BACHELOR RESEARCH PROJECT MATHEMATICS AND PHYSICS

The Kepler Problem and Its Relation to Extremal Black Holes

Author: Wojciech Jan Anyszka

Physics supervisor: Prof. Dr. Diederik Roest
Mathematics supervisor: Prof. Dr. Marcello Seri

November 2023

Abstract

This thesis explores the unique properties of the classical Kepler problem, including Bertrand's theorem [Ber73], the connection between Kepler and harmonic oscillator potentials [Boh11], and the existence of an additional conserved quantity — the Laplace-Runge-Lenz (LRL) vector. The role of symmetries in this context is explored. Then a comprehensive proof of Moser's construction [Mos70], establishing the correspondence between non-constant geodesics on an n-dimensional sphere and Kepler orbits with negative energies in n-dimensions, is presented. This construction demonstrates that the Kepler problem has a larger symmetry group compared to an arbitrary central potential. The relativistic corrections to two-body problems, which generically induce perihelion precession, are then investigated. Notably, a specific relativistic system involving an extremal test particle near an oppositely charged extremal Einstein-Maxwell-dilaton black hole does not exhibit perihelion precession [Nee+23]. However, this phenomenon is limited to a specific value of the dilaton coupling constant, specifically $a = \sqrt{3}$. A generalized theorem based on this construction is established, followed by an examination of cases where $a \neq \sqrt{3}$. It was shown that away from $a = \sqrt{3}$, the test-particle orbits correspond to the orbits of a perturbed Kepler problem.

university of groningen

faculty of science and engineering

Acknowledgements

I want to sincerely thank my supervisors Prof. Dr. Marcello Seri and Prof. Dr. Diederik Roest for their support and guidance throughout the completion of this thesis. They not only introduced me to this interesting topic but also struck the perfect balance between granting me the freedom to explore my ideas and providing me with invaluable suggestions that steered me in the right direction. I would like to also thank Dijs de Neeling for giving me the context behind his work and pointing out the important technical details. Lastly, I would like to thank Prof. Tamás Görbe and Martijn Kluitenberg for their insightful questions during my talk, which encouraged me to elaborate on certain technical details in my thesis.

Contents

Chapter 1

Introduction

The Kepler problem, which involves the motion of two nonrelativistic bodies in a gravitational field, is one of the most thoroughly investigated problems of classical mechanics. In this classical scenario, it is well-known that the orbits are conic sections. However, there are also less known facts about this problem which we will examine in this thesis. For example, Bertrand's Theorem [Ber73] states that the Kepler $1/r$ potential and the harmonic oscillator r^2 potential are the only ones whose bounded orbits are closed. In fact, in two dimensions, these two potentials can be regarded as describing the same problem — this connection goes via Bohlin transformations [Boh11]. The Kepler potential also has an additional conserved quantity — the Laplace-Runge-Lenz (LRL) vector — which other central potentials do not have. It encodes information about the eccentricity and orientation of the orbits. Noether's Theorem then suggests that the Kepler potential should have an additional symmetry as compared to generic central potentials whose symmetry group is $SO(3)$ — the group of rotations in three dimensions. However, this additional symmetry is not obvious from just looking at the potential. For this reason, it is often referred to as a hidden symmetry. There is a fascinating connection between the Kepler orbits with negative energy and a higher-dimensional geometric model [Mos70], which unveils this hidden symmetry. Namely, this construction reduces the negative energy Kepler orbits in n-dimensions to the free motion (geodesic motion) on an n-dimensional sphere. Hence it follows that the symmetry group for negative energy orbits — the set of operations which carry negative energy solutions to the problem to other such solutions — of the n-dimensional Kepler problem is $SO(n+1)$, the group of rotations in $(n+1)$-dimensional space. In particular, in three dimensions this symmetry group is $SO(4)$, which is larger than the symmetry group of a generic central potential.

After closely examining the classical Kepler problem we will analyze the relativistic corrections to two specific two-body problems. We will show that those induce perihelion precession. Remarkably, a recent paper [Nee+23] shows that there are relativistic systems whose bounded orbits nevertheless are elliptical and hence the perihelion doesn't process. Concretely, it was shown that, for a specific value of a coupling constant ($a = \sqrt{3}$), the orbits of an extremal test particle in the background of oppositely charged extremal Einstein-Maxwell-dilaton black hole correspond to classical Kepler orbits. The research objective of this thesis is to investigate what happens when $a \neq \sqrt{3}$.

This thesis is structured as follows. In chapter 2 and chapter 3, we will study the distinct properties of the Kepler potential. We will see that some of them are consequences of the large symmetry group of the Kepler problem. We will display this symmetry in chapter 5 where, following [Mos70], we provide a comprehensive proof that nonconstant geodesics on $\mathbb{S}^n$ corresponds to Kepler orbits with negative energies. However, this construction uses significant mathematical machinery such as differential and symplectic geometry — we will introduce the necessary prerequisites in chapter 4. After displaying the remarkable properties of the classical Kepler problem, we will give a physical intermezzo on relativistic corrections in chapter 6. These will serve as a prelude to the final chapter in which we present a result from [Nee+23]. In the last section, we attempt to extend the existing result.

This is a Physics-Mathematics double bachelor thesis. Purposefully, the topic of this thesis lies at the intersection of these two subjects. Nevertheless, one can regard chapters 2, 3, 6 and 7 to belong to the physics part of the thesis while chapters 2, 3, 4, 5 and 7 to belong to the mathematics part.

Chapter 2

The Classical Kepler Problem and Its Conserved Quantities

In this chapter, we will give a Newtonian treatment of the Kepler Problem in three dimensions, i.e. the problem of finding the trajectories of two gravitationally interacting bodies. More mathematically, we want to solve

$$\begin{aligned} m_1\ddot{\mathbf{x}}_1 &= -\frac{Gm_1m_2}{\|\mathbf{x}_1-\mathbf{x}_2\|^3}(\mathbf{x}_1-\mathbf{x}_2) \\ m_2\ddot{\mathbf{x}}_2 &= -\frac{Gm_1m_2}{\|\mathbf{x}_1-\mathbf{x}_2\|^3}(\mathbf{x}_2-\mathbf{x}_1) \end{aligned}$$

for $\mathbf{x}_1(t),\mathbf{x}_2(t)\in\mathbb{R}^3$. Up to a point our argument works for any force generated by a central potential $V(\|\mathbf{x}_1-\mathbf{x}_2\|)$ — a potential that only depends on the separation of particles. We will make this explicit by first treating this more general case and only when necessary we will specify that $V=-\frac{Gm_1m_2}{\|\mathbf{x}_1-\mathbf{x}_2\|}$. Our treatment is based on the ones given in [Ton13], [Ser23b] and [Bae08].

2.1 Two Body Problem in a Central Potential

The general two-body problem in a shared central potential is to find the trajectories $\mathbf{x}_1(t),\mathbf{x}_2(t)\in\mathbb{R}^3$ obeying

$$\begin{aligned} m_1\ddot{\mathbf{x}}_1 &= -\frac{\partial V(\|\mathbf{x}_1-\mathbf{x}_2\|)}{\partial\mathbf{x}_1} \\ m_2\ddot{\mathbf{x}}_2 &= -\frac{\partial V(\|\mathbf{x}_1-\mathbf{x}_2\|)}{\partial\mathbf{x}_2}. \end{aligned}$$

We can simplify this problem by decomposing the motion into the motion of the centre of mass and the motion of a particle in a stationary potential. First, the center of mass is given by

$$\mathbf{x}_{CoM}=\frac{1}{M}(m_1\mathbf{x}_1+m_2\mathbf{x}_2),$$

where $M=m_1+m_2$ is the total mass. To recover the $\mathbf{x}_1,\mathbf{x}_2$ from $\mathbf{x}_{CoM}$ we need another quantity that is independent of $\mathbf{x}_{CoM}$. A reasonable choice is $\mathbf{x}:=\mathbf{x}_1-\mathbf{x}_2$. Let's also define $r:=\|\mathbf{x}\|$. Now, the center of mass is stationary as

$$\ddot{\mathbf{x}}_{CoM}=\frac{1}{M}(m_1\ddot{\mathbf{x}}_1+m_2\ddot{\mathbf{x}}_2)=-\frac{1}{M}\left(\frac{\partial V(r)}{\partial\mathbf{x}_1}+\frac{\partial V(r)}{\partial\mathbf{x}_2}\right)=0,$$

while the separation vector $\mathbf{x}$ satisfies

$$\ddot{\mathbf{x}}=\ddot{\mathbf{x}}_1-\ddot{\mathbf{x}}_2=-\frac{1}{m_1}\frac{\partial V(r)}{\partial\mathbf{x}_1}+\frac{1}{m_2}\frac{\partial V(r)}{\partial\mathbf{x}_2}=-\frac{m_2+m_1}{m_1m_2}\frac{\partial V(r)}{\partial\mathbf{x}}.$$

Both equations follow from

$$\frac{\partial V(r)}{\partial \mathbf{x}_1} = \frac{\partial V(r)}{\partial r}\frac{\partial r}{\partial \mathbf{x}_1} = \frac{\partial V(r)}{\partial r}\frac{\mathbf{x}_1 - \mathbf{x}_2}{r} = \frac{\partial V(r)}{\partial r}\frac{\partial r}{\partial \mathbf{x}} = \frac{\partial V(r)}{\partial \mathbf{x}}$$

and

$$\frac{\partial V(r)}{\partial \mathbf{x}_2} = \frac{\partial V(r)}{\partial r}\frac{\partial r}{\partial \mathbf{x}_2} = \frac{\partial V(r)}{\partial r}\frac{\mathbf{x}_2 - \mathbf{x}_1}{r} = -\frac{\partial V(r)}{\partial \mathbf{x}}.$$

By solving the two equations

$$\begin{aligned} \ddot{\mathbf{x}}_{CoM} &= 0 \\ \frac{m_1 m_2}{m_1 + m_2}\ddot{\mathbf{x}} &= -\frac{\partial V(r)}{\partial \mathbf{x}}, \end{aligned}$$

we can solve the original problem since

$$\mathbf{x}_1 = \mathbf{x}_{CoM} + \frac{m_2}{M}\mathbf{x} \quad \text{and} \quad \mathbf{x}_2 = \mathbf{x}_{CoM} - \frac{m_1}{M}\mathbf{x}.$$

The solution to $\ddot{\mathbf{x}}_{CoM} = 0$ is straightforward to find, no "twists" or "turns" needed. Thus, we effectively reduced our six-dimensional problem to a three-dimensional one — finding $\mathbf{x}$. A key observation at this point is that $M\dot{\mathbf{x}}_{CoM}$ is the total momentum of our system. What allowed us to reduce the complexity of our problem is precisely the conservation of momentum. This elementary result already shows why it is important to study conserved quantities of a given system.

Moving forward, let us examine

$$\mu\ddot{\mathbf{x}} = -\frac{\partial V(r)}{\partial \mathbf{x}} = -\frac{\partial V(r)}{\partial r}\hat{\mathbf{x}}, \tag{2.1}$$

where we introduced the reduced mass $\mu = m_1 m_2/(m_1 + m_2)$ and the unit vector $\hat{\mathbf{x}} = \mathbf{x}/r$. The above equation tells us that we can forget the definition of $\mathbf{x}$ and treat it as the position of a particle of mass μ moving in a central potential. This is still a non-trivial equation but fortunately, conservation laws again come to our rescue. Namely, the angular momentum $\mathbf{L} = \mu\mathbf{x} \times \dot{\mathbf{x}}$ is conserved since

$$\dot{\mathbf{L}} = \mu\underbrace{\dot{\mathbf{x}} \times \dot{\mathbf{x}}}_{=0} + \mu\mathbf{x} \times \ddot{\mathbf{x}} = -\mu\mathbf{x} \times \frac{\partial V(r)}{\partial r}\hat{\mathbf{x}} = 0.$$

By the properties of cross product, both $\mathbf{x}$ and $\dot{\mathbf{x}}$ are perpendicular to $\mathbf{L} \propto \mathbf{x} \times \dot{\mathbf{x}}$. Constancy of $\mathbf{L}$ then implies that $\mathbf{x}$ and $\dot{\mathbf{x}}$ lie in a plane perpendicular to $\mathbf{L}$. Picking a coordinate system with z-axis along $\mathbf{L}$ allows us to treat $\mathbf{x}$ as a two-dimensional vector in the xy-plane. In particular, we can represent it in polar coordinates as

$$\mathbf{x} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{bmatrix}.$$

Differentiation rules together with equation (2.1) give us that r and θ obey

$$\begin{aligned} \mu(\ddot{r}\cos(\theta) - 2\dot{r}\dot{\theta}\sin(\theta) - r\ddot{\theta}\sin(\theta) - r\dot{\theta}^2\cos(\theta)) &= -\frac{\partial V(r)}{\partial r}\cos(\theta) \\ \mu(\ddot{r}\sin(\theta) + 2\dot{r}\dot{\theta}\cos(\theta) + r\ddot{\theta}\cos(\theta) - r\dot{\theta}^2\sin(\theta)) &= -\frac{\partial V(r)}{\partial r}\sin(\theta) \end{aligned} \tag{2.2}$$

Multiplying the first equation by $\sin(\theta)$ and the second by $\cos(\theta)$ gives

$$\mu(\ddot{r}\cos(\theta)\sin(\theta) - 2\dot{r}\dot{\theta}\sin^2(\theta) - r\ddot{\theta}\sin^2(\theta) - r\dot{\theta}^2\sin(\theta)\cos(\theta)) = -\frac{\partial V(r)}{\partial r}\cos(\theta)\sin(\theta)$$
$$\mu(\ddot{r}\sin(\theta)\cos(\theta) + 2\dot{r}\dot{\theta}\cos^2(\theta) + r\ddot{\theta}\cos^2(\theta) - r\dot{\theta}^2\sin(\theta)\cos(\theta)) = -\frac{\partial V(r)}{\partial r}\sin(\theta)\cos(\theta).$$

Subtracting the first from the second yields

$$0 = 2\mu\dot{r}\dot{\theta} + \mu\ddot{\theta} = \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}(\mu r^2\dot{\theta}).$$

The conserved quantity $\ell := \mu r^2\dot{\theta}$ is nothing but the magnitude of angular momentum

$$\mathbf{L} = \mu\mathbf{x}\times\dot{\mathbf{x}} = \mu\begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 0 \end{bmatrix} \times \begin{bmatrix} \dot{r}\cos(\theta) - r\dot{\theta}\sin(\theta) \\ \dot{r}\sin(\theta) + r\dot{\theta}\cos(\theta) \\ 0 \end{bmatrix} = \mu r^2\dot{\theta}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Multiplying the first equation in (2.2) by $\cos(\theta)$ and the second by $\sin(\theta)$ and adding the results gives

$$\mu\ddot{r} - \mu r\dot{\theta}^2 = -\frac{\partial V(r)}{\partial r}. \tag{2.3}$$

A moment of reflection reveals that to go from a 3D to a 2D problem we only used the direction of $\mathbf{L}$. This suggests that we can squeeze more out of the conservation of angular momentum. Indeed, substituting $\dot{\theta} = \ell/\mu r^2$ into equation (2.3) gives

$$\mu\ddot{r} = \frac{\ell^2}{\mu r^3} - \frac{\partial V(r)}{\partial r} = -\frac{\partial}{\partial r}\left(V(r) + \frac{\ell^2}{2\mu r^2}\right)$$

Therefore conservation of the magnitude of $\mathbf{L}$ further reduces our problem to the motion of a one-dimensional particle in the effective potential

$$V_{\text{eff}}(r) := V(r) + \frac{\ell^2}{2\mu r^2}.$$

2.2 The Kepler Problem

To make further progress, we specialize to the Kepler potential

$$V(r) = -\frac{\mu k}{r}$$

with $k = G(m_1 + m_2)$. Therefore our problem is to find solutions to

$$\ddot{r} = \frac{\ell^2}{\mu^2 r^3} - \frac{k}{r^2}.$$

Let us pursue the generally good idea to first identify the stationary solutions. Using the ansatz $\dot{r}_{stat} = 0$ and $\ell = r_{stat}^2\dot{\theta}_{stat}$ gives us that

$$r_{stat} = \frac{\ell^2}{\mu^2 k} \qquad \dot{\theta}_{stat} = \frac{k^2\mu^4}{\ell^3}$$

is the only stationary solution. This is clearly a circle traversed with constant speed. Another simple solution of interest is when $\ell = 0$ which in 3D Cartesian coordinates corresponds to $\mathbf{x}$ and $\dot{\mathbf{x}}$ being parallel. Qualitatively, we see that this is a collision orbit with the angle θ being constant and r rapidly decreasing to 0.

To tackle the general case with $\ell \neq 0$, we use the remarkable fact that in the case of the inverse square law, there exists another conserved vector quantity. It is the Laplace-Runge-Lenz (LRL) vector defined by

$$\mathbf{A} = \mu \dot{\mathbf{x}} \times \mathbf{L} - \frac{\mu^2 k}{\|\mathbf{x}\|}\mathbf{x}.$$

To check that it is conserved we take its time derivative

$$\dot{\mathbf{A}} = \mu^2 \ddot{\mathbf{x}} \times (\mathbf{x} \times \dot{\mathbf{x}}) + \mu \dot{\mathbf{x}} \times \dot{\mathbf{L}} - \frac{\mu^2 k}{\|\mathbf{x}\|}\dot{\mathbf{x}} + \frac{\mu^2 k}{(\mathbf{x} \cdot \dot{\mathbf{x}})}\|\mathbf{x}\|^3 \mathbf{x}.$$

We can simplify this expression by using the vector identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{c} \cdot \mathbf{a})\mathbf{b} - (\mathbf{b} \cdot \mathbf{a})\mathbf{c}$ and that $\mathbf{L}$ is conserved. This yields

$$\dot{\mathbf{A}} = \mu^2 (\ddot{\mathbf{x}} \cdot \dot{\mathbf{x}})\mathbf{x} - \mu^2 (\ddot{\mathbf{x}} \cdot \mathbf{x})\dot{\mathbf{x}} - \frac{\mu^2 k}{\|\mathbf{x}\|}\dot{\mathbf{x}} + \frac{\mu^2 k (\mathbf{x} \cdot \dot{\mathbf{x}})}{\|\mathbf{x}\|^3}\mathbf{x}.$$

Finally, using the equation of motion (2.1) we get

$$\dot{\mathbf{A}} = -\frac{k\mu^2}{\|\mathbf{x}\|^3}(\mathbf{x} \cdot \dot{\mathbf{x}})\mathbf{x} + \frac{k\mu^2}{\|\mathbf{x}\|^3}(\mathbf{x} \cdot \mathbf{x})\dot{\mathbf{x}} - \frac{k\mu^2}{\|\mathbf{x}\|}\dot{\mathbf{x}} + \frac{k\mu^2 (\mathbf{x} \cdot \dot{\mathbf{x}})}{\|\mathbf{x}\|^3}\mathbf{x} = 0$$

thus $\mathbf{A}$ is conserved. Before we leverage the conservation of $\mathbf{A}$ to solve the Kepler problem, we will compute the norm of $\mathbf{A}$. It is

$$\begin{aligned}
\|\mathbf{A}\|^2 &= \mu^2 \|\dot{\mathbf{x}} \times \mathbf{L}\|^2 + \frac{k^2 \mu^4}{\|\mathbf{x}\|^2}\|\mathbf{x}\|^2 - \frac{2k\mu^3}{\|\mathbf{x}\|}\mathbf{x} \cdot (\dot{\mathbf{x}} \times \mathbf{L}) \\
&= \mu^2 ((\dot{\mathbf{x}} \times \mathbf{L}) \times \dot{\mathbf{x}}) \cdot \mathbf{L} + k^2 \mu^4 - \frac{2k\mu^3}{\|\mathbf{x}\|}\mathbf{L} \cdot (\mathbf{x} \times \dot{\mathbf{x}}) \\
&= \mu^2 \|\dot{\mathbf{x}}\|^2 \ell^2 + k^2 \mu^4 - \frac{2k\mu^2}{\|\mathbf{x}\|}\ell^2 \\
&= 2\mu \ell^2 \left(\frac{\mu \|\dot{\mathbf{x}}\|^2}{2} - \frac{k\mu}{\|\mathbf{x}\|} \right) + k^2 \mu^4 \\
&= 2\mu \ell^2 E + k^2 \mu^4,
\end{aligned}$$

where E is the total energy of our particle. Accidentally, we can deduce from the conservation of $\mathbf{A}$ and $\mathbf{L}$ and the above expression that the total energy E is also a conserved quantity — albeit not independent of $\mathbf{A}$ and $\mathbf{L}$. Moreover, all of the components of $\mathbf{A}$ and $\mathbf{L}$ are not independent either since

$$\mathbf{A} \cdot \mathbf{L} = \mu (\dot{\mathbf{x}} \times \mathbf{L}) \cdot \mathbf{L} - \frac{k\mu^2}{\|\mathbf{x}\|}\mathbf{x} \cdot \mathbf{L} = 0.$$

Thus $\mathbf{A}$ lies in the plane orthogonal to $\mathbf{L}$ and we have 5 independent conserved quantities.

Now recall that when we transformed to polar coordinates we didn't specify the direction of the x-axis. We can use this freedom to postulate that $\mathbf{A}$ points in the positive x-direction. With this choice, we have that

$$\mathbf{A} \cdot \mathbf{x} = \|\mathbf{A}\| r \cos(\theta) \quad \text{and} \quad \mathbf{A} \cdot \mathbf{x} = m (\dot{\mathbf{x}} \times \mathbf{L}) \cdot \mathbf{x} - k\mu^2 \|\mathbf{x}\|.$$

Combining the two equations and permuting the triple product yields

$$\|\mathbf{A}\| r \cos(\theta) = \ell^2 - k\mu^2 r.$$

Therefore r as a function θ is

$$r(\theta) = \frac{\frac{\ell}{k\mu^2}}{\frac{\|\mathbf{A}\|}{k\mu^2}\cos(\theta) + 1}, \tag{2.4}$$

which we recognize as an equation for a conic section with eccentricity

$$e = \frac{\|\mathbf{A}\|}{k\mu^2}. \tag{2.5}$$

Note that when we specified the x-axis we implicitly assumed that $\mathbf{A} \neq 0$. However, we see that the solution $r(\theta)$ is still valid in this case and reduces to the previously found circular stationary solution so all is well. It is possible to find an explicit time parametrization of all solutions but we will not pursue this. We already derived Kepler's first law — the shapes of orbits — from Newtonian mechanics. Finally, we can use our results (2.4), (2.5) to give a physical interpretation of the LRL vector. Namely, observe that the perihelion of the orbit occurs when $\theta = 0$. Since in the chosen coordinates θ measures the angle from the LRL vector, the LRL vector points towards the perihelion. Moreover, its magnitude determines the eccentricity of the orbit. Therefore, the conservation of the LRL vector means that the perihelion does not precess and the eccentricity of the orbit remains constant. A pictorial representation of these facts can be seen in Figure 2.1.

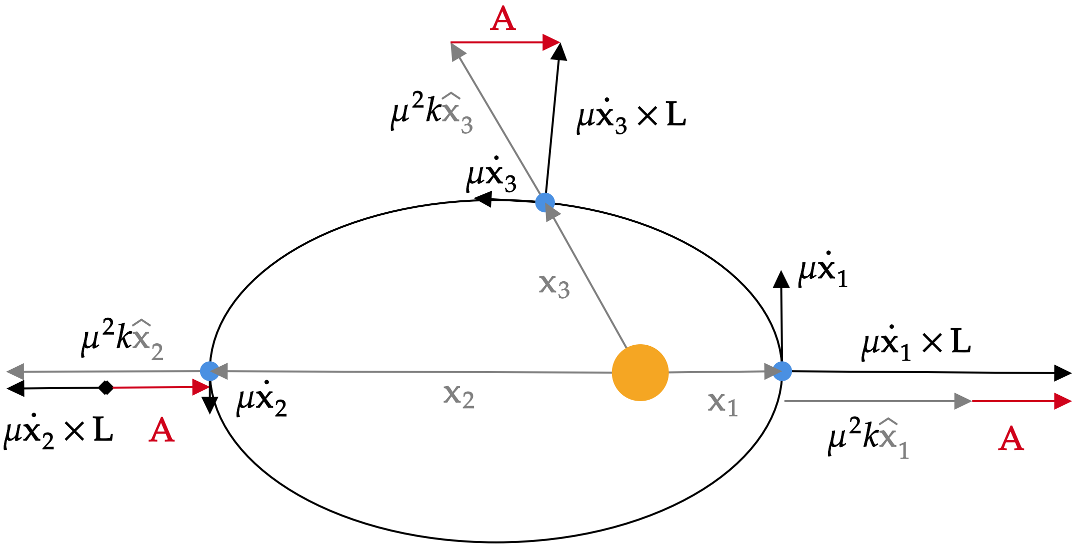

Figure 2.1: The vectors $\mu\dot{\mathbf{x}}$, $\mathbf{L}$, and $\mathbf{A}$ at three positions in an elliptical orbit. The Laplace-Runge-Lenz vector $\mathbf{A}$ always points in the direction of x-axis with a magnitude $k\mu^2 e$ where e is the eccentricity of the orbit. The image was adapted from [GSP14].

Let's reflect on what we did in the previous paragraphs. We started with a two-body problem for an arbitrary central potential. We then used the conservation of angular momentum to show that the motion is constrained to a plane and can be regarded as the motion of a one-dimensional particle in an effective potential. At this point, we specified the potential to be the gravitational one which enjoys an additional conserved quantity — the LRL vector. The conservation of the LRL vector allowed us to quickly derive the shape of the orbits. In the next section, we will see that the Kepler potential is the only one for which an LRL-type vector is conserved.

2.3 Another Look at the Laplace-Runge-Lenz Vector

Recall that we defined the LRL vector by

$$\mathbf{A} = \mu\dot{\mathbf{x}} \times \mathbf{L} - \frac{\mu^2 k}{r}\mathbf{x}.$$

We can rewrite this expression and make it more amendable to a generalization. Namely, observe that

$$\mathbf{A} = \mu\dot{\mathbf{x}} \times \mathbf{L} + \mu V(r)\mathbf{x}.$$

This definition can immediately be applied to any central potential $V(r)$. Let us check for which potentials this LRL-type vector is conserved. Supposing that $\dot{\mathbf{A}} = 0$ for all trajectories, we get

$$0 = \mu\ddot{\mathbf{x}} \times \mathbf{L} + \mu\dot{\mathbf{x}} \times \dot{\mathbf{L}} + \mu V(r)\dot{\mathbf{x}} + \mu\frac{\partial V(r)}{\partial r}\frac{(\mathbf{x}\cdot\dot{\mathbf{x}})}{r}\mathbf{x} = -\frac{\partial V(r)}{\partial r}\frac{\mathbf{x}}{r} \times \mathbf{L} + \mu V(r)\dot{\mathbf{x}} + \mu\frac{\partial V(r)}{\partial r}\frac{(\mathbf{x}\cdot\dot{\mathbf{x}})}{r}\mathbf{x},$$

where we used that for central potentials, $\mathbf{L}$ is conserved. Using the definition of $\mathbf{L}$ and cross product properties, we further get

$$\begin{aligned}
0 &= -\mu\frac{\partial V(r)}{\partial r}\frac{\mathbf{x}}{r} \times (\mathbf{x} \times \dot{\mathbf{x}}) + \mu V(r)\dot{\mathbf{x}} + \mu\frac{\partial V(r)}{\partial r}\frac{(\mathbf{x}\cdot\dot{\mathbf{x}})}{r}\mathbf{x} \\
&= -\mu\frac{\partial V(r)}{\partial r}\frac{1}{r}((\mathbf{x}\cdot\dot{\mathbf{x}})\mathbf{x} - (\mathbf{x}\cdot\mathbf{x})\dot{\mathbf{x}}) + \mu V(r)\dot{\mathbf{x}} + \mu\frac{\partial V(r)}{\partial r}\frac{(\mathbf{x}\cdot\dot{\mathbf{x}})}{r}\mathbf{x} \\
&= \mu\left(r\frac{\partial V(r)}{\partial r} + V(r)\right)\dot{\mathbf{x}}.
\end{aligned}$$

Since this must hold for trajectories and in particular for all initial conditions, we must have that

$$r\frac{\partial V(r)}{\partial r} + V(r) = 0 \quad \Rightarrow \quad \frac{\partial V(r)}{\partial r} = -\frac{1}{r}V(r).$$

The unique solution to this differential equation is $V(r) = -\frac{k}{r}$ for any $k \in \mathbb{R}$. Therefore we see that Kepler potential is the unique central potential for which the LRL-type vector is conserved! The fact that Kepler potential has this additional conserved quantity is closely related to the fact that its symmetry group is $SO(4)$ — rotations in 4D — as opposed to the smaller symmetry group $SO(3)$ —rotations in 3D—of a generic central potential. In the next section, we will show how conservation laws can be derived from symmetries. This will prepare the stage for our analysis of the symmetry group of the Kepler problem that will explain its additional conserved quantity.

