \frac{1}{2\pi}\oint_{\xi} p\,dq$$

and this is the definition of the action variable.

In fact, we can not always construct such a canonical variable. There are some conditions that a Hamiltonian system has to satisfy. Therefore, we will introduce a special kind of dynamical systems called **completely integrable systems**. A system in n dimensions is called completely integrable if one could find n first integrals (constants of motion) which are functionally independent and in involution with each other.

Functions $F_1, F_2, \ldots, F_n$ are said to be functionally independent if the gradient vectors of $F_i' s$ are linearly independent; i,e., $\sum_{i=0}^{n} c_i(\nabla F_i) \neq 0$ unless $c_i = 0$ for all i. Two functions are in involution with each other if and only if their **Poisson bracket** vanishes. Formally, let F_i be n first integrals such that $F_i = F_i(q(t), p(t))$ for all $1 \leq i \leq n$. Their Poisson bracket is,

$$\{F_i, F_j\} = \frac{\partial F_i}{\partial q}\frac{\partial F_j}{\partial p} - \frac{\partial F_i}{\partial p}\frac{\partial F_j}{\partial q} \tag{5.4.3}$$

Where $1 \leq i, j \leq n$. Clearly, when $i = j$, then Poisson bracket always vanishes.

Observe that a Poisson bracket can be used to test a physical quantity whether it is a first integral or not by computing Poisson bracket of this quantity with the Hamiltonian. Formally, a quantity F is conserved if $\{F, H\} = 0$. In fact, this is obvious when we use the definition (5.4.3).

$$\{F, H\} = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q}.$$

Using Hamilton's equations (3.2.4) and (3.2.5), we end up with

$$\{F, H\} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p} = \frac{dF}{dt}.$$

Therefore, vanishing Poisson bracket of a dynamical variable with the Hamiltonian and constancy are two equivalent concepts.

According to Liouville-Arnold theorem, if a system in a 2n-dimensional phase space has n first integrals which are independent of each other and in involution then there exists locally a canonical transformation from $(\mathbf{q}, \mathbf{p})$ phase space to $(\Phi, \mathbf{J})$ phase space such that the transformed Hamiltonian depends only on half of the variables, say, $\mathbf{J}$. As a result, the equations of motion of the new Hamiltonian K leads to the following:

$$\frac{\partial K(\mathbf{J})}{\partial \theta_i} = 0 = -\dot{J}_i$$

which implies that J_i are constants of motion. Also,

$$\frac{\partial K(\mathbf{J})}{\partial J_i} = constant = \dot{\Phi}_i$$

which implies that Φ_i are linearly dependent functions of t. This new set of variables are usually denoted by $(\Phi, \mathbf{J})$ and called **Action-Angle Variables**.

It turns out that in the action-angle variables' space, the bounded trajectory of a system with n degrees of freedom lies on n-dimensional tori $T^n := S^1 \times S^1 \times \cdots \times S^1$ such that each I_i identifies a specific torus and Φ_i identify where at this torus the trajectory is. In other words, think of the rectangular plane of θ, say for simplicity $\Phi = (\Phi_1, \Phi_2)$ such that $0 \leq \Phi_1, \Phi_2 \leq K$. This plane, topologically, is equivalent to a 2-torus, $S^1 \times S^1 = T^2$ and we can determine every point on the torus using (Φ_1, Φ_2), where Φ_1 is the angle in the horizontal direction and Φ_2 is the angle in the vertical direction.

5.5 Action-Angle variables For The Two-Body Problem (The Kepler Problem)

Starting with the Lagrangian of a two-body system with two masses m and M in Cartesian coordinates (x, y, z),

$$L = T - U = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{\gamma}{\sqrt{x^2 + y^2 + z^2}}$$

Where U is the potential energy with $\gamma = GmM = constant$.

Now, applying the transformation function from Cartesian coordinates to spherical coordinates defined by $(x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta)$ We can write the new Lagrangian in terms of the new coordinates (r, θ, ϕ) as the following:

$$L(r, \theta, \phi, \dot{r}, \dot{\theta}, \dot{\phi}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2) + \frac{\gamma}{r} \tag{5.5.1}$$

We have,

$$\frac{\partial L}{\partial \dot{r}} = p_r = m\dot{r} \tag{5.5.2}$$

$$\frac{\partial L}{\partial \dot{\theta}} = p_\theta = mr^2\dot{\theta} \tag{5.5.3}$$

$$\frac{\partial L}{\partial \dot{\phi}} = p_\phi = mr^2 \sin^2\theta\dot{\phi} \tag{5.5.4}$$

Since there is no explicit dependence on ϕ in L, we claim that p_ϕ is constant of motion. This can be easily seen from Euler-Lagrange equation

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{\phi}}) = \frac{\partial L}{\partial \phi} = 0 \tag{5.5.5}$$

The above equation shows that $\frac{\partial L}{\partial \dot{\phi}} = p_\phi = \alpha_\phi = constant$.

Now that we found the formula of the three components of the momentum (p_r, p_θ, p_ϕ), we can easily write the Hamiltonian of the system.

$$H(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\sin^2\theta}p_\phi^2) - \frac{\gamma}{r} \tag{5.5.6}$$

It is clear that the Hamiltonian does not depend explicitly on time which means that the total energy of this system is conserved. In other words, $H = -|E| = constant$. By applying Hamilton-Jacobi equation to our Hamilton function we obtain the following equation for some function $S = S(r,\theta,\phi,t)$ such that $\frac{\partial S}{\partial r} = p_r$, $\frac{\partial S}{\partial \theta} = p_\theta$, $\frac{\partial S}{\partial \phi} = p_\phi$

$$\frac{1}{2m}[(\frac{\partial S}{\partial r})^2 + \frac{1}{r^2}(\frac{\partial S}{\partial \theta})^2 + \frac{1}{r^2\sin^2\theta}(\frac{\partial S}{\partial \phi})^2] - \frac{\gamma}{r} + \frac{\partial S}{\partial t} = 0. \tag{5.5.7}$$

Observe that because of the independence of the time in the total energy, which is the Hamiltonian, we can write S as a function of two separate terms,

$$S = W(r,\theta,\phi,E) + |E|t. \tag{5.5.8}$$

Therefore, from (5.5.7) and (5.5.8) we get

$$H = \frac{1}{2m}[(\frac{\partial W}{\partial r})^2 + \frac{1}{r^2}(\frac{\partial W}{\partial \theta})^2 + \frac{1}{r^2\sin^2\theta}(\frac{\partial W}{\partial \phi})^2] - \frac{\gamma}{r} = -|E|. \tag{5.5.9}$$

Now, we try to find $W(r,\theta,\phi,E) = W_r(r,E) + W_\theta(\theta,E) + W_\phi(\phi,E)$ such that W satisfies (5.5.9).

$$\frac{1}{2m}[(\frac{dW_r}{dr})^2 + \frac{1}{r^2}(\frac{dW_\theta}{d\theta})^2 + \frac{1}{r^2\sin^2\theta}(\frac{dW_\phi}{d\phi})^2] - (\frac{\gamma}{r} - |E|) = 0 \tag{5.5.10}$$

$$\frac{-r^2}{2m}(\frac{dW_r}{dr})^2)+(\gamma r-r^2|E|)=\frac{1}{2m}(\frac{dW_\theta}{d\theta})^2+\frac{1}{2msin^2\theta}(\frac{dW_\phi}{d\phi})^2$$

Since the left hand side equals the right hand side with independent variables, then there exists a constant C so that $L.H.S=R.H.S=\frac{C^2}{2m}$

$$\frac{dW_r}{dr}=\sqrt{\frac{-2m}{r^2}(-\gamma r+r^2|E|)-\frac{C^2}{r^2}}=\sqrt{2m(\frac{\gamma}{r}-|E|)-\frac{C^2}{r^2}}, \tag{5.5.11}$$

$$\frac{dW_\theta}{d\theta}=\sqrt{C^2-\frac{p_\phi^2}{\sin^2\theta}}. \tag{5.5.12}$$

Remember, we showed in (5.5.5) that $p_\phi=\alpha_\phi$ is constant. Here we make an observation.

