Spring 2022

Hidden Symmetries of the Kepler Problem

Julia Kathryn Sheffler
Bard College

Recommended Citation

Hidden Symmetries of the Kepler Problem

A Senior Project submitted to
The Division of Science, Mathematics, and Computing
of
Bard College

by
Julia Sheffler

Annandale-on-Hudson, New York
May, 2022

Acknowledgments

iv

Contents

Abstract

The orbits of planets can be described by solving Kepler's problem which considers the motion due to by gravity (or any inverse square force law). The solutions to Kepler's problem, for energies less then 0, are ellipses, with a few conserved quantities: energy, angular momentum and the Laplace-Runge-Lenz (LRL) vector. Each conserved quantity corresponds to symmetries of the system via Nöther's theorem. Energy conservation relates to time translations and angular momentum to three dimensional rotations. The symmetry related to the LRL vector is more difficult to visualize since it lives in phase space rather than configuration space. To understand the symmetry corresponding to the LRL vector, I use tools from Hamiltonian Mechanics, including the Poisson bracket, flow parameters, and action angle variables to make a visualization of the effect of the symmetry corresponding to the LRL vector. In particular the LRL vector corresponds to four-dimensional rotations in phase space. Though it is beyond the scope of this project I hope to use the solidified understanding of the relationship between conserved quantities and symmetries to simplify the derivation of the probability distribution of semi-major axis given a single direct image of an exoplanet.

1
Introduction

1.1 Discovering Exoplanets

Since the discovery of the first planet outside our solar system in 1992, the number of known exoplanets has increased dramatically with over 4000 confirmed exoplanets (Figure 1.1.1). Kepler is the telescope responsible for the discovery of the majority of the confirmed exoplanets. It uses the transit method which observes how starlight dims as a planet passes in front of its host star, or transits. As the planet crosses its star, it blocks some starlight from reaching our telescopes. So, by detecting a periodic decrease in the intensity of the light received from a given star, we can infer the existence of an exoplanet. The second most common method is the radial velocity method, which measures the gravitational effect of a planet on its star. While we often think of planets orbiting stars, in reality, both planets and stars are orbiting the center of mass of the system. For planets with small mass relative to their stars, the center of mass of of the system is essentially located at the center of the star. However, for large planets the center of mass may instead be outside the star and as a result the star can also be observed orbiting the center of mass. We can detect the motion of the star towards and away from us, called the radial velocity, by measuring the red and blue shift of the starlight.

In Figure 1.1.1, we can see how each method tends to favor particular types of planets. The yellow data points, which indicate planets found by Kepler, tend to have a period shorter than that of Earth at 365 days. Since the transit method requires detecting the planet as it passes in front of its star, we are more likely to find planets that pass in front of their stars more often. Similarly, since the radial velocity method can more easily detect large planets the majority of the blue data points have larger, more Jupiter-like mass.
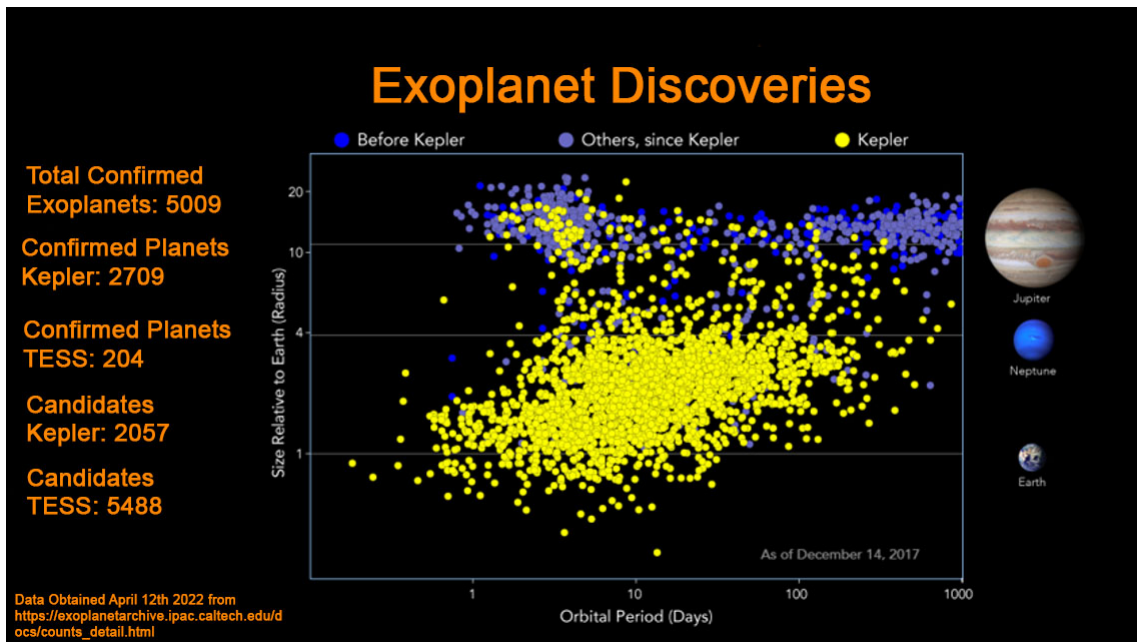
Figure 1.1.1: Graphical representation of the confirmed exoplanets in 2017 distributed based on orbital period (x-axis) and mass (y-axis). While the central panel of the plot only contains data from 2017, the text at left summarizes data on the confirmed and candidate planets as of April 2022 ([12] and [13]). The plot demonstrates the biases of our detection methods. Transit methods, as used by Kepler, tend to find planets with shorter period, while other methods, typically radial velocity, tend to find more massive planets. The summary at left indicates that discoveries from TESS could nearly double the number of confirmed exoplanets.

1.2   Direct imaging of exoplanets

The Transiting Exoplanet Survey Satellite (TESS), which launched in 2018, is a telescope designed specifically to identify transiting planets. So far, TESS has confirmed 204 planets and has detected 5488 candidates. The large and growing record of confirmed exoplanets,

leads many researchers to shift away from searching for exoplanets to studying their properties more directly. Properties of exoplanets including planet-star distance, temperature, mass, chemical composition, and surface albedo (light reflected by the surface), are useful not only for categorizing and comparing various planetary systems, but also for assessing whether a planet may have the potential to develop Earth-like life and provide new laboratories for understanding climate physics. However, both transit and radial velocity methods rely on detecting the effects of planets on their host stars. The indirect nature of such methods means they only provide limited information about the specific characteristics of exoplanets. For instance, transit methods determine the period of a planet's orbit quite precisely, but mass and planet-star distance must be found by other means.

We would like to develop telescopes that can observe exoplanets directly. Rather than inferring their existence via their effects on their stars we would like to be able to take images of the light directly reflected by the planets. Current direct imaging methods are limited to planets extremely far from their stars as the brightness of planets is small in comparison (Figure 1.2.1). In order to see planets closer to their stars we need telescopes with star shields or coronagraphs that can block the light from the stars and observe the light directly reflected by planets (see Figure 1.2.2).

The precise distances at which we can take direct images in visible light currently depend significantly on the method used to block starlight (star shield or coronagraph) and on the distance between us and the planet. To give some scale one of the closest orbit planets detected by direct imaging is Beta Pictoris b which orbits at a distance about nine times that of earth (M. Bruna, Personal Communication, April 28, 2022).

Planets nearer to stars are of particular interest to searches for extraterrestrial life as they will be warmer and potentially be within range for liquid water. The Nancy Grace Roman Space Telescope, which is scheduled for launch in 2027, should be able to find Jupiter like planets at about five times the distance from the sun to Earth. Another telescope, Habex, with a proposed launch date in the mid to late 2030's should have the capability
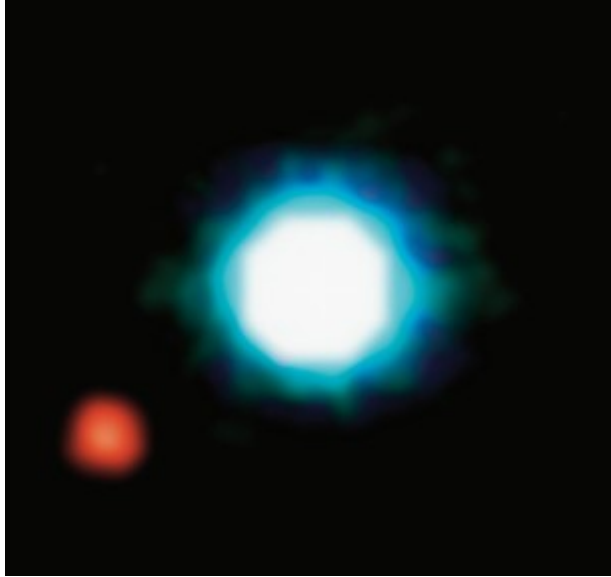
Figure 1.2.1: Direct image of an exoplanet. As mentioned earlier the current direct images all depict planets far from their stars [3]

to detect Earth-like planets in reflected light.(M. Bruna, Personal Communication, April 28, 2022).

1.3   Data from Direct Images

Telescopes equipped to take direct images of Earth-like planets, capture two important quantities for understanding the properties of exoplanets: the total amount of light, or intensity of light, reflected by the planet and a projected distance between the planet and its star.

1.3.1   Intensity of light

The intensity of light varies as the planet orbits. By tracking the intensity over multiple time steps, called epochs, we obtain a plot of intensity of light as a function of time called a light curve. Light curves can be used to approximate the albedo maps of planets, and constrain the orbital parameters that describe the planets orbital path. Albedo maps indicate the reflectivity of different regions of a planet: high albedo indicates more reflective features such as ice, snow, or clouds while low albedo indicates less reflective features such
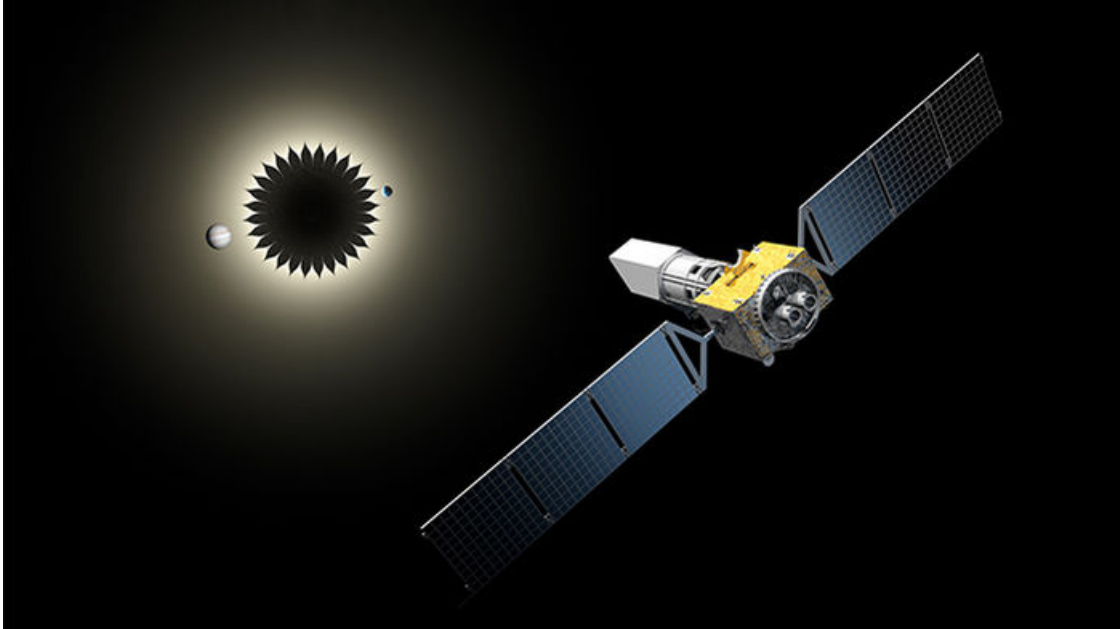
Figure 1.2.2: Depiction of the starshade designed for Habex a proposed telescope designed for studying potentially habitable planets through direct imaging [21]

as bodies of water, or landmasses. We cannot completely reconstruct the albedo map from just a light curve as data is lost in the translation from the two dimensional surface of a planet to the one dimensional light curve, we can generate approximate albedo maps [6]. To find an approximate map we use a Markov Chain Monte Carlo (MCMC) that makes an initial guess for the map, calculates the light curve that would result, compares the guess light curve to the data, and guesses again until the light curves match. Using similar numerical methods and light curve data we can also better constrain the various parameters that describe the path a planet takes along its orbit called orbital parameters. Since a planet will reflect more light towards our telescopes at different positions in its orbit, we can use the change in the intensity of light to constrain its location along its orbit [4].