2.4 Symmetries and Conservation Laws of Central Potentials

It was a great achievement of 20th-century physics to bring to light the intimate connection between symmetries and conservation laws. This was accomplished with the celebrated Noether's theorem which (roughly) states that for any continuous symmetry of the Lagrangian of a system, there is a corresponding quantity conserved in time. In what follows we will give arguments based on Newtonian mechanics to display this general phenomenon in a few special cases.

First, we need to clarify what we mean by a symmetry of a physical system. To do this, let us first examine what statements like "circles have rotational symmetry" mean and then generalize this notion. We say that a circle has rotational symmetry because if I show you a circle and ask you to close your eyes, then after opening your eyes you cannot tell if I rotated the circle or kept it fixed. Thus a symmetry of an object is a transformation that keeps all of its essential features intact. Analogously, imagine that you are closed in an elevator in deep space. If I move the elevator to some other location while you are asleep then after you wake up you cannot tell if it was moved. In other words, the outcomes of all the experiments you can do in your elevator will be the same regardless of whether I move the elevator or not. Since the evolution of physical systems and hence outcomes of experiments are described by differential equations, we say that a transformation of a physical system is its **symmetry** if it carries solutions to the underlying equations to other solutions. With this definition, we will examine a specific class of symmetries — the Newtonian continuous symmetries.

Consider a general Newtonian system of n-particles moving in an arbitrary shared potential $V(\mathbf{x}_1, ..., \mathbf{x}_n)$. The equations of motion are then

$$m_i \ddot{\mathbf{x}}_i = -\frac{\partial V(\mathbf{x}_1, ..., \mathbf{x}_n)}{\partial \mathbf{x}_i} \tag{2.6}$$

for $i = 1, .., n$, where $\mathbf{x}_i \in \mathbb{R}^3$ indicates the position of i-th particle. By a family of Newtonian continuous symmetries of such a system, we mean a family of symmetries $T_\alpha : \mathbb{R}^3 \to \mathbb{R}^3$ indexed by $\alpha \in \mathbb{R}$ such that for all $\mathbf{x}_i \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$ the following properties are satisfied

1. $T_0 = \mathrm{id}_{\mathbb{R}^3}$
2. $V(T_\alpha(\mathbf{x}_1), ..., T_\alpha(\mathbf{x}_n)) = V(\mathbf{x}_1, ..., \mathbf{x}_n)$.

In this section, we will abbreviate Newtonian continuous symmetries to symmetries. We will also say that a system is invariant under some transformations if those transformations are symmetries of this system. Equipped with this definition let us examine some examples of the connection between symmetries and conservation laws.

Translation symmetry Fix $\mathbf{v} \in \mathbb{R}^3$ and suppose $T_\alpha(\mathbf{x}_i) = \mathbf{x}_i + \alpha\mathbf{v}$ is a symmetry. It follows that

$$V(\mathbf{x}_1, ..., \mathbf{x}_n) = V(\mathbf{x}_1 + \alpha\mathbf{v}, ..., \mathbf{x}_n + \alpha\mathbf{v})$$

for all $\mathbf{x}_i \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$. Taking the partial derivative with respect to α of both sides and evaluating at 0 gives

$$\frac{\partial}{\partial \alpha}\bigg|_{\alpha=0} V(\mathbf{x}_1 + \alpha\mathbf{v}, ..., \mathbf{x}_n + \alpha\mathbf{v}) = \frac{\partial}{\partial \alpha}\bigg|_{\alpha=0} V(\mathbf{x}_1, ..., \mathbf{x}_n) = 0$$

and so

$$0 = \frac{\partial}{\partial \alpha}\bigg|_{\alpha=0} V(\mathbf{x}_1 + \alpha\mathbf{v}, ..., \mathbf{x}_n + \alpha\mathbf{v}) = \sum_{i=1}^{n} \frac{\partial V(\mathbf{x}_1, ..., \mathbf{x}_n)}{\partial \mathbf{x}_i} \cdot \frac{\partial(\mathbf{x}_i + \alpha\mathbf{v})}{\partial \alpha}\bigg|_{s=0} = \mathbf{v} \cdot \sum_{i=1}^{n} \frac{\partial V(\mathbf{x}_1, ..., \mathbf{x}_n)}{\partial \mathbf{x}_i}.$$

The above computation shows that the component of the total momentum along $\mathbf{v}$ is conserved since

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{v} \cdot \mathbf{P}) = \mathbf{v} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i}^{n} \mathbf{p}_i = \mathbf{v} \cdot \sum_{i}^{n} \dot{\mathbf{p}}_i = -\mathbf{v} \cdot \sum_{i}^{n} \frac{\partial V(\mathbf{x}_1, ..., \mathbf{x}_n)}{\partial \mathbf{x}_i} = 0.$$

Note that if a system has translation symmetry along three linearly independent axes then the total momentum $\mathbf{P}$ is conserved.

Rotational symmetry For the second example, suppose that $T_\theta(\mathbf{x}_i) = R(\theta)\mathbf{x}_i$ is a symmetry where

$$R(\theta) := \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is rotation around z-axis by angle θ. We follow the same strategy as before. First we observe that

$$V(\mathbf{x}_1, ..., \mathbf{x}_n) = V(R(\theta)\mathbf{x}_1, ..., R(\theta)\mathbf{x}_n)$$

implies that

$$\frac{\partial}{\partial\theta}\bigg|_{\theta=0} V(R(\theta)\mathbf{x}_1, ..., R(\theta)\mathbf{x}_n) = \frac{\partial}{\partial\theta}\bigg|_{\theta=0} V(\mathbf{x}_1, ..., \mathbf{x}_n) = 0.$$

Applying the chain rule to compute the partial derivative then gives

$$0 = \frac{\partial}{\partial\theta}\bigg|_{\theta=0} V(R(\theta)\mathbf{x}_1, ..., R(\theta)\mathbf{x}_n) = \sum_{i=1}^{n} \frac{\partial V(\mathbf{x}_1, ..., \mathbf{x}_n)}{\partial\mathbf{x}_i} \cdot \frac{\partial R(\theta)\mathbf{x}_i}{\partial\theta}\bigg|_{\theta=0}. \tag{2.7}$$

The coordinate representation of $\frac{\partial R(\theta)\mathbf{x}_i}{\partial\theta}\big|_{\theta=0}$ is

$$\frac{\partial R(\theta)\mathbf{x}_i}{\partial\theta}\bigg|_{\theta=0} = \frac{\partial}{\partial\theta}\bigg|_{\theta=0} \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \end{bmatrix} = \frac{\partial}{\partial\theta}\bigg|_{\theta=0} \begin{bmatrix} \cos(\theta)x_{i1} - \sin(\theta)x_{i2} \\ \cos(\theta)x_{i2} + \sin(\theta)x_{i1} \\ 0 \end{bmatrix} = \begin{bmatrix} -x_{i2} \\ x_{i1} \\ 0 \end{bmatrix}.$$

Applying this to equation (2.7) yields

$$\sum_{i=1}^{n} \left(x_{i1} \frac{\partial V(x_1, ..., x_n, t)}{\partial x_{i2}} - x_{i2} \frac{\partial V(x_1, ..., x_n, t)}{\partial x_{i1}} \right) = 0.$$

This equation implies that the z component of the total angular momentum is conserved as

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{L} \cdot \hat{\mathbf{z}}) = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{n} m_i(x_{i1}\dot{x}_{i2} - x_{i2}\dot{x}_{i1}) = \sum_{i=1}^{n} m_i(x_{i1}\ddot{x}_{i2} - x_{i2}\ddot{x}_{i1}) = -\sum_{i=1}^{n} \left(x_{i1}\frac{\partial V}{\partial x_{i2}} - x_{i2}\frac{\partial V}{\partial x_{i1}} \right) = 0.$$

For a rotation around an arbitrary axis $\hat{\mathbf{n}}$ we can first orient our coordinate system so that $\hat{\mathbf{n}}$ points in the positive $\mathbf{z}$-axis and then use the above argument to conclude that $\mathbf{L} \cdot \hat{\mathbf{n}}$ is conserved. Similarly to momentum, if a system is rotationally invariant for three linearly independent rotation axes we can conclude that the total angular momentum $\mathbf{L}$ is conserved.

After these general considerations let us apply them to our initial problem of finding the trajectories of two particles in a shared potential $V(\|\mathbf{x}_1 - \mathbf{x}_2\|)$. In this case, Newton's equations are invariant under both translation by a constant vector and rotations by a fixed angle. Thus, these transformations are symmetries in the general sense. Moreover, the separation vector $\mathbf{x}_1 - \mathbf{x}_2$ is invariant under all space translations, therefore our system enjoys conservation of total momentum. Finally, because V only depends on the length of the separation vector it is invariant under rotations about any axis. Therefore the total angular momentum of our two-body system is conserved. More generally, the same argument shows that in a n-particle system whose shared potential only depends on the separation of particles the total momentum and total angular momentum are conserved.

In this section, we investigated a particular class of symmetries — those that preserve the potential energy — and their connection with conservation laws. The conservation of the LRL vector for the Kepler potential suggests that this system has more symmetry than a generic central potential. However, it is not a symmetry of the kind we investigated in this section. For this reason, this symmetry of the Kepler problem is often referred to as a hidden symmetry. In chapter 5, we will show a particularly elegant approach that brings it to light.

Chapter 3

Exploring the Distinctive Attributes of the Kepler Potential

This chapter will display some of the features that make the Kepler potential unique. Namely, we will discuss Bertrand's, Bohlin's and Henon's Theorems. This will motivate our study in the subsequent chapter that shows that the symmetry group of the Kepler potential is larger than that of an arbitrary central potential which will explain the existence of an additional conserved quantity — the LRL vector. This larger symmetry group also explain the degeneracy of energy levels of hydrogen which we also hint at.

3.1 Bertrand's and Henon's Theorems

It was already Newton who showed in his seminal *Principia* [New87] that there are two central power-law potentials with the property that all their bounded orbits are closed. These are the familiar $1/r$ Kepler potential and the r^2 harmonic oscillator potential. It was not until 1873 that Joseph Bertrand proved what we now call Bertrand's Theorem [Ber73] which states that these are the only central power-law potentials with this property. The strategy used by Bertrand to prove this statement was first to note that any central potential admits a circular orbit. Namely, when the centripetal force exactly matches the central force. He then proceeded by analyzing the perturbations of these solutions and derived the necessary and sufficient conditions for the perturbed solutions to also close. This is the main idea of the proof. A rigorous proof does not bring much more physical insight so we refer the interested reader to see [LMS23] for more details. The significance of Bertrand's Theorem is that it implies that the Kepler potential is unique in that it is the simplest reasonable gravitational potential. It is the simplest because it is a power-law potential. It is the only reasonable among those because Bertrand's Theorem gives us that only it and the r^2 potential have all their bounded orbits close, but the r^2 potential does not decay at infinity leaving us with $1/r$ Kepler potential.

Another consequence of Bertrand's Theorem is that for Kepler and harmonic potentials, the perihelion of orbits does not precess. There are results that quantify the amount of precession for arbitrary power-law central potentials. For example, a special case of Henon's Theorem [Hén77] — proved only in 1977 — relates the change in the perihelion angle, the change in the period at constant energy to angular momentum, energy and the exponent of the potential. While we will not concern ourselves with the details of this theorem — see [DJ22] for an exposition — we want to observe that Henon's Theorem is a relatively recent result which makes Bertrand's Theorem more quantitative. Both Bertrand's and Henon's Theorem hint at the fact that there is some connection between Kepler and harmonic potentials. In the next section, we will see that in a sense the harmonic oscillator orbits are square roots of Kepler orbits.

3.2 Bohlin Transformations and Dual Forces

It is a great surprise that the motion of a body in a gravitational field is closely related to the motion of a bob on a spring. It was Newton who first noticed the duality between those two forces. He also observed that similar relationships hold for other pairs of power-law forces. The duality between Kepler and harmonic potential was later rediscovered by Bohlin [Boh11], while Arnold in [Arn90] generalized the proof to other power-laws which then were called dual forces. Our exposition of these ideas is based on [HJ00; Sag12].

We have shown in chapter 2 that for any central potential, the angular momentum is conserved. Consequently, we can restrict the motion to a plane. The "trick" in the aforementioned theorems is to interpret planar motion as happening in the complex plane. Let us see this trick in action by first displaying (and defining) the duality between the harmonic oscillator and Kepler potentials — the generalization will then follow readily. To avoid case distinction in the theorem, we use the convention that in the collision orbits in the Kepler problem the moving body bounces off the singularity at the origin. Thus the collision trajectory is a half-line. With this convention, we have the following theorem:

Theorem. (Bohlin Theorem) Suppose $w : I \to \mathbb{C}$ follows Hooke's law, i.e.

$$\frac{\mathrm{d}^2 w}{\mathrm{d}t^2} = -\frac{k}{m} w.$$

Then the transformed trajectory $z(\tau) = w(t(\tau))^2$ follows the inverse square law

$$\frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} = -\frac{4E_w}{m}\frac{z}{|z|^3},$$

with $E_w = \frac{m}{2}|w'(0)|^2 + \frac{k}{2}|w(0)|^2$ and the parameters t, τ being implicitly related by

$$\frac{\mathrm{d}\tau(t)}{\mathrm{d}t} = |w(t)|^2.$$

Before diving into a proof let us give the motivation behind the specified parametrization of $w(t)^2$. This reparametrization is necessary because while merely squaring an orbit of a harmonic oscillator does give a trajectory of the right shape, we also need to ensure that angular momentum is conserved. Since the motion is planar we only need to look at the z-component of L. Writing a complex number $z \in \mathbb{C}$ as $z = x + iy$, we have that $x = (z + \bar{z})/2$ and $y = (z - \bar{z})/2i$. This allows us to write

$$L_z = m(x\dot{y} - y\dot{x}) = \frac{m}{4i}((z+\bar{z})(\dot{z} - \dot{\bar{z}}) - (z - \bar{z})(\dot{z} + \dot{\bar{z}})) = \frac{m}{2i}\left(\dot{z}\bar{z} - z\dot{\bar{z}}\right).$$

Therefore imposing the conservation of L_z for both $w(t)$ and $z(\tau)$, and applying the chain rule gives us

$$\text{const.} = \frac{\left(\frac{\mathrm{d}w}{\mathrm{d}t}\bar{w} - w\frac{\mathrm{d}\bar{w}}{\mathrm{d}t}\right)}{\left(\frac{\mathrm{d}z}{\mathrm{d}\tau}\bar{z} - z\frac{\mathrm{d}\bar{z}}{\mathrm{d}\tau}\right)} = \frac{\left(\frac{\mathrm{d}w}{\mathrm{d}t}\bar{w} - w\frac{\mathrm{d}\bar{w}}{\mathrm{d}t}\right)}{\left(2w\frac{\mathrm{d}w}{\mathrm{d}\tau}\bar{w}^2 - 2w^2\bar{w}\frac{\mathrm{d}\bar{w}}{\mathrm{d}\tau}\right)} = \frac{d\tau}{dt}\frac{1}{2w\bar{w}}\frac{\left(\frac{\mathrm{d}w}{\mathrm{d}t}\bar{w} - w\frac{\mathrm{d}\bar{w}}{\mathrm{d}t}\right)}{\left(\frac{\mathrm{d}w}{\mathrm{d}t}\bar{w} - w\frac{\mathrm{d}\bar{w}}{\mathrm{d}t}\right)}.$$

Therefore $\frac{\mathrm{d}\tau}{\mathrm{d}t} \propto \|w\|^2$. The converse also holds, i.e. $\frac{\mathrm{d}\tau}{\mathrm{d}t} = k|w|^2$ ensures that angular momentum is conserved. In the theorem, $k = 1$ was chosen. Now, the proof of the theorem boils down to applying the chain rule.

Proof.

$$\begin{aligned}
\frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} &= \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}z}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) \\
&= \frac{1}{w\bar{w}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}w^2}{\mathrm{d}t}\frac{1}{|w(t(\tau))|^2}\right) \\
&= \frac{1}{w\bar{w}}\frac{\mathrm{d}}{\mathrm{d}t}\left(2w(t(\tau))\frac{\mathrm{d}w(t(\tau))}{\mathrm{d}t}\frac{1}{w(t(\tau))\bar{w}(t(\tau))}\right) \\
&= \frac{2}{w\bar{w}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\frac{\mathrm{d}w(t(\tau))}{\mathrm{d}t}}{\bar{w}(t(\tau))}\right) \\
&= \frac{2}{w\bar{w}^2}\frac{\mathrm{d}^2 w(t(\tau))}{\mathrm{d}t^2} - \frac{2}{w\bar{w}^3}\frac{\mathrm{d}w(t(\tau))}{\mathrm{d}t}\frac{\mathrm{d}\bar{w}(t(\tau))}{\mathrm{d}t} \\
&= -\frac{2k}{m\bar{w}^2} - \frac{2}{w\bar{w}^3}\frac{\mathrm{d}w(t(\tau))}{\mathrm{d}t}\frac{\mathrm{d}\bar{w}(t(\tau))}{\mathrm{d}t} \\
&= -\frac{4}{m|w|^2}\left(\frac{m}{2\bar{w}^2}\frac{\mathrm{d}w}{\mathrm{d}t}\frac{\overline{\mathrm{d}w}}{\mathrm{d}t} + \frac{k}{2}\frac{w}{\bar{w}}\right) \\
&= -\frac{4w^2}{m|w|^6}\left(\frac{m}{2}\left|\frac{\mathrm{d}w}{\mathrm{d}t}\right|^2 + \frac{k}{2}|w|^2\right).
\end{aligned}$$

The term in parentheses is just the energy of the harmonic oscillator and hence is a constant and equal to its initial value. Therefore we conclude that

$$\frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} = -\frac{4E_w}{m}\frac{z}{|z|^3}.$$

□

Moreover, it can be shown that with our convention for singular orbits, all solutions to Kepler problem can be reached from harmonic oscillator solutions in this way. So there is a one-to-one correspondence of solutions to these problems. Now, the Kepler problem is nonlinear so it is much harder to solve than the harmonic oscillator. Thus one can use this theorem together with the properties of $w \mapsto w^2$ map to give a simple proof that the orbits of the Kepler problem are conic sections. Details of these computations can be found in [Sag12; HJ00].

We have shown in chapter 2 that the Kepler problem has an additional conserved quantity — the LRL vector. The duality between Kepler and the harmonic oscillator suggests there exists a dual conserved quantity for the harmonic oscillator. This is indeed true, the classical harmonic oscillator has an additional conserved quantity known as the Fradkin-Jauch-Hill (FJH) tensor. One can derive the LRL vector by starting from FJH tensor and transforming it using the $w \mapsto w^2$ map — for details see [Sag12]. Finally, as we already mentioned, Bohlin's theorem holds for more general pairs of dual forces.

Theorem. (Arnold) Suppose $w : I \to \mathbb{C}$ solves

$$\frac{\mathrm{d}^2 w}{\mathrm{d}t^2} = -C\frac{w}{|w|^{1-a}}.$$

Then the transformed trajectory $z(\tau) = w(t(\tau))^\beta$ solves

$$\frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} = -\bar{C}\frac{z}{|z|^{1-A}},$$

where a, A and β satisfy

$$(a+3)(A+3) = 4 \quad \text{and} \quad \beta = \frac{a+3}{2},$$

while the parameters t, τ being implicitly related by

$$\frac{\mathrm{d}\tau(t)}{\mathrm{d}t} = |w(t)|^{2(\beta-1)}.$$

The strategy of the proof is the same as in the particular case. In the next section, we will collect all the threads and argue about the importance and depth of the Kepler problem.

3.3 The Significance of the Kepler Potential: A Summary

In chapter 2, we have shown that the Kepler potential has an additional conserved quantity that allows us to quickly derive the shape of orbits. We followed this by giving a brief overview of the relation between conserved quantities and symmetries. This relation suggests that the conservation of the LRL vector is a consequence of an additional symmetry of Kepler potential[1]. In chapter 5 we will show that the Kepler potential does indeed have additional symmetries. In this chapter, we have shown that apart from the conservation of the LRL vector, the Kepler potential has other remarkable properties. Bertrand's and Henon's Theorems show that it is the only reasonable power-law gravitational potential. These theorems also show that it is closely related to the harmonic oscillator potential. This connection is explained by Bohlin's Theorem.

All these results show that there is something deep and special about the Kepler potential and its close cousin harmonic oscillator potential. To support this statement, let us quote a remark from [Kot11]:

"The two major theories of theoretical physics, general relativity and quantum field theory, are based, respectively, on geometrization of the $1/r^2$ gravity law (which appears in the weak field limit of GR) and quantization of a collection of harmonic oscillators described by Hooke's law."

This insightful observation gives us further motivation for studying the symmetries and different perturbations of the Kepler problem. We will concern ourselves with this in the subsequent chapters.

[1] In fact, the same manifests itself in quantum mechanics. Namely, after resolving some ordering issues one can promote the LRL vector to an operator and show that it commutes with Hamiltonian. One can use this fact to explain the "accidental" additional degeneracy of the hydrogen atom spectrum — for details see [Jon98]

Chapter 4

Mathematical Preliminaries

In this chapter, we give a brief overview — mostly without proofs — of the mathematical tools used in the subsequent chapters. In the first section, we introduce manifolds and how to do calculus on them. The main references for this section are [Ser23a; Tu10]. However, for our purpose, a mere manifold is too little. We need to add additional structure to it. In the second section, we will introduce metrics and connections. Metrics allow us to talk about lengths and angles of vectors, while a connection is needed to define geodesics. These two independent concepts interplay nicely to give rise to a unique connection on a manifold with a metric — the Riemannian (or Levi-Civita) connection. This section is based on [Tu17]. The subsequent section introduces a selection of tools from symplectic geometry which is a study of manifolds with a special two-form. Metrics or connections are not needed in this setting. For more details, the reader is referred to [Sil01; Ser23b]. In the final section, we will show that symplectic geometry is a particularly elegant framework for doing classical mechanics, as shown in [Ser23b].

Let us indicate what all these tools are going to be used for. In the next chapter, we will show that the non-constant geodesics on $\mathbb{S}^n$ equipped with the Riemannian connection are in one-to-one correspondence with solutions of negative energy to the Kepler problem on a particularly simple symplectic manifold $T^* \mathbb{R}^n$. By doing so we will show that the symmetry group of negative energy n-dimensional Kepler orbits is the group of rotations in $n+1$ dimensions.

4.1 Analysis on Manifolds

In this section, we will introduce the concept and properties of manifolds, tangent vectors, tangent spaces, vector fields and differential forms. We will provide the reader with rigorous definitions of those. However, we will present them in an informal manner that, hopefully, highlights the intuition behind these objects. For more details (and proofs) the reader is referred to [Ser23a; Tu10].

4.1.1 Smooth Manifolds and Tangent Vectors

Let us start with the notion of a topological manifold which is a special kind of topological space. Topological manifolds allow us to restrict the untamed category of topological space to something more well-behaved which has some of the nice properties that the space which surrounds us has — or at least, the properties that we assume the space around us has. The first of those properties is Hausdorffness. The formal definition is that a topological space X is **Hausdorff** if for any distinct points $p, q \in X$ there exists open neighbourhoods of p and q which are disjoint. What this intuitively means is that no matter how closely you zoom in, you can always distinguish between two points of your space[1]. The second property is that topological manifolds look just like Euclidean space if you zoom in close enough — thus the

[1] This conforms to our classical intuition. However, quantum-mechanically this claim is rather dubious.

Earth is a topological manifold because, from a human perspective, it looks flat. More formally, a topological space X is **locally Euclidean** of dimension n if for any point $p \in X$ there exists i) an open neighbourhood U of p and ii) a homeomorphism $\phi : U \to V$ where $V \subset \mathbb{R}^n$ is open. The pair (U, ϕ) is called a **chart** about p. The final property that is (usually) required of topological manifolds is second-countability. A topological space X is called **second-countable** if it admits a countable basis for its topology. This is a rather technical condition which guarantees the existence of a partition of unity which is a tool that allows us to piece together locally defined objects into a global one. Thus we define topological manifolds as follows.

Definition 4.1. A **topological manifold** of dimension n is a Hausdorff, second-countable topological space which is locally Euclidean of dimension n.

Since topological manifolds look locally like Euclidean spaces, one might hope that this already allows us to define differentiability on manifolds. However, to make this definition consistent we must postulate further requirements. This results in the notion of a smooth manifold.

Definition 4.2. A **smooth manifold** is a topological manifold equipped with a maximal atlas.

Where an **atlas** for a locally Euclidean topological space X is a compatible collection of charts $\{(U_\alpha, \phi_\alpha)\}$ such that $\{U_\alpha\}$ cover X. An atlas is called **maximal** if it contains all charts that are compatible with all its charts. While a collection of charts $\{(U_\alpha, \phi_\alpha)\}$ on a locally Euclidean topological space X is **compatible** if for any α and β

$$\phi_\alpha \circ \phi_\beta^{-1} : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$$

is smooth as a function of Euclidean spaces. Intuitively, what this all means is that a smooth manifold is a space with a physical atlas full of charts (maps). This atlas has the property that if you draw a smooth path on one page of the atlas and this page overlaps with another page, then your path will also be smooth when drawn on the other page.

Example. The prototype of a smooth manifold is $\mathbb{R}^n$ equipped with a maximal atlas[2] containing the chart $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$.

In what follows, we will only concern ourselves with smooth manifolds and call them simply manifolds. Moreover, spaces denoted by M, N, P or S are always going to be manifolds. With this let us formalize the notion of smoothness.

Definition 4.3. Let M and N be manifolds and $f : M \to \mathbb{R}^n, g : \mathbb{R}^m \to N, h : M \to N$ functions. Then we say that

1. f is **smooth** if $f \circ \phi^{-1}$ is smooth for any chart ϕ on M
2. g is **smooth** if $\phi \circ g$ is smooth for any chart ψ on N
3. h is **smooth** if $\psi \circ g \circ \phi^{-1}$ is smooth for any charts ψ, ϕ on N and M, respectively.
4. h is a **diffeomorphism** if h is smooth and it has a smooth inverse.

The space of all smooth functions $f : M \to \mathbb{R}$ on a manifold M is denoted by $C^\infty(M)$.

This definition has all the nice properties that one would expect of smooth maps, such as that smooth maps are closed under composition. Now, the above definition enables us to talk about smooth curves $c : I = (a, b) \to M$ on a smooth manifold. However, it is not immediately clear how a tangent vector to a curve could be generalized to this setting. The key insight comes

[2] In fact, one can show that this maximal atlas is unique.

from closely examining the Euclidean case. Namely, if $\gamma : (a, b) \to \mathbb{R}^n$ is a smooth curve and $f : \mathbb{R}^n \to \mathbb{R}$ is smooth real-valued function, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(c(t)) = \gamma'(t) \cdot (\nabla f) = (\gamma'(t) \cdot \nabla) f.$$

Thus the tangent vector $\gamma'(t)$ can be thought of as a directional derivative acting on functions. Using this analogy we can define tangent vectors on smooth manifolds.

Definition 4.4. For any smooth curve $c : (a, b) \to M$ we define its **tangent vector** $c'(t)$ at $c(t) \in M$ to be a derivative operator defined by

$$c'(t) f = \left.\frac{\mathrm{d}(f \circ c)}{\mathrm{d}t}\right|_t$$

for any smooth function $f : M \to \mathbb{R}$. We denote by T_pM the vector space of all tangent vectors at $p \in M$.

One might worry that the space T_pM of all directional derivatives at a point p of an arbitrary manifold is intractable. Fortunately, we have the following result.

Theorem 4.1. Let M be an n-dimensional manifold and $\phi : U \to V$ a chart around $p \in M$. For $i = 1, ..., n$, define the tangent vectors $\partial_i|_p$ at p by

$$\partial_i|_p f := \left.\frac{\partial(f \circ \phi^{-1})}{\partial x^i}\right|_{\phi(p)}.$$

Then $\{\partial_i|_p\}$ forms a basis for T_pM. Consequently, the dimension of T_pM coincides with the dimension of M. Denoting the components of ϕ by $(x^1, ..., x^n)$, we will sometimes use $\frac{\partial}{\partial x^i} := \partial_i$.