Remark 5.3. The constant C is the magnitude value of the angular momentum.

Proof. Let the angular momentum be ξ

$$\begin{aligned}\xi &= (r,0,0)\times(m\dot{r},mr\dot{\theta},mr\sin\theta\dot{\phi})\\ &= (0,-mr^2\sin\theta\dot{\phi},mr^2\dot{\theta})\\ &= (0,\frac{-p_\phi}{\sin\theta},p_\theta),\end{aligned}$$

and the square of the magnitude value $|\xi|^2=\frac{p_\phi^2}{\sin^2\theta}+p_\theta^2$. Therefore, from (5.5.12) we have $C^2=p_\theta^2+\frac{p_\phi^2}{\sin^2\theta}=|\xi|^2$ which means C is the magnitude of the angular momentum. We denote it A. □

From the definition of action variable (5.4.1)

$$J_r=\frac{1}{2\pi}\oint_r p_r\,dr=\frac{1}{2\pi}\oint_r\sqrt{2m(\frac{\gamma}{r}-|E|)-\frac{A^2}{r^2}}\,dr, \tag{5.5.13}$$

$$J_\theta=\frac{1}{2\pi}\oint_\theta p_\theta\,d\theta=\frac{1}{2\pi}\oint_\theta\sqrt{A^2-\frac{\alpha_\phi^2}{\sin^2\theta}}\,d\theta, \tag{5.5.14}$$

$$J_\phi=\frac{1}{2\pi}\oint_\phi\alpha_\phi\,d\phi. \tag{5.5.15}$$

Using the method of complex integration to compute the Integrals (5.5.13) and (5.5.14), we obtain, (appendix .2),

$$J_r = -A + \frac{m\gamma}{\sqrt{2m|E|}} \tag{5.5.16}$$

$$J_\theta = A - \alpha_\phi. \tag{5.5.17}$$

Now, we write $|E|$ as a function of the action variables J_r and J_θ,

$$J_r + J_\theta = \frac{m\gamma}{\sqrt{2m|E|}} - \alpha_\phi.$$

We know that $|E| = -H$. Therefore,

$$H = -\frac{m\gamma^2}{2(J_r + J_\theta + \alpha_\phi)^2}.$$

Now, we have the Hamiltonian depends only on the action variables. The equations of motion can be used at this moment to obtain the frequencies.

$$\frac{\partial H}{\partial J_r} = \frac{m\gamma^2}{(J_r + J_\theta + \alpha_\phi)^3} = \omega_r. \tag{5.5.18}$$

Note that $\omega_r = \omega_\theta$ which means that the system is **degenerate**.

Definition 5.2. A system is said to be non-degenerate if the frequencies of it are functionally independent, that is if $\det \mathbf{N} \neq 0$, for

$$\mathbf{N} = \begin{pmatrix} \frac{\partial\omega_1}{\partial J_1} & \frac{\partial\omega_2}{\partial J_1} \cdots & \frac{\partial\omega_n}{\partial J_1} \\ \\ \frac{\partial\omega_1}{\partial J_2} & \frac{\partial\omega_2}{\partial J_2} \cdots & \frac{\partial\omega_n}{\partial J_2} \\ \vdots & & \\ \frac{\partial\omega_1}{\partial J_n} & \frac{\partial\omega_2}{\partial J_n} \cdots & \frac{\partial\omega_n}{\partial J_n} \end{pmatrix}.$$

At the end of the next chapter we will investigate in detail the meaning and the importance of having a nondegenerate system. In the previous problem, we have seen that the frequencies are equal so the system is actually strongly degenerate.

Chapter 6

The Two-Body Problem Given By The Kepler Potential Perturbed By $\Delta U(r) = \frac{\delta}{r^3}$

In this section, we will study the system in the case where the potential energy is the sum of Kepler's potential and a perturbation term $\frac{\delta}{r^3}$, for δ small enough. We, also, will study the motion of the particles with this potential and compare it with the previous case when we had $U(r) = \frac{-\gamma}{r}$. Moreover, we will find the formula of the action variables.

6.1 Analysis of The Motion of The system with Potential Energy $U(r) = -\frac{\gamma}{r} \pm \frac{\delta}{r^3}$

Starting with $U(r) = -\frac{\gamma}{r} \pm \frac{\delta}{r^3}$, leads to the following Lagrangian:

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + \frac{\gamma}{r} \pm \frac{\delta}{r^3}.$$

From remark 5.1, we can conclude that the angular momentum for the system is conserved and it can be computed the same way

$$A = mr^2\dot{\phi}. \tag{6.1.1}$$

Therefore, the total energy can be written in terms of r and $\dot{r}$ as

$$E = \frac{1}{2}m(\dot{r}^2 + \frac{A^2}{m^2r^2}) - \frac{\gamma}{r} \pm \frac{\delta}{r^3}.$$

As we did in section 5.3, it is straightforward to obtain the following

$$\dot{r} = \frac{dr}{dt} = \sqrt{\frac{2}{m}(E_1 + \frac{\gamma}{r} + \frac{\delta}{r^3}) - \frac{A_1^2}{m^2r^2}}. \tag{6.1.2}$$

Again, with the same procedure we followed in section 5.3, we end up with the following expression

$$\Delta\phi(r) = 2(\phi(r_+) - \phi(r_-)) = 2\int_{r_-}^{r_+} \frac{A_1}{mr^2\sqrt{\frac{2}{m}(E + \frac{\gamma}{r} + \frac{\varepsilon}{r^3}) - \frac{A_1^2}{m^2r^2}}} dr. \quad (6.1.3)$$

With the help of some approximation formulas, we can find an approximated solution for this integral, (appendix .3). Unfortunately, this is not sufficient to conclude that the path is open. However, according to Landau and Lifshitz [**9**], there are only two cases where the bounded motion in a central field is closed. These two cases are when the potential energy is in the form of $\frac{1}{r}$ or r^2. based on that, we say that the motion with our new potential is open (figure 6.1).