Additionally, by measuring not only the total intensity, but the intensity at each wavelength, or energy of light, we may be able to use spectral analysis to deduce the chemical composition of the atmosphere of a planet. Each molecule has characteristic energies at which it will reflect light which act a spectral signature. By analysing the intensity of each

wave length we can search for these signatures and potentially determine the presence of various molecules.

1.3.2 Projected Radius

The projected radius of the planet is the measure, in the plane of the telescope between the planet and its star (see Figure 1.3.1). We say this radius is project because it does not necessarily reflect the true three-dimensional distance between the planet and the host star. The planet could be located anywhere along the line of sight of the telescope and give the same projected radius and different three-dimensional distances. The projected distance is particularly useful for understanding the semi-major axis of the orbit which gives a sort of average of the distance a planet is from its star. A more rigorous definition of the semi-major axis is given in Chapter 2. The semi-major axis is particularly important for determining habitability. One of the biggest factors that effects the habitability of a planet is temperature which is determined in large part by the distance between the planet and the star.
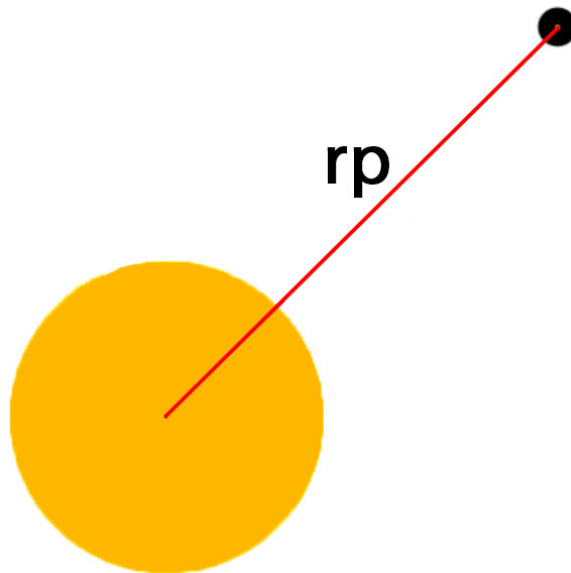
Figure 1.3.1: Depiction of the projected radius r_p of a planet. The star is shown in orange and the planet in black. Notice that the projected radius does not correspond directly to the three dimensional planet-star distance. If the planet is located any where into or out of the page we would still measure the same r_p.

1.4 Single Direct Images

By observing a planet for long periods of time we could reconstruct its orbit precisely. However, the resource intensive nature of direct imaging and the growing number of exoplanets, means we will want to spend only as much time as is necessary observing any given planet. In order to better understand how each image of a planet narrows the possibilities for the orbit of a planet, I have been studying what we can learn about planetary orbits from a single image.

The intensity of reflected light provides insight into many fascinating properties of exoplanets, however analysis methods for intensity requires at least two data points and often more. Brightness information, without a reference, cannot distinguish between a planet that is super reflective, large, or in a position that reflects more light towards our telescope. Such ambiguity makes intensity data essentially useless for analysis of the information contained in a single image. Therefore, in my research, I focus only on the projected radius.

1.5 Probability Distribution for Semi-Major Axis

Since we cannot reconstruct the orbits of planets entirely from only one image, I have instead been working to derive a probability distribution for the semi-major axis denoted a. A probability distribution for semi-major axis gives the probability of that the planet has a semi-major axis between a and $a + da$. In other words by integrating the probability density over a range of a we would obtain the probability that the planet's semi-major axis lies between the bounds of integration.

To find the probability distribution we use Bayesian statistics. In particular, we are looking for the probability of a given some projected radius r_p. We begin with our prior assumptions for the distribution of various parameters of the orbits including inclination, true anomaly, eccentricity, argument of periapsis. The precise definitions of the orbital parameters are discussed in greater detail in Chapter 2. We then use the relationship

between r_p and the orbital parameters to change variables. The result is a joint probability density for each parameter of the orbit. Since we want to know the probability density for only a, we marginalize or integrate over all values of the remaining parameters. We have been able to integrate successfully over two of the parameters. For a ore detailed discussion of our current results see Chapter 6.

Unfortunately, we have been unable to execute some of the integrals necessary to derive the full probability distribution. However, it's possible that our integrals would be more tractable in a nicer set of variables; in particular, a set of variables that respect the symmetries of planetary orbits.

1.6   The Kepler Problem

Fundamental to understanding the symmetries of the orbits of exoplanets is the Kepler problem, which allows us to derive planetary motion from Newton's laws. The Kepler problem considers the motion created by a central inverse square force law which applies to the electromagnetic force and the gravitational force. In the case of planets we are interested in the the gravitational force specifically. In either case, the solutions to Kepler's problem are ellipses, meaning that the motion of two objects interacting via the gravitational force trace elliptical paths. While the position and velocity of the planet will change as it orbits, there are a few quantities that do not change in time, called conserved quantities. For the Kepler problem energy, angular momentum and the Laplace-Runge-Lenz vector are conserved.

The conserved quantities of the Kepler problem can be related to symmetries via Nöther's theorem, which proves that for every conserved quantity there is a corresponding symmetry and for every symmetry there is a corresponding conserved quantity. Since coordinates systems that respect the symmetries of the system could simplify our calculations, we want to understand in-depth each conserved quantity and its corresponding symmetry. Conservation of energy corresponds to a symmetry in time translations, mean-

ing that at a later time, the description for the motion of the planet has not changed. Angular momentum corresponds to a three-dimensional rotational symmetry, meaning we can rotate the orbit in three dimensions and not change its properties. However, for the third conserved quantity, the Laplace-Runge-Lenz (LRL) vector the corresponding symmetry is not obvious. For my project, I explore in-depth the extra and hidden symmetry associated with the LRL vector and generate a visualization of its effect on an orbit.

1.7 Outline

In Chapter 2, I give a brief introduction to Kepler's laws, derive Kepler's laws from Newton's laws, solve the Kepler problem and describe the important parameters for defining planetary orbits. For readers already familiar with the Kepler problem and the geometry of orbits Chapter 2 may be productively skipped. I review the conserved quantities, energy and angular momentum, discuss their associated symmetries in the Kepler problem, define and describe the Laplace-Runge-Lenz (LRL) vector, and state a few useful relationships between the LRL vector, energy and angular momentum in Chapter 3. In Chapter 4, I introduce useful tools from Hamiltonian Mechanics, in particular the Poisson bracket and flow variables. In Chapter 5, I apply the tools described in Chapter 4 to the particular case of the LRL vector and present a visualization of symmetry of the LRL vector that results. In Chapter 6, I return to my discussion of direct imaging and provide a detailed derivation of the partial results for the probability distribution of semi-major axis. Chapter 7 is my conclusion. In the Appendix, I derive the action angle variables for the Kepler problem, give some intuition for generating functions, and give a brief overview and connection to group theory.

2

The Kepler Problem

Johannes Kepler, a German-born theologian turned mathematician, developed Kepler's laws by analyzing an extensive and precise collection of astronomical data[1]. He sought access to the data in order to test his theory of the motion of the planets and quickly learned his initial theories were not accurate. After 8 years of mathematical analysis, he determined two laws that did precisely describe the motion of planets. It takes another nine years to find his third law.

2.1 Kepler's laws

For all of the following laws see Figure 2.1.1.

Kepler's First Law Planets move in elliptical orbits around the sun.

Kepler's Second Law The area of the sector between the planet and the sun traversed in equal time is constant.

Kepler's Third Law The period of a planet's orbit squared is proportional to the distance between a planet and the sun cubed. Mathematically, Kepler's Third Law implies that $T^2 \propto a^3$ where a is the semi-major axis of the orbit.

---

[1]For more information on Kepler's history and background see pp. 62-74 of [20]
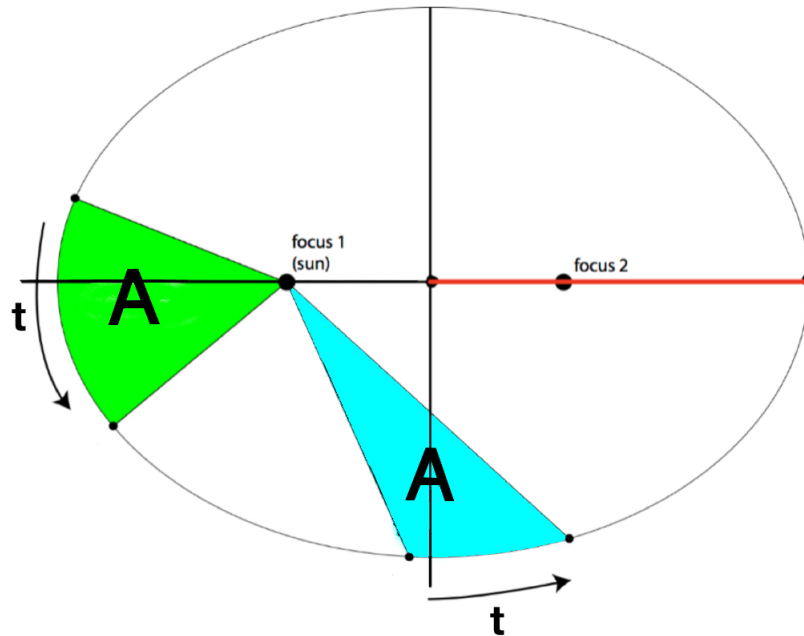
Figure 2.1.1: Elliptical orbit of a planet around the Sun. The two areas traversed in time t labeled A are equal as per Kepler's second law. The semi-major axis, a is indicated in red. This is an edited version of the image found [16].

Kepler's three laws challenged the Aristotelian assumption that planets and stars moved in perfect circles as well as the geocentric view of the universe, revolutionizing astronomy.

2.2 Deriving Kepler's laws from Newton's laws

Kepler's laws precisely predicted planetary motion but without a physical explanation. Issac Newton's three laws of motion and law of universal gravitation created a theoretical framework from which Kepler's laws can be derived[2]. The following derivation is based on work by Carl D. Murray and Alexandre C.M. Correia [11].