Having defined tangent vectors, let us proceed and define the derivative of a map $F : M \to N$. Recall, from multivariable analysis that the (total) derivative of a map $f : \mathbb{R}^m \to \mathbb{R}^n$ is represented by its Jacobian matrix. Thus at each point $x \in \mathbb{R}^m$, the derivative of f is just a matrix, i.e. a linear map on vectors. This suggests to define the derivative of $F : M \to N$ at a point $p \in M$ to be some linear map on tangent vectors to curves through p. In other words, given a tangent vector $c'(t)$ on M we want to use F to define a tangent vector on N. Since tangent vectors on N originate from curves on N, a natural choice is to consider the curve $F \circ c$ and its tangent vectors. These considerations give a clear motivation for the following definition.

Definition 4.5. The **differential** dF_p of a smooth map $F : M \to N$ at $p \in M$ is the linear map

$$dF_p : T_pM \to T_{F(p)}N,$$

defined by

$$(dF_pX_p)(h) := (F \circ c)'(0)h = X_p(h \circ F)$$

for $h \in C^\infty(N)$ and $c'(0) = X_p \in T_pM$ with $c(t) : I = (a, b) \to M$ a curve.

Smooth manifolds give rise to such a rigid structure that this somewhat unusual but natural definition still enjoys the usual properties of derivatives such as the chain rule. Moreover, when we consider $\mathbb{R}^n, \mathbb{R}^m$ as manifolds our definition coincides with the familiar Jacobian matrix.

Theorem 4.2. The differential satisfies the **chain rule**, i.e.

$$d(G \circ F)_p = dG_{F(p)} \circ dF_p,$$

where $F : M \to N, G : N \to P$ are smooth.

Theorem 4.3. By making the canonical identification of $T_p\mathbb{R}^n$ with $\mathbb{R}^n$, the coordinate representation of the differential of a smooth map $F : \mathbb{R}^m \to \mathbb{R}^n$ coincides with its Jacobian matrix.

4.1.2 Submanifolds

The goal of this section is to develop theory that will allow us to quickly identify subsets P of $\mathbb{R}^n$ as manifolds and determine the smoothness of maps $f : P \to \mathbb{R}^n, g : \mathbb{R}^n \to P$ by looking at the corresponding extensions $\hat{f} : \mathbb{R}^p \to \mathbb{R}^n, \hat{g} : \mathbb{R}^n \to \mathbb{R}^p$. The notion that we are looking for is that of a regular submanifold.

Definition 4.6. Let N be an n-dimensional manifold and $P \subset M$. We say that P is a **regular submanifold** of dimension k (or **codimension** $n-k$) if for every $p \in P$ there exists a chart (U, ϕ) on M around p such that

$$U \cap P = \phi^{-1}(\{x \in \phi(U) \mid x^{k+1} = ... = x^n = 0\}).$$

Such a chart is called an **adapted chart** relative to P and we denote by $\phi_P := \pi_k \circ \phi|_P$ the restriction of ϕ to $P \cap U$ composed with the Euclidean projection to the first k coordinates.

Theorem 4.4. Equipping a regular submanifold P of dimension k with the subspace topology and atlas $\{(U \cap P, \phi_P)\}$ makes it into a manifold of dimension k.

Note that our definition concerns subsets of a general manifold M — not necessarily $\mathbb{R}^n$. Informally, a regular submanifold is just a manifold that sits in a larger manifold and has a nice position there so that its manifold structure is induced by the ambient manifold. These manifolds sitting inside larger manifolds have the properties that we sought.

Theorem 4.5. Let $f : M \to N$ be smooth and P be a regular submanifold of M. Then the restriction $f|_P : P \to N$ is smooth.

Theorem 4.6. Let $f : M \to N$ be smooth and suppose $f(M) \subset S$. If S is a regular submanifold of N then the induced map $\tilde{f} : M \to S$ is smooth.

In particular, if $M = \mathbb{R}^m, N = \mathbb{R}^n$ the smoothness of f is easily determined and so the above theorems give us a shortcut for checking the smoothness of $\tilde{f}$ and f_P. To complete the goal that we set for this section, we only need a way of showing that a subset of $\mathbb{R}^n$ is its regular submanifold. This is achieved with the following definition and the regular level set theorem.

Definition 4.7. Let $F : M \to N$ be smooth. We call $c \in N$ a **regular value** of F if $c \notin F(M)$ or for all $p \in F^{-1}(c)$ the differential dF_p is surjective. If $c \in N$ is a regular value we call $F^{-1}(c)$ a **regular level set**.

Theorem 4.7. (Regular level set theorem) Let $F : M \to N$ be a smooth map and $\dim M = m, \dim N = n$. Then a nonempty regular level set $P = g^{-1}(c)$ is a regular submanifold of M of codimension n.

With these general tools, we can specialize to $M = \mathbb{R}^n, N = \mathbb{R}$ and get the following useful corollary.

Corollary 4.8. Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is smooth and for any $p \in g^{-1}(c)$ there exists $i \in \{1, ..., n\}$ such that $\frac{\partial g}{\partial x^i}\big|_p \neq 0$. Then $P = g^{-1}(c)$ is a regular submanifold of $\mathbb{R}^n$ of codimension 1.

Proof. This follows from the regular level set theorem by recalling that the differential of a map between Euclidean spaces is just its Jacobian matrix, see Theorem 4.3. □

Example. Let us use the above corollary to show that the n-dimensional unit sphere $\mathbb{S}^n$ is a regular submanifold of $\mathbb{R}^{n+1}$. To this end, define $g : \mathbb{R}^{n+1} \to \mathbb{R}$ by $g(x) = \|x\|^2 - 1$ and observe that $\mathbb{S}^n = g^{-1}(0)$. Thus for any $p \in \mathbb{S}^n$ we have that $g(p) = 0$. Moreover since

$$\left.\frac{\partial g}{\partial x_i}\right|_p = 2p_i$$

there must exists $i \in \{1, ..., n\}$ such that $\frac{\partial g}{\partial x^i}\big|_p \neq 0$, otherwise $p = 0$ and so $p \notin \mathbb{S}^n$. Thus, our corollary is applicable and hence we conclude that $\mathbb{S}^n$ is a regular submanifold of $\mathbb{R}^{n+1}$. We will use this fact in chapter 5.

4.1.3 Vector Fields

We already saw that to each point p of a manifold M we can attach the vector space T_pM of tangent vectors. Thus a vector field on a manifold should be a map that assigns to each point of a manifold a single tangent vector in the corresponding tangent space. To formalize this idea we introduce the tangent bundle of a manifold.

Definition 4.8. Let M be a manifold, the set

$$TM := \bigsqcup_{p \in M} T_pM$$

together with the map $\pi : TM \to M$ defined by $\pi(X_p) = p$ for $X_p \in T_pM$ is called the **tangent bundle** of M. The smooth structure on M naturally makes TM into a manifold so that π is smooth.

Intuitively, TM is just the collection of all tangent spaces of a manifold glued in such a way that the origins of those vector spaces form our original manifold M. Using TM the notion of a vector field is defined as follows.

Definition 4.9. A **vector field** X on M is a smooth map $X : M \to TM$ such that $\pi \circ X = \mathrm{id}_M$. The space of all vector fields on a manifold is denoted by $\mathfrak{X}(M)$ and for $X \in \mathfrak{X}(M)$ the vector at a point p is denoted by $X_p := X(p)$.

Now, recall that tangent vectors act on functions as directional derivatives. Thus for $X_p \in T_pM$ and $f : M \to \mathbb{R}$ a smooth function, X_pf is a number. This suggest that we can act with a vector field X on a function f to produce another function Xf defined by

$$(Xf)(p) = X_pf.$$

In fact, we can uniquely specify a vector field by knowing how it acts on all functions. This property allows us to define a product of two vector fields.

Definition 4.10. A **Lie bracket** of vector fields $X, Y \in \mathfrak{X}(M)$ is the vector field $[X, Y]$ defined pointwise by

$$[X, Y]_pf := X_p(Yf) - Y_p(Xf)$$

for $f \in C^\infty(M)$.

One can check that so-defined $[X, Y]_p$ is indeed a tangent vector. Sometimes we will also look at vector fields that are not defined over the whole manifold. Formally, these are defined as follows.

Definition 4.11. A **local vector field** V of TM over U is defined as a map $V : U \to TM$ such that $\pi \circ V = \mathrm{id}_U$, where $U \subset M$ is open.

Example. The primary example of a local vector field is $\partial_i : U \to TM$ defined by

$$\partial_i(p) := \partial_i|_p,$$

where U is the domain of a chart.

We see from this the above example, that $\partial_1, ..., \partial_n$ can be used to express vector fields in local coordinates. This phenomenon is captured in the next definition.

Definition 4.12. A **local frame** over U for the tangent bundle $\pi : TM \to M$ is a set of local vector fields $s_1, ..., s_m$ over U such that for all $p \in U$, the set $\{s_1(p), ..., s_k(p)\}$ is a basis for T_pM. A global frame is a local frame with $U = M$.

Example. Thus on a chart domain U a vector field X can be expressed as

$$X_p = \sum_{i=1}^{n} X^i(p)\partial_i|_p$$

for some smooth functions X^i.

Now, let us go back to the definition of a tangent vector. If we treat t as an independent variable, what kind of object is $c'(t)$? It is not a local vector field since in general $c([a,b])$ is not open in M. This motivates us to give the last definition regarding vector fields.

Definition 4.13. Let $c : (a,b) \to M$ be a curve. A **vector field along** $c(t)$ is a map

$$V : (a,b) \to \bigsqcup_{t\in(a,b)} T_{c(t)}M$$

such that $V(t) \in T_{c(t)}M$ and for any $f \in C^\infty M$ the function $V(t)f : (a,b) \to \mathbb{R}$ is smooth. The space of vector fields along a curve $c(t)$ is denoted by $\Gamma(TM|_{c(t)})$.

With this definition, it is clear that $c'(t)$ is a vector field along $c(t)$ since $c'(t)f = \frac{\mathrm{d}}{\mathrm{d}t}\big|_t f(c(t))$ is smooth as a derivative of the composition of smooth functions.

4.1.4 Differential Forms

In this section, we introduce the dual objects to tangent vectors, i.e. one-forms. A map that assigns to each point of a manifold a one-form will be called a differential one-form (often abbreviated to just a one-form). Despite the close relation between one-forms and vector fields, it turns out that the former have a much richer algebraic structure. We will see a glimpse of this. Unfortunately, the rich algebraic properties of differential forms come at the price of an obscured geometric intuition. As a result, this section will be more abstract than the former more geometric ones.

First, recall from linear algebra that the dual vector space V^* of the vector space V is the vector space of all the linear maps $f : V \to \mathbb{R}$. It has the same dimension as V and given a basis $e_1, ..., e_n$ for V the linear maps $f^1, ..., f^n$ defined by $f^j(e_i) = \delta_i^j$ form the dual basis for V^*. Carrying out this construction with $V = T_pM$ gives the following.

Definition 4.14. The **cotangent space** at a point $p \in M$ is defined to be $T_p^*M := (T_pM)^*$, its elements are called one-forms at p. The basis for T_p^*M dual to $\partial_1|_p, ..., \partial_m|_p$ is denoted by $dx_p^1, ..., dx_p^n$.

Thus a one-form at p is a linear map $\omega_p : T_pM \to \mathbb{R}$. A k-form is defined as an alternating linear map which takes k vectors.

Definition 4.15. For $k \geq 1$, a k-linear function

$$\omega_p : \underbrace{T_pM \times ... \times T_pM}_{k\text{-times}} \to \mathbb{R}$$

is called a k**-form** on M at p if for any vectors $v_1, ..., v_k \in T_pM$ and permutation of k elements $\pi \in S_k$ it holds that

$$\omega_p(v_{\pi(1)}, ..., v_{\pi(k)}) = \mathrm{sgn}(\pi)\omega_p(v_1, ..., v_k),$$

i.e. ω_p is **alternating**. We denote the vector space of all k-forms at p by $\Lambda^k(T_pM)$ and define $\Lambda^0(T_pM) := \mathbb{R}$.

Observe that $T_p^*M = \Lambda^1 T_pM$ since any linear map is alternating. Now, it is possible to define a wedge product $\wedge$ that takes an ℓ-form and a k-form to a $k+\ell$-form. However, we don't need such a generality so we will only define it for $\ell = k = 1$.

Definition 4.16. Let $\omega, \eta \in T^*M$ and $v_1, v_2 \in T_pM$. We define the wedge product $\omega \wedge \eta \in \Lambda^2(T_pM)$ to be

$$(\omega \wedge \eta)(v_1, v_2) := \omega(v_1)\eta(v_2) - \omega(v_2)\eta(v_1).$$

It is clear that $\omega \wedge \eta = -\eta \wedge \omega$.

In analogy to the tangent bundle, we can piece together all the cotangent spaces and $\Lambda^k(T_pM)$ spaces to obtain the cotangent bundle T^*M and $\Lambda^k M$.

Theorem 4.9. Let M be an n-dimensional manifold and define

$$\Lambda^k M := \bigsqcup_{p \in M} \Lambda^k(T_pM).$$

Then M naturally induces a manifold structure on $\Lambda^k M$, making the projection to the base point $\pi : \Lambda^k M \to M$ smooth. We define $T^*M := \Lambda^1 M$ and call $\pi : T^*M \to M$ the **cotangent bundle**.

The objects analogous to vector fields are differential k-forms. One can also define objects analogous to local vector fields and local frames quite easily as well.

Definition 4.17. A **differential k-form** ω on M is a map $\omega : M \to \Lambda^k M$ such that $\pi \circ \omega = \mathrm{id}_M$. For short we will also say that ω is a k-form. The space of k-forms is denoted by $\Omega^k(M)$. Note that $\Omega^0(M) = C^\infty(M)$, so 0-forms are just functions.

Definition 4.18. A local k-form is k-form defined only on open $U \subset M$. And a local frame for $\Lambda^k M$ over U is a collection of local k-forms $\omega_1, ..., \omega_k$ over U such that for each $p \in U$ the set $\{\omega_1(p), ..., \omega_k(p)\}$ is a basis for $\Lambda^k T_pM$.

Theorem 4.10. Let M be a manifold and (U, ϕ) a chart on it. Define the local sections $\partial_i : U \to TM, dx^i : U \to T^*M$ by $\partial_i(p) = \partial_i|_p$ and $dx^i(p) = dx^i_p$. Then $\{\partial_i\}$ and $\{dx^i\}$ are frames over U for TM and T^*M, respectively.

Example. Thus any one-form ω on M can be expressed on a chart domain U as

$$\omega_p = \sum_i^n a_i(p) dx^i$$

for some smooth functions a_i.

The algebraic richness of the theory of differential forms comes from the algebraic operations that can be defined on them, and the way these operations interact. First, we can extend the wedge product of forms at a point to differential forms by defining it pointwise. Another operation which allows us to increase the degree of a form is the exterior derivative. We will only define it for 0- and 1-forms but the construction can be extended to any k-form, see the references [Tu10; Ser23a] for details.

Definition 4.19. The **exterior derivative** is the map which takes k-forms to $k+1$-forms. For $k = 0$, it is defined locally by

$$df = \sum_{i=1}^n (\partial_i f) dx^i,$$

while for $k = 1$ it is defined by

$$d(f dx^j) := \sum_{i=1}^n (\partial_i f) dx^i \wedge dx^j$$

and extended linearly to arbitrary elements of $\Omega^1(M)$.

Another important operation on k-forms is the pullback by a smooth map. After defining it, we summarize how pullback, exterior derivative and the wedge product interact with each other. We will state these properties in full generality but we will only use them for $0, 1$ or 2-forms.

Definition 4.20. Let $F : N \to M$ be smooth and $\omega \in \Omega^k(M)$. Then the **pullback** $F^* : \Omega^k M \to \Omega^k N$ of ω by F is the k-form on N defined by

$$(F^*\omega)_p(v_1, ..., v_k) := \omega_{F(p)}(dF_p v_1, ..., dF_p v_k)$$

for $v_1, ..., v_k \in T_pN$.

Theorem 4.11. The pullback, the wedge product and the exterior derivative satisfy the following properties

1. Both F^* and d are $\mathbb{R}$-linear
2. $d \circ F^* = F^* \circ d$
3. $F^*(\omega \wedge \eta) = F^*\omega \wedge F^*\eta$
4. $d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge d\eta$
5. $d \circ d = 0$
6. $df(X) = Xf$ for $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$.

Both exterior derivative and wedge product can be used to increase the degree of a form. Given a vector field X, we can define the interior product which allows us to decrease the degree.

Definition 4.21. For $k \geq 1$. The **interior product** of a k-form ω and a vector field X is a $k-1$-form $\iota_X\omega$ defined pointwise by

$$(\iota_X\omega)_p(v_1, ..., v_{k-1}) = \omega(X_p, v_1, ..., v_{k-1}),$$

for $v_1, ..., v_{k-1} \in T_pM$. Note that for $k = 1$, the interior product is simply the natural pairing $\iota_X\omega = \omega(X)$ of vector fields and one forms.

Finally, let us define some important classes of k-forms.

Definition 4.22. Let $\omega \in \Omega^k(M)$ be a k-form. Then

1. ω is **closed** if $d\omega = 0$
2. ω is **exact** if $\omega = d\theta$ for some $\theta \in \Omega^{k-1}(M)$.
3. ω is **non-degenerate** if $\omega_p \neq 0$ for all $p \in M$.

Note that by the preceding theorem, $d^2 = 0$ and so every exact form is closed. The converse is not true in general and the answer to the question of how badly closed forms fail to be exact turns out to be an important topological invariant called de Rham cohomology.

4.2 Differential Geometry

In this section, we will define a Riemannian metric and then an affine connection on a manifold M. The former is necessary to talk about the lengths or angles of tangent vectors, as well as the lengths of curves. On the other hand, a connection allows us to differentiate vector fields along some other vector field, and using a connection we can define geodesics which can be interpreted as "straight" curves on our manifold. We will also see how these two notions interplay.

4.2.1 Riemannian Metric

Informally, a Riemannian metric on a manifold M is a smooth assignment of an inner product to each of the tangent spaces T_pM. Having an inner product, we can talk about lengths and angles of tangent vectors. Thus a Riemannian metric adds a geometric structure to our manifold. Intuitively, it also specifies the "shape" of a manifold. For example, a sphere is diffeomorphic to any potato. Therefore, through the lens of manifolds with no further structure, spheres and potatoes are indistinguishable. However, once we equip a manifold with a metric it fixes its "shape" in the sense that the metric-preserving diffeomorphisms, i.e. isometries, preserve the curvature of a manifold. Therefore, in this context, a sphere is no longer the same as an arbitrary potato — only the special curvature one potatoes! With this motivation, let us give a rigorous definition of a metric.

Definition 4.23. A **Riemannian metric** g on a manifold M is a map

$$g : \bigcup_{p \in M} \{p\} \times T_pM \times T_pM \to \mathbb{R}$$

such that

1. For each $p \in M$, $g_p(-,-)$ is an inner product on T_pM
2. g is smooth in the sense that for all vector fields $X, Y \in \mathfrak{X}(M)$ the map $g(X,Y) : M \to \mathbb{R}$ defined by

$$g(X,Y)(p) := g_p(X_p, Y_p)$$

is smooth.

We call a manifold with a Riemannian metric a **Riemannian manifold**.

Now one might wonder how common are such structures. In other words, can any manifold be equipped with some Riemannian metric? The answer is affirmative as stated in the following theorem.

Theorem 4.12. On every manifold, there exists a Riemannian metric.

We know from the previous sections that on a coordinate chart (U, ϕ) of M, we have a local frame $\partial_1, ..., \partial_m$ for TM over U. Thus on a coordinate chart, a metric g on M can be represented as a matrix of functions $[g_{ij}]$ defined by

$$g_{ij}(p) := g_p(\partial_i|_p, \partial_j|_p).$$

Around every point of a Riemannian manifold one can find a coordinate chart such that the metric matrix $[g_{ij}]$ is diagonal. In such a case, the coordinate representation of g is commonly denoted by

$$ds^2 = \sum_{i=1}^{n} g_{ii}(dx^i)^2.$$

We will encounter this notation when discussing the Schwarzschild solution in chapter 6.

In addition to measuring vector lengths and angles, the metric can be used to induce a canonical isomorphism between T_pM and T_p^*M for each $p \in M$. This isomorphism together with its inverse are called musical isomorphisms.

Definition 4.24. Let M be Riemannian manifold and take $p \in M$. Then the map $\flat : T_pM \to T_p^*M$ defined by

$$\flat(X) := X^\flat := g_p(X, -)$$

is an isomorphism. Its inverse is denoted by $\sharp(\omega) := \omega^\sharp$.

Finally, as advertised, let us define the length of a curve $c : (a, b) \to \mathbb{R}$. The intuition behind this definition is the same as for the usual calculus definition. Namely, the tangent vector to a curve is its best linear approximation, adding up (integrating) the lengths of all these linear pieces gives us the length of the whole curve.

Definition 4.25. Let M be a Riemannian manifold and $c : (a, b) \to M$ be a curve. The **speed** of $c(t)$ at t is defined to be $\sqrt{g_{c(t)}(c'(t), c'(t))}$ while **the length of the curve** c is

$$\ell(c) = \int_a^b \sqrt{g_{c(t)}(c'(t), c'(t))}dt.$$

4.2.2 Connections

An affine connection is a structure that enables us to differentiate vector fields along other vector fields. It directly generalizes the directional derivative on $\mathbb{R}^n$ which is a map $D : \mathfrak{X}(\mathbb{R}^n) \times \mathfrak{X}(\mathbb{R}^n) \to \mathfrak{X}(\mathbb{R}^n)$ defined by

$$D_X(Y) = D_X\left(\sum_{i=1}^n Y^i \partial_i\right) = \sum_{i=1}^n (XY^i)\partial_i.$$

This definition makes sense because ∂_i's form a global canonical frame on $\mathbb{R}^n$. However, on an arbitrary manifold, we do not have such a frame. Hence on the overlap of two "framed"[3] open sets the above definition would be generally inconsistent. Thus to have an operator on a manifold analogous to D, we need to manually add it. This is what an affine connection is.

Definition 4.26. An **affine connection** on a tangent bundle $\pi : TM \to M$ is a map

$$\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$$

such that for $X, Y \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$

1. ∇ is $C^\infty(M)$-linear in the first argument
2. While in the second argument ∇ is only $\mathbb{R}$-linear but it satisfies the Leibniz rule, i.e.

$$\nabla_X(fY) = (Xf)Y + f\nabla_X Y.$$

One can check, the the D operator on $\mathbb{R}^n$ satisfies $D_XY - D_YX = [X, Y]$. An arbitrary affine connection will no longer satisfy this identity and the object in the next definition detects whenever this identity fails.

Definition 4.27. The **torsion** of an affine connection is the map $T : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ defined as

$$T(X, Y) := \nabla_X Y - \nabla_Y X - [X, Y].$$

If $T(X, Y) = 0$ for all $X, Y \in \mathfrak{X}(M)$ the affine connection is said to be **torsion-free**.

[3] Meaning that there is a frame over them

We will now see how an affine connection can be used to define "straight" curves (geodesics) on a manifold. To this end, first recall that curve $\gamma : (a, b) \to \mathbb{R}^n$ is a straight curve (line) if and only if

$$\gamma''(t) = \frac{\mathrm{d}}{\mathrm{d}t}\gamma'(t) = 0.$$

In coordinates, this reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n} \dot{\gamma}^i \partial_i = \sum_{i=1}^{n} \left(\frac{\mathrm{d}}{\mathrm{d}t}\dot{\gamma}^i\right) \partial_i = 0.$$

However, to generalize this definition we run into the same issue as in generalizing the D operator. Namely, on an arbitrary manifold M there is no canonical global frame and so differentiating each component function with respect to t is a coordinate-dependent definition. Fortunately, we don't need another structure to carry out this generalization. An affine connection uniquely defines an operator $\frac{D}{dt}$ along a curve $c(t)$ which generalizes the $\frac{\mathrm{d}}{\mathrm{d}t}$ operator on $\mathbb{R}^n$.

Definition 4.28. Let M be a manifold with an affine connection ∇ and $c : (a, b) \to M$ a curve. The **covariant derivative along** $c(t)$ is a map

$$\frac{D}{dt} : \Gamma(TM|_{c(t)}) \to \Gamma(TM|_{c(t)})$$

satisfying

1. $\frac{D}{dt}$ is $\mathbb{R}$-linear.
2. The Leibniz rule is satisfied, i.e.

   $$\frac{D(fV)}{dt} = Vf + f\frac{DV}{dt}$$

   for all $f \in C^\infty(M), V \in \Gamma(TM|_{c(t)})$.
3. If $V \in \Gamma(TM|_{c(t)})$ is such that $V(t) = \tilde{V}_{c(t)}$ for some $\tilde{V} \in \mathfrak{X}(M)$ then

   $$\frac{DV}{dt} = \nabla_{c'(t)}\tilde{V}.$$

Theorem 4.13. On a manifold with an affine connection, for each curve, there exists a unique covariant derivative.

With this technical result, we can now easily generalize the condition of "straightness" and define geodesics on an arbitrary manifold M.

Definition 4.29. Let M be manifold with an affine connection. A **geodesic** $c : (a, b) \to M$ is a smooth curve such that $\frac{Dc'(t)}{dt} = 0$ for all $t \in (a, b)$.

Thus to speak about straight curves, we only need an affine connection on our manifold. A Riemannian metric is not needed for that. However, if one has a manifold with both a connection and a metric and these two structures are compatible with each other then certain nice properties hold, such as that geodesics have constant speed. Formally, the compatibility between a metric and a connection is defined as follows.

Definition 4.30. An affine connection is said to be **compatible with the metric** if for all $X, Y, Z \in \mathfrak{X}(M)$

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Theorem 4.14. Let M be a Riemannian manifold with an affine connection compatible with the metric $\langle -, - \rangle$ and $c : (a, b) \to M$ be a curve. Then for any $V, W \in \Gamma(TM_{c(t)})$ we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle V, W \rangle = \left\langle \frac{DV}{dt}, W \right\rangle + \left\langle V, \frac{DW}{dt} \right\rangle .$$

In particular, if c is a geodesic then it has constant speed.

Now, in a way, a connection is a coarser notion than a metric. This a consequence of the following theorem which states that given a metric we can always construct a certain unique affine connection. On the other hand, given an affine connection, it is in general not possible to extract from it a unique Riemannian metric. Thus there is only a one-way "connection" between those concepts.

Theorem 4.15. Let M be a Riemannian manifold. Then there exists a unique torsion-free affine connection ∇ on M that is compatible with the metric. Such a connection is called the **Riemannian connection** on M. Moreover, the Riemannian connection is characterized by the Koszul formula

$$2g(\nabla_X Y, Z) = Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) - g(X, [Y, Z]) + g(Y, [Z, X]) + g(Z, [X, Y]).$$

Note that the torsion-free requirement makes Riemannian connections more similar to the D operator. In fact, the Riemannian connection on a hypersurface in $\mathbb{R}^n$ and their associated covariant derivatives along curves can be expressed in terms of the D and $\frac{\mathrm{d}}{\mathrm{d}t}$ operators.

Theorem 4.16. The Riemannian connection on a regular submanifold M of $\mathbb{R}^n$ of codimension 1 is given by

$$\nabla_X Y := (D_X Y)_{tan} ,$$

where $X, Y \in \mathfrak{X}(M)$, D is the direction derivative on $\mathbb{R}^n$ and $()_{tan}$ denotes the component of a vector field tangential to M. Moreover, if $V(t)$ is a vector field along a curve $c(t)$ in M then

$$\frac{DV}{dt} = \left(\frac{\mathrm{d}V}{\mathrm{d}t} \right)_{tan} ,$$

with $\frac{\mathrm{d}}{\mathrm{d}t}$ the derivative of components of V with respect to the canonical frame on $\mathbb{R}^n$.

4.3 Symplectic Geometry and Hamiltonian Systems

In this section, we give an overview of the elements of symplectic geometry that we will use in the subsequent section to give a rigorous framework for Hamiltonian mechanics. The central objects of study of symplectic geometry are symplectic manifolds.

Definition 4.31. A manifold M together with a closed, non-degenerate two-form ω is called a **symplectic manifold** and ω its **symplectic form**.

Theorem 4.17. A symplectic manifold has an even dimension.