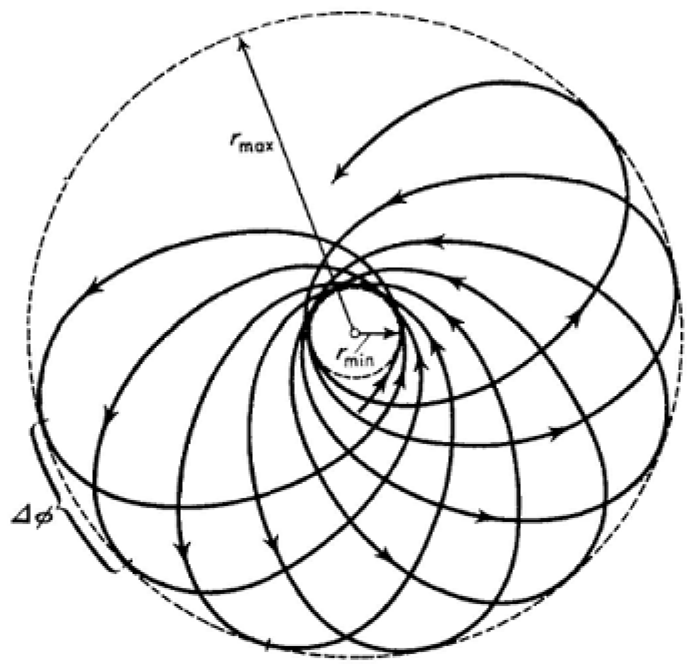

Figure 6.1: Open path of a particle[**9**]

6.2 Action-Angle Variables For the Perturbed Hamiltonian

The Hamiltonian with our new potential can be written in the following form,

$$H(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}(p_r^2 + \frac{1}{r^2}p_\theta^2 + \frac{1}{r^2\sin^2\theta}p_\phi^2) - \frac{\gamma}{r} - \frac{\delta}{r^3}. \quad (6.2.1)$$

Notice that the system's total energy is still conserved. For this reason, we write $H = -|E|$. The reason we take $H < 0$ is that it gives a bounded state for the system. It can be clearly seen from figures 5.1 and 6.1 in section 5.3.

At this point, we see that our perturbation term, $U_1(r)$, has nothing to do with p_θ and p_ϕ and that is there is no impact on the action variables J_θ and J_ϕ. Since the perturbation term depends only on the distance r, we then expect that we will have some changes in J_r. By applying Hamilton-Jacobi equation to our system, analogous to (5.5.7) and (5.5.9), we end up with

$$\frac{1}{2m}[(\frac{dW_r}{dr})^2 + \frac{1}{r^2}(\frac{dW_\theta}{d\theta})^2 + \frac{1}{r^2\sin^2\theta}(\frac{dW_\phi}{d\phi})^2] - (\frac{\gamma}{r} + \frac{\delta}{r^3} - |E|) = 0. \tag{6.2.2}$$

Separation of variables leads to the following two ordinary differential equations:

$$\frac{dW_r}{dr} = \sqrt{\frac{-2m}{r^2}(-\gamma r - \frac{\delta}{r} + r^2|E|) - \frac{A^2}{r^2}} = \sqrt{2m(\frac{\gamma}{r} + \frac{\delta}{r^3} - |E|) - \frac{A^2}{r^2}} \tag{6.2.3}$$

$$\frac{dW_\theta}{d\theta} = \sqrt{A^2 - \frac{\alpha_\phi^2}{\sin^2\theta}}. \tag{6.2.4}$$

Compare (6.2.3) and (6.2.4) to (5.5.11) and (5.5.12), respectively. Now we formulate the actions of the system using the definition of action variable (5.4.1)

$$J_r = \frac{1}{2\pi}\oint_r \sqrt{2m(\frac{\gamma}{r} + \frac{\delta}{r^3} - |E|) - \frac{A^2}{r^2}}\, dr. \tag{6.2.5}$$

$$J_\theta = \frac{1}{2\pi}\oint_\theta \sqrt{A^2 - \frac{\alpha_\phi^2}{\sin^2\theta}}\, d\theta. \tag{6.2.6}$$

As expected, the perturbation does not effect the action about θ. However, we need to treat the action J_r to see what impact the perturbation has on the system. An approximated value of J_r is computed in appendix D [1], for a sufficiently small δ, can be used here to express the transformed Hamiltonian $H = H(J_r, J_\theta)$. We

[1] This integral can be computed using contour integration over a Riemann surface but it was not needed here.

have

$$J_r \approx -A_1 + \frac{m\gamma}{\sqrt{2m|E|}} + \frac{m\delta}{A_1 F} \tag{6.2.7}$$

which leads to the following:

$$H(J_r, J_\theta) \approx \frac{-m\gamma^2}{2(J_r + J_\theta + \alpha_\phi - \frac{m\delta}{A_1 F})^2}. \tag{6.2.8}$$

6.3 Non-degeneracy of The Frequency Map

In this section we look into the non-degeneracy of the frequency map of the system. Our goal is to find some r_- and r_+ such that our system is non-degenerate in the sense of 5.2 and then, by analyticity of $\det N$ as a function of r_- and r_+, we conclude that $\det N \neq 0$ on an open dense set. Before we start looking for the non-degeneracy, we first see what the frequency map is. As we have seen, the second Hamilton's equation for the Hamiltonian expressed in action variables is as follows,

$$\frac{\partial H(\mathbf{J})}{\partial J_i}\Big|_{\mathbf{J}=\mathbf{J_0}} = \omega_i(\mathbf{J_0}). \tag{6.3.1}$$

This equation in fact defines a map $\mathbf{J} \mapsto \omega$. Therefore, we study this map to determine the degeneracy of the system. In order to make this job hopeful, we study the case where $\bar{r} = r_-$, figure (6.2).

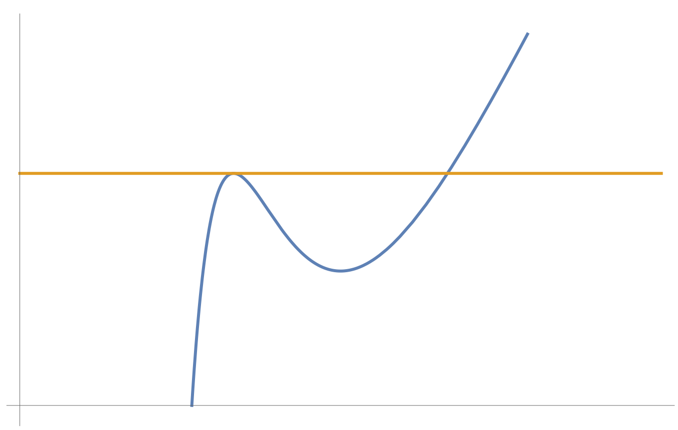

Figure 6.2: $\bar{r} = r_-$

Here, we have the action about r expressed as the following[2]

$$J_r = \frac{\sqrt{2m|E(r_-, r_+)|}}{2\pi} \oint_r (r - r_-) \sqrt{\frac{(r_+ - r)}{r^3}}\, dr \tag{6.3.2}$$

$$= \sqrt{2mE}\left(-\frac{(2r_- + r_+)\pi}{2} - 3\sqrt{(r_+ - r_-)r_-} - (2r_- + r_+) \arctan\sqrt{\frac{r_-}{r_+ - r_-}}\right). \tag{6.3.3}$$

Since r_- and r_+ are both real numbers, we write r_+ in the following form $r_+ = kr_-$ for $k > 1$. According to Vieta's theorem, total energy and total angular momentum can be written in terms of the zeros of p_r.