To derive Kepler's laws we begin by considering a masses m_1 and m_2 interacting only via the gravitational force (Figure 2.2.1). Combining Newton's second law, $\mathbf{F} = m\mathbf{a}$, and his law of gravity, we have equations for the forces, $\mathbf{F}_1$ and $\mathbf{F}_2$ felt by the planet and star respectively. We define the vector $\mathbf{r}_1$ to be the position of the star and, $\mathbf{r}_2$ to be the position of the planet. The forces $\mathbf{F}_1$ and $\mathbf{F}_2$ represent the gravitational force felt by the

---

[2]For more information on Newton's history and background see pp. 82-89 of [20]
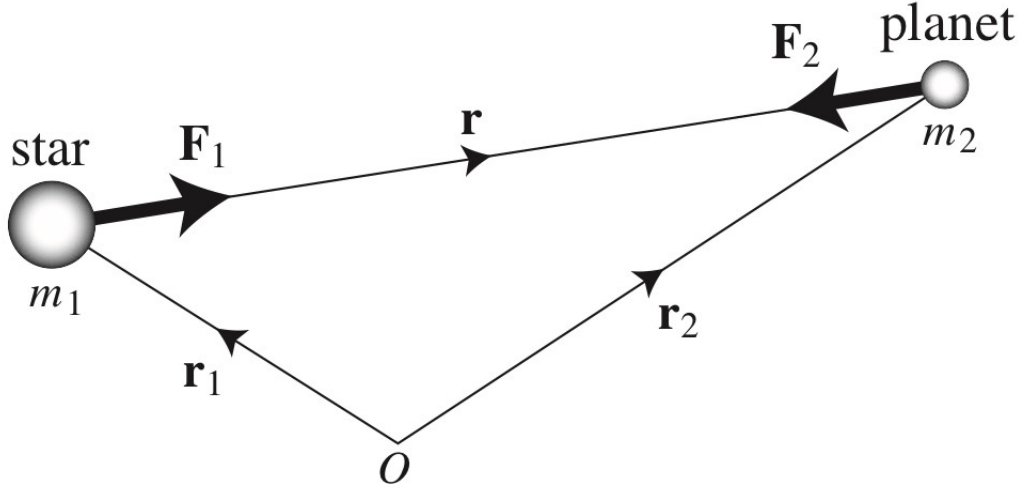
Figure 2.2.1: Diagram of a star and planet positions. The point O represents the origin, the vector $\mathbf{r}_1$ is the position of the star, $\mathbf{r}_2$ the position of the planet. The forces $\mathbf{F}_1$ and $\mathbf{F}_2$ represent the gravitational force felt by the star and planet respectively. The vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ is the position of the planet relative to the star [11].

star and planet respectively. The vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ is the position of the planet relative to the star. Then, we have

$$\mathbf{F}_1 = m_1 \ddot{\mathbf{r}}_1 = +\frac{Gm_1m_2}{r^3}\mathbf{r} \quad \text{and,} \quad \mathbf{F}_2 = m_2 \ddot{\mathbf{r}}_2 = -\frac{Gm_1m_2}{r^3}\mathbf{r}. \tag{2.2.1}$$

Solving for $\ddot{\mathbf{r}}_1$ and $\ddot{\mathbf{r}}_2$, in Eqn. 2.2.1 we have

$$\ddot{\mathbf{r}}_1 = +\frac{Gm_1m_2}{m_1r^3}\mathbf{r} = +\frac{Gm_2}{r^3}\mathbf{r} \quad \text{and,} \quad \ddot{\mathbf{r}}_2 = -\frac{Gm_1m_2}{m_2r^3} = -\frac{Gm_1}{r^3}\mathbf{r}. \tag{2.2.2}$$

Since we want to understand the motion of the planet around the star, we are interested in $\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_2 - \ddot{\mathbf{r}}_1$. Substituting in Eqn. 2.2.2, simplifying and rearranging we have,

$$\ddot{\mathbf{r}} + \frac{G(m_1 + m_2)}{r^3}\mathbf{r} = 0, \tag{2.2.3}$$

which is a differential equation with respect to time. It's easier to solve the differential equation if we first change to a polar coordinate system (r, θ) centered on the star and with an arbitrary choice for the position of $\theta = 0$. In this coordinate system,

$$\mathbf{r} = r\hat{\mathbf{r}}, \tag{2.2.4}$$

$$\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\theta}, \tag{2.2.5}$$

$$\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\mathbf{r}} + \left(\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta}) \right)\hat{\theta}, \tag{2.2.6}$$

where $\hat{\mathbf{r}}$ is the unit vector along the radius vector and $\hat{\theta}$ is the unit vector perpendicular to the radius vector.

2.2.1   Kepler's Second Law

Since Kepler's laws originated from data analysis it is not necessarily easiest to derive them in numbered order. We will first derive his second law, then his first law and finally his third law.

To derive Kepler's Second Law we need to show that the planet traverses equal area in equal time (Figure 2.1.1). The area element in the polar coordinate system is $dA = r'dr'd\theta$ (Figure 2.2.2). The area swept out by the planet in a given time will be a sector centered at the star. The area of a thin sector, dA_s, that spans from $r' = 0$ to $r' = r$, is found by integrating dA with respect to r' from 0 to r:

$$dA_s = \int_0^r r'dr'd\theta = \frac{1}{2}r^2 d\theta, \tag{2.2.7}$$

which also implies,

$$\frac{dA_s}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt}. \tag{2.2.8}$$

Taking the cross product of $\mathbf{r}$ with our inital differential equation, Eqn 2.2.3, we see that $\mathbf{r} \times \ddot{\mathbf{r}} = 0$. Since $\frac{d}{dt}(\mathbf{r} \times \dot{\mathbf{r}}) = \mathbf{r} \times \ddot{\mathbf{r}} = 0$ we have that

$$\mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h}, \tag{2.2.9}$$

where $\mathbf{h}$ is a constant vector. Notice that $\mathbf{h}$ is proportional to the angular momentum $\mathbf{L} = m(\mathbf{r} \times \dot{\mathbf{r}})$.

By substituting the polar definition of $\dot{\mathbf{r}}$, Eqn.2.2.5, into our definition of $\mathbf{h}$ given by Eqn. 2.2.9 we have that

$$\mathbf{h} = r^2\dot{\theta}\hat{\mathbf{z}}. \tag{2.2.10}$$
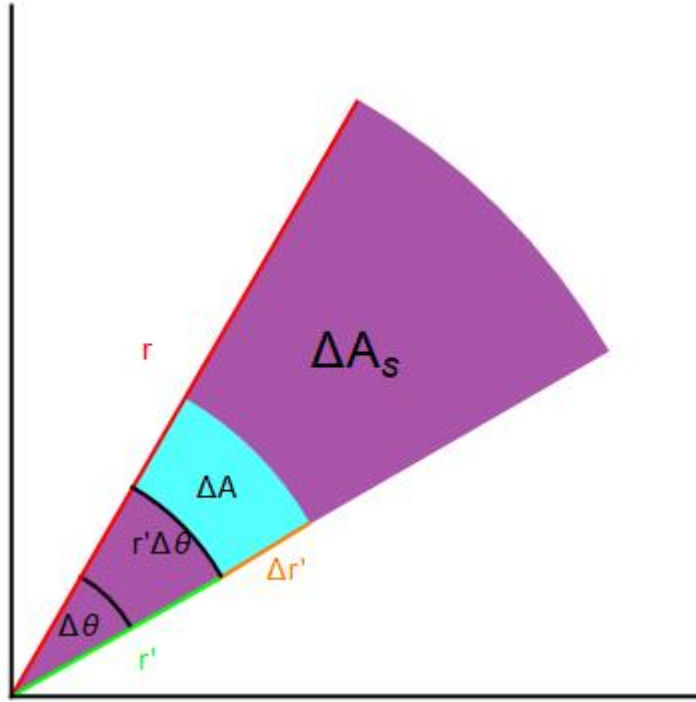
Figure 2.2.2: The area element in polar coordinates, shown in blue, is given by $dA = r'dr'd\theta$. The area of a thin sector, shown in purple, is given by $dA_s = \frac{1}{2}r^2d\theta$.

is constant. So, the magnitude of $\mathbf{h}$, $h = r^2\dot{\theta}$, is also constant. We then substitute h into our equation for rate of change of the area traversed by a planet and find that

$$\frac{dA_s}{dt} = \frac{1}{2}r^2\dot{\theta} = \frac{1}{2}h, \qquad (2.2.11)$$

is a constant, which is exactly Kepler's Second Law.

2.2.2 Kepler's First Law

To derive Kepler's first law, we need to find $r = r(\theta)$. Substituting $\mathbf{r} = r\hat{\mathbf{r}}$ into the initial differential equation, Eqn. 2.2.3, and rearranging we have

$$\ddot{\mathbf{r}} = -\frac{G(m_1 + m_2)}{r^2}\hat{\mathbf{r}}. \qquad (2.2.12)$$

Comparing the $\hat{\mathbf{r}}$ components of $\ddot{\mathbf{r}}$ given in Eqn. 2.2.6, and Eqn. 2.2.12, we have a scalar differential equation,

$$\ddot{r} - r\dot{\theta}^2 = -\frac{G(m_1 + m_2)}{r^2}. \qquad (2.2.13)$$

We now make a clever substitution, $u = 1/r$, in order to reduce the differential equation to that of a forced harmonic oscillator. Solving for r and differentiating with respect to t gives

$$r = \frac{1}{u}, \quad \dot{r} = -\frac{1}{u^2}\dot{u} \quad \text{and,} \quad \ddot{r} = \frac{2\dot{u}^2}{u^3} - \frac{1}{u^2}\ddot{u}. \tag{2.2.14}$$

Using the chain rule we rewrite $\dot{u}$ and $\ddot{u}$ in terms of $\dot{\theta}$ and $\ddot{\theta}$

$$\dot{u} = \frac{du}{dt} = \frac{du}{d\theta}\frac{d\theta}{dt} = \frac{du}{d\theta}\dot{\theta} \quad \text{and,} \quad \ddot{u} = \dot{\theta}^2\frac{d^2u}{d\theta^2} + \frac{du}{d\theta}\ddot{\theta}. \tag{2.2.15}$$

Substituting Eqn 2.2.15 into 2.2.14 we have,

$$\ddot{r} = \frac{2\dot{u}^2}{u^3} - \frac{1}{u^2}\ddot{u}, \tag{2.2.16}$$

$$= \frac{2(\frac{du}{d\theta}\dot{\theta})^2}{u^3} - \frac{1}{u^2}(\dot{\theta}^2\frac{d^2u}{d\theta^2} + \frac{du}{d\theta}\ddot{\theta}), \tag{2.2.17}$$

$$= \frac{2(\frac{du}{d\theta}\dot{\theta})^2}{u^3} - \frac{1}{u^2}\dot{\theta}^2\frac{d^2u}{d\theta^2} - \frac{1}{u^2}\frac{du}{d\theta}\ddot{\theta}. \tag{2.2.18}$$

From the equation for h given by Eqn. 2.2.9 and $r = 1/u$ and since h is constant we know,

$$h = r^2\dot{\theta} = \frac{1}{u^2}\dot{\theta} \quad \text{and,} \quad 0 = \frac{dh}{dt} = \frac{-2}{u^3}\dot{u}\dot{\theta} + \frac{1}{u^2}\ddot{\theta}. \tag{2.2.19}$$

Taking Eqn 2.2.19, we can exchange the derivatives with respect to time in equation 2.2.18, for derivatives with respect to θ to give,

$$\ddot{r} = 2u^2h(d\theta\dot{\theta})^2 - u^2h^2\frac{d^2u}{d\theta^2} - 2uh^2\frac{du}{d\theta}, \tag{2.2.20}$$

$$= -u^2h^2\frac{d^2u}{d\theta^2}. \tag{2.2.21}$$

Now we can substitute equation 2.2.21 into equation 2.2.13 to obtain a differential equation for u with respect to θ,

$$\frac{d^2u}{d\theta^2} + u = \frac{G(m_1 + m_2)}{h^2}. \tag{2.2.22}$$

The differential equation is a second order linear differential equation that also describes the motion of a forced harmonic oscillator. The general solution to the differential equation is,

$$u = \frac{G(m_1 + m_2)}{h^2}(1 + e\cos\theta - \omega). \tag{2.2.23}$$

where e and ω are arbitrary constants that we will later interpret as the eccentricity, e, and the argument of periapsis ω.