Example. The prototype of a symplectic manifold is $T^* \mathbb{R}^n$ with the standard symplectic form ω. To define ω, we first define the cotangent coordinates $(q^1, ..., q^n, p_1, ..., p_n)$ on $T^* \mathbb{R}^n$. Let us be slightly more general and define cotangent coordinate on any cotangent bundle T^*M. So suppose $(U, q^1, ..., q^n)$ is a coordinate chart on M. Then for $x \in U$, any one-form $\alpha \in T_x^* M$ can be expressed as

$$\alpha = \sum_{i=1}^{n} p^i dq_i|_x ,$$

for some numbers p^i. The cotangent coordinates of $(x, \alpha) \in T^*U$ are defined to be $(q^1, .., q_n, p_1, ..., p_n)$, and $(T^*U, q_1, ...q^n, p_1, .., p_n)$ is a chart on T^*M. Now by picking the standard Cartesian coordinates on $\mathbb{R}^n$, we obtain standard global cotangent coordinates on $T^*\mathbb{R}^n$. Using these global coordinates, we define the standard symplectic form ω by

$$\omega := \sum_{i=1}^{n} dp_i \wedge dq^i.$$

Since this is a global definition, it is readily verified that ω is non-degenerate and closed — making $T^*\mathbb{R}^n$ into a symplectic manifold.

Note that a symplectic manifold need not have a metric, nor a connection. These notions are independent. However, similarly to a metric, the symplectic form induces an isomorphism of tangent and cotangent spaces.

Theorem 4.18. The symplectic form induces an isomorphism between T_pM and T_p^*M denoted by $\iota(\omega_p)$ and defined by $X_p \mapsto \iota_{X_p}\omega_p$ with $(\iota_{X_p}\omega_p)(Y_p) := \omega_p(X_p, Y_p)$. This isomorphism induces a bijection $\iota(\omega)$ between $\mathfrak{X}(M)$ and $\Omega^1(M)$.

The structure-preserving maps for symplectic manifold are symplectomorphism, as defined below.

Definition 4.32. A **symplectomorphism** $\phi : M \to N$ between two symplectic manifolds is a diffeomorphism such that ϕ^* takes the symplectic form on N to the symplectic form on M.

The following important theorem by Darboux states that every symplectic manifold is locally symplectomorphic to $T^*\mathbb{R}^n$ with $\omega = \sum_{i=1}^{n} dp_i \wedge dq^i$.

Theorem 4.19 (Darboux)**.** Let (M, ω) be symplectic manifold of dimension $2n$. Then around any point $x \in M$ there exist cotangent coordinates $(U, q^1, ..., q^n, p_1, ..., p_n)$ such that $\omega|_U = \sum_i^n dp_i \wedge dq^i$.

Having a symplectic form allows us to associate with each function f on M a vector field X_f. In other words, each function defines a dynamical system on our manifold. The following three definitions state what we mean by a Hamiltonian system, a Hamiltonian vector field and a solution to a Hamiltonian system.

Definition 4.33. A **Hamiltonian system** is a symplectic manifold M together with a smooth function $H : M \to \mathbb{R}$.

Definition 4.34. A **Hamiltonian vector field** of a Hamiltonian system (M, ω, H) is the unique vector field X_H on M such that $\iota_{X_H}\omega = -dH$.

Definition 4.35. An ***H*-solution** of a Hamiltonian system (M, ω, H) is a curve $c : (a, b) \to M$ such that $c'(t) = (X_H)_{c(t)}$.

Note that in local coordinates $c'(t) = (X_H)_{c(t)}$ is just an ODE. Thus the existence theorem for ODEs tells us that for any initial condition $c(0), (X_H)_{c(0)}$ there exists a curve c that is an H-solution satisfying it. The following theorem states that symplectomorphisms preserve H-solutions.

Theorem 4.20. Suppose $f : M \to N$ is a symplectomorphism between Hamiltonian systems $(M, \omega, H), (N, \eta, (f^{-1})^*H)$. Then $c : (a, b) \to M$ is an H-solution if and only if $f \circ c : (a, b) \to N$ is a $(f^{-1})^*H$-solution.

Using a symplectic form on M we can define an additional "product" operation of functions on M, called Poisson bracket. After stating the definition, we will prove a theorem which states that, remarkably, we can detect if a function f is constant along H-solutions, for some other function H, by looking at their Poisson bracket. Thus the conserved quantities of a Hamiltonian system (M, ω, H), can be intrinsically characterized using only the symplectic form ω.

Definition 4.36. Suppose (M, ω) is a symplectic manifold. The **Poisson bracket** of two functions $f, g \in C^\infty(M)$ is defined by

$$\{f, g\} = -\omega(X_f, X_g).$$

Theorem 4.21. A function $f \in C^\infty(M)$ is constant along all H-solutions if and only if $\{f, H\} = 0$.

Proof. Let $c(t)$ be an H-solution. Then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t (f \circ c) &= c'(t)f = (X_H)_{c(t)} f = (X_H f) \circ c(t) = (df(X_H)) \circ c(t) \\ &= (-\iota_{X_f} \omega(X_H)) \circ c(t) = \{f, H\} \circ c(t). \end{aligned}$$

Thus if f is constant along all H-solutions then $\{f, H\}(p) = 0$ for all $p \in M$ since for every initial condition there is a local H-solution. Conversely, if $\{f, H\} = 0$ then f is constant along any H-solution. □

We already showed that $T^* \mathbb{R}^n$ can be made into a symplectic manifold. In fact, any cotangent space can be equipped with a symplectic form. To do so we first introduce a one form θ on T^*M which is then used to define a symplectic form ω on T^*M by $\omega := d\theta$.

Definition 4.37. Let $\pi : T^*M \to M$ be the cotangent bundle. Then a **tautological one-form** $\theta \in \Omega^1(T^*M)$ is defined pointwise by

$$\theta_{(p,\xi)}(X_{(p,\xi)}) = \xi(d\pi_{(p,\xi)} X_{(p,\xi)})$$

for $X_{(p,\xi)} \in T_{(p,\xi)}(T^*M)$.

Observe that this definition is completely coordinate-free. In other words, the tautological one-form is an intrinsic object associated with the cotangent bundle. Moreover, as promised, we can construct a symplectic form out of it.

Theorem 4.22. Let $\pi : T^*M \to M$ be the cotangent bundle, then $\omega := d\theta$ makes T^*M into a symplectic manifold.

In the case the symplectic form is induced from the tautological form, we have the following alternative characterization of symplectomorphisms.

Theorem 4.23. Let $f : T^*M \to T^*N$ be a diffeomorphism and α, β be the tautological one forms on T^*M and T^*N, respectively. Then f is a symplectomorphism if and only if $f^*\beta = \alpha$.

Any diffeomorphism $f : M \to N$ between two manifolds M and N can be extended to the so-called cotangent lift $f_\sharp : T^*M \to T^*N$ between the respective cotangent spaces. Remarkably, if the symplectic forms on cotangent bundles are induced from fundamental forms, the cotangent lift is a symplectomorphism. More precisely, we have the following.

Definition 4.38. Let $f : M \to N$ be a diffeomorphism of manifolds and $(x, \alpha) \in T^*M$. Define the **cotangent lift** $f_\sharp : T^*M \to T^*N$ by

$$f_\sharp(x, \alpha) = (f(x), \beta),$$

where $\beta \in T_{f(x)}N$ satisfies $\beta \circ df_x = \alpha$.

Theorem 4.24. Let $f : M \to N$ be a diffeomorphism and α, β the tautological one-forms on M and N, respectively. Then the cotangent lift $f_\sharp$ is a diffeomorphism and satisfies $f_\sharp^*\beta = \alpha$

Corollary 4.25. The cotangent lift $f_\sharp$ is a symplectomorphism.

4.4 Hamiltonian Mechanics

With all the abstract tools introduced in the last section, let us examine where all of those definitions lead in the case of $T^*\mathbb{R}^n$ equipped with the standard symplectic form ω. Throughout, we will work in the global cotangent coordinates $(q, p) = (q^1, ..., q^n, p_1, ..., p_n)$ on $T^*\mathbb{R}^n$.

In this setting, a Hamiltonian system on $(T^*\mathbb{R}^n, \omega)$ is just a function $H : T^*\mathbb{R}^n \to \mathbb{R}$ which in coordinates can be simply written as $H(q, p)$. Recalling that the standard symplectic form ω is given by $\omega = \sum_{i=1}^n dp_i \wedge dq^i$, we can compute the coordinate representation of a Hamiltonian vector field X_H. By definition, we have that $\omega(X_H, -) = \iota_{X_H}\omega = -dH$ which in coordinates reads

$$\sum_{i=1}^n dp_i(X_H)dq^i - dq^i(X_H)dp_i = \sum_{i=1}^n \left(-\frac{\partial H}{\partial q^i}dq^i - \frac{\partial H}{\partial p_i}dp_i\right)$$

Comparing the coefficients gives us $dp_i(X_H) = -\frac{\partial H}{\partial q^i}, dq^i(X_H) = \frac{\partial H}{\partial p_i}$. Therefore the coordinate representation of X_H is

$$X_H = \sum_{i=1}^n \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i}. \tag{4.1}$$

Now let us give a coordinate characterization of H-solutions. First, suppose $c(t)$ is an H-solution and let us write its components as $c(t) = (q^1(t), ..., q^n(t), p_1(t), ..., p_n(t))$. Then using that $(X_H)_{c(t)} = c'(t)$, we can write $(\iota_{X_H})_{c(t)}$ as

$$\begin{aligned}
(\iota_{X_H}\omega)_{c(t)} &= \sum_{i=1}^n dp_i|_{c(t)}(X_H|_{c(t)})dq^i|_{c(t)} - dq^i|_{c(t)}(X_H|_{c(t)})dp_i|_{c(t)} \\
&= \sum_{i=1}^n dp_i|_{c(t)}(c'(t))dq^i|_{c(t)} - dq^i|_{c(t)}(c'(t))dp_i|_{c(t)} \\
&= \sum_{i=1}^n ((c'(t)p_i) \circ c(t))dq^i|_{c(t)} - ((c'(t)q^i) \circ c(t))dp_i|_{c(t)} \\
&= \sum_{i=1}^n ((p_i \circ c)' \circ c(t))dq^i|_{c(t)} - ((q^i \circ c)' \circ c(t))dp_i|_{c(t)} \\
&= \sum_{i=1}^n \left.\frac{\mathrm{d}p_i(t)}{\mathrm{d}t}\right|_{c(t)} dq^i|_{c(t)} - \left.\frac{\mathrm{d}q^i(t)}{\mathrm{d}t}\right|_{c(t)} dp_i|_{c(t)}.
\end{aligned}$$

On the other hand,

$$(\iota_{X_H}\omega)_{c(t)} = -dH_{c(t)} = -\sum_{i=1}^n \left.\frac{\partial H}{\partial q^i}\right|_{c(t)} dq^i|_{c(t)} + \left.\frac{\partial H}{\partial p_i}\right|_{c(t)} dp_i|_{c(t)}.$$

Comparing the coefficients of the above equations gives us that along H-solution $c(t)$ the following differential equations are satisfied for $i = 1, ...n$

$$\begin{aligned}
\dot{p}_i &= -\frac{\partial H}{\partial q^i} \\
\dot{q}^i &= \frac{\partial H}{\partial p_i}.
\end{aligned}$$

These equations are known as Hamilton's equations. Now we will show that conversely if $c(t) = (q^1(t), ..., q^n(t), p_1(t), ..., p_n(t))$ is a curve whose coefficients satisfy Hamilton's equations then

$c(t)$ is an H-solution. To this end, Hamilton's equations give us that

$$-dH_{c(t)} = -\sum_{i=1}^{n} \frac{\partial H}{\partial q^i}\bigg|_{c(t)} dq^i|_{c(t)} + \frac{\partial H}{\partial p_i}\bigg|_{c(t)} dp_i|_{c(t)} = \sum_{i=1}^{n} \frac{\mathrm{d}p_i(t)}{\mathrm{d}t}\bigg|_{c(t)} dq^i|_{c(t)} - \frac{\mathrm{d}q^i(t)}{\mathrm{d}t}\bigg|_{c(t)} dp_i|_{c(t)}$$
$$= \sum_{i=1}^{n} dp_i|_{c(t)}(c'(t))dq^i|_{c(t)} - dq^i|_{c(t)}(c'(t))dp_i|_{c(t)}.$$

But we also have that

$$-dH_{c(t)} = \iota_{(X_H\omega)_{c(t)}} = \sum_{i=1}^{n} dp_i|_{c(t)}(X_H|_{c(t)})dq^i|_{c(t)} - dq^i|_{c(t)}(X_H|_{c(t)})dp_i|_{c(t)}.$$

Looking at the coefficients of the above expressions, we see that all the components of $c'(t)$ and $(X_H)_{c(t)}$ agree and so we conclude that $c'(t) = (X_H)_{c(t)}$. Therefore $c(t)$ is an H-solutions and, consequently, we have the following theorem.

Theorem 4.26. Let $(T^*\mathbb{R}^n, \omega, H)$ be a Hamiltonian system, where ω is the standard symplectic form on $T^*\mathbb{R}^n$. Then a curve $c : (a, b) \to T^*\mathbb{R}^n$ is an H-solution if and only if its components (with respect to the global cotangent coordinates) satisfy Hamilton's equations.

Finally, let us compute a formula for the Poisson bracket in the cotangent coordinates. Using equation (4.1), we obtain

$$\{f, g\} = -\omega(X_f, X_g) = -\sum_{i=1}^{n} \big(dp^i(X_f)dq^i(X_g) - q_i(X_f)dp_i(X_g)\big) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}\right).$$

Having translated all of this general theory to $T^*\mathbb{R}^n$ let us see how it relates to classical mechanics.

Example. Suppose our Hamiltonian is given by $H(q, p) = \|p\|^2/2m + U(q)$ for some smooth function $U(q)$. Then the components of H-solutions satisfy

$$\dot{q}^i = \frac{\partial H}{\partial p_i} = p_i/m$$
$$\dot{p}_i = -\frac{\partial H}{\partial q^i} = \frac{\partial U}{\partial q^i}.$$

The first equation tells us that p_i is the i-th component of the usual momentum, while the second equation is just Newton's second law for a particle moving in a potential $U(q)$. We see that by taking the Hamiltonian to be the total energy of a system, we can describe its dynamics using Hamilton's equation. Therefore projecting H-solutions to the first n coordinates gives us the physical trajectories of a system whose total energy is H. As a simple corollary of Theorem 4.21, we have that $\{H, H\} = 0$ and so the energy is conserved along trajectories.

We will use the above formalism in the next chapter to show that non-constant geodesics on $\mathbb{S}^n$ correspond to H-solution on $T^*\mathbb{R}^n$ with H being the Kepler Hamiltonian, whose properties and solutions we studied extensively in the previous chapters.

Chapter 5

Kepler Problem as a Geodesic Flow on a Sphere

Equipped with the tools from chapter 4, we are now ready to give a detailed account of the construction from [Mos70]. We will show that the Kepler problem in $\mathbb{R}^n$ for negative energies (closed orbits) corresponds to non-constant geodesics on an n-dimensional sphere $\mathbb{S}^n$. Showing this will allow us to conclude that the symmetry group of the negative energy n-dimensional Kepler orbits is $SO(n+1)$ — the group of rotations in $n+1$ dimensions. This is because the non-constant geodesics on $\mathbb{S}^n$ are great circles and rotating a great circle gives another great circle. Thus acting with $SO(n+1)$ on a geodesic produces another geodesic. The equivalence shown by [Mos70] then implies that $SO(n+1)$ can also act on negative energy Kepler orbits in this way and hence is a symmetry group of this problem as well. From this, we will conclude that the Kepler potential has a larger symmetry group than a generic central potential — $SO(n)$ — which explains the conservation of the LRL vector.

Let us remark that it can be proven that for the n-dimensional Kepler problem, the symmetry group of zero-energy orbits is the Euclidean group $E(n)$ while for positive-energy orbits it is $SO(1,n)$ [Osi77; Bel77]. Both of those groups are larger than $SO(n)$ thus the Kepler problem has a larger symmetry group than a generic power-law central force problem in all parts of the phase space. These symmetry properties follow from the fact that the zero-energy orbits correspond to geodesic flow in Euclidean space while positive-energy orbits correspond to geodesics on a two-sheeted hyperboloid embedded in a Lorentz space [Osi77; Bel77]. It has been shown that the three constructions — for negative, positive and zero energy — can be unified with a single one-parameter family of maps and surfaces where the parameter is energy [Bel81].

In this chapter, we will focus on the negative energy case and refer the reader to the literature for other cases [Osi77; Mil83; Bel77] as well as for the general approach [Bel81]. We will follow the general structure of Moser's construction quite closely. However, our account spells out more details and brings to light some of the technicalities that Moser leaves implicit. Some of those details can also be found in [GS77; GS90].

5.1 Geodesic Flow as a Restricted Hamiltonian System

In this section, we will first derive the geodesic equation on $\mathbb{S}^n$. We will then find a Hamiltonian $\Phi : T^*\mathbb{R}^{n+1} \to \mathbb{R}$ such that the restriction of the corresponding Hamilton's equations to $\mathbb{S}^n$ is equivalent to the geodesic equation.

Remark. We will number the standard coordinates in $\mathbb{R}^{n+1}$ from 0, i.e. $(x^0, ..., x^n) \in \mathbb{R}^{n+1}$, while those in $\mathbb{R}^n$ from 1, i.e. $(y^1, ..., y^n) \in \mathbb{R}^{n+1}$. This will simplify some of the expressions later on. Moreover, for a curve $c : I \to \mathbb{R}^{n+1}$ we will denote by $c'(t)$ its velocity vector field. The dot denotes the differentiation of functions with respect to time. Thus the components of $c'(t)$ with respect to the standard frame ∂_i are $\sum_{i=0}^{n} \dot{c}^i(t)\partial_i$. We will also denote by $c''(t)$ the vector field $\sum_{i=0}^{n} \ddot{c}^i \partial_i$.

A problem that we will frequently encounter in what follows is that $\mathbb{R}^n$ is too nice. Namely, using the standard basis $e_1, ..., e_n$ and standard frames $\partial_1, .., \partial_n$, $dx^1, ..., dx^n$ for $T\,\mathbb{R}^n$ and $T^*\,\mathbb{R}^n$ respectively, we have isomorphism between $\mathbb{R}^n, T_p\,\mathbb{R}^n, T_q^*\,\mathbb{R}^n$ for all $p, q \in \mathbb{R}^n$. To make matters worse, both the Euclidean metric and the standard symplectic form on $T^*\,\mathbb{R}^n$ induce the same isomorphism between vectors and covectors. On top of that, we can apply the metric to points, vectors, covectors, or combinations thereof. This will make keeping track of what lives where quite cumbersome. Fortunately, if leveraged cautiously this difficulty is also a blessing — we can seemingly transfer objects between those spaces. To keep things as clear as possible we will stick to the following rules:

1. We will never implicitly change the identity of a given object
2. We will exploit all the isomorphism, and denote the metric and the associated norm on all the spaces by the same symbols, i.e. by $\langle -, - \rangle$ and $\|-\|$.
3. We will never take mixed inner products, for example, of a point with a vector.
4. We will never take inner products of vectors at different points or covectors at different points.
5. All the isomorphism between points, vectors and covectors will be inner product preserving.
6. In this chapter, we won't use the Einstein summation convention. Moreover, all components will have lower indices.

With these precautions, we are ready to start. First, let us define the notation for the isomorphism between points and vectors at this point.

Definition 5.1. Let $p = (p_0, ..., p_n) \in \mathbb{R}^{n+1}$. We will denote by $\bar{p} \in T_p\,\mathbb{R}^{n+1}$ the vector

$$\bar{p} := \sum_{i=0}^{n} p_i \partial_i.$$

Note that $\langle p, q \rangle = \langle \bar{p}, \bar{q} \rangle$.

The make a vector into a covector we will use the musical isomorphism with one modification accounting for the velocity vector field of a curve.

Definition 5.2. Let $v = \sum_{i=0}^{n} v_i \partial_i \in T_q\,\mathbb{R}^{n+1}$. We will denote by $v^\flat \in T_q^*\,\mathbb{R}^{n+1}$ the covector

$$p^\flat = \sum_{i=0}^{n} v_i dx^i.$$

Additionally, for a curve $c : I \to \mathbb{R}^{n+1}$ we will denote by $c^\flat(t)$ the covector field associated to the velocity vector field $c'(t)$.

With these definitions, we will use the standard embedding of $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$ to characterize $T\,\mathbb{S}^n$ as a subset of $T\,\mathbb{R}^{n+1}$. As a side note, recall from chapter 4 that $\mathbb{S}^n$ is a regular submanifold of $\mathbb{R}^{n+1}$.

Lemma 5.1. Let $f(x) := \|x\|^2 - 1$ and consider $T\,\mathbb{S}^n$ to be standardly embedded in $T\,\mathbb{R}^{n+1}$. Take $(\xi, \eta) \in T\,\mathbb{S}^n$. Then the following are true:

1. $\nabla f(\xi) = 2\bar{\xi}$
2. $\langle \bar{\xi}, \eta \rangle = 0$, i.e. $\bar{\xi}$ is orthogonal to $T_\xi\,\mathbb{S}^n$
3. There is an orthogonal decomposition $T_\xi\,\mathbb{R}^{n+1} = T_\xi\,\mathbb{S}^n \oplus \operatorname{span}(\bar{\xi})$.
4. For $(x, y) \in T\,\mathbb{R}^{n+1}$, $(x, y) \in T\,\mathbb{S}^n$ if and only if $\|x\| = 1$ and $\langle \bar{x}, y \rangle = 0$.

Proof. 1. This follows immediately since

$$\nabla f(\xi) = \sum_{i=0}^{n} 2\xi_i \partial_i = 2\bar{\xi}.$$

2. Due to 1 this claim is equivalent to showing that $\nabla f(\xi)$ is orthogonal to $T_\xi\,\mathbb{S}^n$. To this end, take $X_\xi = \sum_{i=0}^{n} X_i \partial_i \in T_\xi\,\mathbb{S}^n$ then there exists a curve $c(t) \in \mathbb{S}^n$ such that $c(0) = \xi$ and $c'(0) = X_\xi$. Observing that $\mathbb{S}^n = f^{-1}(0)$, it follows that $f(c(t)) = 0$. Differentiating this equation gives us

$$0 = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_0 f(c(t)) = c'(0)f = X_\xi f = \sum_{i=0}^{n} X_i \partial_i f = \langle X_\xi, \nabla f \rangle\,,$$

which proves our claim.

3. This follows from the fact that $\operatorname{span}(\bar{\xi})$ and $T_\xi\,\mathbb{S}^n$ are subspaces of $T_\xi\,\mathbb{R}^{n+1}$ which are orthogonal to each other and their dimensions add up to the dimension of $T_\xi\,\mathbb{R}^{n+1}$.

4. Take $(x, y) \in T\,\mathbb{R}^{n+1}$. First suppose $(x, y) \in T\,\mathbb{S}^n$. Then $x \in \mathbb{S}^n$ so $\|x\| = 1$ and by 2. $\langle \bar{x}, y \rangle = 0$. Conversely, suppose $\|x\| = 1$ and $\langle \bar{x}, y \rangle = 0$. Then $x \in \mathbb{S}^n$ and since $T_x\,\mathbb{R}^{n+1} = T_x\,\mathbb{S}^n \oplus \operatorname{span}(\bar{x})$ orthogonality of $\bar{x}$ and y implies that $y \in T_x\,\mathbb{S}^n$. Thus $(x, y) \in T\,\mathbb{S}^n$.

□

Note that since the musical isomorphism $\flat$ preserves the inner product — or rather the inner product on the cotangent space is defined using the musical isomorphism and inner product on the tangent space — we have exactly the same characterization of the cotangent space.

Corollary 5.1. For $(x, y) \in T^*\,\mathbb{R}^{n+1}$ it holds that $(x, y) \in T^*\,\mathbb{S}^n$ if and only if $\|x\| = 1$ and $\langle \bar{x}^\flat, y \rangle = 0$.

Before we derive the geodesic equation on $\mathbb{S}^n$, let us collect some identities that we will use in the proof.

Lemma 5.2. Let $\xi \in \mathbb{S}^n$ and $c : I \to \mathbb{S}^n$ be a curve. Then

1. $\langle \bar{c}(t), c'(t) \rangle = 0$
2. $\langle \bar{c}(t), c''(t) \rangle = -\|c'(t)\|^2$

Proof. 1. Since $c'(t) \in T_{c(t)}\,\mathbb{S}^n$, Lemma 5.1.2 implies that $\langle \bar{c}(t), c'(t) \rangle = 0$.

2. Differentiating the previous identity gives

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\langle \bar{c}(t), c'(t) \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=0}^{n} c_i(t)\dot{c}_i(t) = \sum_{i=0}^{n} \dot{c}_i(t)\dot{c}_i(t) + c_i(t)\ddot{c}_i(t) = \left\|c'(t)\right\|^2 + \langle \bar{c}(t), c''(t) \rangle.$$

Consequently, $\langle \bar{c}(t), c''(t) \rangle = -\|c'(t)\|^2$.

□

With all these tools, we are ready to derive the geodesic equation on $\mathbb{S}^n$.

Theorem 5.2. Let $\mathbb{S}^n$ denote the n-dimensional unit sphere embedded in $\mathbb{R}^{n+1}$ and equipped with the induced metric and the corresponding Riemannian connection. Then $c : I \to \mathbb{S}^n \subset \mathbb{R}^{n+1}$ is a geodesic if and only if

$$\ddot{c}_i + \|c'(t)\|^2 c_i(t) = 0 \quad \text{for all} \quad i = 0, ..., n \quad \text{and} \quad t \in I, \tag{5.1}$$

where $c_i(t)$'s are the component functions of $c(t)$ with respect to the standard basis on $\mathbb{R}^{n+1}$.

Proof. Recall that a curve $c : I \to \mathbb{S}^n$ is a geodesic if the covariant derivative along $c(t)$ of its velocity vector field $c'(t)$ vanishes, i.e. if

$$\frac{Dc'(t)}{dt} = 0 \quad \text{for all} \quad t \in I.$$

Moreover, we know from Theorem 4.16 that the covariant derivative on $\mathbb{S}^n$ of a vector field $V(t)$ along some curve is the tangential component of its covariant derivative on $\mathbb{R}^{n+1}$, that is

$$\frac{DV}{dt} = \left(\frac{dV}{dt}\right)_{tan}.$$

Thus $c(t)$ is a geodesic if

$$\left(\frac{dc'(t)}{dt}\right)_{tan} = 0 \quad \text{for all} \quad t \in I.$$

To proceed, observe that since $c(t) \in \mathbb{S}^n$ it follows from Lemma 5.1.1 that $\bar{c}(t)$ is a unit normal vector. Thus we can use it to compute the tangential component of $dc'(t)/dt$. Namely, we get that

$$0 = \left(\frac{dc'(t)}{dt}\right)_{tan} = \frac{dc'(t)}{dt} - \left\langle \frac{dc'(t)}{dt}, \bar{c}(t) \right\rangle \bar{c}(t).$$

Writing both vectors in the standard frame $\{\partial_i\}$, we obtain that

$$\sum_{i=0}^{n} \left(\ddot{c}_i - \sum_{j=0}^{n} \ddot{c}_j(t) c_j(t) c_i(t) \right) \partial_i = 0.$$

By linear independence of ∂_i's, it follows that $c(t)$ is a geodesic if and only if its component functions satisfy

$$\ddot{c}_i(t) - \langle c''(t), \bar{c}(t) \rangle c_i(t) = 0 \tag{5.2}$$

for all $i = 0, ..., n$ and $t \in I$. Applying Lemma 5.2.3 yields

$$\ddot{c}_i(t) + \|c'(t)\|^2 c_i(t) = 0 \quad \text{for all} \quad i = 0, ..., n \quad \text{and} \quad t \in I,$$

which is the desired form of the geodesic equation. □

As a side note, recall from Theorem 4.14 that $\|c'(t)\|^2$ is a constant. Thus the derived geodesic equation (5.1) is a second-order linear equation. Moreover, it only uses the components of $c'(t)$ and $c(t)$ thus the above theorem can be equivalently stated in terms of $c(t)$ and $c^\flat(t)$. We will use this fact in the next theorem where we express the geodesic equation as the restriction of Hamilton's equation.

Theorem 5.3. Let $\Phi : T^*\mathbb{R}^{n+1} \to \mathbb{R}$ be the Hamiltonian defined by $\Phi(\xi,\eta) = \frac{1}{2}\|\xi\|^2\|\eta\|^2$ and $c : I \to \mathbb{R}^{n+1}$ be a curve. Then $c(t)$ is a geodesic on $\mathbb{S}^n$ if and only if $(c(t), c^\flat(t)) \in T^*\mathbb{R}^{n+1}$ is a solution to Hamilton's equation corresponding to the Hamiltonian Φ that satisfies $\|c(0)\| = 1$ and $\langle c'(0), \bar{c}(0)\rangle = 0$.