$$2r_- + r_+ = \frac{\gamma}{|E|}, \tag{6.3.4}$$

$$r_-^2 + 2r_- r_+ = \frac{A^2}{2m|E|}. \tag{6.3.5}$$

Solving for E and A gives the following:

$$|E| = \frac{\gamma}{(2+k)r_-}, \tag{6.3.6}$$

$$A^2 = \frac{2m\gamma(1+2k)r_-}{2+k}. \tag{6.3.7}$$

With the help of the last equation, we rewrite the action.

$$J_r = \sqrt{2m|E|}\left(-\frac{\pi}{2}(2+k)r_- - 3\sqrt{(k-1)}r_- - a(2+k)r_-\right)$$

$$= \xi A^2 \sqrt{2m|E|}, \tag{6.3.8}$$

where, for shortness, $\xi = \frac{2+k}{(2m\gamma)(1+2k)}\left(-\frac{\pi}{2}(2+k) - 3\sqrt{(k-1)} - a(2+k)\right)$ and $a = \arctan\sqrt{\frac{1}{k-1}}$. Solving (6.3.8) for $|E|$,

$$H = -|E| = -\frac{{J_r}^2}{2m\xi^2 A^4} = -\frac{{J_r}^2}{2m\xi^2 (J_\theta + \alpha_\phi)^4} \tag{6.3.9}$$

Now that we have the Hamiltonian expressed in terms of the actions, it is straightforward to test the non-

[2] The integral was computed with the help of *Mathematica*.

degeneracy of the frequencies. First, let's write the precise formula for the frequencies.

$$\omega_r = \frac{\partial H}{\partial J_r} = -\frac{J_r}{m\xi^2(J_\theta + \alpha_\phi)^4} \tag{6.3.10}$$

$$\omega_\theta = \frac{\partial H}{\partial J_\theta} = \frac{2{J_r}^2}{m\xi^2(J_\theta + \alpha_\phi)^5} \tag{6.3.11}$$

Recall (5.2) and apply the definition to these frequencies.

$$\det M = \frac{-6J_r}{(m\xi)^2(J_\theta + \alpha_\phi)^{10}}. \tag{6.3.12}$$

Alternatively, define $\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ such that $(A,E) \to (J_r(A,E), J_\theta(A,E))$. Now, we look at the derivative.

$$D_{\alpha^{-1}} = \begin{pmatrix} \frac{\partial A}{\partial J_r} & \frac{\partial A}{\partial J_\theta} \\ \\ \frac{\partial E}{\partial J_r} & \frac{\partial E}{\partial J_\theta} \end{pmatrix} \tag{6.3.13}$$

Using inverse function theorem, we can write

$$D_{\alpha^{-1}} = (D_\alpha)^{-1} = \begin{pmatrix} \frac{\partial J_r}{\partial A} & \frac{\partial J_r}{\partial E} \\ \\ \frac{\partial J_\theta}{\partial A} & \frac{\partial J_\theta}{\partial E} \end{pmatrix}^{-1} = \frac{1}{\frac{\partial J_r}{\partial A}\frac{\partial J_\theta}{\partial E} - \frac{\partial J_r}{\partial E}\frac{\partial J_\theta}{\partial A}} \begin{pmatrix} \frac{\partial J_\theta}{\partial E} & -\frac{\partial J_r}{\partial E} \\ \\ -\frac{\partial J_\theta}{\partial A} & \frac{\partial J_r}{\partial A} \end{pmatrix} \tag{6.3.14}$$

Equating the right hand side of the previous two equations leads to the following equations.

$$\omega_r = \frac{\partial E}{\partial J_r} = \frac{\frac{-\partial J_\theta}{\partial A}}{\frac{\partial J_r}{\partial A}\frac{\partial J_\theta}{\partial E} - \frac{\partial J_r}{\partial E}\frac{\partial J_\theta}{\partial A}} \tag{6.3.15}$$

$$\omega_\theta = \frac{\partial E}{\partial J_\theta} = \frac{\frac{\partial J_r}{\partial A}}{\frac{\partial J_r}{\partial A}\frac{\partial J_\theta}{\partial E} - \frac{\partial J_r}{\partial E}\frac{\partial J_\theta}{\partial A}} \tag{6.3.16}$$

Now, we express the actions as functions of r_- and r_+ [**13**]. The reason we want to write the actions in terms of the minimum and the maximum values of r is that it gives some useful relations for the second derivatives in the non-degeneracy condition (5.2) as we will see.

$$\begin{aligned} \omega_r &= \omega_r(J_r(r_-, J_+), J_\theta(r_-, r_+)) \\ \omega_\theta &= \omega_\theta(J_r(r_-, J_+), J_\theta(r_-, r_+)) \end{aligned} \tag{6.3.17}$$

Taking the derivatives of the frequencies with respect to r_- and r_+ and write the four resulting equations in the matrix form.

$$\underbrace{\begin{pmatrix} \frac{\partial \omega_r}{\partial r_-} & \frac{\partial \omega_r}{\partial r_+} \\ \\ \frac{\partial \omega_\theta}{\partial r_-} & \frac{\partial \omega_\theta}{\partial r_+} \end{pmatrix}}_{N} = \underbrace{\begin{pmatrix} \frac{\partial \omega_r}{\partial J_r} & \frac{\partial \omega_r}{\partial J_\theta} \\ \\ \frac{\partial \omega \theta}{\partial J_r} & \frac{\partial \omega_\theta}{\partial J_\theta} \end{pmatrix}}_{M} \begin{pmatrix} \frac{\partial J_r}{\partial r_-} & \frac{\partial J_r}{\partial r_+} \\ \\ \frac{\partial J \theta}{\partial r_-} & \frac{\partial J_\theta}{\partial r_+} \end{pmatrix} \tag{6.3.18}$$

$$M = N \frac{1}{\frac{\partial J_r}{\partial r_-}\frac{\partial J_\theta}{\partial r_+} - \frac{\partial J_r}{\partial r_+}\frac{\partial J_\theta}{\partial r_-}} \begin{pmatrix} \frac{\partial J_\theta}{\partial r_+} & -\frac{\partial J_r}{\partial r_+} \\ \\ -\frac{\partial J_\theta}{\partial r_-} & \frac{\partial J_r}{\partial r_-} \end{pmatrix}. \tag{6.3.19}$$

From (6.3.19), we conclude that

$$\det(M) = \frac{\det(N)}{\frac{\partial J_r}{\partial r_-}\frac{\partial J_\theta}{\partial r_+} - \frac{\partial J_r}{\partial r_+}\frac{\partial J_\theta}{\partial r_-}} \tag{6.3.20}$$

From (6.3.15) and (6.3.16), we have

$$\omega_r = \frac{-1}{D} \tag{6.3.21}$$

$$\omega_\theta = \frac{C}{D}. \tag{6.3.22}$$

Where

$$D = \frac{\partial J_r}{\partial A}\frac{\partial J_\theta}{\partial E} - \frac{\partial J_r}{\partial E}\frac{\partial J_\theta}{\partial A} = -\frac{\partial J_r}{\partial E} = \frac{m}{\pi\sqrt{-2mE}} \int_{r_-}^{r_+} \frac{\sqrt{r^3}}{\sqrt{(r-\bar{r})(r-r_-)(r_+-r)}}\, dr \tag{6.3.23}$$

$$C = \frac{\partial J_r}{\partial A} = -\frac{A}{\pi\sqrt{-2mE}} \int_{r_-}^{r_+} \frac{1}{\sqrt{r(r-\bar{r})(r-r_-)(r_+-r)}}\, dr \tag{6.3.24}$$