Defining $p = h^2/G(m_1 + m_2)$ and substituting $r = 1/u$ we have an equation for r given by

$$r = \frac{p}{1 + e\cos(\theta - \omega)}. \tag{2.2.24}$$

Equation 2.2.24 is the general description of the conic sections which include ellipses, circles, hyperbolas and parabolas. While hyperbolas and parabolas are solutions to our differential equation, they are unbounded. An object that takes a hyperbolic or parabolic path does not orbit but rather is flung out of the system. So, for our purposes we only consider the elliptical (and circular as a special case) solutions.

In the elliptical case, $p = a(1-e^2)$ where a is the semi-major axis and e is the eccentricity. Using our definition $p = h^2/G(m_1 + m_2)$ we then have that

$$h = \sqrt{G(m_1 + m2a(1 - e^2)} \tag{2.2.25}$$

We can also define $\nu = \theta - \omega$ where ν is the true anomaly. A more detailed discussion of the definition of these quantities and some common alternatives is available in section 2.4. We have therefore found an equation for $r(\nu)$:

$$r(\nu) = \frac{a(1 - e^2)}{1 + e\cos\nu}. \tag{2.2.26}$$

We have thus derived Kepler's first law.

2.2.3 Kepler's Third Law

We can now relate the semi-major axis (a) to the period of the orbit (T). We know that the planet sweeps out the entire area of the ellipse, $A = \pi ab$ in time T, where $b^2 = a^2(1-e^2) = ap$. From equation 2.2.11 we know that $hT/2 = \pi ab$. Since $p = h^2/G(m_1 + m_2)$, we have that

$$T^2 = \left(\frac{2\pi ab}{h}\right)^2 = \frac{4\pi^2 a^2 ap}{h^2} = \frac{4\pi^2 a^3}{G(m_1 + m_2)}, \tag{2.2.27}$$

which is exactly Kepler's third law. The period of the planet squared is proportional to the semi-major axis cubed. It's also useful to rewrite Kepler's third law by defining $n = 2\pi/T$. Then, we know that

$$G(m_1 + m_2) = n^2 a^3. \tag{2.2.28}$$

We have now derived all three of Kepler's laws beginning with Newton's laws of motion and law of universal gravitation. However, we have not finished solving the differential equation we began with. Notice that while we have found and equation for r in terms of ν we began with a differential equation with respect to time t. We will see how to find r given t in the next section.

2.3   Solving Kepler's Problem

In the previous section we found r as a function of θ or ν. However, we began our derivation with a differential equation with respect to time. In order to understand how to find the position of a planet at a given time, $r(t)$. It's useful to understand how it's velocity changes as it orbits. To find the velocity as a function of r we note that since $\nu = \theta - \omega$ and ω is a constant we know that $\dot{\nu} = \dot{\theta}$. Then, from 2.2.5, we know that

$$v^2 = \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} = \dot{r}^2 + r^2 \dot{\nu} \tag{2.3.1}$$

Now, we need to find $\dot{r}$ and $r\dot{\nu}$. We start by taking the derivative of equation 2.2.26 to give that

$$\dot{r} = \frac{r\dot{\nu}e \sin\nu}{1 + e\cos\nu}. \tag{2.3.2}$$

Then, since $h = r^2\dot{\theta}$ and $\dot{\nu} = \dot{\theta}$, we see that $h = r^2\dot{\nu}$. Using equation 2.2.28, we also know that $h = \sqrt{G(m_1 + m2a(1 - e^2)} = na^2\sqrt{1 - e^2}$. So,

$$\dot{r} = \frac{na}{\sqrt{1 - e^2}}(e \sin\nu), \tag{2.3.3}$$

and

$$r\dot{\nu} = \frac{na}{\sqrt{1 - e^2}}(1 + e\cos\nu). \tag{2.3.4}$$

We can now substitute equations 2.3.3 and 2.3.4 into equation 2.3.1 to give,

$$v^2 = \frac{n^2 a^2}{1 - e^2}(1 + 2e\cos\nu + e^2) \tag{2.3.5}$$

Finally, we can simplify further to find v^2 as a function of r,

$$v^2 = G(m_1 + m_2)(\frac{2}{r} - \frac{1}{a}) \tag{2.3.6}$$

We can now substitute equations 2.2.26, 2.3.6 and 2.3.4 into equation 2.3.1 and rearrange to get

$$\dot{r}^2 = n^2 a^3 (\frac{2}{r} - \frac{1}{a}) - \frac{n^2 a^4 (1 - e^2)}{r^2}. \tag{2.3.7}$$

Then, simplifying

$$\dot{r}^2 = \frac{na}{r}\sqrt{a^2 e^2 - (r - a^2)} \tag{2.3.8}$$

Notice that 2.3.8 is a differential equation for r with respect to t. To solve this differential equation we introduce a new variable E called the eccentric anomaly, defined by

$$r = a(1 - e\cos E) \tag{2.3.9}$$

Making the change of variables our differential equation becomes

$$\dot{E} = \frac{n}{1 - e\cos E} \tag{2.3.10}$$

the solutions to which can be found separation of variables giving,

$$n(t - t_0) = E - e\sin E \tag{2.3.11}$$

where t_0 is an arbitrary constant of integration. We use the boundary condition $E = 0$ when $t = t_0$. We now define yet another variable M called the mean anomaly, defined by

$$M = n(t - t_0) = E - e\sin E \tag{2.3.12}$$

where t_0 is the time at which the planet is closest to its star. The equation $M = E - e\sin E$ is called Kepler's equation and must be solved numerically.

We can therefore find r at some time t by finding M at that time, solving for E via equation 2.3.12 and then using equation 2.3.9 to find r.

We now have an algorithm for determining the position of the planet as a function of time. In the next section I will overview the parameters we use to define the orbits of planets.

2.4   Parameters of Ellipses

It's now useful to explain the parameter's we use to describe the elliptical orbit of a planet. Here I describe the typical sets of variables we use and a few common alternatives. There are 2 intrinsic parameters which are unaffected by the location of the observer, semi-major axis (a) and eccentricity (e). We also have 3 extrinsic variables, inclination (i), argument of periapsis (ω), and longitude of ascending node Ω. Finally, we have the true anomaly (ν) which describes where in its orbit the planet is located. Any readers already familiar with orbital parameters may skip this section. It's also useful to note that while I describe all the parameters along with a few alternatives those used most often in the rest of my project are semi-major axis, true anomaly, and eccentricity.

2.4.1   Intrinsic Parameters

The intrinsic parameters define the particular shape of the elliptical orbit and are independent of the position of the observer. As mentioned briefly earlier the semi-major axis, (a), is the distance from the center of the ellipse to the farthest point on the ellipse. The semi-minor axis is the distance from the center to the closest point on the ellipse (see Figure 2.4.1). The eccentricity gives us a measure of how close an ellipse is to a circle with $e = 0$ corresponding to a circular orbit and $e = 1$ corresponding to a linear "orbit". We can calculate the eccentricity from the semi-major and semi-minor axes using the formula $e = \sqrt{1 - \frac{b^2}{a^2}}$, where a is the semi-major and b is the semi-minor axis. Two of these three parameters are necessary to pick out precisely the shape of a particular orbit.
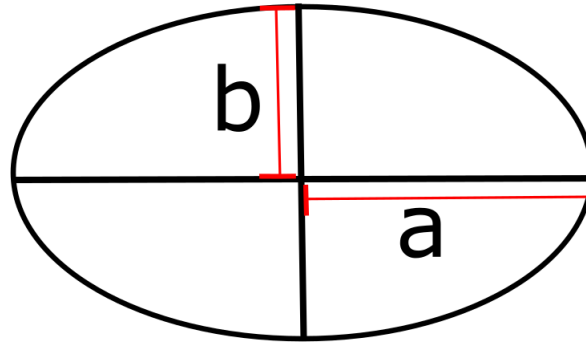
Figure 2.4.1: Diagram of the semi-major axis (a) and the semi-minor axis (b) of an ellipse.

2.4.2 Extrinsic parameters

The extrinsic parameters indicate how the orbit is oriented in three dimensional space with respect to the observer. In order to define the extrinsic parameters we must first choose a reference plane and a reference direction. For studying the orbits of exoplanets, we typically choose the plane visible to the observer (called the sky plane) as the reference plane. In other words, the vector that points along the line of sight of the observer is perpendicular to the reference plane. The reference direction is chosen within the reference plane and is typically take to be in the direction of the north pole, the y-axis in the sky plane. It's also useful to define periapsis which is the point at which the planet is closest to the star, and apoapsis where the planet is farthest form its star.

Inclination: The inclination determines the plane in which the orbit lies. In particular, it is the angle between the reference plane and the plane of the orbit (see Figure 2.4.2). To give a few concrete examples, an orbit with $i = 0$ is face-on, meaning the entire orbit lies in the reference plane. On the other extreme, $i = \frac{\pi}{2}$ is edge on, meaning the orbit lies in a plane perpendicular to the reference plane.

Longitude of the Ascending Node and Argument of Periapsis Given our reference directions and the inclination, we know the plane in which the ellipse lies. However, we do not know how it's oriented within that plane. For instance the ellipse could be such that the longer end or shorter end is angled toward us. We first identify two the two
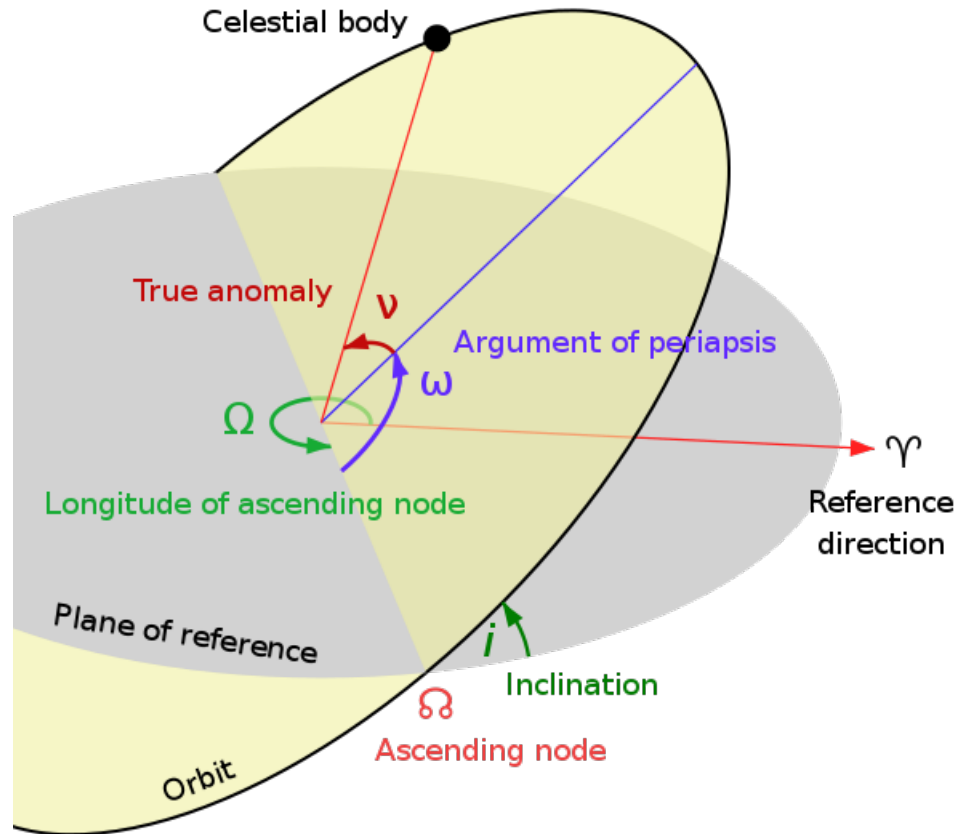
Figure 2.4.2: Illustration of all extrinsic orbital parameters and the true anomaly. We can see that $i = 0$ means the orbit lies entirely in the reference plane, while $i = \frac{\pi}{2}$ would give an orbit perpendicular to the reference plane. We can see that if we shift the orbit within its own plane, we would shift Ω along the orbit, and by rotating the orbit within its own plane, we would change ω. Moving the planet along the orbit changes ν [10].

locations (assuming non-zero inclination), at which the planet passes through the reference plane. The ascending node is where the planet moves through the reference planet in the same direction as the reference direction and the descending node where it moves opposite the reference direction. The longitude of the ascending node, Ω, fixes how the orbit is shifted with respect to reference plane. Precisely, it is the angle from the given reference direction to the ascending node (see Figure 2.4.2). The argument of periapsis (ω) indicates how the ellipse is oriented within the orbital plane. It is the angle, within the orbital plane from the ascending node to periapsis (see Figure 2.4.2).