Proof. Let us start by computing the Hamilton's equation corresponding to $\Phi(\xi,\eta)$. They are given by

$$\begin{cases} \dot{\xi}_i &= \frac{\partial\Phi}{\partial\eta_i} = \|\xi\|^2\eta_i \\ \dot{\eta}_i &= -\frac{\partial\Phi}{\partial\xi_i} = -\|\eta\|^2\xi_i, \end{cases} \tag{5.3}$$

where $i = 0, ..., n$. Note that $\eta \in T_\xi\mathbb{R}^{n+1}$.

($\Rightarrow$) : Suppose $\xi(t)$ is a geodesic on $\mathbb{S}^n$ thus by Theorem 5.2 it satisfies

$$\ddot{\xi}_i(t) + \left\|\xi'(t)\right\|^2\xi_i(t) = 0 \quad \text{for all} \quad i = 0, ..., n \quad \text{and} \quad t \in I.$$

Defining $\eta(t) := \xi^\flat(t) \in T^*_{\xi(t)}\mathbb{S}^n \subset T^*_{\xi(t)}\mathbb{R}^{n+1}$ and using the above equation we get that the components of $\xi(t)$ and $\eta(t)$ satisfy

$$\begin{cases} \dot{\xi}_i(t) &= \eta_i(t) \\ \dot{\eta}_i(t) &= \ddot{\xi}_i(t) = -\|\xi'(t)\|^2\xi_i(t). \end{cases}$$

Further, observe that $\|\xi(t)\|^2 = 1$ and by Lemma 5.2 $\left\langle\bar{\xi}(t), \xi'(t)\right\rangle = 0$. Hence we conclude that $(\xi(t), \eta(t)) = (c(t), c^\flat(t)) \in T^*\mathbb{R}^{n+1}$ solves

$$\begin{cases} \dot{\xi}_i(t) &= \|\xi(t)\|^2\eta_i(t) \\ \dot{\eta}_i(t) &= -\|\xi'(t)\|^2\xi_i(t) \end{cases}$$

with $\|\xi(0)\| = 1$ and $\left\langle\bar{\xi}(0), \xi'(0)\right\rangle = 0$.

($\Leftarrow$): Conversely, suppose $(c(t), c^\flat(t))$ is a solution to the system (5.3) satisfying $\|c(0)\|^2 = 1$ and $\langle\bar{c}(0), c'(0)\rangle = 0$. Hence

$$\begin{cases} \dot{c}_i(t) &= \|c(t)\|^2\dot{c}_i(t) \\ \ddot{c}_i(t) &= -\|c'(t)\|^2c_i(t), \end{cases} \tag{5.4}$$

for $i = 1, ..., n$. The first equation implies that $\|c'(t)\|^2 = \|c(t)\|^2\|c'(t)\|^2$ while the second can be rewritten as $c''(t) = -\|c'(t)\|^2\bar{c}(t)$. Using these we get that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\bar{c}(t), c'(t)\right\rangle = \left\|c'(t)\right\|^2 + \left\langle\bar{c}(t), c''(t)\right\rangle = \left\|c'(t)\right\|^2 - \left\|c'(t)\right\|^2\|c(t)\|^2 = 0.$$

Since $\langle\bar{c}(0), c'(0)\rangle = 0$ the above implies that $\langle\bar{c}(t), c'(t)\rangle = 0$ for all t. From this, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|c(t)\|^2 = 2\left\langle\bar{c}(t), c'(t)\right\rangle = 0$$

which together with $\|c(0)\|^2 = 1$ implies that $\|c(t)\|^2 = 1$. Now $\|c(t)\|^2 = 1$, $\langle\bar{c}(t), c'(t)\rangle = 0$ together with Lemma 5.1.4 implies that $(c(t), c'(t)) \in T\mathbb{S}^n$ so $c(t)$ can be viewed as a curve $c : I \to \mathbb{S}^n$. Finally, the second equation in (5.4) gives that

$$\ddot{c}_i(t) + \left\|c'(t)\right\|^2c_i(t) = 0 \quad \text{for all} \quad i = 0, ..., n \quad \text{and} \quad t \in I,$$

and so we conclude by Theorem 5.2 that $c(t)$ is a geodesic. □

It follows from the above theorem that we can identify geodesics on $\mathbb{S}^n$ with Φ-solutions which start on $T^*\mathbb{S}^n \subset T^*\mathbb{R}^{n+1}$. The next step in the construction is to apply the cotangent lift of the stereographic projection to solutions of this Hamiltonian system and obtain an equivalent description in terms of a Hamiltonian system on $T^*\mathbb{R}^n$. In the next section, we will define the stereographic projection, and prove some of its properties that are necessary to compute its cotangent lift which we will do in the subsequent section.

5.2 Intermezzo on the Stereographic Projection

Let $\hat{\mathbb{S}}^n = \mathbb{S}^n - \{(1, 0, ..., 0)\}$ denote the n-sphere with its north pole $N = (1, 0, ...0)$ removed. Note that $\hat{\mathbb{S}}^n$ is a regular submanifold of $\mathbb{R}^{n+1}$. This follows because we can create adapted charts relative to $\hat{\mathbb{S}}^n$ around every $p \in \hat{\mathbb{S}}^n$ by taking the corresponding adapted chart relative to $\mathbb{S}^n$ and intersecting its domain with an open subset of $\mathbb{R}^{n+1}$ that contains p but does not contain N (such subset exists because $\mathbb{R}^{n+1}$ is Hausdorff). With this, the stereographic projection $\sigma : \hat{\mathbb{S}}^n \to \mathbb{R}^n$ is defined by

$$\sigma(x_0, ..., x_n) := \left(\frac{x_1}{1 - x_0}, \dots, \frac{x_n}{1 - x_0}\right).$$

There is an intuitive geometric interpretation of this definition, as shown in the following lemma.

Lemma 5.3. For any $x \in \hat{\mathbb{S}}^n$, the stereographic projection maps x to $u \in \mathbb{R}^n$ where $(0, u) \in \mathbb{R}^{n+1}$ is the point of intersection of the line passing through N and x with the hyperplane $\{x_0 = 0\}$. This geometric construction is depicted in Figure 5.1.

Proof. Well indeed, take $x \in \hat{\mathbb{S}}^n$ then a parametrization of the line connecting x and $N = (1, 0, ..., 0)$ is given by

$$r(t) = (1 - t)x + t(1, 0, ..., 0) = ((1 - t)x_0 + t, (1 - t)x_1, ..., (1 - t)x_n).$$

This line intersect the plane $\{x_0 = 0\}$ when

$$(1 - t')x_0 + t' = 0 \quad \Leftrightarrow \quad t'(1 - x_0) = -x_0 \quad \Leftrightarrow \quad t' = -\frac{x_0}{1 - x_0}.$$

Note that since $x \in \hat{\mathbb{S}}^n$, we have that $x_0 \neq 1$ so we are not dividing by 0. Now we can obtain the point of intersection of our line with the plane $\{x_0 = 0\}$ by plugging t' into the parametrization $r(t)$. Before we do that observe that

$$1 - t' = 1 + \frac{x_0}{1 - x_0} = \frac{1 - x_0 + x_0}{1 - x_0} = \frac{1}{1 - x_0}.$$

Using this, we get that the desired intersection point is

$$r(t') = ((1 - t')x_0 + t', (1 - t')x_1, ..., (1 - t')x_n) = \left(\frac{x_0}{1 - x_0} - \frac{x_0}{1 - x_0}, \frac{x_1}{1 - x_0}, ..., \frac{x_n}{1 - x_0}\right) = (0, \sigma(x)),$$

as desired. □

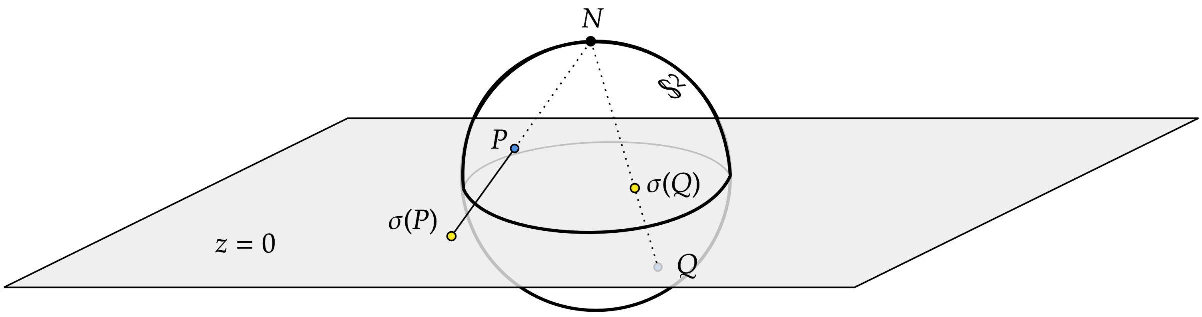

Figure 5.1: The geometric representation of the stereographic projection $\sigma : \mathbb{S}^2 \to \mathbb{R}^2$.

Now, in order for the cotangent lift of σ to be well-defined, we must first check that σ is a diffeomorphism.

Lemma 5.4. The stereographic projection σ is bijective and

$$\sigma^{-1}(u) = \left(\frac{\|u\|^2 - 1}{\|u\|^2 + 1}, \frac{2u_1}{\|u\|^2 + 1}, \dots, \frac{2u_n}{\|u\|^2 + 1} \right).$$

(Keep in mind our convention for denoting points in $\mathbb{R}^{n+1}$ and $\mathbb{R}^n$) Moreover, when $\hat{\mathbb{S}}^n$ is considered as a regular submanifold of $\mathbb{R}^{n+1}$ the stereographic projection is a diffeomorphism.

Proof. Let us verify that the given σ^{-1} is the inverse of σ. Well indeed, take $x \in \hat{\mathbb{S}}^n$ then

$$(\sigma^{-1} \circ \sigma)(x) = \sigma^{-1} \left(\frac{x_1}{1 - x_0}, \dots, \frac{x_n}{1 - x_0} \right) = \left(\frac{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} - 1}{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} + 1}, \frac{\frac{2x_1}{1-x_0}}{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} + 1}, \dots, \frac{\frac{2x_n}{1-x_0}}{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} + 1} \right).$$

Now using that $\|x\| = 1$, we get

$$\frac{\frac{2x_j}{1-x_0}}{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} + 1} = \frac{2x_j(1 - x_0)}{\sum_{i=1}^n x_i^2 + (1 - x_0)^2} = \frac{2x_j(1 - x_0)}{\sum_{i=1}^n x_i^2 + x_0^2 - 2x_0 + 1} = \frac{2x_j(1 - x_0)}{1 - 2x_0 + 1} = x^j,$$

and

$$\frac{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} - 1}{\sum_{i=1}^n \frac{x_i^2}{(1-x_0)^2} + 1} = \frac{\sum_{i=1}^n x_i^2 - (1 - x_0)^2}{\sum_{i=1}^n x_i^2 + (1 - x_0)^2} = \frac{(1 - x_0^2) - (1 - x_0)^2}{(1 - x_0^2) + (1 - x_0)^2} = \frac{2x_0 - 2x_0^2}{2 - 2x_0} = x_0.$$

Therefore for any $x \in \hat{\mathbb{S}}^n$

$$(\sigma^{-1} \circ \sigma)(x) = x.$$

Conversely, take $u \in \mathbb{R}^n$ and consider $(\sigma \circ \sigma^{-1})(u)$. First let us check that $\sigma^{-1}(u) \in \hat{\mathbb{S}}^n$. Well indeed,

$$\left\|\sigma^{-1}(u)\right\|^2 = \frac{(\|u\|^2 - 1)^2}{(\|u\|^2 + 1)^2} + \sum_{j=1}^n \frac{4u_j^2}{(\|u\|^2 + 1)^2} = \frac{1}{(\|u\|^2 + 1)^2} \left(\|u\|^4 - 2\|u\| + 1 + 4\|u\| \right) = 1.$$

Moreover, $\sigma^{-1}(u) \neq (1, 0, ..., 0)$ as otherwise the last n coordinates would have to be zero, which would imply that u is 0 and hence that the first coordinate of $\sigma^{-1}(u)$ is -1. With this out of the way, let us compute $(\sigma \circ \sigma^{-1})(u)$

$$(\sigma \circ \sigma^{-1})(u) = \sigma \left(\frac{\|u\|^2 - 1}{\|u\|^2 + 1}, \frac{2u_1}{\|u\|^2 + 1}, \dots, \frac{2u_n}{\|u\|^2 + 1} \right) = \left(\frac{\frac{2u_1}{\|u\|^2+1}}{1 - \frac{\|u\|^2-1}{\|u\|^2+1}}, \dots, \frac{\frac{2u_n}{\|u\|^2+1}}{1 - \frac{\|u\|^2-1}{\|u\|^2+1}} \right) = u.$$

Therefore we conclude that $\sigma \circ \sigma^{-1} = \mathrm{id}_{\mathbb{R}^n}$ and $\sigma^{-1} \circ \sigma = \mathrm{id}_{\hat{\mathbb{S}}^n}$. So σ^{-1} is the inverse of σ and as a consequence σ is bijective.

Since σ is not defined in terms of coordinates on $\hat{\mathbb{S}}^n$, we cannot immediately conclude that it is smooth. However, σ can be extended to a well-defined function on the open set $\mathbb{R}^{n+1} \setminus \{x^0 = 1\} \subset \mathbb{R}^{n+1}$, as such σ is smooth as a component-wise rational function. Since $\hat{\mathbb{S}}^n$ is a regular submanifold of $\mathbb{R}^{n+1} \setminus \{x^0 = 1\}$ [1] we can conclude by Theorem 4.5 that σ is smooth. Finally, the smoothness of σ^{-1} follows from Theorem 4.6 since it is induced from a smooth map on Euclidean spaces and $\hat{\mathbb{S}}^n$ is a regular submanifold of $\mathbb{R}^{n+1}$. □

[1] This follows from the fact that a regular submanifold of a manifold M is a regular submanifold of any open set $U \subset M$ that contains it.

5.3 The Cotangent Lift of the Stereographic Projection

In this section, we will compute the cotangent lift $\sigma_\sharp : T^*\hat{\mathbb{S}}^n \to T^*\mathbb{R}^n$ in terms of coordinates on $T\mathbb{R}^{n+1}$. Recall that the cotangent lift is defined by

$$\sigma_\sharp(\xi,\eta) = (\sigma(\xi),\beta),$$

where $\beta \in T_{\sigma(\xi)}N$ satisfies $\beta \circ d\sigma_\xi = \eta$. Because σ is a diffeomorphism, $d\sigma_\xi$ is an isomorphism with inverse $d(\sigma^{-1})_{\sigma(\xi)}$. Therefore $\beta = \eta \circ d(\sigma^{-1})_{\sigma(\xi)}$. Let us compute the components of β with respect to the basis $dx^1,...,dx^n$ for $T^*_{\sigma(\xi)}\mathbb{R}^n$. We get

$$\beta_i = \beta(\partial_i) = \eta(d(\sigma^{-1})_{\sigma(\xi)}(\partial_i))$$

so let us first compute $d(\sigma^{-1})_{\sigma(\xi)}(\partial_i)$ in terms of $\frac{\partial}{\partial\xi_0},...,\frac{\partial}{\partial\xi_n}$, we get

$$d(\sigma^{-1})_{\sigma(\xi)}(\partial_i) = \sum_{j=0}^{n} d\xi_j(d(\sigma^{-1})_{\sigma(\xi)}(\partial_i))\frac{\partial}{\partial\xi_j} = \sum_{j=0}^{n}(d(\xi_j\circ\sigma^{-1})_{\sigma(\xi)}\partial_i)\frac{\partial}{\partial\xi_j} = \sum_{j=0}^{n}\frac{\partial(\xi_j\circ\sigma^{-1})}{\partial x_i}\bigg|_{\sigma(\xi)}\frac{\partial}{\partial\xi_j}.$$

Substituting the formula for $\sigma^{-1}(x)$ yields

$$\begin{aligned} d(\sigma^{-1})_{\sigma(\xi)}(\partial_i) &= \frac{\partial}{\partial x_i}\bigg|_{\sigma(\xi)}\left(\frac{\|x\|^2-1}{\|x\|^2+1}\right)\frac{\partial}{\partial\xi_0} + \sum_{j=1}^{n}\frac{\partial}{\partial x_i}\bigg|_{\sigma(\xi)}\left(\frac{2x_j}{\|x\|^2+1}\right)\frac{\partial}{\partial\xi_j} \\ &= \frac{4\sigma(\xi)_i}{(\|\sigma(\xi)\|^2+1)^2}\frac{\partial}{\partial\xi_0} + \sum_{j=1}^{n}\left(\delta_{ij}\frac{2}{\|\sigma(\xi)\|^2+1} - \frac{4\sigma(\xi)_i\sigma(\xi)_j}{(\|\sigma(\xi)\|^2+1)^2}\right)\frac{\partial}{\partial\xi_j}. \end{aligned}$$

Going back to β_i, we write $\eta = \sum_{i=0}^{n}\eta_i d\xi_i \in T^*_\xi\mathbb{S}^n \subset T^*_\xi\mathbb{R}^{n+1}$ and obtain that

$$\beta_i = \eta(d(\sigma^{-1})_{\sigma(\xi)}(\partial_i)) = \frac{4\sigma(\xi)_i}{(\|\sigma(\xi)\|^2+1)^2}\eta_0 + \sum_{j=1}^{n}\left(\delta_{ij}\frac{2}{\|\sigma(\xi)\|^2+1} - \frac{4\sigma(\xi)_i\sigma(\xi)_j}{(\|\sigma(\xi)\|^2+1)^2}\right)\eta_j.$$

Observing that

$$\|\sigma(\xi)\|^2+1 = \frac{\|\xi\|^2-(\xi_0)^2}{(1-\xi_0)^2}+1 = \frac{1-(\xi_0)^2}{(1-\xi_0)^2}+1 = \frac{1+\xi_0}{1-\xi_0}+1 = \frac{2}{1-\xi_0} \tag{5.5}$$

and using the formula for $\sigma(\xi)$ gives us

$$\begin{aligned} \beta_i &= \frac{4\xi_i}{1-\xi_0}\frac{(1-\xi_0)^2}{4}\eta_0 + \sum_{j=1}^{n}\left(\delta_{ij}(1-\xi_0) - \frac{4\xi_i\xi_j}{(1-\xi_0)^2}\frac{(1-\xi_0)^2}{4}\right)\eta_j \\ &= \xi_i(1-\xi_0)\eta_0 + (1-\xi_0)\eta_i - \xi_i\sum_{j=1}^{n}\xi_j\eta_j. \end{aligned}$$

Finally, notice that

$$\sum_{j=1}^{n}\xi_j\eta_j = \left\langle\vec{\xi}^\flat,\eta\right\rangle - \xi_0\eta_0,$$

and, moreover, by Corollary 5.1 we have $\left\langle\vec{\xi}^\flat,\eta\right\rangle = 0$. Therefore

$$\beta_i = \xi_i(1-\xi_0)\eta_0 + (1-\xi_0)\eta_i + \xi_i\xi_0\eta_0 = \xi_i\eta_0 + (1-\xi_0)\eta_i.$$

Therefore, the coordinate representation of the cotangent lift $\sigma_\sharp$ with respect to the standard coordinates $\xi_0, ..., \xi_n, \eta_0, ..., \eta_n$ on $T^* \mathbb{R}^{n+1}$ is given by

$$\sigma_\sharp(\xi, \eta) := (\sigma(\xi), g(\xi, \eta)),$$

where

$$g(\xi, \eta) := (\eta_1(1 - \xi_0) + \xi_1 \eta_0, ..., \eta_n(1 - \xi_0) + \xi_n \eta_0) \,.$$

Applying Corollary 4.25 to our case gives us that the cotangent lift $\sigma_\sharp$ is a symplectomorphism. Thus, $\sigma_\sharp$ is a diffeomorphism and in particular it has an inverse. Let us compute it.

Lemma 5.5. The inverse of the cotangent lift $\sigma_\sharp$ is given by

$$\sigma_\sharp^{-1}(x, y) = (\sigma^{-1}(x), f(x, y)),$$

where in terms of the components with respect to the basis $dx^1, ..., dx^n$ of $T_x^* \mathbb{R}^n$ the function f is given by

$$f(x, y) = \left(\langle x, y \rangle \,, \frac{\|x\|^2 + 1}{2} y_1 - \langle x, y \rangle \, x_1, ..., \frac{\|x\|^2 + 1}{2} y_n - \langle x, y \rangle \, x_n \right).$$

Proof. First, observe that $\left\|\sigma^{-1}(x)\right\| = 1$ and

$$\begin{aligned} \left\langle \bar{\sigma}^{-1}(x), f(x, y) \right\rangle &= \frac{\|x\|^2 - 1}{\|x\|^2 + 1} \langle x, y \rangle + \sum_{i=1}^n \frac{2x_i}{\|x\|^2 + 1} \left(\frac{\|x\|^2 + 1}{2} y_i - \langle x, y \rangle \, x_i \right) \\ &= \frac{\|x\|^2 - 1}{\|x\|^2 + 1} \langle x, y \rangle + \langle x, y \rangle - 2 \langle x, y \rangle \frac{\|x\|^2}{\|x\|^2 + 1} = 0. \end{aligned}$$

Consequently, Corollary 5.1 gives us that $\sigma_\sharp^{-1}(x, y) \in T^* \mathbb{S}^n$. Since $\sigma^{-1}(x) \neq (1, 0, ..., 0)$ we have that $\sigma_\sharp^{-1}(x, y) \in T^* \hat{\mathbb{S}}^n$ and so $\sigma_\sharp^{-1}$ is well-defined. Now direct computation gives us

$$\begin{aligned} (\sigma_\sharp^{-1} \circ \sigma_\sharp)(\xi, \eta) &= \sigma_\sharp{}^{-1}(\sigma(\xi), g(\xi, \eta)) = (\sigma^{-1}(\sigma(\xi)), f(\sigma(\xi), g(\xi, \eta)) \\ &= \left(\xi, \langle \sigma(\xi), g(\xi, \eta) \rangle \,, ..., \frac{\|\sigma(\xi)\|^2 + 1}{2} g_k(\xi, \eta) - \langle \sigma(\xi), g(\xi, \eta) \rangle \, \sigma_k(\xi), ... \right). \end{aligned}$$

Using that $\left\langle \bar{\xi}^\flat, \eta \right\rangle = 0$ and $\|\xi\| = 1$ we get

$$\langle \sigma(\xi), g(\xi, \eta) \rangle = \sum_{k=1}^n \frac{\xi_k}{1 - \xi_0} (\eta_k(1 - \xi_0) + \xi_k \eta_0) = -\xi_0 \eta_0 + \frac{1 - (\xi_0)^2}{1 - \xi_0} \eta_0 = \eta_0.$$

Moreover, reusing the computation of $\|\sigma(\xi)\|^2 + 1$ from equation (5.5) we obtain

$$\frac{\|\sigma(\xi)\|^2 + 1}{2} g_k(\xi, \eta) - \langle \sigma(\xi), g(\xi, \eta) \rangle \, \sigma_k(\xi) = \frac{1}{1 - \xi_0} (\eta_k(1 - \xi_0) + \xi_k \eta_0) - \eta_0 \frac{\xi_k}{1 - \xi^0} = \eta_k.$$

Hence, we conclude that

$$(\sigma_\sharp{}^{-1} \circ \sigma_\sharp)(\xi, \eta) = (\xi, \eta).$$

Conversely,

$$(\sigma_\sharp \circ \sigma_\sharp{}^{-1})(x, y) = \sigma_\sharp(\sigma^{-1}(x), f(x, y)) = (\sigma(\sigma^{-1}(x)), g(\sigma^{-1}(x), f(x, y))) = (x, y)$$

since

$$f_k(x,y)(1-(\sigma^{-1})_0(x)) + (\sigma^{-1})_k(x) f_0(x,y) = \left(\frac{\|x\|^2+1}{2} y_k - \langle x,y\rangle x_k\right)\left(1 - \frac{\|x\|^2-1}{\|x\|^2+1}\right) + \frac{2x_k}{\|x\|^2+1}\langle x,y\rangle$$
$$= y_k - \langle x,y\rangle x_k \frac{2}{\|x\|^2+1} + \frac{2x_k}{\|x\|^2+1}\langle x,y\rangle = y_k.$$

□

Lemma 5.6. With $f(x,y)$ defined as in Lemma 5.5 we have that

$$\|f(x,y)\| = \frac{\|x\|^2+1}{2}\|y\|.$$

Proof. Direct computation gives

$$\|f(x,y)\|^2 = \langle x,y\rangle^2 + \sum_{i=1}^{n}\left(\frac{\|x\|^2+1}{2} y_i - \langle x,y\rangle x_i\right)^2 = \left(\frac{\|x\|^2+1}{2}\right)^2 \|y\|^2.$$

□

Equipped with all these tools let us see how geodesic flow transforms under the cotangent lift of the stereographic projection.

5.4 Geodesic Flow under the Cotangent Lift of the Stereographic Projection

To transform the differential equation we pullback our Hamiltonian Φ on $T^*\mathbb{R}^{n+1}$ first by the inclusion $\mathfrak{i} : T^*\hat{\mathbb{S}}^n \hookrightarrow T^*\mathbb{R}^{n+1}$ and then by $\sigma_\sharp^{-1}$ and get a Hamiltonian $F := (\sigma_\sharp^{-1})^*\mathfrak{i}^*\Phi$ on $T\mathbb{R}^n$ given by

$$F(x,y) = ((\sigma_\sharp{}^{-1})^*\mathfrak{i}^*\Phi)(x,y) = \Phi(\mathfrak{i}(\sigma_\sharp{}^{-1}(x,y))) = \Phi(\sigma^{-1}(x), f(x,y)) = \frac{1}{2}\left\|\sigma^{-1}(x)\right\|^2 \|f(x,y)\|^2.$$

Using that $\left\|\sigma^{-1}(x)\right\| = 1$ and an expression for $\|f(x,y)\|$ from Lemma 5.6 gives

$$F(x,y) = \frac{(\|x\|^2+1)^2}{8}\|y\|^2.$$

If it was not for the inclusion, we could have used that $\sigma_\sharp$ is a symplectomorphism to conclude that Φ-solutions are mapped to F-solutions. However, since we are looking at a *restricted* Hamiltonian system we cannot blindly apply this theorem. Fortunately, not all hope is lost — we can still use that $\sigma_\sharp$ is a symplectomorphism. However, this requires some preliminary results regarding regular submanifolds.

Theorem 5.4. Suppose N is a regular submanifold of M and θ is the tautological one form on T^*M. Then $\mathfrak{i}^*\theta = \theta$ is the tautological one form on T^*N where $\mathfrak{i} : T^*N \hookrightarrow T^*M$ is the inclusion.

Proof. Let (U,ϕ) be an adapted chart relative to N, thus $N \cap U$ is defined by the vanishing of the last $m-n$ coordinates so $N \cap U = \phi^{-1}(\{x \in \mathbb{R}^m \,|x^{n+1} = ... = x^m = 0\})$. Hence $x^1,...,x^n$ are local coordinates on N while $x^1,...,x^m$ are local coordinates on M. Moreover, for any $p \in N$ the differentials $(dx^1)_p,...,(dx^n)_p$ form a basis of T_p^*N and extending them to $(dx^1)_p,...,(dx^m)_p$ gives a basis for T_p^*M. Let y_k denote the coefficient function in front of (dx^k) in decomposition of an arbitrary element of T_p^*M, i.e. $\sum_{i=1}^m y_k(\alpha)(dx^k)_p = \alpha$ for any $\omega \in T_p^*M$.

Then $(x^1, ..., x^m, y^1, ..., y_m)$ is a local chart for T^*M while $(x^1, ..., x^n, y^1, ..., y^n)$ is a local chart for T^*N. In these coordinates, the tautological one-form on T^*M is given by

$$\theta = \sum_{i=1}^{m} y_i dx^i.$$

But then on $T^*(U \cap N)$

$$\mathfrak{i}^*\theta = \theta|_{T^*(U\cap N)} = \sum_{i=1}^{m} y_i|_{T^*(U\cap N)} d\left(x^i|_{T^*(U\cap N)}\right) = \sum_{i=1}^{n} y_i dx^i.$$

Since we worked with an arbitrary adapted chart, we conclude that $i^*\theta$ is the tautological one form on the whole T^*N. □

Corollary 5.5. *Suppose N is a regular submanifold of M and ω is the symplectic form on T^*M. Then $\mathfrak{i}^*\omega$ is the symplectic form on T^*N.*

Proof. Using that $\omega = d\theta$ and commutativity of d with pullback we get that

$$\mathfrak{i}^*\omega = \mathfrak{i}^* d\theta = d(\mathfrak{i}^*\theta).$$

Applying Theorem 5.4 then gives that $\mathfrak{i}^*\omega$ is the symplectic form on T^*N. □

With these two tools, we are ready to prove a theorem that generalizes the theorem that states that symplectomorphisms carry solutions to Hamiltonian systems to solutions to some other Hamiltonian systems.