Having shown that $\det M$ is proportional to $\det N$, it is sufficient to study the determinant of the matrix

N.

$$\det(N) = \frac{1}{D^4}\begin{vmatrix} \partial_{r_-}D & \partial_{r_+}D \\ \partial_{r_-}(\partial_A J_r)D - (\partial_A J_r)\partial_{r_-}D & \partial_{r_+}(\partial_A J_r)D - (\partial_A J_r)\partial_{r_+}D \end{vmatrix} \tag{6.3.25}$$

$$= \frac{1}{D^3}\left(\partial_{r_+}(\partial_A J_r)\partial_{r_-}D - \partial_{r_-}(\partial_A J_r)\partial_{r_+}D\right) \tag{6.3.26}$$

Here, we want to show that the determinant does not vanish for some r_- and r_+.

First, we rewrite (6.3.25)

$$\det(N) = \frac{-1}{D^3}\left(\partial_A(\partial_{r_+}J_r)\partial_E(\partial_{r_-}J_r) - \partial_A(\partial_{r_-}J_r)\partial_E(\partial_{r_+}J_r)\right). \tag{6.3.27}$$

Recall,

$$J_r = \frac{\sqrt{2m|E|}}{2\pi}\oint_r (r-r_-)\sqrt{\frac{(r_+-r)}{r^3}}\,dr \tag{6.3.28}$$

$$\det M = \frac{-6J_r}{(m\xi)^2(J_\theta+\alpha_\phi)^{10}}. \tag{6.3.29}$$

It should be noted here that (6.3.21) and (6.3.22) lead to the same formulas as (7.0.1) and (7.0.2). Moreover, According to the work shown above, the determinant expressed in terms of the the actions is related to the determinant expressed in terms of the turning points r_- and r_+ by (6.3.20).

To conclude, for the case $\bar{r} = r_-$ we have $\det M = 0$ if and only if $\xi = 0$. Now we want to show that $\xi \neq 0$ for all r_- and r_+ such that $r_+ = kr_-$ for $k > 1$ and, as a result, we are guaranteed that the frequency map is non-degenerate.

We set $\xi = 0$ and see for which values of k this equation holds.

$$\xi = \frac{2+k}{(2m)(1+2k)}\left(-\frac{\pi}{2}(2+k) - 3\sqrt{(k-1)} - a(2+k)\right) = 0,$$

where $a = \arctan\sqrt{\frac{1}{k-1}}$. This implies

$$a = -\frac{\pi}{2} - 3\frac{\sqrt{(k-1)}}{2+k} < -\frac{\pi}{2}.$$

But $0 < a < \frac{\pi}{2}$ for $k > 1$. Hence, $\xi \neq 0$. Consequently, $J_r \neq 0$.

Chapter 7

KAM Theorem

Now that we have discussed in detail our system and were able to write it in terms of the action variables for a specific case, we are prepared to, at least, state and say something about the celebrated theorem, that was named after Kolomgrov, Arnold, and Moser, KAM theorem.

In our system, we have $\omega = (\omega_r, \omega_\theta)$. First we notice that the angle variables (Φ_r, Φ_θ) are linear in time. This can be immediately seen from the equations of motion of the system

$$\dot{\Phi}_r = \omega_r = \frac{\partial H}{\partial J_r} = -\frac{J_r}{m\xi^2(J_\theta + \alpha_\phi)^4} \tag{7.0.1}$$

$$\dot{\Phi}_\theta = \omega_\theta = \frac{\partial H}{\partial J_\theta} = \frac{2J_r{}^2}{m\xi^2(J_\theta + \alpha_\phi)^5}. \tag{7.0.2}$$

Also, we can see that the motion will be closed if

$$\frac{\omega_r}{\omega_\theta} = \frac{\kappa_1}{\kappa_2} \tag{7.0.3}$$

for κ_1 and κ_2 integers. In the case when the ratio of the frequencies is irrational, the motion is called **quasiperiodic** and (Φ_r, Φ_θ) plane will be densely filled with lines.

The non-resonant trajectories are those with rationally independent frequencies; i.e.,

$$\langle \mathbf{n}, \omega \rangle \neq 0 \tag{7.0.4}$$

$\forall \mathbf{n} \in \mathbb{Z}^2 \setminus \{0\}$ and $\langle .,. \rangle$ is the dot product.

Theorem 7.1. *If the unperturbed Hamiltonian system H_0 is non-degenerate, then under sufficiently small* *perturbation such that*

$$H(J,\Phi) = H_0(J) + \varepsilon H_1(J,\Phi) \tag{7.0.5}$$

for $\varepsilon << 1$, most of the non-resonant invariant tori do not disappear but only slightly deformed. In other words, the invariant tori are still there filled with phase curves winding around them conditionally periodically.

Now let's look at our new system (7.0.5). First, we write the equations of motion

$$\frac{\partial H}{\partial \Phi} = -\varepsilon \frac{\partial H_1}{\partial \Phi} \quad , \quad \frac{\partial H}{\partial J} = \omega_0 + \varepsilon \frac{\partial H}{\partial J}. \tag{7.0.6}$$

Clearly, $(\mathbf{J},\Phi)$ are not considered to be action-angle variables for the system, neither $\mathbf{J}$ are constants nor Φ linear in time. Therefore, with the same type of generating function we used to construct $(\mathbf{J},\Phi)$, we again use the same type, namely, $F = F(\mathbf{I},\Phi)$, where $\mathbf{I}$ are the new actions. The associated transformation equations with F are

$$\frac{\partial F(\mathbf{I},\Phi)}{\partial I} = \Psi \quad , \quad \frac{\partial F(\mathbf{I},\Phi)}{\partial \Phi} = J \tag{7.0.7}$$

The goal now is to find a generating function F such that the system H can be written as a function of the $I's$.

Expand F as a power series of ε:

$$F = F_0 + \varepsilon F_1 + \varepsilon^2 F_2 + \dots, \tag{7.0.8}$$

where $F_0 = \Phi\mathbf{I}$ the identity transformation. As we did before, we use Hamilton-Jacobi equation.

$$H_0\big(\frac{\partial F}{\partial \Phi}\big) + \varepsilon H_1\big(\frac{\partial F}{\partial \Phi}, \Phi\big) = K(\mathbf{I}). \tag{7.0.9}$$

Again, we expand K:

$$K(\mathbf{I}) = K_0(\mathbf{I}) + \varepsilon K_1(\mathbf{I}) + \varepsilon^2 K_2(\mathbf{I}) + \dots \tag{7.0.10}$$

Using (7.0.9) and (7.0.10) and equating powers of ε

$$o(1) : H_0(\mathbf{I}) = K_0(\mathbf{I}) \tag{7.0.11}$$

$$o(\varepsilon) : \omega_0(\mathbf{I})\frac{\partial F_1}{\partial \Phi} + H_1(\mathbf{I},\Phi) = K_1(\mathbf{I}) \tag{7.0.12}$$

As we have seen, F_1 is periodic in Φ. We require also that its derivative with respect to Φ is also periodic. As a result, averaging (7.0.12) over Φ gives:

$$\bar{K}_1(I) = K_1(I) = \bar{H}_1(I,\Phi) \tag{7.0.13}$$

Where $\bar{g} = \frac{1}{2\pi}\int_0^{2\pi} g\, d\Phi$.