2.4.3 Time parameters

With the above parameters we have fixed the orbit the planet takes but have yet to describe where the planet is located along its orbit. The position of the planet could be tracked in time, but we know the position of the planet is periodic. So, it is often more convenient to choose a parameter that is also periodic. There are three common conventions, true anomaly (ν), eccentric anomaly (E), mean anomaly (M).

True Anomaly: The easiest of the three parameters to understand geometrically is the the true anomaly (ν), occasionally in other papers denoted f. It is the angle from periapsis, about the star, to the current location of the planet (see Figure 2.4.3).
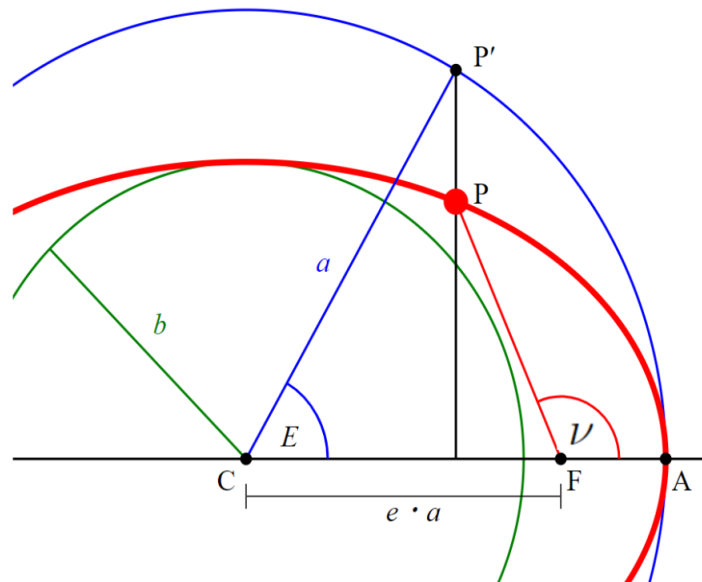
Figure 2.4.3: A diagram to show the geometric interpretation of ν, and E. The mean anomaly does not have a simple geometric interpretation so it does not appear in this figure. The red ellipse represents the orbit of the planet and point P is the current location of the planet. The point P' is on a circle with radius equal to the semi-major axis of the orbit and is positioned vertically above (or below) the point P. We see that ν is the angle from periapsis (point A) to the location of the planet [5]

Eccentric Anomaly The eccentric anomaly (E) can also be seen geometrically but is more difficult to understand intuitively. To define the eccentric anomaly, we consider a circle with radius equal to the semi-major axis of our ellipse and centered at the center of the ellipse. We then consider a vertical projection of our planet's position onto the

circle and call E the angle from periapsis to the projected point on the circle 2.4.3. Mathematically, the relationship between ν and E is given by the equation

$$\nu = 2\arctan\left(\sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right)\right) \tag{2.4.1}$$

where e is the eccentricity of the elliptical orbit.

Mean Anomaly: The mean anomaly (M) is mathematically simpler, but difficult to illustrate geometrically. It directly relates time to the periodicity of the orbit, in particular, $M = \frac{2\pi}{T}(t-\tau)$ where T is the period of the elliptical orbit, t_0 is the time at periapsis, and t is the current time. Intuitively, M describes how much of the full orbital period the planet has completed. Imagining moving around a circle with the same period as the ellipse and at constant speed, then we call M the angle from periapsis to our location on the circle.

While the definition of the mean anomaly is simple, the relationship between M and E (and therefore ν) is transcendental meaning it cannot be solved algebraically. In particular,

$$E = M + e\sin(E) \tag{2.4.2}$$

where e is the eccentricity. Note we could find the relationship between M and ν by solving equation 2.4.1 for E and substituting into equation 2.4.2.

In this chapter we have derived Kepler's laws, solved Kepler's problem, and understood the parameters we typically use to describe the orbits of planets. In Chapter 3 we begin our discussion about the relationship between conserved quantities and symmetries.

3
Symmetries and Conserved Quantities

The quantity h, fundamental to deriving Kepler's laws in Chapter 2, exemplifies the more general category of conserved quantities which do not change in time. In other words, their derivatives with respect to time are zero. Conserved quantities are useful not only for deriving Kepler's laws, but also for understanding the symmetries of the Kepler problem. Up to this point I have left the understanding of symmetries up to intuition but in this chapter will give a more mathematical description of symmetries alongside some examples via the harmonic oscillator. Since symmetries can simplify calculations and generate a better understanding of the solutions, we start with a more rigorous definition of symmetry. We then discuss the conserved quantities of the Kepler problem and their corresponding symmetries. The symmetries corresponding to angular momentum and energy can be intuitively visualized without extensive mathematics so I give only an informal explanation. The full mathematics, explained in depth in Chapter 4 and applied explicitly in Chapter 5 to the Laplace-Runge-Lenz (LRL) vector, can be used analogously to derive the symmetries of energy and angular momentum. For readers already familiar with conserved quantities and their associated symmetries skipping this chapter may be productive.

3.1   Symmetries

When we discuss the symmetries of a system in physics we refer to sets of transformations that leave the "orbit" unchanged. Since transformations imply that I changed something we need to understand what we mean when we say the "orbit" is left unchanged.

In order to illustrate what we do mean by unchanging transformations, we will look first at the simpler case of the one-dimensional harmonic oscillator. The dynamics of the harmonic oscillator is given by the function,

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{3.1.1}$$

The function H is called the Hamiltonian. We will discuss it more in-depth in Chapter 4 but for now, we can think of it as a function whose output is the total energy.

We take the simplest case where $m = \omega = 1$ giving

$$H(x, p) = \frac{1}{2}(x^2 + p^2). \tag{3.1.2}$$

Notice that Eqn. 3.1.2 is the equation for a circle, not in our typical Cartesian (x, y)-plane but rather an (x, p)-space. The (x, p) space is called the phase space and turns out to be a very important notion for understanding the symmetries of the Kepler problem, as well as dynamical (time-dependent) systems in general. Here the phase space trajectory is simply a circle about the origin as seen in Figure 3.1.1.

We now consider making the following transformation from the coordinates x, p to $\tilde{x}, \tilde{p}$ such that

$$\tilde{x} = x \cos\phi + p \sin\phi \quad \text{and} \quad \tilde{p} = p \cos\phi - x \sin\phi. \tag{3.1.3}$$

Notice that this transformation is a rotation in phase space, which comparing to the plot should intuitively be a symmetry of the system.

Let's find the energy of our transformed orbit. First we will need to invert the transformation above to find x and p in terms of $\tilde{x}$ and $\tilde{p}$ giving

$$x = \tilde{x} \cos\phi - \tilde{p} \sin\phi \quad \text{and} \quad p = \tilde{p} \cos\phi + \tilde{x} \sin\phi. \tag{3.1.4}$$
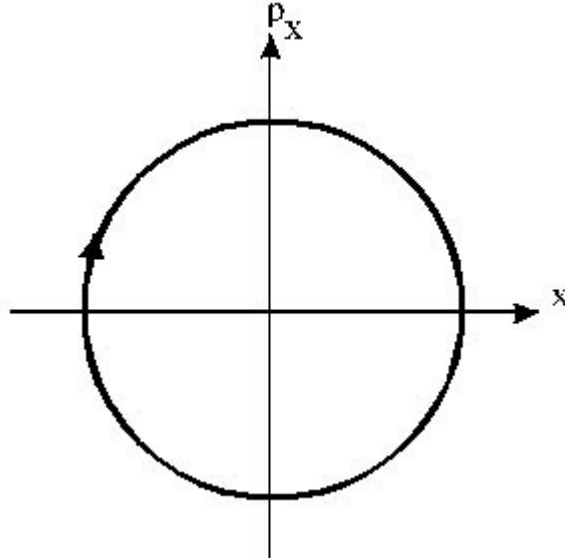
Figure 3.1.1: The phase space of the harmonic oscillator with m=ω=1. It is important to note that as this is a one dimensional example. The choice of energy fixes the exact path in phase space. However, in higher dimensional examples the surface of constant energy E may be more then one dimensional and more variables would need to be fixed to determine the exact phase space trajectory [19].

We can then find the energy of this transformed harmonic oscillator by substituting in our transformations of p and x to get,

$$H(x,p) = \frac{1}{2}(x^2 + p^2) \tag{3.1.5}$$

$$= \frac{1}{2}((\tilde{x}\cos\phi - \tilde{p}\sin\phi)^2 + (\tilde{p}\cos\phi + \tilde{x}\sin\phi)^2) \tag{3.1.6}$$

$$= \frac{1}{2}(\tilde{x}^2\cos^2\phi - 2\tilde{x}\tilde{p}\cos\phi + \tilde{p}^2\sin^2\phi) + (\tilde{p}2\cos^2\phi + 2\tilde{x}\tilde{p}\cos\phi\sin\phi\tilde{x}^2\sin^2\phi) \tag{3.1.7}$$

$$= \frac{1}{2}(\tilde{x}^2 + \tilde{p}^2). \tag{3.1.8}$$

Notice that the equation that describes the dynamics of the system hasn't been altered. We have just transformed the names of our coordinates. For any arbitrary transformation this would not be the case. For instance, if instead we took the transformation such that

$$x = x'\cos\phi + p'\sin\phi \quad \text{and} \quad p = p'\cos\phi + x'\sin\phi \tag{3.1.9}$$

Then, we see that

$$H(x, p) = \frac{1}{2}(x^2 + p^2) \tag{3.1.10}$$

$$= \frac{1}{2}((x' \cos \phi + p' \sin \phi)^2 + (p' \cos \phi + x' \sin \phi)^2) \tag{3.1.11}$$

$$= \frac{1}{2}(x'^2 \cos^2 \phi + 2x'p' \cos \phi + p'^2 \sin^2 \phi) + (p'^2 \cos^2 \phi + 2x'p' \cos \phi \sin \phi x'^2 \sin^2 \phi) \tag{3.1.12}$$

$$= \frac{1}{2}(x'^2 + 4x'p' \cos \phi + p'^2). \tag{3.1.13}$$

The Hamiltonian, which defines our dynamics, does not keep the same form in the x', p' transformation. In particular we have an additional $4x'p' \cos \phi$ term. Transformations like $\tilde{x}, \tilde{p}$ are symmetries of the dynamical system.