Theorem 5.6. *Let N be a regular submanifold of M and H a Hamiltonian on T^*M. Suppose $f : T^*N \to T^*P$ is a symplectomorphism and $c : I \to T^*N \subset T^*M$ a curve. Then $c(t)$ is a solution to the H-system on T^*M if and only if $f(c(t))$ is a solution to the $H \circ f^{-1}$ system on T^*P. Additionally, all solutions to the $H \circ f^{-1}$ system on T^*P are of this form.*

Proof. We will prove the theorem by first showing that for $c(t) \in T^*N$ an H-solution, $f(c(t))$ is a $H \circ f^{-1}$-solution and then that for $x(t)$ a $H \circ f^{-1}$-solution, $f^{-1}(x(t))$ is an H-solution on T^*N. Additionally, we will not work in charts with Hamilton's equations but rather directly with the Hamiltonian vector fields.

($\Rightarrow$): Suppose $c(t) \in T^*M$ is an H-solution on M. Thus

$$dc_t\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t\right) = X_H\bigg|_{c(t)},$$

where X_H is the unique vector field on T^*M satisfying $\iota_{X_H}\omega = dH$ with ω the symplectic form on T^*M. Now observe that by the chain rule

$$d(f \circ c)_t\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t\right) = df_{c(t)}\left(dc_t\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t\right)\right) = df_{c(t)}\left(X_H\bigg|_{c(t)}\right).$$

Therefore, if we show that $df_{c(t)}\left(X_H\big|_{c(t)}\right) = X_{H\circ f^{-1}}\big|_{f(c(t))}$ then we can conclude that $f \circ c$ is an $H \circ f^{-1}$ solution. To this end, let ω and η be the symplectic forms on M and P, respectively. By Corollary 5.5, $\mathfrak{i}^*\omega$ is the symplectic form on N and since f is a symplectomorphism

$$\eta = (f^{-1})^*\mathfrak{i}^*\omega = (\mathfrak{i} \circ f^{-1})^*\omega.$$

With this, take $Y \in T_{f(c(t))}T^*P$ and consider the following

$$\begin{aligned}\left(\iota_{df_{c(t)}\left(X_H\big|_{c(t)}\right)}\eta_{f(c(t))}\right)(Y) &= \eta_{c(t)}\left(df_{c(t)}\left(X_H\Big|_{c(t)}\right), Y\right)\\ &= \left((\mathfrak{i}\circ f^{-1})^*\omega\right)_{f(c(t))}\left(df_{c(t)}\left(X_H\Big|_{c(t)}\right), Y\right)\\ &= \omega_{c(t)}\left(d(\mathfrak{i}\circ f^{-1})_{f(c(t))}\left(df_{c(t)}\left(X_H\Big|_{c(t)}\right)\right), d(\mathfrak{i}\circ f^{-1})_{f(c(t))}Y\right).\end{aligned}$$

Applying the chain rule gives

$$\begin{aligned}\left(\iota_{df_{c(t)}\left(X_H\big|_{c(t)}\right)}\eta_{f(c(t))}\right)(Y) &= \omega_{c(t)}\left(d(\mathfrak{i}\circ f^{-1}\circ f)_{c(t)}\left(X_H\Big|_{c(t)}\right), d\mathfrak{i}_{c(t)}\circ df^{-1}_{f(c(t))}Y\right)\\ &= \omega_{c(t)}\left(d\mathfrak{i}_{c(t)}\left(X_H\Big|_{c(t)}\right), d\mathfrak{i}_{c(t)}\circ df^{-1}_{f(c(t))}Y\right).\end{aligned}$$

Now, both $X_H|_{c(t)}$ and $df^{-1}_{f(c(t))}Y$ belong to $T^*_{c(t)}N$. Thus $d\mathfrak{i}_{c(t)}$ acts on them trivially and so

$$\left(\iota_{df_{c(t)}\left(X_H\big|_{c(t)}\right)}\eta_{f(c(t))}\right)(Y) = \omega_{c(t)}\left(X_H\Big|_{c(t)}, df^{-1}_{f(c(t))}Y\right) = \left(\iota_{(X_H)_{c(t)}}\omega_{c(t)}\right)\left(df^{-1}_{f(c(t))}Y\right) = dH_{c(t)}(df^{-1}_{f(c(t))}Y),$$

where we used the definition of X_H. Finally, applying the chain rule yield

$$\left(\iota_{df_{c(t)}\left(X_H\big|_{c(t)}\right)}\eta_{f(c(t))}\right)(Y) = dH_{c(t)}(df^{-1}_{f(c(t))}Y) = d\left(H\circ f^{-1}\right)_{f(c(t))}(Y) = (\iota_{(X_{f^{-1}\circ H})_{f(c(t))}}\eta_{f(c(t))})(Y).$$

By non-degeneracy of η it follows that

$$df_{c(t)}\left(X_H\Big|_{c(t)}\right) = X_{H\circ f^{-1}}\Big|_{f(c(t))}.$$

Thus we have shown that $f(c(t))$ is an $(H\circ f^{-1})$-solution.

$(\Leftarrow)$: Conversely, suppose $x(t)$ is an $(H\circ f^{-1})$-solution. Then (omitting the subscripts)

$$dx\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) = X_{H\circ f^{-1}}.$$

Now the chain rule gives

$$d(f^{-1}\circ x)\frac{\mathrm{d}}{\mathrm{d}t} = df^{-1}X_{H\circ f^{-1}}.$$

But from the previous direction we know that $df(X_H) = X_{H\circ f^{-1}}$. Therefore another application of the chain rule gives

$$d(f^{-1}\circ x)\frac{\mathrm{d}}{\mathrm{d}t} = df^{-1}(df(X_H)) = d(f^{-1}\circ f)X_H = X_H,$$

which allows us to conclude that $f^{-1}(x(t))$ is an H-solution. □

By picking $N = \hat{\mathbb{S}}^n, M = \mathbb{R}^{n+1}, P = \mathbb{R}^n$ and $\sigma_\sharp$ to be the symplectomorphism, we can apply the above theorem to the Φ and $F = \Phi\circ\sigma_\sharp^{-1}$ Hamiltonians and obtain the following.

Corollary 5.7. We have that $(x(t), y(t))$ is a Φ-solution satisfying $\|x(0)\| = 1$ and $\langle x(0), y(0)\rangle = 0$ if and only if $\sigma_\sharp(x(t), y(t))$ is an F-solution. Moreover, all F-solutions are of that form, i.e. $\sigma_\sharp$ induces a bijection

$$\{\Phi\text{-solutions with } \|x(0)\| = 1, \langle x(0), y(0)\rangle = 0\} \xrightarrow{\sigma_\sharp} \{F\text{-solutions}\}$$
$$(x(t), y(t)) \mapsto \sigma_\sharp(x(t), y(t)).$$

Proof. The only detail that we need to prove is that $(x(t), y(t)) \in T^*\hat{\mathbb{S}}^n$ for all t. This follows from Theorem 5.3. □

With this corollary and Theorem 5.3 the following corollary is immediate.

Corollary 5.8. Let $c : I \to \mathbb{S}^n$ be a curve. Then $c(t)$ is a geodesic on $\mathbb{S}^n$ if and only if $\sigma_\sharp(c(t), c^\flat(t))$ is an F-solution.

5.5 Obtaining the Kepler Hamiltonian

In this section, we will show that the Hamiltonian

$$F(x, y) = \frac{(\|x\|^2 + 1)^2}{8}\|y\|^2,$$

is equivalent to the Kepler Hamiltonian H when considered on a specific energy surface. In particular, we will first show that the F-solutions on the energy surface $F = 1/2$ correspond to Kepler orbits on the energy surface $H = -1/2$. We will then generalize this to other energy surfaces.

Remark. Note that $c(t)$ is a geodesic with unit speed, i.e. $\|c'(t)\| = 1$, if and only if it is a Φ-solution on the surface

$$\Phi(c(t), c^\flat(t)) = \frac{1}{2}\|c(t)\|^2\left\|c^\flat(t)\right\|^2 = \frac{1}{2},$$

satisfying $\|c(0)\| = 1$ and $\langle c'(0), \bar{c}(0)\rangle = 0$. This follows from Theorem 5.3. But then $\sigma_\sharp(c(t), c^\flat(t))$ lies on the surface

$$F(\sigma_\sharp(c(t), c^\flat(t))) = \Phi(i({\sigma_\sharp}^{-1}(\sigma_\sharp(c(t), c^\flat(t))))) = \Phi(c(t), c'(t)) = 1/2.$$

Hence geodesics with unit speed correspond to F-solutions on the surface $F = 1/2$.

To start transforming F-solutions to H-solutions, we first define an intermediate Hamiltonian G and show its equivalence to F on an energy surface. We will then show equivalence of G to H on an energy surface.

Lemma 5.7. Let $G : T^*\mathbb{R}^n \to \mathbb{R}$ be the Hamiltonian defined by $G(x, y) = \sqrt{2F(x, y)} - 1$. Then $(x(t), y(t))$ is an F-solution on $F = 1/2$ if and only if it is a G-solution on $G = 0$. *Remark:* Note that $F = 1/2$ implies that $y(t)$ never vanishes.

Proof. First, observe that $G(x, y) = 0$ if and only if $F(x, y) = 1/2$. Moreover, note that

$$\frac{\partial G}{\partial x} = \frac{1}{\sqrt{2F(x, y)}}\frac{\partial F}{\partial x}$$
$$\frac{\partial G}{\partial y} = \frac{1}{\sqrt{2F(x, y)}}\frac{\partial F}{\partial y}.$$

Consequently, $(x(t), y(t))$ solves

$$\dot{x} = \frac{\partial G}{\partial y}$$
$$\dot{y} = -\frac{\partial G}{\partial x}$$

with $G(x(t), y(t)) = 0$ if and only if $F(x(t), y(t)) = 1/2$ and

$$\dot{x} = \frac{1}{\sqrt{2F(x,y)}}\frac{\partial F}{\partial y} = \frac{\partial F}{\partial y}$$
$$\dot{y} = -\frac{1}{\sqrt{2F(x,y)}}\frac{\partial F}{\partial x} = -\frac{\partial F}{\partial x}.$$

In other words, $(x(t), y(t))$ is a G-solution on $G = 0$ if and only if it is an F-solution on $F = 1/2$. □

Lemma 5.8. Let G be defined as in the previous lemma and define the Kepler Hamiltonian $H : T^*\mathbb{R}^n \to \mathbb{R}$ by

$$H(q,p) = \frac{1}{2}\|p\|^2 - \frac{1}{\|q\|}.$$

Then $(x(s), y(s))$ is a G-solution on $G = 0$ if and only if $(q(t), p(t)) = (\bar{y}(t), -\bar{x}(t))$ is an H-solution on $H = -1/2$, where $\bar{x}(t)$ and $\bar{y}(t)$ are reparametrizations of $x(s)$ and $y(s)$ defined by

$$\bar{x}(t) := x(s(t))$$
$$\bar{y}(t) := y(s(t)),$$

with the parameters implicitly related by

$$t(s) = \int_0^s \|y(r)\| dr.$$

Proof. First, recall that

$$G(x,y) = \sqrt{2F(x,y)} - 1 = \frac{(\|x\|^2 + 1)\|y\|}{2} - 1.$$

Additionally, observe that the parameters t and s satisfy

$$\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = \|y(s)\|,$$

and, assuming that $y(s(t)) \neq 0$, the inverse function theorem gives

$$\left.\frac{\mathrm{d}s}{\mathrm{d}t}\right|_t = \frac{1}{\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_{s(t)}} = \frac{1}{\|y(s(t))\|}.$$

Moreover, in our case $y(s(t))$ never vanishes since otherwise $H \neq -1/2$ and $G \neq 0$. Therefore we can always use the inverse function theorem as above. Equipped with those facts, let us proceed to the main proof.

$(\Rightarrow)$: Suppose $(x(s), y(s))$ is a G-solution on $G = 0$ and consider $(q(t), p(t)) = (\bar{y}(t), -\bar{x}(t))$. Then we have

$$\left.\frac{\mathrm{d}q}{\mathrm{d}t}\right|_t = \left.\frac{\mathrm{d}\bar{y}}{\mathrm{d}t}\right|_t = \left.\frac{\mathrm{d}(y \circ s)}{\mathrm{d}t}\right|_t = \left.\frac{\mathrm{d}y}{\mathrm{d}s}\right|_{s(t)}\left.\frac{\mathrm{d}s}{\mathrm{d}t}\right|_t = -\left.\frac{\partial G}{\partial x}\right|_{(x(s(t)),y(s(t)))}\frac{1}{\|y(s(t))\|}.$$

Using the expression for $G(x, y)$ we obtain

$$\left.\frac{\mathrm{d}q}{\mathrm{d}t}\right|_t = -\|y(s(t))\|x(s(t))\frac{1}{\|y(s(t))\|} = -\bar{x}(t) = p(t) = \left.\frac{\partial H}{\partial p}\right|_{p(t),q(t)},$$

thus the first of H-equations is satisfied. Similarly,

$$\left.\frac{\mathrm{d}p}{\mathrm{d}t}\right|_t = -\left.\frac{\mathrm{d}\bar{x}}{\mathrm{d}t}\right|_t = -\left.\frac{\mathrm{d}x}{\mathrm{d}s}\right|_{s(t)}\left.\frac{\mathrm{d}s}{\mathrm{d}t}\right|_t = -\left.\frac{\partial G}{\partial y}\right|_{(x(s(t)),y(s(t)))}\frac{1}{\|y(s(t))\|} = -\frac{\|x(s(t))\|^2+1}{2}\frac{1}{\|y(s(t))\|^2}y(s(t)). \tag{5.6}$$

Now recall that $(x(s), y(s))$ satisfy $G(x(s), y(s)) = 0$ for all s thus

$$0 = G(x(s(t)), y(s(t))) = \frac{(\|x(s(t))\|^2+1)\|y(s(t))\|}{2} - 1.$$

Using $y(s(t)) \neq 0$ we obtain

$$\frac{\|x(s(t))\|^2+1}{2} = \frac{1}{\|y(s(t))\|}.$$

Using this fact, we get that

$$H(p(t), q(t)) = \frac{\|p(t)\|^2}{2} - \frac{1}{\|q(t)\|} = \frac{\|x(s(t))\|^2}{2} - \frac{1}{\|y(s(t))\|} = -1/2$$

and applying it to (5.6)

$$\left.\frac{\mathrm{d}p}{\mathrm{d}t}\right|_t = -\frac{1}{\|y(s(t))\|^3}y(s(t)) = -\frac{1}{\|q(t)\|^3}q(t) = -\left.\frac{\partial H}{\partial q}\right|_{(q(t),p(t))}.$$

Thus $(q(t), p(t))$ is an H-solution on $H = -1/2$.

($\Leftarrow$) : Conversely suppose that $(q(t), p(t)) = (\bar{y}(t), -\bar{x}(t))$ is an H-solution on $H = -1/2$ and consider $(x(s), y(s))$. First, observe that $H = -1/2$ implies that

$$-1/2 = H(q(t), p(t)) = \frac{\|p(t)\|^2}{2} - \frac{1}{\|q(t)\|} = \frac{\|x(s(t))\|^2}{2} - \frac{1}{\|y(s(t))\|},$$

for all t and so

$$G(x(s(t)), y(s(t))) = \frac{(\|x(s(t))\|^2+1)\|y(s(t))\|}{2} - 1 = 0,$$

for all t. As $s(t)$ is a bijection, it follows that $(x(s), y(s))$ lies on $G = 0$. Moreover, denoting the inverse of $s(t)$ by $t(s)$, we get

$$\left.\frac{\mathrm{d}x}{\mathrm{d}s}\right|_s = \left.\frac{\mathrm{d}(x\circ s\circ t)}{\mathrm{d}s}\right|_s = \left.\frac{\mathrm{d}(x\circ s)}{\mathrm{d}t}\right|_{t(s)}\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = -\left.\frac{\mathrm{d}p}{\mathrm{d}t}\right|_{t(s)}\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = \left.\frac{\partial H}{\partial q}\right|_{(q(t(s)),p(t(s)))}\|y(s)\|.$$

Differentiating $H(p, q)$ with respect to q gives

$$\left.\frac{\mathrm{d}x}{\mathrm{d}s}\right|_s = \frac{1}{\|q(t(s))\|^3}q(t(s))\|y(s)\| = \frac{1}{\|y(s(t(s)))\|^3}y(s(t(s)))\|y(s)\| = \frac{1}{\|y(s)\|^2}y(s).$$

Using that $G(x(s), y(s)) = 0$ we obtain that

$$\left.\frac{\mathrm{d}x}{\mathrm{d}s}\right|_s = \frac{\|x(s)\|^2+1}{2}\frac{1}{\|y(s)\|}y(s) = \left.\frac{\partial G}{\partial y}\right|_{(x(s),y(s))},$$

so the first G-equation is satisfied. Similarly,

$$\left.\frac{\mathrm{d}y}{\mathrm{d}s}\right|_s = \left.\frac{\mathrm{d}(y\circ s\circ t)}{\mathrm{d}s}\right|_s = \left.\frac{\mathrm{d}(y\circ s)}{\mathrm{d}t}\right|_{t(s)}\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = \left.\frac{\mathrm{d}q}{\mathrm{d}t}\right|_{t(s)}\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = \left.\frac{\partial H}{\partial p}\right|_{(q(t(s)),p(t(s)))}\|y(s)\|.$$

Differentiating $H(p,q)$ with respect to p gives

$$\left.\frac{\mathrm{d}y}{\mathrm{d}s}\right|_s = \|y(s)\|p(t(s)) = -\|y(s)\|x(s) = -\left.\frac{\partial G}{\partial x}\right|_{(x(s),y(s))}.$$

We conclude that $(x(s),y(s))$ is a G-solution on $G=0$. □

Applying Lemma 5.7 and Lemma 5.8 to our initial observation that geodesics with unit speed correspond to F-solution on $F=1/2$, allows us to conclude that the unit speed geodesics on $\mathbb{S}^n$ correspond to Kepler orbits on $\mathbb{R}^n$ with $H=-1/2$. Let us extend this construction to other negative energies. To this end, suppose $c(t)$ is an arbitrary non-constant geodesic on $\mathbb{S}^n$ and define $\lambda := \|c'(t)\| \in \mathbb{R}_{>0}$. Then we can write $c'(t)=\lambda\hat{c}(t)$ where $\hat{c}(t)$ is a unit vector. By Corollary 5.8, $\sigma_\sharp(c(t),c^\flat(t))$ is an F-solution lying on the energy surface

$$F(\sigma_\sharp(c(t),c^\flat(t))) = \Phi(c(t),c^\flat(t)) = \frac{\|c(t)\|^2\left\|c^\flat(t)\right\|^2}{2} = \frac{\lambda^2}{2}.$$

Defining

$$K(x,y) := \sqrt{2\lambda^2F(x,y)} - \lambda^2 = \lambda\frac{\|x\|^2+1}{2}\|y\| - \lambda^2,$$

we have that $K(x,y)=0$ if and only if $F(x,y)=\lambda^2/2$. Moreover, on $F(x,y)=\lambda^2/2$ we have

$$\frac{\partial K}{\partial x} = \sqrt{\frac{\lambda^2}{2F(x,y)}}\frac{\partial F}{\partial x} = \frac{\partial F}{\partial x}$$
$$\frac{\partial K}{\partial y} = \sqrt{\frac{\lambda^2}{2F(x,y)}}\frac{\partial F}{\partial y} = \frac{\partial F}{\partial y}.$$

Thus F-solutions of $F=\lambda^2/2$ are precisely the K-solutions with $K=0$. Now analogously to Lemma 5.8 we have the following.

Lemma 5.9. With K and H defined above, we have that $(x(s),y(s))$ is a K-solution on $K=0$ if and only if $(q(t),p(t)) = (\lambda\bar{y}(t), -\lambda^{-1}\bar{x}(t))$ is an H-solution on $H=-\frac{1}{2\lambda^2}$, where $\bar{x}(t)$ and $\bar{y}(t)$ are reparametrizations of $x(s)$ and $y(s)$ defined by

$$\bar{x}(t) := x(s(t))$$
$$\bar{y}(t) := y(s(t)),$$

with the parameters implicitly related by

$$t(s) = \int_0^s \lambda^3\|y(r)\|dr.$$

Proof. Following the same steps as before we have

$$\left.\frac{\mathrm{d}t}{\mathrm{d}s}\right|_s = \lambda^3\|y(s)\| \quad\text{and}\quad \left.\frac{\mathrm{d}s}{\mathrm{d}t}\right|_t = \frac{1}{\lambda^3\|y(s(t))\|},$$

where, again, $y(s(t))$ never vanishes as otherwise $K\neq 0$ or $H\neq -\frac{1}{2\lambda^2}$.

($\Rightarrow$) : Suppose $(x(s), y(s))$ is a K-solution on $K = 0$. Then we have

$$\frac{\mathrm{d}q}{\mathrm{d}t}\bigg|_t = \lambda\frac{\mathrm{d}\bar{y}}{\mathrm{d}t}\bigg|_t = \lambda\frac{\mathrm{d}y}{\mathrm{d}s}\bigg|_{s(t)}\frac{\mathrm{d}s}{\mathrm{d}t}\bigg|_t = -\frac{\partial K}{\partial x}\bigg|_{(x(s(t)),y(s(t)))}\frac{1}{\lambda^2\|y(s(t))\|}.$$

Using the expression for $K(x, y)$ we obtain

$$\frac{\mathrm{d}q}{\mathrm{d}t}\bigg|_t = -\lambda\|y(s(t))\|x(s(t))\frac{1}{\lambda^2\|y(s(t))\|} = -\lambda^{-1}\bar{x}(t) = p(t) = \frac{\partial H}{\partial p}\bigg|_{p(t),q(t)},$$

thus the first of H-equations is satisfied. Similarly,

$$\frac{\mathrm{d}p}{\mathrm{d}t}\bigg|_t = -\lambda^{-1}\frac{\mathrm{d}\bar{x}}{\mathrm{d}t}\bigg|_t = -\lambda^{-1}\frac{\partial K}{\partial y}\bigg|_{(x(s(t)),y(s(t)))}\frac{1}{\lambda^3\|y(s(t))\|} = -\frac{\|x(s(t))\|^2+1}{2}\frac{1}{\lambda^3\|y(s(t))\|^2}y(s(t)). \tag{5.7}$$

Now recall that $(x(s), y(s))$ satisfy $K(x(s), y(s)) = 0$ for all s thus

$$0 = K(x(s(t)), y(s(t))) = \lambda\frac{(\|x(s(t))\|^2+1)\|y(s(t))\|}{2} - \lambda^2.$$

Assuming that $y(s(t)) \neq 0$ gives

$$H(p(t), q(t)) = \frac{\|p(t))\|^2}{2} - \frac{1}{\|q(t)\|} = \lambda^{-2}\frac{\|x(s(t))\|^2}{2} - \frac{1}{\lambda\|y(s(t))\|} = -\frac{1}{2\lambda^2}.$$

Moreover, using $K(x, y) = 0$ also gives

$$\frac{\mathrm{d}p}{\mathrm{d}t}\bigg|_t = -\frac{1}{\lambda^2\|y(s(t))\|^3}y(s(t)) = -\frac{1}{\|\lambda y(s(t))\|^3}\lambda y(s(t)) = -\frac{1}{\|q(t)\|^3}q(t) = -\frac{\partial H}{\partial q}\bigg|_{(q(t),p(t))},$$

thus $(q(t), p(t))$ is an H-solution on $H = -\frac{1}{2\lambda^2}$.

($\Leftarrow$) : Conversely suppose that $(q(t), p(t)) = (\lambda\bar{y}(t), -\lambda^{-1}\bar{x}(t))$ is an H-solution on $H = -\frac{1}{2\lambda^2}$. Note that $H = -\frac{1}{2\lambda^2}$ implies

$$-\frac{1}{2\lambda^2} = H(q(t), p(t)) = \frac{\|p(t)\|^2}{2} - \frac{1}{\|q(t)\|} = \lambda^{-2}\frac{\|x(s(t))\|^2}{2} - \frac{1}{\|\lambda y(s(t))\|},$$

for all t and so

$$K(x(s), y(s)) = \lambda\frac{(\|x(s)\|^2+1)\|y(s)\|}{2} - \lambda^2 = 0,$$

for all s. Moreover, denoting the inverse of $s(t)$ by $t(s)$, we have

$$\frac{\mathrm{d}x}{\mathrm{d}s}\bigg|_s = \frac{\mathrm{d}(x\circ s\circ t)}{\mathrm{d}s}\bigg|_s = \frac{\mathrm{d}(x\circ s)}{\mathrm{d}t}\bigg|_{t(s)}\frac{\mathrm{d}t}{\mathrm{d}s}\bigg|_s = -\lambda\frac{\mathrm{d}p}{\mathrm{d}t}\bigg|_{t(s)}\frac{\mathrm{d}t}{\mathrm{d}s}\bigg|_s = \lambda^4\frac{\partial H}{\partial q}\bigg|_{(q(t(s)),p(t(s)))}\|y(s)\|.$$

Differentiating $H(p, q)$ with respect to q gives

$$\frac{\mathrm{d}x}{\mathrm{d}s}\bigg|_s = \lambda^4\frac{1}{\|q(t(s))\|^3}q(t(s))\|y(s)\| = \frac{\lambda^2}{\|y(s(t(s)))\|^3}y(s(t(s)))\|y(s)\| = \frac{\lambda^2}{\|y(s)\|^2}y(s).$$

Using that $K(x(s), y(s)) = 0$ we obtain that

$$\frac{\mathrm{d}x}{\mathrm{d}s}\bigg|_s = \frac{\|x(s)\|^2+1}{2}\frac{\lambda}{\|y(s)\|}y(s) = \frac{\partial K}{\partial y}\bigg|_{(x(s),y(s))},$$

so the first K-equation is satisfied. Similarly,

$$\frac{\mathrm{d}y}{\mathrm{d}s}\bigg|_s = \frac{\mathrm{d}(y\circ s\circ t)}{\mathrm{d}s}\bigg|_s = \frac{\mathrm{d}(y\circ s)}{\mathrm{d}t}\bigg|_{t(s)}\frac{\mathrm{d}t}{\mathrm{d}s}\bigg|_s = \lambda^{-1}\frac{\mathrm{d}q}{\mathrm{d}t}\bigg|_{t(s)}\frac{\mathrm{d}t}{\mathrm{d}s}\bigg|_s = \lambda^2\frac{\partial H}{\partial p}\bigg|_{(q(t(s)),p(t(s)))}\|y(s)\|.$$

Differentiating $H(p, q)$ with respect to p gives

$$\frac{\mathrm{d}y}{\mathrm{d}s}\bigg|_s = \lambda^2\|y(s)\|p(t(s)) = -\lambda\|y(s)\|x(s) = -\frac{\partial K}{\partial x}\bigg|_{(x(s),y(s))}.$$

We conclude that $(x(s), y(s))$ is a K-solution on $K = 0$. $\square$

5.6 Collision Orbits and Extension of the Stereographic Projection

In the preceding sections, we have proven that any non-constant geodesic $c(t)$ on $\mathbb{S}^n$ corresponds in bijective fashion to a Kepler orbit with energy $H = -\frac{1}{2\|c'(t)\|^2}$. The transformations that achieved this can be summarized as

$$(c(t), c'(t)) \mapsto (\sigma(c(t)), g(c(t), c^\flat(t))) \mapsto \left(\left\|c^\flat(t)\right\| g(c(t), c^\flat(t)), -\left\|c^\flat(t)\right\|^{-1} \sigma(c(t))\right). \tag{5.8}$$

We can see that there is an issue with the current construction. Namely, if $c(t)$ is a geodesic that goes through the north pole $(1, 0, ..., 0)$ then the first step of our construction doesn't work since $\sigma(c(t))$ is undefined at the north pole.