Back to (7.0.12),

$$\frac{\partial}{\partial \Phi}F_1(I,\Phi) = -\frac{1}{\omega_0(I)}\tilde{H}_1(I,\Phi) \tag{7.0.14}$$

Where $\tilde{H}_1 = H_1 - K_1$. We can think of it as the periodic part of H_1; i.e.,

$$\bar{\tilde{H}}_1 = \bar{H}_1 - \bar{K}_1 = K_1 - K_1 = 0$$

which means $\tilde{H}_1$ is periodic in Φ. Hence, we have (7.0.14) periodic in Φ. Writing F_1 and $\tilde{H}_1$ as Fourier series in variables Φ,

$$F_1(I,\Phi) = \sum_{k\in\mathbb{Z}^2\setminus\{(0,0)\}}^{\infty} N_{1,k}(I)e^{i\langle k\cdot\Phi\rangle} \tag{7.0.15}$$

$$\tilde{H}_1(I,\Phi) = \sum_{k\in\mathbb{Z}^2\setminus\{(0,0)\}}^{\infty} M_{1,k}(I)e^{i\langle k\cdot\Phi\rangle} \tag{7.0.16}$$

using (7.0.15) and (7.0.16) to write the derivative of F_1 and equating with (7.0.14),

$$\frac{\partial F_1}{\partial \Phi} = \sum_k ikN_{1,k}(I)e^{i\langle k\cdot\Phi\rangle} = \frac{-1}{\omega_0(I)}\sum_k M_{1,k}(I)e^{i\langle k\cdot\Phi\rangle} \tag{7.0.17}$$

We obtain

$$N_k(I) = \frac{i}{\langle k\cdot\omega_0(I)\rangle}M_{1,k}(I). \tag{7.0.18}$$

Consequently, F_1 is

$$F_1(I,\Phi) = \sum_k \frac{iM_{1,k}(I)}{\langle k \cdot \omega_0(I) \rangle} e^{i\langle k\cdot\Phi\rangle}, \tag{7.0.19}$$

where $M_{1,k}(I)$ is supposed to be known; that is,

$$M_{1,k}(I) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{H}_1 e^{i\langle k\cdot\Phi\rangle}\, d\Phi. \tag{7.0.20}$$

From the definition of $\tilde{H}_1$ and the fact that K_1 is independent of the angle variables, we write $M_k(I)$ as

$$M_{1,k}(I) = \frac{1}{2\pi}\int_0^{2\pi} H_1(I,\Phi) e^{i\langle k\cdot\Phi\rangle}\, d\Phi. \tag{7.0.21}$$

Our goal to finding the generating function that we need to write the new system in terms of the $I's$ is achieved. In fact we were able to find the ε order.

$$F(I,\Phi) = I\Phi + \varepsilon \sum_k \frac{iM_{1,k}(I)}{\langle k \cdot \omega_0(I) \rangle} e^{i\langle k\cdot\Phi\rangle} + o(\varepsilon^2). \tag{7.0.22}$$

Of course, we can carry out the same procedure to find higher orders of ε. However, it is worth mentioning that the formula of F_1 has a problem regarding vanishing denominators. Remember that For the original system H_0, we have two different kinds of motion. Namely, the resonant and the non-resonant trajectories. In the resonant case, we see that the denominator in (7.0.19) vanishes for some k. Not only this but also for the non-resonant case the denominator could be arbitrary close to zero for some k. This problem is known as existence of **small divisors**.

In order to overcome this problem, we need to find frequencies $\omega = (\omega_r, \omega_\theta)$ such that the smallness of the denominator in (7.0.22) is controlled. This kind of frequency is called the **Diophantine frequency**.

Definition 7.1. A vector $u = (u_1, u_2)$ is called Diophantine if there exist constants $\lambda > 0$ and $\tau > 0$ such that

$$|\langle u, l\rangle| \geq \frac{\lambda}{|l|^\tau}, \tag{7.0.23}$$

$\forall \mathbf{l} \in \mathbb{Z}^2 \setminus \{0\}$ where $|\mathbf{l}| = |l_1| + |l_2|$.

Of course, this can be generalized to n dimensions. Now, we want to investigate the size of the set that contains the Diophantine frequencies. Let

$$\chi_c^\tau = \{\omega \in \mathbb{R}^2 : |\langle l, \omega\rangle| \geq \frac{c}{|l|^\tau}, \forall l \in \mathbb{Z}^2 \setminus \{(0,0)\}\}, \tag{7.0.24}$$

for fixed c and τ. This is the complement set of $\Gamma_c^\tau = \cup_{l \neq (0,0)} \Gamma_{c,l}^\tau = \Gamma_c^\tau$ where

$$\Gamma_{c,l}^\tau = \{\omega \in \mathbb{R}^2 : |\langle l, \omega\rangle| < \frac{c}{|l|^\tau}, \text{for some } l \in \mathbb{Z}^2 \setminus \{(0,0)\}\}. \tag{7.0.25}$$

Now, we see that for any bounded $\Omega \subset \mathbb{R}^2$ there exist constants d, c' such that the Lebesgue measure estimation

$$m(\Gamma_{c,l}^\tau \cap \Omega) \leq \frac{dc}{|l|^{\tau+1}} \tag{7.0.26}$$

for all $c < c'$. In other words, for $0 < c << 1$ there exists $d > 0$ such that (7.0.26) holds. Thus,

$$m(\Gamma_{c,l}^\tau \cap \Omega) = \mathscr{O}(\frac{c}{|l|^{\tau+1}}). \tag{7.0.27}$$

By countable subadditivity,

$$m(\Gamma_c^\tau \cap \Omega) \leq \sum_l m(\Gamma_{c,l}^\tau \cap \Omega) = \mathscr{O}(c), \tag{7.0.28}$$

since $\tau + 1 > 1$. For fixed τ, define $\Gamma^\tau = \cap_{c>0} \Gamma_c^\tau$. This is the set of all $\omega \in \mathbb{R}^2$ satisfying (7.0.25) for some $c > 0$ and fixed τ. This set has measure zero and the complement $\chi^\tau = \cup_{c>0} \chi_c^\tau$ is of full measure.

Hence, the set of the frequencies that satisfy Diophantine condition exists and, in fact, almost every vector $\omega \in \mathbb{R}^2$ lies in χ^τ. This guarantees the existence of the Fourier series and give, at least, a hope that the series converges, and consequently, the averaging technique can be applied. In our system we were able to show the existence of such frequencies and in fact they are relatively easy to construct. Notice that this condition is actually stronger than the non-resonance condition, it does not only require independence of frequencies but also sufficiently rational independence. In other words, it guarantees that the smallness of the denominator is controlled.