We can understand the symmetries of the Kepler problem in a similar way. The energy that defines the dynamics of the orbit in phase space for a planet in three-dimensions in Cartesian coordinates is

$$H(\boldsymbol{x}, \boldsymbol{p}) = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} - \frac{k}{\sqrt{x^2 + y^2 + z^2}}. \tag{3.1.14}$$

We first notice that the phase space for the 3D Kepler problem is 6 dimensional, consisting of three position axes, x, y, and, z, and three momentum axes p_x, p_y, and p_z. The high dimensions makes it difficult to picture the surface of constant energy or the phase space orbits directly. However, the symmetries of the Kepler problem will have the same property as the 1-d harmonic oscillator. When we transform the phase space coordinates via a symmetry, say to the new coordinate system with coordinates $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{p}_x, \tilde{p}_y$, and $\tilde{p}_z$, the description for the energy remains in the same form:

$$H(\boldsymbol{x}, \boldsymbol{p}) = \frac{\tilde{p}_x^2}{2m} + \frac{\tilde{p}_y^2}{2m} + \frac{\tilde{p}_z^2}{2m} - \frac{k}{\sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2}}. \tag{3.1.15}$$

It should now be clearer why understanding the symmetries of particular systems can be useful in simplifying mathematical descriptions. The symmetries allow us to choose the most convenient coordinate systems while preserving our descriptions of the system.

Finding the symmetries of a system can be difficult; it's not obvious looking at just the energy what sorts of transformations would keep the form of the energy the same. However, Nöther's theorem states that for every symmetry there is a corresponding conserved quantity and for every conserved quantity there is a corresponding symmetry. In

We can therefore use our understanding of conserved quantities to find the symmetries of the Kepler problem. In the next section we will discuss the conserved quantities and their corresponding symmetries.

3.2 Conserved Quantities

In the following section we consider intuitively the symmetries corresponding to angular momentum and energy and see why the symmetry corresponding to the LRL vector is not as obvious.

3.2.1 Energy

Now that we have seen what we mean mathematically by a symmetry we can return to the Kepler problem. As with the harmonic oscillator the energy for the Kepler problem is given by summing the kinetic and potential energies. The energy for the Kepler problem is thus given by

$$H(\boldsymbol{x}, \boldsymbol{p}) = \frac{\mathbf{p}^2}{2m} - \frac{Gm_1 m_2}{r}, \tag{3.2.1}$$

where r is the magnitude of $\boldsymbol{x}$. Since we know that energy is conserved we also know that it has a corresponding symmetry. Deriving rigorously the symmetry is not necessary for understanding the symmetry that corresponds to energy conservation. So, I do not derive it here. However, the process we use in Chapter 5 to find the symmetry of the LRL vector can be analogously used to see the symmetry in that corresponds to conservation of energy. The symmetry corresponding to energy is time translations. As we move forward in time from t_0 to t_1, we do not change our description of the energy of the orbit.

To understand why it's special for time translations to be symmetries, it may be helpful to consider how they behave in a case where energy is not conserved. Returning to the harmonic oscillator once more; if we allow the energy to change in time (for instance not ignoring friction), the trajectory in phase space will spiral towards the center (see Figure 3.2.1). The description of the dynamics is then also changing in time. So, in the case where energy is not conserved, time translations are not symmetries.
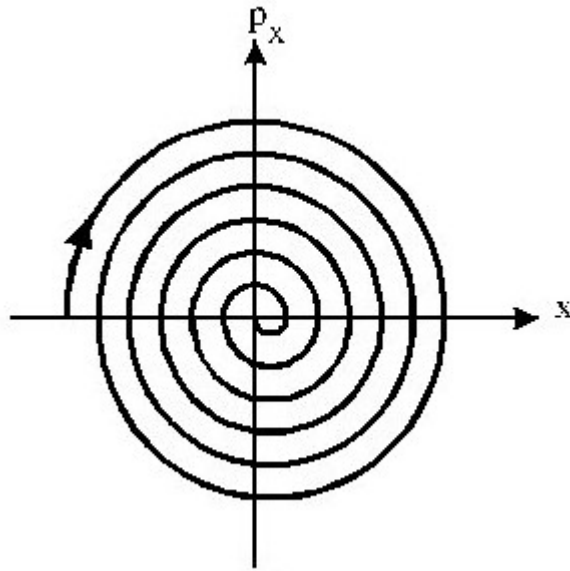
Figure 3.2.1: The phase space of the harmonic oscillator with m=ω=1 allowing the energy to change in time. As the oscillator slows down, it loses energy, so time translations are not a symmetry of the system [19].

When energy is conserved, the time we choose to look at doesn't change the equation we use to describe the dynamics, it just changes the position of the planet. So, time translations are a symmetry of the Kepler problem and correspond to conservation of energy.

3.2.2 Angular Momentum

In chapter two we saw one particular conserved quantity, **h**, which is proportional to the angular momentum, assuming mass is constant. Recalling equation 2.2.9, and multiplying **h** by mass we have

(a) Orbit 1   (b) Orbit 2
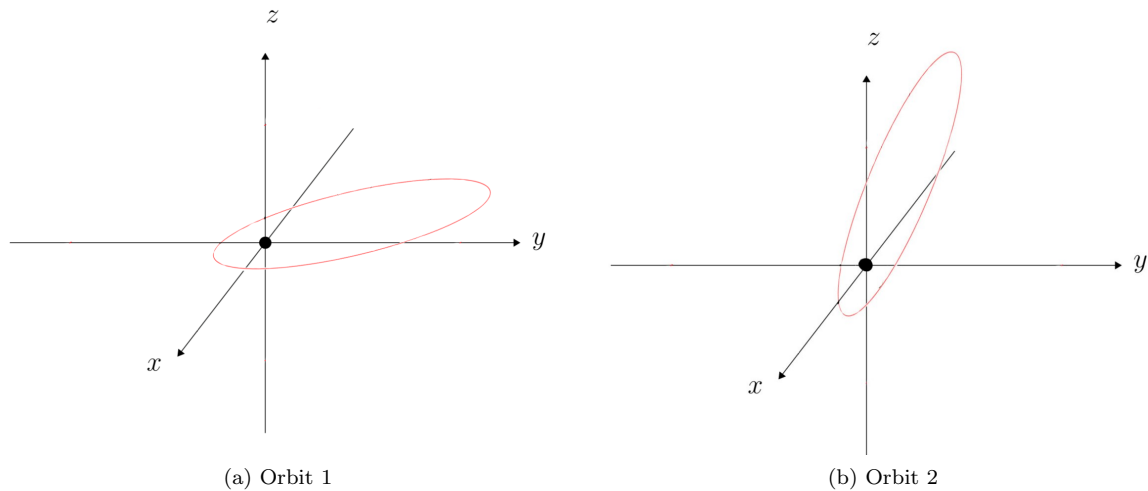
Figure 3.2.2: Looking at the difference between the two orbits above, we see that we haven't changed the description of the orbits. We have just changed coordinates systems. We could transform orbit 1 into orbit 2 by changing the orientation of my axes. Notice that this is a plot in configuration space with coordinates x,y, and z rather then phase space. These are edited versions of the images here [18]

$$m\mathbf{h} = m(\mathbf{r} \times \dot{\mathbf{r}}) = \mathbf{r} \times m\dot{\mathbf{r}} = \mathbf{r} \times \mathbf{p} = \mathbf{L}, \tag{3.2.2}$$

Since we have already seen that $\mathbf{h}$ is constant and we assume mass is constant, the angular momentum is also constant in time and therefore a conserved quantity.

The conservation of angular momentum corresponds to the symmetry of three-dimensional rotations. Intuitively, rotating the orbit about its star doesn't change how we describe the orbit, we have simply changed coordinates systems (see Figure 3.2.2).

The symmetry of three dimensional rotations allows us to make a major simplification. Thinking about the phase space of the Kepler problem, as mentioned previously, we would have a 6 dimensional phase space space, with coordinates x, y, z, p_x, p_y, and p_z. However, since we have rotational symmetry, we can consider the dynamics in a coordinate system such that the orbital plane lies only in the xy-plane. We would then be able to retrieve any orbit by rotating to back to the initial coordinate system. Along with this transformation its useful to shift to using polar coordinates r, ν, p_ν, and p_r (see Figure 3.2.3). Note that p_ν
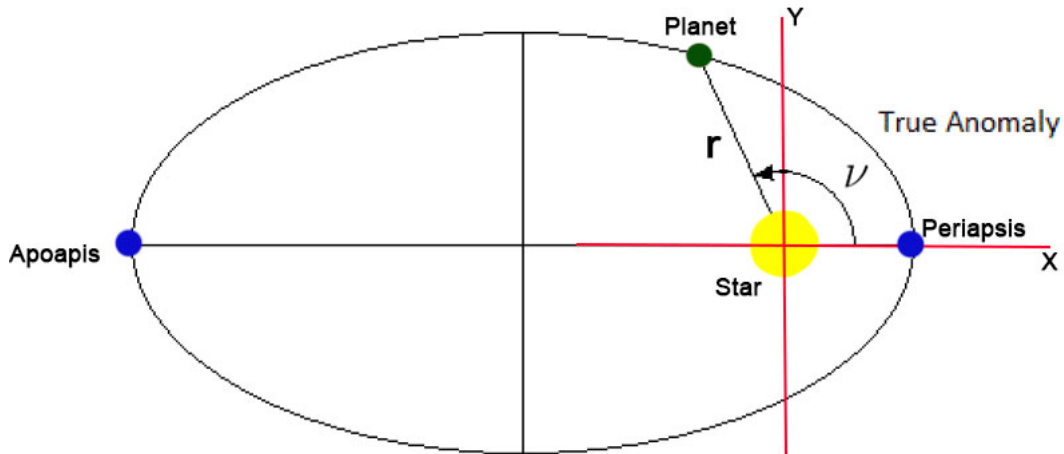
Figure 3.2.3: Using the symmetry corresponding to angular momentum, we can reduce the size the phase space. In the new system, the configuration space is conveniently described in terms of polar coordinates r and ν as shown. Edited version of image found here [1]

is the angular momentum. We now have a four dimensional as opposed to 6 dimensional phase space.

3.2.3 Laplace-Runge-Lenz Vector

Along with energy and angular momentum, the Laplace-Runge-Lenz (LRL) vector is a conserved quantity of the Kepler problem. The LRL vector is defined as

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - mk\hat{\mathbf{r}},$$

based on Rogers [14] conventions. There are also a few other useful normalization's of $\mathbf{A}$. One that ensures that the magnitude of $|\mathbf{A}| = e$ where e is the eccentricity,

$$\mathbf{A} = \frac{\mathbf{r} \times \mathbf{L}}{mk} - \hat{\mathbf{r}},$$

and one that simplifies the Lie algebra relationships as we use in Appendix C,

$$\mathbf{A} = \frac{1}{\sqrt{-2mE}}(\mathbf{p} \times \mathbf{L} - mk\hat{\mathbf{r}}).$$

Introducing the conservation of the LRL vector may appear to suggest we have seven conserved quantities: energy, and the three components each of $\mathbf{L}$ and $\mathbf{A}$. However, the Kepler problem is a six dimensional phase space, meaning we should only be able fix six

quantities. To demonstrate why this might be concerning, let's take a phase space that is slightly easier to conceptualize. Suppose we have a four dimensional phase space with coordinates, x, y, p_x, and p_y. If we have no fixed quantities, can choose any four arbitrary numbers for each of the four phase space coordinates. In other words, it has four degrees of freedom. The choice of the previous coordinates has no effect on the allowed values for the last coordinate. However, if we know that one quantity is fixed, say the energy, we can only choose three coordinates arbitrarily and the last will be fixed by their relationship to energy. So, the surface with constant energy would be a three dimensional surface inside of the initial four-dimensional phase space. Fixing another quantity, say the magnitude of the angular momentum, reduces the allowed phase space coordinates down to a two-dimensional plane, and similarly, one more constraint or conserved quantity would give a one dimensional trajectory. So, for every conserved quantity the possible dimension of the trajectories is decreased by one.