To analyze this issue, let us first recall that

$$\sigma(\xi) = \left(\frac{\xi^1}{1-\xi^0}, ..., \frac{\xi^n}{1-\xi^0}\right)$$
$$g(\xi, \eta) = \left(\eta_1(1-\xi^0) + \xi^1\eta_0, ..., \eta_n(1-\xi^0) + \xi^n\eta_0\right).$$

With the help of the geometric interpretation of σ from Lemma 5.3, we see that if $c(t)$ is a geodesic (i.e. great circle) through the north pole $N = (1, 0, ..., 0)$ then the image of $c(t)$ (excluding N) under σ is a line in $\mathbb{R}^n$ that goes through the origin. As $c(t)$ approaches N its image in $\mathbb{R}^n$ travels to infinity along the given line. While when $c(t)$ goes through the south pole $S = (-1, 0, ..., 0)$ its image goes through the origin $0 \in \mathbb{R}^n$. Moreover, keeping in mind that $\|c'(t)\| = \left\|c^\flat(t)\right\|$ is constant, we also have that as $c(t) \to N$ the image $g(c(t), c^\flat(t)) \to 0$ and as $c(t) \to S$ the image $g(c(t), c^\flat(t)) \to 2c^\flat(t)$. Applying the second transformation in (5.8) to $(c(t), c'(t))$ and denoting its image by $(q(t), p(t))$, we see that

$$c(t) \to N \quad \Longrightarrow \quad \begin{cases} q(t) & = \|c'(t)\| g(c(t), c'(t)) \to 0 \\ \|p(t)\| & = -\|c'(t)\|^{-1} \|\sigma(c(t)))\| \to \infty \end{cases}$$
$$c(t) \to S \quad \Longrightarrow \quad \begin{cases} q(t) & = \|c'(t)\| g(c(t), c'(t)) \to 2\|c'(t)\| c'(t) \\ p(t) & = -\|c'(t)\|^{-1} \sigma(c(t))) \to 0 \end{cases}$$

But this behaviour of $(q(t), p(t))$ corresponds exactly to collision orbits of the Kepler problem! To see this more clearly, suppose S is the initial position of $c(t)$ and we watch it travel to N (in any direction). Then we get that our (q, p) system starts at rest, i.e. $p = 0$, and a distance $\|q\| = 2\|c'(t)\|^2$ from the origin. As $c(t) \to N$ the position $q(t)$ approaches 0 while the momentum $p(t)$ diverges in such a way so that $H(q, p) = \|p\|^2/2 - 1/\|q\|$ remains constant and equal to

$$H(q(0), p(0)) = -\frac{1}{\|q(0)\|} = -\frac{1}{2\|c'(t)\|^2}.$$

We conclude that the problematic geodesics $c(t)$ which go through N correspond to collision orbits of the Kepler problem with energy $H = -\frac{1}{2\|c'(t)\|^2}$, where the north pole N maps to the point of the orbit where the particle collides with its centre of attraction causing the momentum to diverge.

We conclude that there is a bijection between non-constant geodesic on $\mathbb{S}^n$ and completed negative energy Kepler orbits on $\mathbb{R}^n$, in which "completed" means that we include the collision point in the corresponding collision orbit.

Chapter 6

Relativistic Corrections: A Prelude to Relativistic Kepler

This chapter serves as a prelude to the final chapter in which we will present some of the results from [Nee+23]. One of them is that, for a specific value of a coupling constant, the orbits of an extremal test particle moving in the background of an oppositely charged extremal Einstein-Maxwell-dilaton black hole are equivalent to Kepler orbits. Therefore bounded orbits are elliptical and hence there is no perihelion precession. What makes this result interesting is that generically, relativistic corrections to two-body problems cause perihelion precession and bounded orbits do not close on themselves. In this chapter, we will see two examples of this phenomenon. Along the way, we will also show how to obtain relativistic corrections to classical systems. The general strategy is to consider a solution of the system in a relativistic theory and then analyze the limit of this solution for small velocities or $c \to \infty$.

The first example we will consider is the electrostatic two-body problem, and the second is the relativistic Kepler problem. In the first problem, we will analyze a charge moving in a stationary Coulomb potential and look at corrections due to special relativity. For that we will only need standard classical electrostatics and basic notions from special relativity — for example, as presented in [Gri17; Zan12]. To solve the second problem, we will first give an overview of the necessary elements of General Relativity (GR) — the reference for this section is [Ton19] — and then consider the Schwarzschild solution in the limit of small velocities. In both examples, we will see that the relativistic corrections cause the perihelion of orbits to processes. We will quantify this effect.

6.1 Relativistic Corrections to Electrostatic Two-body Problem

In this section, we will derive and analyze the solution to the (special) relativistic electrostatic two-body problem in which the first charge is much heavier and fixed at the origin while the second opposite charge moves in the static Coulomb potential. Classically, the equation of motion is

$$\dot{\mathbf{p}} = -\frac{\alpha}{r^3}\mathbf{x}$$

with $\alpha = q^2/4\pi\epsilon_0$. This corresponds exactly to what we did in chapter 2 with $\alpha = k$. Thus the solutions to this problem are conic sections. In particular, bounded orbits are ellipses and so there is no perihelion precession. We will see that making this problem consistent with special relativity induces perihelion precession. Our account is based only on elementary special relativity. More sophisticated approaches which use the Hamilton-Jacobi equation or action-angle coordinates are presented in [LL80] and [Thi13], respectively.

To make the equations of motion consistent with special relativity, we simply replace classical momentum $\mathbf{p} = m\mathbf{v}$ with relativistic momentum $\mathbf{p} = m\gamma\dot{\mathbf{x}}$, where $\gamma = \frac{1}{\sqrt{1-\|\dot{\mathbf{x}}\|^2/c^2}}$. Thus we want to solve the following equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\gamma m\dot{\mathbf{x}}\right] = -\frac{\alpha}{r^3}\mathbf{x},$$

where as usual $r = \|\mathbf{x}\|$. Similarly to the classical case, the relativistic angular momentum $\mathbf{L} = \mathbf{x} \times (m\gamma\dot{\mathbf{x}})$ is conserved as

$$\dot{\mathbf{L}} = \mathbf{x} \times \frac{\mathrm{d}}{\mathrm{d}t}\left[\gamma m\dot{\mathbf{x}}\right] + \dot{\mathbf{x}} \times (\gamma m\dot{\mathbf{x}}) = -\mathbf{x} \times \frac{k}{r^3}\mathbf{x} = 0.$$

Since $\mathbf{L}$ is conserved and both $\mathbf{x}$ and $\dot{\mathbf{x}}$ are perpendicular to it, the motion is planar. Thus we can use polar coordinates and write

$$\mathbf{x} = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 0 \end{bmatrix}.$$

In these coordinates, $\mathbf{L}$ is

$$\mathbf{L} = \gamma m x \times \dot{x} = \gamma m \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 0 \end{bmatrix} \times \begin{bmatrix} \dot{r}\cos(\theta) - r\sin(\theta)\dot{\theta} \\ \dot{r}\sin(\theta) + r\cos(\theta)\dot{\theta} \\ 0 \end{bmatrix} = \gamma m r^2\dot{\theta} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Hence $\dot{\mathbf{L}} = 0$ also implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m\gamma r^2\dot{\theta}\right] = 0.$$

Thus, we can write

$$m\gamma r^2\dot{\theta} = \ell$$

for some constant ℓ. To proceed, we will first write down the total energy solely in terms of the radial variable r. Then we will rewrite it in terms of the Binet variable $u(\theta) = 1/r(\theta)$. Differentiating the resulting energy equation with respect to θ will give us an ODE for u which we will then analyze. To this end, recall that the total relativistic energy E is conserved and given by

$$E = c\sqrt{\gamma^2 m^2\|\dot{\mathbf{x}}\|^2 + m^2c^2} - \frac{\alpha}{r}.$$

Using the conservation of angular momentum and

$$\|\dot{\mathbf{x}}\|^2 = \dot{r}^2 + r^2\dot{\theta}^2,$$

we can write the total energy as

$$E = c\sqrt{\gamma^2 m^2(\dot{r}^2 + r^2\dot{\theta}^2) + m^2c^2} - \frac{\alpha}{r} = c\sqrt{\gamma^2 m^2\left(\dot{r}^2 + \frac{\ell^2}{m^2\gamma^2 r^2}\right) + m^2c^2} - \frac{\alpha}{r}.$$

Note that the γ factor still depends on the velocity $\dot{\mathbf{x}}$ — fortunately, as we will see, this factor cancels out. Now, let us introduce $u(\theta) = 1/r(\theta)$. Observe that

$$\dot{u} = \frac{\mathrm{d}u}{\mathrm{d}\theta}\dot{\theta} = \frac{\mathrm{d}u}{\mathrm{d}\phi}\frac{\ell}{m\gamma r^2} = \frac{\ell}{\gamma m}u^2\frac{\mathrm{d}u}{\mathrm{d}\phi},$$

where we used conservation of angular momentum and chain rule. Consequently,

$$\dot{r} = -\frac{\dot{u}}{u^2} = -\frac{\ell}{\gamma m}\frac{\mathrm{d}u}{\mathrm{d}\phi}.$$

Denoting differentiation with respect to θ by u', the total energy in terms of the Binet variable reads

$$E = c\sqrt{\gamma^2 m^2\left(\frac{\ell^2 u'^2}{m^2\gamma^2} + \frac{\ell^2 u^2}{m^2\gamma^2}\right) + m^2c^2} - \alpha u = c\sqrt{\ell^2\left(u'^2 + u^2\right) + m^2c^2} - \alpha u$$

Taking the potential energy to the left side and squaring the result yields

$$(E + \alpha u)^2 = c^2\ell^2(u'^2 + u^2) + m^2c^4.$$

Expanding the square gives

$$c^2\ell^2u'^2 + (c^2\ell^2 - \alpha^2)u^2 - 2E\alpha u = E^2 - m^2c^4.$$

Finally, by differentiating with respect to θ and dividing by $2u'$ we get

$$c^2\ell^2u'' + (c^2\ell^2 - \alpha^2)u - E\alpha = 0,$$

and so

$$u'' + \left(1 - \frac{\alpha^2}{c^2\ell^2}\right)u - \frac{E\alpha}{c^2\ell^2} = 0.$$

For $|\alpha| \geq c\ell$ the solutions of this equation are unbounded. Concretely, for $|\alpha| > c\ell$ the solution is

$$u = A\cosh\left(\theta\sqrt{\frac{\alpha^2}{c^2\ell^2} - 1} + \theta_0\right) + \frac{E\alpha}{c^2\ell^2}\frac{1}{1 - \alpha^2/c^2\ell^2}$$

while for $|\alpha| = c\ell$ it is

$$u = \frac{E}{\alpha}\theta^2 + B\theta + C,$$

where A, B, C, θ_0 are integration constants. On the other hand, for $|\alpha| < c\ell$ we get bounded solutions that are given by

$$u = A\cos\left(\theta\sqrt{1 - \frac{\alpha^2}{c^2\ell^2}} + \theta_0\right) + \frac{E\alpha}{c^2\ell^2 - \alpha^2}.$$

We can use the freedom to rotate the x-axis to set $\theta_0 = 0$ in which case

$$u = A\cos\left(\theta\sqrt{1 - \frac{\alpha^2}{c^2\ell^2}}\right) + \frac{E\alpha}{c^2\ell^2 - \alpha^2}.$$

Note that in the non-relativistic limit $c^2 \to \infty$ this corresponds to familiar conic sections[1] and so there is no perihelion precession. But, for finite c, after rotating by 2π the orbit doesn't close. In fact, u is maximized and hence r minimized (perihelion) first at $\theta = 0$ but then not at $\theta = 2\pi$ but rather at

$$\theta = \frac{2\pi}{\sqrt{1 - \frac{\alpha^2}{c^2\ell^2}}} \approx 2\pi + \frac{2\pi\alpha^2}{2c^2\ell^2}.$$

In the last step, we applied the binomial approximation and assumed that $\frac{\alpha^2}{c^2\ell^2} \ll 1$. Therefore, to first-order[2], the perihelion precesses by

$$\delta \approx \frac{\pi\alpha^2}{c^2\ell^2}.$$

[1] Keep in mind that E is the relativistic energy so $E/c^2 \to m$ as $c \to \infty$.
[2] We say that a term has n-th order if it is proportional to $(1/c^2)^n$.

6.2 Relativistic Corrections to Kepler Problem

As already mentioned, before diving into the analysis of the Schwarzschild solution let us give a brief overview of GR.

6.2.1 A Crash Course in General Relativity

During undergraduate studies, one often hears about the beauty of the theory of General Relativity (GR). Fortunately, only knowledge of metrics — that we saw in chapter 4 — is needed to glimpse this beauty. Namely, the first main idea behind GR is encapsulated in the innocuous statement "Gravity is geometry". This means that the metric describing the universe's geometry is the gravitational field's source. In other words, what we perceive as gravitational attraction is just objects following geodesics — nothing more. Notably, the fact that geodesics are independent of the rest mass of a particle corresponds to the well-known (weak) equivalence principle — inertial mass equals gravitational mass. Now, if that was it GR wouldn't be also known for its infamous difficulty. The mathematical difficulty comes from Einstein's second great insight[3] that the metric is a dynamic entity governed by the matter content of the universe. This relationship is described by the seemingly simple Einstein field equation which states that

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu},$$

where $G_{\mu\nu}$ is the Einstein tensor, $T_{\mu\nu}$ is the stress-energy tensor, $g_{\mu\nu}$ the metric and Λ is the cosmological constant. This an elegant coordinate-free[4] formula but deriving it takes significant mathematical machinery — it took Einstein 10 years to translate his physical insights into a mathematically rigorous framework — and so we won't pursue it.

As was already mentioned, another thing that one often hears is that Einstein's equations are notoriously difficult to solve. The reason is that the elegant equation that we gave contains, in fact, 16 different nonlinear, coupled partial differential equations. When all the definitions are unpacked, the Einstein equation in terms of the metric $g_{\mu\nu}$ reads

$$\begin{aligned}
&\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu} + \frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\alpha\partial_\nu g_{\mu\beta} - \frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} - \frac{3}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\mu\partial_\nu g_{\alpha\beta} \\
&-\frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3} g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\mu g_{\beta\nu} - \frac{1}{2}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3} g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\nu g_{\mu\beta} + \\
&\frac{1}{4}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3}\sum_{\rho=0}^{3}\sum_{\lambda=0}^{3} g^{\beta\lambda}g^{\alpha\rho}\partial_\nu g_{\alpha\lambda}\partial_\mu g_{\rho\beta} + \frac{1}{4|g|}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\beta|g|\partial_\nu g_{\mu\alpha} - \frac{1}{4|g|}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\beta|g|\partial_\alpha g_{\mu\nu} \\
&-\frac{1}{4|g|}\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} g^{\alpha\beta}\partial_\beta|g|\partial_\mu g_{\alpha\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}.
\end{aligned}$$

Mind you, this needs to hold for all $\mu, \nu = 1, 2, 3, 4$ and so we have 16 such equations that are moreover coupled to each other. It goes without saying that there are only a handful of known solutions to Einstein's equation. One of them is the Schwarzschild metric which is the unique asymptotically flat, spherically symmetric vacuum solution (i.e. $T_{\mu\nu} = 0$) — as stated in Birkhoff's Theorem [Ton19]. We will give it a closer look in the next subsection.

Finally, once one has a solution to Einstein's equation one can check how a test particle moves in such a universe by solving the geodesic equation we already encountered in chapter 5.

[3]The first one — gravity is geometry — can be implemented in Newtonian mechanics. This is known as the Newton-Cartan gravity. Thus it is the second insight that makes GR truly different from Newtonian gravity.

[4]The abstract index notation is used.

However, for a specific metric, it is easier to derive the geodesic equation using variational calculus rather than first to compute the Riemannian connection[5] and then solve $\frac{Dc'(t)}{dt} = 0$. Without going into mathematical details, we state that the geodesic equation is equivalent to Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{x}^\rho}\right) - \frac{\partial L}{\partial x^\rho} = 0,$$

with the Lagrangian $L = \frac{1}{2}g_{\mu\nu}(x)\dot{x}^\mu\dot{x}^\nu$ and supplied with the constraint $L = -\frac{c^2}{2}$ where the dot represents differentiation with respect to the proper time τ. This Lagrangian formalism can also be applied to the example from the previous section. We opted against using it since — contrary to the present case — we already knew the equations of motion. For details on Lagrangian mechanics, the reader is referred to [Ser23b; Ton04] while its relation to GR is shown in [Ton19].

6.2.2 The Schwarzschild Solution and Perihelion Precession

The Schwarzschild metric reads

$$ds^2 = -A(r)dt^2 + A(r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2),$$

where $A(r) := 1 - \frac{2GM}{rc^2}$ and (r, θ, ϕ) are the usual spherical coordinates. This metric is most known for its relation to black holes. We can see that something unusual is happening by observing that the metric diverges at $r = R_s = \frac{2GM}{c^2}$ — this corresponds to the black hole event horizon. It is possible to analyze the behaviour of spacetime around this apparent singularity but this is not of interest to us. We will only look at the region where $r > R_s$, where the Schwarzschild metric correctly describes the curved spacetime created by a star of mass M. Therefore the relativistic counterpart of the Kepler problem is to find the orbit of a test particle moving in this spacetime. In other words, we have to solve the geodesic equation. In what follows, we will first derive this geodesic equation. Then we will check that in the limit $c \to \infty$ it coincides with the Kepler problem. Finally, we will use perturbation theory in some small parameter $\beta \propto 1/c^2$ to find relativistic corrections to the Kepler problem.

To find the associated geodesic equation we first look at the Lagrangian

$$\begin{aligned} L = \frac{1}{2}g_{\mu\nu}(x)\dot{x}^\mu\dot{x}^\nu &= \frac{1}{2}\left(g_{tt}c^2\dot{t}\dot{t} + g_{rr}\dot{r}\dot{r} + g_{\theta\theta}\dot{\theta}\dot{\theta} + g_{\phi\phi}\dot{\phi}\dot{\phi}\right) \\ &= \frac{1}{2}\left(-A(r)c^2\dot{t}\dot{t} + A(r)^{-1}\dot{r}\dot{r} + r^2\left(\dot{\theta}\dot{\theta} + \sin^2(\theta)\dot{\phi}\dot{\phi}\right)\right). \end{aligned}$$

Its Euler-Lagrange equations are

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{t}}\right) - \frac{\partial L}{\partial t} = -\frac{\mathrm{d}}{\mathrm{d}\tau}\left(c^2 A(r)\dot{t}\right) = 0 \tag{6.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(A(r)^{-1}\dot{r}\right) - \frac{\partial L}{\partial r} = 0 \tag{6.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(r^2\dot{\theta}\right) - r^2\sin(\theta)\cos(\theta)(\dot{\phi})^2 = 0 \tag{6.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial \dot{\phi}}\right) - \frac{\partial L}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(r^2\sin^2(\theta)\dot{\phi}\right) = 0. \tag{6.4}$$

To solve these equations, let us follow a strategy similar to the classical Kepler problem. First, we can constrain the motion to the plane $\theta = \pi/2$. We can do this because given any initial

[5] Technically, we would need the Lorentzian connection.

conditions we can rotate our coordinate system so that $\dot{\theta}(0) = 0$ and $\theta(0) = \pi/2$, and then equation (6.3) is solved by $\dot{\theta} = 0, \theta = \pi/2$. Thus the motion remains planar. With this choice, equation (6.4) becomes

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(r^2\dot{\phi}\right) = 0.$$

Thus

$$\ell := r^2\dot{\phi}$$

is conserved. Note that ℓ is just the angular momentum (divided by mass) in polar coordinates (r, ϕ). Similarly, equation (6.1) gives that

$$E = c^2 A(r)\dot{t}$$

is conserved. To proceed, recall that to get the geodesic equation we also have to impose the constraint

$$L = \frac{1}{2}\left(-A(r)c^2\dot{t}\dot{t} + A(r)^{-1}\dot{r}\dot{r} + r^2 + \dot{\phi}\dot{\phi}\right) = -\frac{c^2}{2},$$

where we already used that $\theta = \pi/2, \dot{\theta} = 0$. Using the conserved quantities ℓ and E, we can rewrite the above as

$$-c^2 = -A(r)c^2\left(\frac{E}{c^2A(r)}\right)^2 + A(r)^{-1}\dot{r}^2 + r^2\left(\frac{\ell}{r^2}\right)^2.$$

Multiplying both sides by $\frac{1}{2}A(r)$ gives

$$\frac{1}{2}\dot{r}^2 + \frac{A(r)}{2}\frac{\ell^2}{r^2} + \frac{A(r)}{2}c^2 = \frac{E^2}{2c^2}. \tag{6.5}$$

But this is just the energy (divided by mass) of a particle moving in the effective potential

$$V_{\text{eff}}(r) = \frac{A(r)}{2}\frac{\ell^2}{r^2} + \frac{A(r)}{2}c^2.$$

Since this particle is constrained to a plane and moves in a central potential, we can follow the same strategy as in the previous section and solve for $r(\phi)$ by deriving the Binet equation for this potential. To this end, introduce $u(\phi) = 1/r$ and observe that

$$\dot{u} = \frac{\mathrm{d}u}{\mathrm{d}\phi}\dot{\phi} = \frac{\mathrm{d}u}{\mathrm{d}\phi}\frac{\ell}{r^2} = \ell u^2\frac{\mathrm{d}u}{\mathrm{d}\phi},$$

where we used conservation of angular momentum and chain rule. Consequently,

$$\dot{r} = -\frac{\dot{u}}{u^2} = -\ell\frac{\mathrm{d}u}{\mathrm{d}\phi}.$$

Substituting this into equation (6.5) and multiplying the result by $2/\ell^2$ gives

$$\frac{E^2}{c^2\ell^2} = \left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 + A(1/u)\left(u^2 + \frac{c^2}{\ell^2}\right) = \left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 + \left(1 - \frac{2GM}{c^2}u\right)\left(u^2 + \frac{c^2}{\ell^2}\right).$$

We can simplify this equation by differentiating it with respect to ϕ and subsequently dividing by $2\frac{\mathrm{d}u}{\mathrm{d}\phi}$. Namely, we obtain

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u - \frac{GM}{\ell^2} - \frac{3GM}{c^2}u^2 = 0. \tag{6.6}$$

The non-relativistic limit corresponds to $c^2 \to \infty$ in which case our equation becomes

$$\frac{\mathrm{d}^2 u_0}{\mathrm{d}\phi^2} + u_0 - \frac{GM}{\ell^2} = 0.$$

It is easy to check that a solution to this equation is

$$u_0(\phi) = \frac{GM}{\ell^2}(1 + e\cos\phi).$$

But recalling that $u_0(\phi) = 1/r_0(\phi)$ we see that u_0 describes a conic section! This is a good sign since we recover the classical two-body problem by taking the limit $c^2 \to \infty$ in the relativistic two-body problem. In contrast to the electrostatic two-body problem, to obtain relativistic corrections we will have to resort to perturbative methods since equation (6.6) doesn't have simple exact solutions[6]. To this end, we look at solutions

$$u \approx u_0 + \beta u_1$$

with $\beta \ll 1$ a dimensionless[7] parameter defined by $\beta = \frac{3G^2M^2}{\ell^2c^2}$. If $\beta = 0$ we recover our non-relativistic solution u_0, thus u_1 is a relativistic correction to our problem. To find it we plug in our ansatz $u \approx u_0 + \beta u_1$ into (6.6) and obtain that

$$\left(\frac{\mathrm{d}^2 u_0}{\mathrm{d}\phi^2} + u_0 - \frac{GM}{\ell^2}\right) + \beta\frac{\mathrm{d}^2 u_1}{\mathrm{d}\phi^2} + \beta u_1 - \frac{\beta\ell^2}{GM}(u_0^2 + 2\beta u_0 u_1 + \beta^2 u_1^2) = 0.$$

The expression in the first parentheses vanishes since u_0 is the classical solution. Since u_1 is approximation at order β, we can neglect the terms proportional to β^2 and β^3. Plugging in the previously obtained expression for u_0 we get that u_1 satisfies

$$\frac{\mathrm{d}^2 u_1}{\mathrm{d}\phi^2} + u_1 = \frac{\ell^2}{GM}u_0^2 = \frac{GM}{\ell^2}\left(1 + 2e\cos\phi + e^2\cos^2\phi\right).$$

This equation is solved by

$$u_1 = \frac{GM}{\ell^2}\left(\left(1 + \frac{e^2}{2}\right) + e\phi\sin(\phi) - \frac{e^2}{6}\cos(2\phi)\right),$$

which is the sought-after (first-order) relativistic correction to the classical two-body problem.

The $e\phi\sin\phi$ term makes u_1 non-periodic in ϕ. Consequently, u is not periodic as well and hence the orbit no longer closes. We can see this clearly by investigating the perihelion of the new orbit. Recalling that the perihelion is the point of closest approach, we want to find ϕ such that $1/r(\phi) = u(\phi) \approx u_0 + \beta u_1$ is maximized. To this end, we set the first derivative of $u_0 + \beta u_1$ to 0 and solve for ϕ, this yields

$$-e\sin(\phi) + \beta\left(e\sin(\phi) + e\phi\cos\phi + \frac{e^2}{3}\sin 2\phi\right) = 0. \tag{6.7}$$

We see that $\phi = 0$ solves this equation. Moreover, in the limit $\beta \to 0$ the next solution is $\phi = 2\pi$ ($\phi = \pi$ correspond to minimum of u_0). Therefore for $\beta \ll 1$ the solution to the perturbed problem should be $\phi = 2\pi + \delta$ for some small angle δ. Plugging this ansatz into equation (6.7),

[6] This new difficulty originates from the fact that in (6.6) a relativistic factor comes in front of an additional nonlinear u^2 term while previously relativistic factors only modified the coefficients of the constant and linear terms.

[7] Keep in mind that ℓ is angular momentum per mass.

expressing sines and cosines as Taylor series around $\phi = 2\pi$ and ignoring terms of order $\delta^2, \beta\delta$ or higher, we obtain

$$0 \approx -e\delta + \beta\left(e\delta + e(2\pi+\delta)(1-\delta^2/2) + \frac{e^2}{3}2\delta\right) \approx -e\delta + 2\pi e\beta.$$

We conclude that the perihelion precesses by

$$\delta \approx 2\pi\beta = \frac{6\pi G^2 M^2}{\ell^2 c^2}$$

for each revolution around the centre, as shown in Figure 6.1. Prediction of this orbit precession explained the anomalous precession of Mercury[8] and was one of the first experimental pieces of evidence in favour of General Relativity.

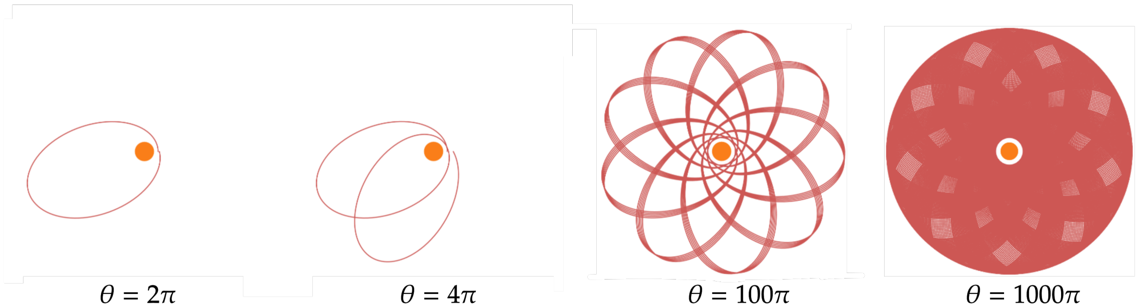

Figure 6.1: Depiction of perihelion precession

Recall that in the previous section, we obtained that perihelion precession in the electrostatic two-body problem due to special relativity is

$$\delta \approx \frac{\pi\alpha^2}{\ell^2 c^2}.$$

Forgetting about the coupling constants, we see that the gravitational precession is 6 times stronger than the electrostatic one. Finally, observe that in the previous section, we obtained exact orbits and only then used $c \to \infty$ to calculate the first-order correction to perihelion precession. On the other hand, in this section, we didn't have exact solutions and worked perturbatively from the beginning.

[8]It was known that considering the pull of other planets can explain the precession but the results obtained in this way were smaller than the experimentally observed precession.

Chapter 7

The Relativistic Kepler Problem

In this final chapter, we will examine the construction given in [Nee+23] and attempt to extend a particular aspect of it. The main idea presented in [Nee+23] is to generalize the Kepler Hamiltonian in such a way that on the energy surface, the resulting Hamiltonian is equivalent to the classical Kepler Hamiltonian. However, the energy surfaces of this generalized Hamiltonian might decompose the phase space differently than those of the classical Kepler Hamiltonian. The authors of [Nee+23] have shown that the generalized Kepler Hamiltonian applies to an extremal test particle in the background of an oppositely charged extremal dilaton-coupled Einstein-Maxwell black hole. But, this statement only holds for a specific value of the dilaton coupling constant — $a = \sqrt{3}$. After giving an overview of this construction, we will prove a generalization of the key theorem from [Nee+23] and then use it in an attempt to characterize the behaviour of solutions to this problem away from the special value of the coupling constant.