To conclude, on the surface of the torus the trajectories with Diophantine frequencies are, in fact, those which will not be destroyed by a sufficiently small perturbation. It can be shown that in our case the distance between r_- and r_+ plays an important role in determining the frequencies. The trajectories that are given by values of k such that the frequencies are not rationally independent will be destroyed once we turn on the perturbation. On the other hand, the frequencies that satisfy Diophantine condition will persist and stay on the torus but will be only slightly deformed. Now the word "most" in the theorem is meaningful. The trajectories on the torus with rationally independent frequencies are dense and have full measure, exactly like measure of rationals and irrationals in the real line. Moreover, we see that not all the non-resonant tori will persist but only a subset of the set of non-resonant tori will do. Formally,

$$\chi_c^\tau = \{\omega : |\langle l, \omega \rangle| > \frac{c}{|l|^\tau}, \forall l \in \mathbb{Z}^2 \setminus \{(0,0)\}\} \subseteq \{\langle l, \omega \rangle| \neq 0 \forall l \in \mathbb{Z}^2 \setminus \{(0,0)\}\}. \tag{7.0.29}$$

It is, however, worth mentioning that although almost all frequencies are Diophantine, it is important to know that not almost all tori will persist for a given perturbation εH_1. In fact, given a perturbation εH_1 only a subset χ_c with $c \gg \sqrt{\varepsilon}$ will persist [**11**]. Meaning, the perturbation limits the size of c or one might say as the perturbation gets bigger the number of Diophantine frequencies gets smaller, which makes sense.

Now we are ready for a more precise statement of the theorem.

Theorem 7.2. *Suppose the integrable hamiltonian H_0 is non-degenerate and $H = H_0 + \varepsilon H_1$ is real analytic. Then there exists a constant $\delta > 0$ such that for*

$$|\varepsilon| < \delta c^2 \tag{7.0.30}$$

all tori of the unperturbed system with frequencies $\omega \in \chi_c^{\tau,\varepsilon}$ persist with a slight deformation.

The set $\chi_c^{\tau,\varepsilon}$ is a subset of frequencies satisfying (7.0.29). Namely, the set of frequencies satisfying (7.0.29) and $c >> \sqrt{\varepsilon}$ simultaneously.

Appendices

Complex integration is widely used in physics and it is usually the most convenient method. Many of the integrals we have done in this paper were computed using complex integration. More specifically, using residue theorem.

.1 Proof of Closeness of The Path In A Central Field With Potential $U(r) = -\frac{|\gamma|}{r}$

Here, we want to show that the path of a particle in a system with a potential $U(r) = -\frac{\gamma}{r}$, $\gamma > 0$, is closed. This can be done by showing that $\Delta\phi(r) = \frac{m}{n}2\pi$ for m, n integers.

$$\begin{aligned}\Delta\phi(r) &= 2\int_{r_-}^{r_+} \frac{A}{mr^2\sqrt{\frac{2}{m}(E+\frac{\gamma}{r}) - \frac{A^2}{m^2r^2}}}\,dr \\ &= 2\int_{r_-}^{r_+} \frac{A}{r\sqrt{2mEr^2 + 2m\gamma r - A^2}}\,dr.\end{aligned}$$

Let $\Gamma(z) = \frac{A}{z\sqrt{2mEz^2+2m\gamma z-A^2}}$. Now, we investigate the singularities of $\Gamma(z)$. Clearly, $z = 0$ is a singular point of $\Gamma(r)$. Therefore, the residue of Γ at $z = 0$ is

$$\lim_{z\to 0}(z-0)\frac{A}{z\sqrt{2mEz^2+2m\gamma z-A^2}} = \lim_{z\to 0}\frac{A}{\sqrt{2mEz^2+2m\gamma z-A^2}} = \frac{1}{i}.$$

Hence, the residue, $R(0)$, of Γ at $z = 0$ is $\frac{1}{i}$. Therefore, applying residue theorem,

$$\Delta\phi = 2\pi i\frac{1}{i} = 2\pi.$$

Which can be interpreted, as mentioned in section 5.3, as when ϕ completes one period, the radius r takes one complete revolution. Note that the whole revolution of r here is taken to be from r_- to r_+ and back to r_-.

.2 Explicit Formula For Action Variables J_r and J_θ

Again, using complex integration, we can compute the two integrals (5.5.13) and (5.5.14) in section 5.5.

$$\begin{aligned}J_r &= \frac{1}{2\pi}\oint_r \sqrt{2m(\frac{\gamma}{r} - |E|) - \frac{A^2}{r^2}}\,dr \\ &= \frac{1}{2\pi}\oint_r \frac{1}{r}\sqrt{2m\gamma r - 2mEr^2 - A^2}\,dr\end{aligned}$$

Let $\psi(z) = \frac{1}{z}\sqrt{2m\gamma z - 2mEz^2 - A^2}$. It can be easily seen that $z = 0$ is a singular point. Residue, $R(0)$, of $\psi(z)$ at $z = 0$ is

$$\lim_{z\to 0}(z-0)\frac{1}{z}\sqrt{2m\gamma z - 2mEz^2 - A^2} = Ai$$

Therefore, we have that $R(0) = Ai$.

Now, let's check if ψ has a singularity at $z = \infty$. Introduce the new variable $\rho = \frac{1}{r}$ so that when ρ approaches 0 ρ approaches ∞. After applying this change of variables, we have the following integral

$$J_\rho = -\frac{1}{2\pi}\oint_\rho \frac{1}{\rho^2}\sqrt{2m\gamma\rho - 2mE - A\rho^2}\,d\rho.$$

According to residue theorem, when a function $\psi(z)$ has a singular point z_0 of order $n > 1$, we multiply $\psi(z)$ by $(z - z_0)^n$, differentiate the result $n-1$ times, divide by $(n-1)!$ and then compute the limit at $z = z_0$. In our case here, residue, $R(\infty)$, of $\psi(z)$ at $z = \infty$ is

$$\frac{d}{d\rho}[(\rho - o)^2\frac{1}{\rho^2}\sqrt{2m\gamma\rho - 2mE - A\rho^2}] = \frac{m\gamma - A\rho}{\sqrt{2m\gamma\rho - 2mE - A\rho^2}}.$$

At $\rho = 0$, we have that

$$R(\infty) = \frac{m\gamma}{\sqrt{-2mE}} = \frac{m\gamma}{i\sqrt{2mE}}.$$

Hence,

$$J_r = \frac{1}{2\pi}2\pi i(Ai + \frac{m\gamma}{i\sqrt{2mE}}) = -A + \frac{m\gamma}{\sqrt{2mE}}$$

Now we look at integral (5.5.14),

$$J_\theta = \frac{1}{2\pi}\oint_\theta \sqrt{A^2 - \frac{\alpha_\phi^2}{\sin^2\theta}}\,d\theta.$$

With the help of the following substitutions:

$$z = \sin\theta \qquad \text{and} \qquad d\theta = \frac{dz}{\sqrt{1-z^2}},$$

we can rewrite the integral as follows,

$$J_z = \frac{A}{2\pi}\oint \frac{1}{z}\sqrt{\frac{z^2 - \frac{\alpha_\phi^2}{A^2}}{1-z^2}}\,dz. \tag{.2.1}$$