Applying similar logic to the full six-dimensional phase space of the Kepler problem, we might be concerned. Since we seem to have seven conserved quantities and 6 degrees of freedom, by taking the dimension down for each conserved quantity we would end up with an over constrained system. However, there are also two relationships between $\mathbf{A}, E$ and $\mathbf{L}$ that bring the total number of independent conserved quantities down to five. With five independent conserved quantities and 6 degrees of freedom, we have exactly the number of fixed variables we need to find the one-dimensional elliptical trajectory in phase space.

The first relationship says that the dot product of $\mathbf{A}$ and $\mathbf{L}$ is zero,

$$
\begin{aligned}
\mathbf{A} \cdot \mathbf{L} &= (\mathbf{p} \times \mathbf{L} - mk\hat{\mathbf{r}}) \cdot \mathbf{L} \\
&= (\mathbf{p} \times \mathbf{L}) \cdot \mathbf{L} - (mk\hat{\mathbf{r}}) \cdot \mathbf{L} \\
&= \mathbf{p} \cdot (\mathbf{L} \times \mathbf{L}) - mk(\hat{\mathbf{r}} \cdot \mathbf{L}) \\
&= 0.
\end{aligned}
$$

In other words, **A** is always perpendicular to the angular momentum, and thus lies in the plane of the orbit.

There is also a relationship between the magnitude of **A**, the magnitude of **L** and E,

$$|\mathbf{A}|^2 = 2mL^2E + m^2k^2. \tag{3.2.3}$$

Since we see that $|A|$ is equivalent to mke and all of these quantities are constant in time, the $|A|$ is also a constant. Notice that the addition of these two conditions, resolves the concern regarding having more independent conserved quantities then dimensions of phase space. The components of the LRL vector are not all completely independent of the choice of angular momentum and energy.

It is also interesting to note that with proper normalization we can rewrite the magnitude in terms of eccentricity. Noting that from Kepler's second law $L = \frac{2mk\sqrt{1-e^2}}{4E^2\sqrt{\frac{-m}{2E^3}}}$

$$|A|^2 = 2mL^2E + m^2k^2 = m^2k^2e^2.$$

The LRL vector symmetry is not nearly as obvious as with the other conserved quantities. It is not clear what other sorts of transformations would preserve our dynamics. For both energy and angular momentum we were able to see the symmetry of the system in configuration space, as the they only change the coordinates x and y rather then p_x and p_y. However for the LRL vector the transformation will affect both position and momentum coordinates. In particular it will turns out that the symmetry of the LRL vector corresponds to 4 dimensional rotations and we will see how these transformations affect the configuration space in chapter 5 [8].

4
Hamiltonian Mechanics

In chapter 3, we considered the symmetries related by Nöther's theorem to the conserved quantities of energy and angular momentum. For the Laplace-Runge-Lenz vector, I suggested that the symmetry was not obvious in configuration space, and we should instead consider the phase space of the orbit. So far we have only dealt with problems from the Newtonian perspective of classical mechanics which is not conducive to systems with arbitrary coordinate systems and works best with Cartesian coordinates. However there are two other, yet equivalent formulations, Lagrangian and Hamiltonian mechanics. Both Lagrangian and Hamiltonian mechanics work in arbitrary coordinates systems. Lagrangian mechanics focuses on generalized position and generalized velocity $(q, \dot{q})$ and Hamiltonian Mechanics looks at generalized position and generalized momentum (q, p). Since we are interested in the phase space trajectories, the most convenient system for studying the symmetry of the LRL vector is via Hamiltonian mechanics. In this section, I introduce the action and the Lagrangian, the Hamiltonian, Poisson brackets, and flow variables based on Goldstein [7]. Any readers already familiar these concepts may want to skip this chapter.

4.1   Lagrangian vs Hamiltonian Mechanics

For readers more familiar with Lagrangian Mechanics I compare them briefly. In Classical Mechanics we define two new quantities that allow us to use more generalized coordinates, the Lagrangian and the action. The Lagrangian is defined by $\mathcal{L} = \mathcal{L}(q, \dot{q}, t) = T - V$ where T is the kinetic energy and V is the potential energy, and the action is defined as

$$S = \int \mathcal{L} dt.$$

In Lagrangian Mechanics, the Euler-Lagrange equations give the equations of motion:

$$\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^i} = 0; \quad i = 1, 2, 3 \ldots n,$$

where n is the dimension of the space. For each i, the Euler-Lagrange equation gives a different partial differential equation. For some systems these may be coupled and, as a result, difficult to solve.

In Hamiltonian Mechanics we define the Hamiltonian, $H = H(q, p, t) = T + U$, which gives the total energy. We then have Hamilton's equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \text{ and } \dot{p}_i = -\frac{\partial H}{\partial q_i}; \quad i = 1, 2, 3 \ldots n.$$

Again, these are true for the q's and p's in each direction and again describe a set of differential equations.

The main distinction between Lagrangian and Hamiltonian mechanics is that the Lagrangian, $\mathcal{L} = \mathcal{L}(q, \dot{q})$, is a function of position (q) and velocity ($\dot{q}$), whereas the Hamiltonian, $H = H(q, p)$ is a function of the generalized coordinate (q) and the generalized momentum (p). We define the canonical momentum to be $p = \frac{\partial \mathcal{L}}{\partial \dot{q}}$. The dependence of H on p, as opposed to $\dot{q}$, makes Hamiltonian mechanics well adapted to studying phase space.

4.2 Poisson bracket

In order to understand the symmetry of the LRL vector we need to understand an important tool from Hamiltonian mechanics called the Poisson bracket. The Poisson bracket is an operation defined between two functions $f(p, q, t)$ and $g(p, q, t)$ by

$$\{f, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}.$$

I could define the Poisson bracket for any number of dimensions n, i.e. p_i, q_i which would include a summation over i from 1 to n. To keep the definition simple, I'll stick to one dimension here. For any readers with background in quantum mechanics, the Poisson bracket is the classical mechanical analog of the commutator.

Example: To see how the Poisson bracket is useful, let's look at some interesting examples directly relate to the Kepler problem. In the Kepler problem we have two coordinates r, and ϕ with r the distance from the center and ϕ is angle from periapsis about the star and thus two corresponding momenta, p_r, and p_ϕ. To match the source notation I have switched from ν to ϕ, but they are the same variable. If we look at the Poisson bracket of $r(r, p, t) = r$ with $H(r, p, t) = \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} - \frac{k}{r}$, we get

$$\{r, H\} = \frac{\partial r}{\partial r}\frac{\partial H}{\partial p_r} - \frac{\partial r}{\partial p_r}\frac{\partial H}{\partial r} + \frac{\partial r}{\partial \phi}\frac{\partial H}{\partial p_\phi} - \frac{\partial r}{\partial p_\phi}\frac{\partial H}{\partial \phi}. \tag{4.2.1}$$

Noting that r does not depend on ϕ, p_r, p_ϕ we see that all but the first term go to 0. Additionally $\frac{\partial r}{\partial r} = 1$, giving

$$\{r, H\} = \frac{\partial H}{\partial p_r}. \tag{4.2.2}$$

Substituting in H we have

$$\{r, H\} = \frac{\partial}{\partial p_r}\left(\frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} - \frac{k}{r}\right). \tag{4.2.3}$$

Distributing the derivative across the sum, we see that the second and third terms do not depend on p_r giving,

$$\{r, H\} = \frac{\partial}{\partial p_r} \frac{p_r^2}{2m} = \frac{p_r}{m}. \tag{4.2.4}$$

Since $p = mv$ we can rewrite the equation above,

$$\{r, H\} = v_r = \frac{\partial r}{\partial t}. \tag{4.2.5}$$

Thus we see that the Poisson bracket of r with H gives us the time derivative of r and it turns out this is a general result.

The Poisson bracket of any function with H, gives the time derivative of that function:

$$\frac{df}{dt} = \{f, H\}. \tag{4.2.6}$$

Using the Poisson bracket we can efficiently check if a quantity is conserved, i.e. if $\{f, H\} = 0$ then f is conserved. In particular, as an exercise one could check that the magnitude of the LRL vector is conserved by finding that $\{A, H\} = 0$.

4.3   Flow Variables

Even more generally, for any choice of f and g, the Poisson bracket tells us how f changes with respect to the flow variable γ of g defined by

$$\frac{df}{d\gamma} = \{f, g\}.$$

In the case of the Hamiltonian the flow variable is time,t. and the Poisson bracket then shows how any function f changes with time. Given some direction $\mathbf{c}$, the flow variable of the $\mathbf{c}$ component of the angular momentum is the angle θ about the $\mathbf{c}$ direction. So, the output of the Poisson bracket of f with the component $\mathbf{c} \cdot \mathbf{L}$, tells us how f changes with the angle θ about the $\mathbf{c}$ direction. By combining the flow variables about the components x, y and z, we obtain all three dimensional rotations.

We can find flow variables for any arbitrary function of the phase space coordinates. However, since both angular momentum and energy are conserved quantities, their flow variables also represent the symmetry of the system to which they correspond. As before, the conservation of energy is related to a time translation symmetry, and the conservation of angular momentum gives a three dimensional rotation symmetry.

5
Symmetry of the Laplace-Runge-Lenz vector

In this chapter we apply the tools described in chapter 4 to the specific case of the LRL vector as done by Joanna Gonera, Pitor Kosínski, and Patryk Michel in [8]. We use the notion of flow variables to find a differential equation that describes how an arbitrary function $f(q, p, t)$ changes along the flow of the LRL vector. In other words as ϵ changes. Using the solution to the differential equation I create an animation of the orbit as ϵ changes.

5.1 Flow variable of LRL

To identify the symmetry associated with the LRL vector, $\mathbf{A}$, as we saw in Chapter 4, it's helpful to consider the Poisson bracket relation between f and the magnitude of the LRL vector, A, or more generally, the symmetry about any direction in our orbital plane. Just as with the angular momentum vector we chose a specified direction, we also choose a direction for the LRL vector. To allow for the general case, we define $A_c = \mathbf{A} \cdot \hat{c}$ where $\hat{c} = (\cos \chi, \sin \chi)$ is a chosen direction. We are able to simplify the formulas by considering the special case $\chi = 0$ and hence $A_c = A_1$. We can recover the general case by replacing Θ_1 with $\Theta_1 - \chi$. For the LRL vector we call the flow variable ϵ allowing us to set up a system of four differential equations:

$$\frac{df_i}{d\epsilon} = \{f_i, A\} = \frac{\partial f_i}{\partial r}\frac{\partial A}{\partial p_r} - \frac{\partial f_i}{\partial p_r}\frac{\partial A}{\partial r} + \frac{\partial f_i}{\partial \phi}\frac{\partial A}{\partial p_\phi} - \frac{\partial f_i}{\partial p_\phi}\frac{\partial A}{\partial \phi}; \quad i = 1, 2, 3, 4, \tag{5.1.1}$$

where $f_1 = r$, $f_2 = \phi$, $f_3 = p_r$, and $f_4 = p_\phi$.