7.1 Generalizing the Kepler Hamiltonian

Recall that the usual Kepler Hamiltonian on $T^* \mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$ in terms of the standard cotangent coordinates is given by

$$H_{\text{usual}}(q,p) = \frac{\|p\|^2}{2} - \frac{1}{\|q\|}.$$

The generalization $H(q,p)$ of this Hamiltonian that we will be analyzing is implicitly defined by

$$f(H(q,p)) = \frac{\|p\|^2}{2} - \frac{g(H(q,p))}{\|q\|}, \tag{7.1}$$

for some sufficiently smooth functions $f, g : \mathbb{R} \to \mathbb{R}$. The first thing to observe is that with $f = \mathrm{id}_{\mathbb{R}}$ and $g \equiv 1$, $H(p,q)$ is the usual Kepler Hamiltonian. The main result presented in [Nee+23] is that on each regular level set $H^{-1}(E)$ the flow of $H(q,p)$ is parallel to the usual Kepler flow with a modified gravitation constant. In particular, the bounded orbits of $H(q,p)$ close and are elliptical. Let us present a proof of this statement. To this end, define an auxiliary Hamiltonian

$$K(q,p) := \frac{\|p\|^2}{2} - \frac{g(H(q,p))}{\|q\|}.$$

Its Hamiltonian vector field relates to that of H in the following way.

Lemma 7.1. Let $f, g : \mathbb{R} \to \mathbb{R}$ be smooth functions and suppose there exists a Hamiltonian $H : T^* \mathbb{R}^n \to \mathbb{R}$ satisfying equation (7.1). Then for any regular value E of H the Hamiltonian vector fields X_K and X_H are parallel on $H^{-1}(E)$.

Proof. Well indeed, since (7.1) is satisfied we have that $K = f(H)$. Consequently,

$$X_K = -\frac{\partial K}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial K}{\partial p_i}\frac{\partial}{\partial q^i} = f'(H)\left(-\frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i}\right) = f'(H)X_H.$$

Therefore, X_K and X_H are parallel on $H^{-1}(E)$ with the proportionality constant being $f'(E)$. □

Now fix a regular value E of H. Then $K(p,q)$ restricted to $H^{-1}(E)$ is just the usual Kepler Hamiltonian but with a different gravitational constant, i.e.

$$K(q,p)\Big|_{H^{-1}(E)} = \frac{\|p\|^2}{2} - \frac{g(E)}{\|q\|}.$$

Let us denote the Kepler Hamiltonian with gravitational constant $g(E)$ by $\Phi(q,p)$. Thus the above equation tells us that

$$K(q,p)|_{H^{-1}(E)} = \Phi(q,p)|_{H^{-1}(E)}.$$

However, this is not enough to prove that the Hamiltonian vector fields of K and Φ coincide on $H^{-1}(E)$. For that, we need the following theorem.

Theorem 7.1. Set $\mathcal{E} = \{E \in \mathbb{R} \mid E \text{ is a regular value of } H\}$ and define $J : T^*\mathbb{R}^n \times \mathcal{E} \to \mathbb{R}$ by

$$J(q,p,E) = \frac{p^2}{2} - \frac{g(E)}{\|q\|}.$$

Take any $E \in \mathcal{E}$ then $J_E(q,p) := J(q,p,E)$ is the Kepler Hamiltonian with gravitational constant $g(E)$. Moreover, the Hamiltonian vector fields X_{J_E} and X_H are parallel on $H^{-1}(E)$ and hence X_H is equivalent to the Kepler problem on each of its regular energy surfaces.

Proof. For any $E \in \mathcal{E}$, J_E is indeed equal to the Kepler Hamiltonian with gravitational constant $g(E)$. Moreover,

$$X_{J_E} = -\frac{\partial J_E}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial J_E}{\partial p_i}\frac{\partial}{\partial q^i}.$$

Now observe that $K(q,p) = J(q,p,H(q,p))$. The chain rule then gives

$$\begin{aligned}
X_K &= -\left(\frac{\partial J}{\partial q^i} + \frac{\partial J}{\partial H}\frac{\partial H}{\partial q_i}\right)\frac{\partial}{\partial p_i} + \left(\frac{\partial J}{\partial p_i} + \frac{\partial J}{\partial H}\frac{\partial H}{\partial p^i}\right)\frac{\partial}{\partial q^i} \\
&= -\frac{\partial J}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial J}{\partial p_i}\frac{\partial}{\partial q^i} + \frac{\partial J}{\partial H}\left(-\frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i}\right) \\
&= X_{J_{H(q,p)}} + \frac{\partial J}{\partial H}X_H.
\end{aligned}$$

Restricting to $H^{-1}(E)$ and using Lemma 7.1 we then get

$$f'(E)X_H|_{H^{-1}(E)} = X_{J_E}|_{H^{-1}(E)} + \frac{\partial J}{\partial E}X_H|_{H^{-1}(E)}.$$

We conclude that

$$X_{J_E}|_{H^{-1}(E)} = \left(f'(E) - \frac{\partial J}{\partial E}\right)X_H|_{H^{-1}(E)}.$$

□

Thus we have proven that on each energy surface, the flow of H is parallel to the Kepler flow. In particular, the bounded orbits close and are elliptical.

7.2 The Einstein-Maxwell-dilaton Two Body Problem

This section will briefly discuss the Einstein-Maxwell-dilaton (EMD) theory. For more details, the reader is referred to [HW92]. The EMD theory is a theory that combines general relativity, electromagnetism and the dilaton field, where the dilaton field is just a scalar field. This theory is mostly of theoretical interest combining integer — 0, 1 and 2 — spin fields. For completeness, its Lagrangian is

$$\mathcal{L} = \sqrt{-g}\left(R - 2(\partial\phi)^2 - e^{-2a\phi}F^2\right),$$

where g is the determinant of the metric, R is the Ricci scalar curvature, F is the electromagnetic field strength, ϕ the dilaton field and a a coupling constant. To derive the corresponding equations governing $g_{\mu\nu}, \phi$ and F using variational calculus would take us too far afield. Nevertheless, let us mention that the equation obtained from varying the metric can be seen as the Einstein field equation that we encountered earlier with the stress-energy tensor $T_{\mu\nu}$ being dependent on ϕ, F and $g_{\mu\nu}$. For our purposes these equations are not needed, the interested reader is referred to [HW92]. We are only interested in the fact in the Lagrangian, F and ϕ are coupled via the exponential term with the coupling constant a.

Now, there is a particularly interesting class of solutions to the EMD theory known as extremal black holes. These extremal black holes correspond to the situation where the attractive gravitational and scalar forces balance the repulsive electromagnetic force. Given such an extremal solution, we can obtain the equations of motion of an extremal test particle moving in the background of such a black hole by calculating the Euler-Lagrange equations of an appropriate Lagrangian L. The only difference compared to what we did when examining the Schwarzschild solution in chapter 7 is that now $L \neq \frac{1}{2}g_{\mu\nu}x^\mu x^\nu$ since we also need to take into account the scalar and electromagnetic fields. However, it would be convenient if we could express this problem within the framework of Hamiltonian mechanics which we already acquainted ourselves with in chapter 4 and chapter 5. Fortunately, the problem of solving Euler-Lagrangian equations is equivalent to solving Hamilton's equations with the Hamiltonian being the Legendre transform of the Lagrangian [Ser23b]. For computational details on how to derive the desired Hamiltonian starting from extremal EMD solutions, we refer the reader to [Nee+23]. The result of this derivation is that in the test-mass limit ($m_1 \ll m_2$), the motion of an extremal test particle of mass m_1 moving in the background of extremal EMD black hole with opposite extremal charge and mass m_2 is described by Hamiltonian

$$H(q,p) = m_1 U^{-1}\left(\sqrt{1 + U^{2(a^2-1)/(1+a^2)}\frac{\|p\|^2}{m_1^2}} + 1\right),$$

where

$$U(q) = 1 + (1+a^2)\frac{m_2}{\|q\|}.$$

In the next section, we will show that for $a = \sqrt{3}$ this Hamiltonian is a particular case of the generalization introduced in the previous section.

7.3 The Relativistic Kepler Problem

For $a = \sqrt{3}$ the Hamiltonian for the orbit of a test particle in the background of an extremal black hole is given by

$$H(q,p) = m_1 U^{-1}\left(\sqrt{1 + U\frac{\|p\|^2}{m_1^2}} + 1\right).$$

For $a = \sqrt{3}$, $U(q) = 1 + 4m_2/\|q\|$ and $U^{-1}(q) = \|q\|/(\|q\| + 4m_2)$. Using these relations, our Hamiltonian satisfies

$$\frac{1}{2}\left(\frac{H^2(q,p)}{m_1} - 2H(q,p)\right) = \frac{1}{2}\left(m_1 U^{-2}\left(2 + U\frac{\|p\|^2}{m_1^2} + 2\sqrt{1 + U\frac{\|p\|^2}{m_1^2}}\right) - 2H(q,p)\right)$$
$$= \frac{1}{2}\left(U^{-1}\frac{\|p\|^2}{m_1} + 2(U^{-1} - 1)H(q,p)\right) = \frac{p^2}{2m_1}U^{-1} - \frac{4m_2}{\|q\| + 4m_2}H(q,p).$$

Similarly,

$$\frac{p^2}{2m_1} - \frac{2m_2 H^2(q,p)}{m_1\|q\|} = \frac{p^2}{2m_1} - \frac{2m_2}{m_1\|q\|}m_1^2 U^{-2}\left(2 + U\frac{\|p\|^2}{m_1^2} + 2\sqrt{1 + U\frac{\|p\|^2}{m_1^2}}\right)$$
$$= \frac{p^2}{2m_1} - \frac{2m_2}{m_1\|q\|}U^{-1}\|p\|^2 - \frac{4m_2}{\|q\|}U^{-1}H(q,p) \; = \frac{p^2}{2m_1}U^{-1} - \frac{4m_2}{\|q\| + 4m_2}H(q,p).$$

We conclude that

$$\frac{m_1}{2}\left(\frac{H^2(q,p)}{m_1} - 2H(q,p)\right) = \frac{p^2}{2} - \frac{2m_2 H^2(q,p)}{\|q\|}.$$

Thus by picking $f(x) = \frac{m_1}{2}(\frac{x^2}{m_1} - 2x)$ and $g(x) = 2m_2x^2$, we can conclude from Theorem 7.1 that the flow of $H(q,p)$ with $a = \sqrt{3}$ is parallel to the Kepler flow on each energy surface. As a side note, we are working with units where $c = 1, G = 1$ so f and g are dimensionally consistent. We conclude that a test particle moving in this particular extremal EMD background has conic sections as its orbits.

7.4 Beyond the Existing Results

7.4.1 Generalizing the Generalization of the Kepler Hamiltonian

Before we attempt to analyze what happens in the $a \neq \sqrt{3}$ case, let us first generalize the results from the first section of this chapter. We will use this generalization in our attempts. The Hamiltonian $H(q,p)$ that we will look at is implicitly defined by

$$f(H(q,p)) = \Phi(q,p) + g(H(q,p))\Psi(q,p),$$

for some sufficiently smooth functions $f, g : \mathbb{R} \to \mathbb{R}$ and Hamiltonians Φ, Ψ. In other words, we replace the kinetic and potential energy parts in eq. (7.1) by arbitrary Hamiltonians. Now, define an auxiliary Hamiltonian K as

$$K(q,p) := \Phi(q,p) + g(H(q,p))\Psi(q,p).$$

Then we have that, whenever such implicitly defined H exists, its Hamiltonian vector field X_H is parallel to X_K on each $H^{-1}(E)$. Well indeed, observing that $K = f(H)$ gives us

$$X_K = f'(H)\left(-\frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i}\right) = f'(H)X_H.$$

Therefore, X_K and X_H are parallel on $H^{-1}(E)$ with the proportionality constant $f'(E)$. This result is completely analogous to Lemma 7.1. Similarly, we have the following theorem that we will use in the next section.

Theorem 7.2. Set $\mathcal{E} = \{E \in \mathbb{R} \mid E \text{ is a regular value of } H\}$ and define $\Lambda : T^*\mathbb{R}^n \times \mathcal{E} \to \mathbb{R}$ by

$$\Lambda(q,p,E) = \Phi(q,p) + g(E)\Psi(q,p).$$

Take any $E \in \mathcal{E}$ and define $\Lambda_E(q,p) = \Lambda(q,p,E)$. Then the Hamiltonian vector fields X_{Λ_E} and X_H are parallel on $H^{-1}(E)$ and hence H and Λ_E are equivalent on each $H^{-1}(E)$.

Proof. For any $E \in \mathcal{E}$,

$$X_{\Lambda_E} = -\frac{\partial \Lambda_E}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial \Lambda_E}{\partial p_i}\frac{\partial}{\partial q^i}.$$

Now observe that $K(q,p) = \Lambda(q,p,H(q,p))$. The chain rule then gives

$$\begin{aligned} X_K &= -\left(\frac{\partial \Lambda}{\partial q^i} + \frac{\partial \Lambda}{\partial H}\frac{\partial H}{\partial q_i}\right)\frac{\partial}{\partial p_i} + \left(\frac{\partial \Lambda}{\partial p_i} + \frac{\partial \Lambda}{\partial H}\frac{\partial H}{\partial p^i}\right)\frac{\partial}{\partial q^i} \\ &= -\frac{\partial \Lambda}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial \Lambda}{\partial p_i}\frac{\partial}{\partial q^i} + \frac{\partial \Lambda}{\partial H}\left(-\frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i}\right) \\ &= X_{\Lambda_{H(q,p)}} + \frac{\partial \Lambda}{\partial H} X_H. \end{aligned}$$

Restricting to $H^{-1}(E)$ and using our preliminary results, we get

$$f'(E) X_H|_{H^{-1}(E)} = X_{\Lambda_E}|_{H^{-1}(E)} + \frac{\partial \Lambda}{\partial E} X_H|_{H^{-1}(E)}.$$

We conclude that

$$X_{\Lambda_E}|_{H^{-1}(E)} = \left(f'(E) - \frac{\partial J}{\partial E}\right) X_H|_{H^{-1}(E)}.$$

□

7.4.2 Attempts at Extending Existing Results

Using Theorem 7.2, we will now examine what happens when $a \neq \sqrt{3}$. To do that, we came up with two general approaches. In the first approach, we will try to find a correspondence between $H(q,p)$ for any a and some classical system with potential energy dependent on a. In particular, we will examine the case when the potential energy is a power-law potential with the exponent related to a. An investigation of other a dependent potentials is left for future research. The second approach is to leverage the known relation for $a = \sqrt{3}$ and work perturbatively around it. Apart from the usual perturbation analysis, we will also try to describe this perturbation in terms of Moser's construction — how the geodesics on $\mathbb{S}^n$ behave once we move away from $a = \sqrt{3}$.

Power-law potentials: We were led to investigate power-law potentials because Bertrand's Theorem states that the $1/r$ and r^2 potentials are the only power-law potentials for which bounded orbits are closed. While the analysis of perihelion precession in extremal EMD two-body problem from [Nee+23] shows that the perihelion does not precess only for $a = \sqrt{3}$. This result hints that in the test-particle limit, we might have the same situation. For this reason, we attempted to identify the $a \neq \sqrt{3}$ Hamiltonian with an arbitrary power-law potential Hamiltonian. However, thus far, we haven't found the required functions f, g that would satisfy the needed relation from Theorem 7.2. It might be that finding such a relation is impossible. Our initial attempts at proving such an impossibility theorem failed. Thus, future research is required to determine if the test-particle Hamiltonian $H(q,p)$ corresponds to an a dependent power-law potential.

Taylor expansion around $a = \sqrt{3}$: We know that with $f(x) = \frac{1}{2}(\frac{x^2}{m_1} - 2x)$ and $g(x) = 2\frac{m_2}{m_1}x^2$ the relation

$$f(H(q,p)) = \frac{p^2}{2m_1} - \frac{g(H(q,p)}{\|q\|}$$

is satisfied for $a = \sqrt{3}$. Let us make a Taylor expansion in a of $f(H) + \frac{g(H)}{\|q\|}$ around $a = \sqrt{3}$. With the help of Mathematica, we get

$$\begin{aligned} f(H(q,p)) = &-\frac{g(H(q,p))}{\|q\|} + \frac{p^2}{2m_1} \\ &+ \frac{\sqrt{3}\left(-4m_2\left(4m_2p^2 + q\left(p^2 + 2m_1^2\left(1 + \sqrt{1 + \frac{p^2(4m_2+q)}{m_1^2 q}}\right)\right)\right) + p^2(4m_2+q)^2 \ln\left(1 + \frac{4m_2}{q}\right)\right)}{4m_1(4m_2+q)^2}\delta \\ &+ \mathcal{O}(\delta^2) \\ =: &-\frac{g(H(q,p))}{\|q\|} + \frac{p^2}{2m_1} + e(q,p)\delta + \mathcal{O}(\delta^2), \end{aligned}$$

where $\delta = a - \sqrt{3}$ is the expansion parameter. Applying our Theorem 7.2, we get that to first order in δ the Hamiltonian vector fields X_H and X_{Λ_E} are parallel on $H^{-1}(E)$, where

$$\Lambda_E(q,p) = \frac{p^2}{2m_1} + e(q,p)\delta - \frac{g(E)}{\|q\|}.$$

Hence, away from $a = \sqrt{3}$, the extremal EMD Hamiltonian $H(q,p)$ is parallel to the perturbed Kepler Hamiltonian. Due to the high non-linearity of the perturbation term, we didn't succeed when applying the perturbation techniques from chapter 7.

Moser's construction: In the final attempt, we investigated the flow of J_E (defined in the previous paragraph) by applying Moser's construction from the opposite end. In other, words we pretend that J_E is the Kepler Hamiltonian and see what happens if we run the construction backwards. First,

$$G(x,y) = \|y\|\left(J_E(y,-x) + \frac{1}{2}\right) = \left(\frac{\|x\|^2 + 1}{2} + \delta e(y,-x)\right)\|y\| - g(E)$$

Then to linear order in δ

$$F(x,y) = \frac{(G + g(E))^2}{2} = \frac{(\|x\|^2+1)^2}{8}\|y\|^2 + \delta\frac{\|x\|^2+1}{2}\|y\|^2 e(y,-x).$$

Finally

$$\begin{aligned} \Phi(\xi,\eta) &= \sigma_\sharp^* F = F(\sigma_\sharp(\xi,\eta)) = F(\sigma(\xi), g(\xi,\eta)) \\ &= \frac{1}{2}\|\eta\|^2\|\xi\|^2 + \delta\|\eta^2\|\xi_0 e(g(\xi,\eta), -\sigma(\xi)). \end{aligned}$$

The first part of this Hamiltonian describes geodesic motion on a sphere. Because of the error term, we are no longer guaranteed the restriction of the vector field X_Φ to $T^*\mathbb{S}^n$ belongs to $TT^*\mathbb{S}^n$. Therefore, the trajectory may leave the surface of the sphere even if it started on it. Thus away from $a = \sqrt{3}$, the motion is no longer geodesic and may even leave the sphere. We hoped that this new Hamiltonian would correspond to geodesic motion on some other surface. However, while trying to pursue this idea we run into difficulties stemming from the non-linearity of the error term and the rigidity of Moser's construction which is tailored to spheres — not any other surfaces.

Chapter 8

Conclusion

In conclusion, the research presented in this thesis explored various aspects of the Kepler problem, providing insights into classical mechanics, symmetries, and their connections to conserved quantities. The investigation began by examining the well-established properties of the Kepler problem, demonstrating that its orbits are conic sections, examining Bohlin transformations and introducing Bertrand's Theorem, which highlights the uniqueness of the Kepler potential and the harmonic oscillator potential in having closed bounded orbits.

Furthermore, the thesis delved into the Kepler potential's hidden symmetry, beyond the generic central potential's $SO(3)$ rotational symmetry. This hidden symmetry was revealed by closely examining a higher-dimensional geometric model due to Moser [Mos70] that establishes that the nonconstant geodesics on $\mathbb{S}^n$ correspond to the orbits of an n-dimensional Kepler problem with negative energies. As a result, for the n-dimensional Kepler problem the symmetry group of negative energy orbits is $SO(n+1)$ as compared to $SO(n)$ for a generic central potential in n-dimensions. To prove these results, significant mathematical machinery was introduced. Namely, notions from Differential Geometry such as metrics and connections, as well as tools from Symplectic Geometry and Hamiltonian Mechanics.

Afterwards, we proceeded to explore the relativistic corrections to two specific two-body problems, revealing how these corrections induce perihelion precession. It was then noted that, nevertheless, there are relativistic systems whose bounded orbits remain elliptical, preventing perihelion precession. Specifically, for the coupling constant $a = \sqrt{3}$, the orbits of an extremal test particle in the background of an oppositely charged extremal Einstein-Maxwell-dilaton black hole correspond to classical Kepler orbits [Nee+23]. Attempts at extending these results beyond $a = \sqrt{3}$ led us to a generalization of the key theorem presented in this paper. Using this result, we have tried to find a correspondence between the test-particle Hamiltonian $H(q, p)$ for any a and a Hamiltonian of a classical particle moving in a dependent power-low potential. Thus far, we haven't established whether such a correspondence exists. This problem as well as investigations of other a dependent potentials are left for future research. In our second attempt, we have shown that away from $a = \sqrt{3}$, on each energy surface $H^{-1}(E)$, the orbits of the test-particle Hamiltonian are the same as orbits of a perturbed Kepler Hamiltonian. Finally, applying Moser's construction to this perturbed Hamiltonian gave us that the motion is no longer geodesic and might even leave the sphere. Further research — using more powerful techniques — is needed to extract more information from the highly nonlinear perturbation term and determine if the perturbed motion corresponds to geodesics on some other surface.

Bibliography

[Arn90] Vladimir I. Arnold. *Huygens and Barrow, Newton and Hooke*. Jan. 1990. DOI: https://doi.org/10.1007/978-3-0348-9129-5.

[Bae08] John Baez. *"The Kepler Problem Revisited: The Laplace–Runge–Lenz Vector"*. Accessed: 21-09-2023. 2008. URL: https://math.ucr.edu/home/baez/classical/runge.pdf.

[Bel77] Edward A. Belbruno. "Two-body motion under the inverse square central force and equivalent geodesic flows". In: *Celestial mechanics* 15.4 (1977), pp. 467–476.

[Bel81] Edward A. Belbruno. "Regularization and geodesic flows". In: *Celestial Mechanics and Dynamical Systems (Lecture Notes in Pure and Applied Mathematics)* (1981), pp. 1–11.

[Ber73] Joseph Bertrand. "Théorème relatif au mouvement d'un point attiré vers un centre fixe." French. In: *C. R. Acad. Sci., Paris* 77 (1873), pp. 849–853. ISSN: 0001-4036.

[Boh11] Karl Bohlin. "Note sur le problème des deux corps et sur une intégration nouvelle dans le problème des trois corps". In: *Bulletin astronomique* 28.1 (Jan. 1911), pp. 113–119. DOI: https://doi.org/10.3406/bastr.1911.12893.

[DJ22] Veljko Dmitrašinović and Marija Jankovic. "Perihelion precession in power-law potentials: Hénon's theorem". In: *American Journal of Physics* 90.8 (Aug. 2022), pp. 580–586. DOI: https://doi.org/10.1119/5.0080691.

[Gri17] David J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, June 2017.

[GS77] Victor Guillemin and Shlomo Sternberg. *Geometric asymptotics*. Vol. 14. Mathematical Surveys and Monographs. American Mathematical Society, 1977. DOI: https://doi.org/10.1090/surv/014.

[GS90] Victor Guillemin and Shlomo Sternberg. *Variations on a theme by Kepler*. Vol. 42. Colloquium Publications. American Mathematical Society, 1990.

[GSP14] Herbert Goldstein, John L. Safko, and Charles P. Poole. *Classical Mechanics*. 3rd ed. Pearson Higher Ed, 2014.

[Hén77] Michel Hénon. "A relation in families of periodic solutions". In: *Celestial mechanics* 15.1 (Feb. 1977), pp. 99–105. DOI: 10.1007/bf01229050. URL: https://doi.org/10.1007/bf01229050.

[HJ00] Rachel W. Hall and Kresimir Josic. "Planetary Motion and the Duality of Force Laws". In: *SIAM Review* 42.1 (2000), pp. 115–124. DOI: https://doi.org/10.1137/S0036144598346005.

[HW92] Christoph F.E. Holzhey and Frank Wilczek. "Black holes as elementary particles". In: *Nuclear Physics B* 380.3 (1992), pp. 447–477. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(92)90254-9. URL: https://www.sciencedirect.com/science/article/pii/0550321392902549.

[Jon98] Hugh F. Jones. *Groups, representations and physics*. CRC Press, Jan. 1998.

[Kot11] Dawood Kothawala. “Duality of force laws and conformal transformations”. In: *American Journal of Physics* 79.6 (May 2011), pp. 624–630. DOI: https://doi.org/10.1119/1.3553231.

[LL80] Lev D. Landau and Evgeny M. Lifshitz. *The classical theory of fields*. 4th ed. Vol. 2. Course of theoretical physics. Butterworth-Heinemann, 1980.

[LMS23] Patrick De Leenheer, John Musgrove, and Tyler Schimleck. “A Comprehensive Proof of Bertrand’s Theorem”. In: *SIAM Review* 65.2 (2023), pp. 563–588. DOI: https://doi.org/10.1137/21M1436658.

[Mil83] John Milnor. “On the Geometry of the Kepler Problem”. In: *The American Mathematical Monthly* 90.6 (1983), pp. 353–365. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/2975570 (visited on 10/28/2023).

[Mos70] Jürgen Moser. “Regularization of Kepler’s problem and the averaging method on a manifold”. In: *Communications on Pure and Applied Mathematics* 23.4 (1970), pp. 609–636. DOI: https://doi.org/10.1002/cpa.3160230406.

[Nee+23] Dijs de Neeling et al. *Extremal Black Holes as Relativistic Systems with Kepler Dynamics*. 2023. arXiv: 2303.13291 [gr-qc].

[New87] Isaac Newton. *”Philosophiae Naturalis Principia Mathematica”*. Royal Society, London, 1687.

[Osi77] Yu S. Osipov. “The Kepler problem and geodesic flows in spaces of constant curvature”. In: *Celestial mechanics* 16.2 (1977), pp. 191–208.

[Sag12] Maria Saggio. “Bohlin transformation: The hidden symmetry that connects Hooke to Newton”. In: *European Journal of Physics* 34 (Dec. 2012), p. 129. DOI: https://doi.org/10.1088/0143-0807/34/1/129.

[Ser23a] Marcello Seri. *Analysis on Manifolds*. Version 1.2.5. Lecture Notes. 2023. URL: https://github.com/mseri/AoM/releases.

[Ser23b] Marcello Seri. *Hamiltonian Mechanics*. Version 1.11. Lecture Notes. 2023. URL: https://github.com/mseri/hammech20/releases.

[Sil01] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Vol. 1764. Lecture Notes in Mathematics. Springer, 2001. DOI: https://doi.org/10.1007/978-3-540-45330-7. URL: https://people.math.ethz.ch/~acannas/Papers/lsg.pdf.

[Thi13] Walter Thirring. *Classical Mathematical Physics*. Springer Science & Business Media, Dec. 2013.

[Ton04] David Tong. *Classical Dynamics*. Lecture Notes. 2004. URL: http://www.damtp.cam.ac.uk/user/tong/dynamics/clas.pdf.

[Ton13] David Tong. *Dynamics and Relativity*. Lecture Notes. 2013. URL: http://www.damtp.cam.ac.uk/user/tong/relativity/dynrel.pdf.

[Ton19] David Tong. *General Relativity*. Lecture Notes. 2019. URL: http://www.damtp.cam.ac.uk/user/tong/gr/gr.pdf.

[Tu10] Loring W. Tu. *An Introduction to Manifolds*. 2nd ed. Universitext. Springer New York, 2010. DOI: https://doi.org/10.1007/978-1-4419-7400-6.

[Tu17] Loring W. Tu. *Differential Geometry: Connections, Curvature, and Characteristic Classes*. Graduate Texts in Mathematics. Springer International Publishing, 2017. DOI: https://doi.org/10.1007/978-3-319-55084-8.

[Zan12] Andrew Zangwill. *Modern Electrodynamics*. Cambridge University Press, Dec. 2012. DOI: https://doi.org/10.1017/cbo9781139034777.