Since zero is a singular point of the integrand, then we have

$$Res(0) = \lim_{z\to 0}(z-0)\frac{1}{z}\sqrt{\frac{z^2-a^2}{1-z^2}},$$

where $a = \frac{\alpha_\phi^2}{A^2}$. As a result, we obtain $Res(0) = i\frac{\alpha_\phi}{A}$. Now, we need to check if the integrand in (.2.1) has singularity at $z = \infty$. Let $f(z) = \frac{1}{z}\sqrt{\frac{z^2A^2-\alpha_\phi{}^2}{1-z^2}}$. Now, define a new function G such that $G(x) = f(\frac{1}{x})$. This new function gives the following expression for J_x:

$$J_x = -\frac{A}{2\pi}\oint \frac{1}{x}\sqrt{\frac{A^2 - x^2\alpha_\phi^2}{A^2(x^2-1)}}\,dx. \tag{.2.2}$$

From (.2.2), one can easily see that as x tends to zero, the integrand tends to ∞. Since G has singularity at zero, then f has singularity at ∞ and $Res(f(x),\infty) = Res(G(x),0) = -i$. Hence, according to residue theorem,

$$J_\theta = \frac{A}{2\pi}2\pi i(i\frac{\alpha_\phi}{A} - i) = A - \alpha_\phi.$$

.3 Path In A Central Field With Potential $U_1(r) = -\frac{\gamma}{r} - \frac{\varepsilon}{r^3}$

For the integral (6.1.3),

$$\Delta\phi(r) = 2\int_{r_-}^{r_+} \frac{A_1}{mr^2\sqrt{\frac{2}{m}(E_1 + \frac{\gamma}{r} + \frac{\varepsilon}{r^3}) - \frac{A_1^2}{m^2r^2}}}dr,$$

we will use some basic tricks rather than complex integration. Unfortunately, the solution of this integral will be approximated. First, we rewrite the integral to be as follows:

$$\begin{aligned}\Delta\phi(r) &= 2\int_{r_-}^{r_+} \frac{A_1}{r^2\sqrt{2m(E_1+\frac{\gamma}{r}+\frac{\varepsilon}{r^3})-\frac{A_1^2}{r^2}}}dr \\ &= 2\int_{r_-}^{r_+} \frac{A_1}{r^2\sqrt{2m(E_1+\frac{\gamma}{r})-\frac{A_1^2}{r^2}+2m\frac{\varepsilon}{r^3}}}dr\end{aligned}$$

Observe that the integrand is actually the partial derivative with respect to A_1 of the following expression:

$$\sqrt{2m(E_1+\frac{\gamma}{r})-\frac{A_1^2}{r^2}+2m\frac{\varepsilon}{r^3}}.$$

Therefore, we can write the integral as

$$\Delta\phi = -2\frac{\partial}{\partial A_1}\int_{r_-}^{r_+}\sqrt{2m(E_1+\frac{\gamma}{r})-\frac{A_1^2}{r^2}+2m\frac{\varepsilon}{r^3}}.$$

Now, let us denote $2m(E_1+\frac{\gamma}{r})-\frac{A_1^2}{r^2} = \chi$, so we have

$$\begin{aligned}\Delta\phi &= -2\frac{\partial}{\partial A_1}\int_{r_-}^{r_+}\sqrt{\chi+2m\frac{\varepsilon}{r^3}}\,dr. \\ &= -2\frac{\partial}{\partial A_1}\int_{r_-}^{r_+}\sqrt{\chi}\sqrt{1+\frac{2m}{\chi}\frac{\varepsilon}{r^3}}\,dr.\end{aligned}$$

Call $\frac{2m\varepsilon}{r^3\chi} = \mu$, then for $2m\varepsilon < r^3\chi$, we can use Taylor expansion to obtain

$$\sqrt{1+\mu} = 1+\frac{1}{2}\mu-\frac{1}{8}\mu^2+\cdots \approx 1+\frac{1}{2}\mu.$$

Plugging this into our original integral gives

$$\Delta\phi \approx -2\frac{\partial}{\partial A_1}\int_{r_-}^{r_+}\sqrt{\chi}\,dr + 2\frac{\partial}{\partial A_1}\int_{r_-}^{r_+}\frac{m\varepsilon}{r^3\sqrt{\chi}}\,dr.$$

Now, the first term after differentiating is integral (5.3.6) which we already have done in appendix .1. We are left with the second term. With the help of (5.3.1) and (5.3.3) we get

$$\frac{d\theta}{dr} = \frac{A_1}{mr^2\sqrt{\frac{2}{m}(E_1 + \frac{\gamma}{r}) - \frac{A_1^2}{m^2r^2}}} = \frac{A_1}{r^2\sqrt{\chi}}. \tag{.3.1}$$

We also have the following relation:

$$\frac{F}{r} = 1 + e\cos\theta. \tag{.3.2}$$

Where F is what is so called the **semilatus rectum** of the orbit and e is the **eccentricity**. [**6**]

Using (.3.1) and (.3.2), we obtain the integral in the following form

$$2\frac{\partial}{\partial A_1}\int_0^{\pi} \frac{m\varepsilon}{A_1}\frac{1+e\cos\theta}{F}\,d\theta$$

which can be easily integrated. Finally, integral (6.1.3) is approximately

$$4\pi - \frac{2\pi m\varepsilon}{FA_1^2}$$

Bibliography

[1] J. L. E. Dreyer, *History of Planetary Systems*, Cambridge University Press, 1906

[2] R. Fitzpatrick, *Introduction to Celestial Mechanics*, Cambridge University Press, 2012

[3] H. Dumas, *The KAM Story: A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov-Arnold-Moser Theory*, World Scientific, New Jersey, 2014

[4] B. V. BRUNT, *The Calculus of Variations*, Springer,NY, 2004.

[5] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer-Verlag, NY, 1978,1989.

[6] M. L. BOAS, *Mathematical Methods in the Physical Sciences*, 3rd ed., John Wiley and Sons, Hoboken, NJ, 2006.

[7] G. TESCHI, *Ordinary Differential Equations and Dynamical Systems*, Vol. 140, American Mathematical Society, Providence, RI, 2012

[8] M. BORN, *The Mechanics of The Atom*, G. Bell and sons, LTD, London, 1927.

[9] L. D. LANDAU AND E. M. LIFSHITZ, *Mechanics*, Vol. 1, Pergamon Press, Elmsford, NY, 1976.

[10] H. GOLDSTEIN, C. POOLE, AND J. SAFKO, *Classical Mechanics*, 3rd ed., Addison Wesley, San Francisco, 2000.

[11] J. KOVALEVSKY, *Introduction To Celestial Mechanics*, Vol. 7, D. Reidel Publishing Company, Dordrecht, Holland, 1967.

[12] M. TABOR, *Chaos And Integrability In Nonlinear Dynamics*, John Wiley and Sons, New York, NY

[13] G. Popov, P. Topalov, *On the Integral Geometry of Liouville Billiard Tables*, https://arxiv.org/abs/0906.0451, 2009

[14] J. POSCHEL, *A Lecture On The Classical KAM Theorem*, https://arxiv.org/abs/0908.2234, 2009