In general equation 5.1.1 is difficult to solve, as it turns out to be non-linear. However, we can change variables to simplify the equation. The particular set of variables we want are the action-angle variables. The origin and reason for switching to the action angle variables is a complicated and very interesting aside that I explain much more in-depth in Appendix A. For now, I will simply quote the results for the action angle variables for the Kepler problem:

$$I_\phi = p_\phi,$$

$$H = \frac{-mk^2}{2(I_r + I_\phi)^2},$$

$$\Theta_r = \sqrt{2mkr - \frac{m^2k^2r^2}{2(I_r + I_\phi)^2} - I_\phi^2} - \arcsin\frac{1 - \frac{mkr}{2(I_r+I_\phi)^2}}{\sqrt{1 - \frac{I_\phi^2}{2(I_r+I_\phi)^2}}}, \quad \text{and}$$

$$\Theta_\phi = \phi - \arcsin\frac{1 - \frac{I_\phi^2}{mkr}}{\sqrt{1 - \frac{I_\phi^2}{2(I_r+I_\phi)^2}}} + \Theta_r.$$

Changing from these to one more set of variables $I_2 = I_r + I_\phi$ and $\Theta_1 = \Theta_\phi - \Theta_r$ and $\Theta_2 = \Theta_r$, we can write an explicit solution to the differential equation:

$$\sin^2\Theta_1(\epsilon) = \frac{2\alpha}{(1+\alpha) - (1-\alpha)\cos\left(\frac{2mk\epsilon}{\sqrt{amk}} + \sigma_1\right)}, \quad \text{and}$$

$$\Theta_2(\epsilon) = \arcsin\left(\frac{\cos\Theta_1}{\sqrt{1-\alpha}}\right) - \frac{mk\sqrt{\alpha}}{\sqrt{amk}}\epsilon + \sigma_2,$$

where $\alpha \equiv A_1^2/(mk^2)$. In these variables we have that $A_c = mk\sqrt{1 - I_1^2/I_2^2}\sin(\Theta_1 - \chi)$.

Now that we have the solutions to the differential equation, we have a description for how Θ_1 and θ_2 change as functions of ϵ. By inverting the transformation I retrieve the semi-major axis, and the eccentricity as functions of ϵ. Given the semi-major axis and the eccentricity I plot the orbits as we change ϵ as seen in Figure 5.1.1.
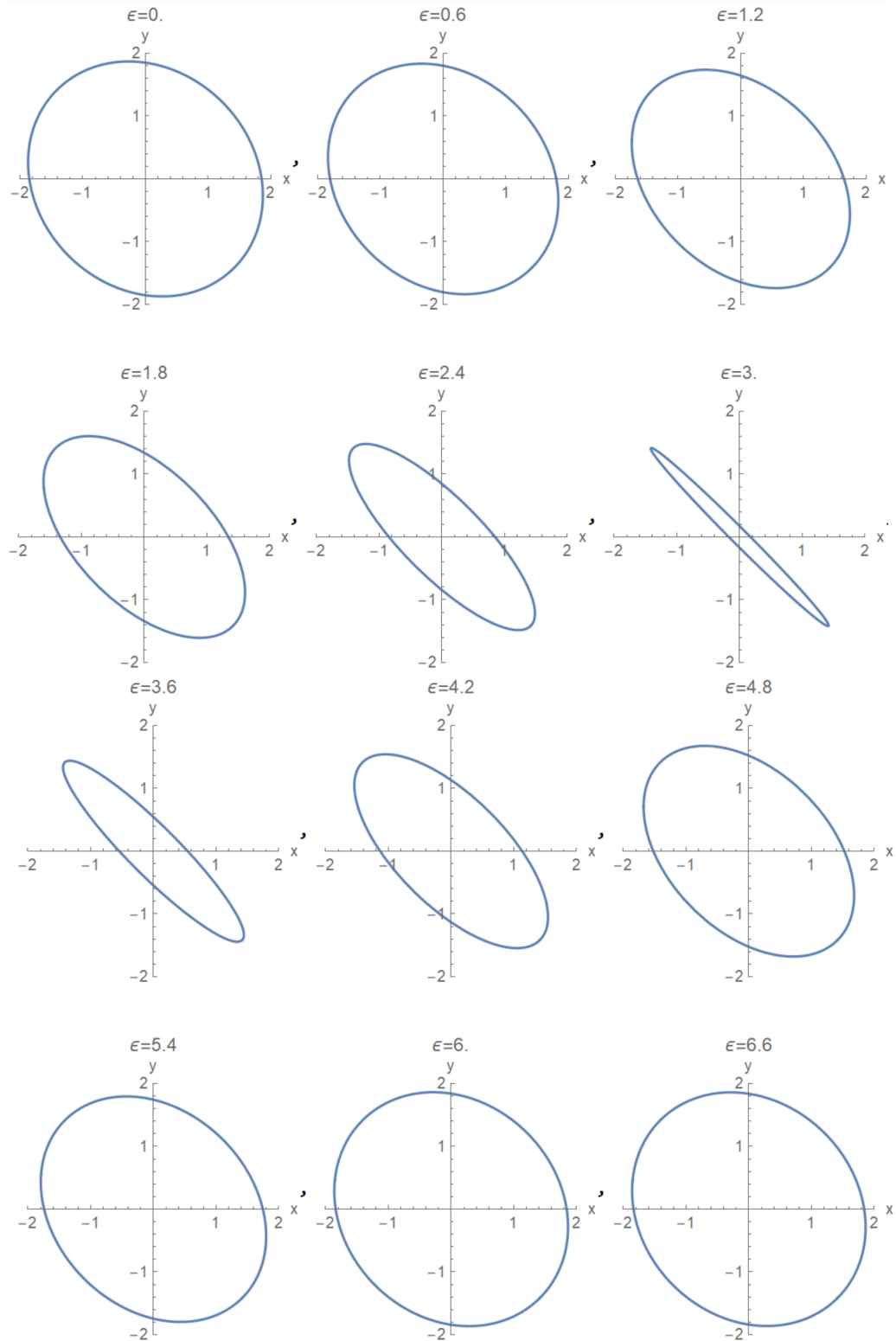
Figure 5.1.1: In the plots above we see that as epsilon changes the orbits first become more eccentric and then less eccentric. We can also view these transformations as three-dimensional rotations in phase space that have been projected on to the configuration space. Since ϵ is our flow parameter for the LRL vector, the transformations associated with the LRL vector result in a change in eccentricity. These plots were found with $A = 0.5$, $\phi = \pi/4$ and, $\chi = \pi/3$. We range ϵ from zero to seven in steps of .06

In the Figure 5.1.1, we can see clearly why the symmetry of the LRL vector is not obvious. In our previous examples, for energy and angular momentum, the shape of the orbit did not change. Here we see that along the flow of the LRL vector, orbits change eccentricities. The symmetries corresponding to energy and angular momentum can be seen entirely in configuration space, while the symmetry corresponding to the LRL vector requires phase space. Rather then thinking of the transformation as one that compresses or expands the orbit, we can consider it as a rotation in phase space. Considering the orbits plotted in figure 5.1.1, we can think of them as changing eccentricity, or we could instead picture the ellipses being rotated in three dimensions, and then viewing them in two dimensions. In that picture, the change in shape of the two-dimensional configuration space is a result of the projection from the phase space onto the configuration space.

At this point it's important to remember that we simplified our calculations and plots by fixing angular momentum. If we once again allow angular momentum to vary, we have a three-dimensional configuration space, and a six-dimensional phase space. The symmetry corresponding to angular momentum is three-dimensional rotations which can all occur in configuration space. However, the symmetry corresponding to the LRL vector is rotations in four dimensions which will also rotate one momentum direction. Similar to the lower dimensional case the projection into configuration space results in a change in eccentricity.

While the symmetry corresponding to the LRL vector is part of a four-dimensional rotation, when we fix angular momentum, we break its symmetry. The result is that for our two-dimensional plots, the LRL vector will correspond to three dimensional rotations. Once again, since the configuration space is only two-dimensional, the rotations will rotate at least one momentum component. More precise mathematics regarding the rotation groups and symmetry breaking can be found in Appendix C on Group Theory.

6
Deriving Probability Distributions

The motivation for understanding the relationship between symmetries and conserved quantities comes from my work to derive the probability distribution for semi-major axis given a single direct image of an exoplanet. In this chapter I introduce the important terms for understanding Bayesian statistics, derive the circular case completely as an example and then give a more detailed mathematical description of the work I have done for the elliptic case up to this point. I will then indicate where I have run into difficulties and our planned next steps.

6.1 Bayesian Statistics

Before I describe my work in-depth, I want to define a few useful terms from Bayesian statistics which we use to derive the probability distribution.

First, a **probability distribution** of a variable x is a function that when summed over some bounds, say a to b, gives the probability that x lies between the values a and b. The **probability density** is the function that when integrated over a and b give the probability that x lies between a and b. Probability density then applies only to continuous variables. Probability distributions can describe discrete variables. All of the variables in the Kepler problem are continuous, so I often use distribution and density interchangeably.

We have a **prior** probability distribution that gives our initial understanding of how a particular variable is distributed. A **uniform** prior means we assume no prior understanding of the distribution of a particular variable. In other words we treat all values of that variable with equal probability. Another common prior is to assume that a variable is Gaussian or normally distributed. A Gaussian prior on exam scores for instance would suggest that there is a peak score that is most likely, and the probabilities drop off symmetrically and more quickly the farther you get from the peak. I also refer to priors that are **log-uniform** or **uniform in cosine**. Both imply that some variable x is not itself uniform, but rather $\ln x$ or $\cos x$ are uniform. In the case of log uniform, the variable favors small values.

Given priors for each of the relevant variables the **posterior** distribution gives the probability distribution after a measurement of one of the variables. In our case, we measure the projected radius from the direct image and we want to find the posterior distribution for the semi-major axis.

While, for simplicity, I have only given examples of probability distributions that are functions of only one variable, we can have probability distributions that depend on multiple variables. Multivariate probability distributions are called **joint** probability distributions or densities.

As we have seen the Kepler problem has many variables and as a result will have joint prior and posterior distributions. We can find single variable probability distributions from joint distributions by **marginalizing** over the other variables. We marginalize by integrating over all possible values of the variable. Since there is 100% chance that the marginalized variable fell within the range of all its values, we are assuming we know nothing more then our prior assumption about the distribution of that variable. Marginalization results in a new probability distribution that no longer depends on the value of the marginalized variable.

6.2   Circular Case

I will begin by deriving the probability distribution of the circular case to demonstrate the method, as it is much simpler and completely solved.

The geometry of a circular planetary orbit is completely determined by the orbit radius a, the orbital inclination i, and the true anomaly ν. Notice that a, which is the semi-major axis in the elliptical case, is now simply the radius of the orbit. The inclination is the same as before, and ν is now easily interpreted as the angle around the orbit from some reference direction. A general prior on these parameters would have the form of a joint distribution $f(a, \nu, \cos i) da d\nu d \cos i$, however, physically, we expect these variables to be independent. For instance, we do not expect the value of a to affect the probabilities of any given i. The mathematical consequence is that we can rewrite the joint prior distribution as a product of single variable priors:

$$f_\nu(\nu) f_i(\cos i) f_a(a) \mathrm{d}\nu \ \mathrm{d} \cos i \ \mathrm{d}a.$$

Since the planet is on a circular orbit, we expect that each value of ν should be equally likely, so ν has a uniform prior. We have a prior on a, $f_a(a)$, that we can take as either uniform, in the most pessimistic case, or as log uniform in the more realistic case. The uniform case is pessimistic since it assumes we know nothing about the typical distribution of exoplanets. The derivation for the prior on inclination is a bit non-trivial so we first take an aside to derive the prior for i.

6.2.1   Prior on Inclination

To find the prior on i, we take the possible directions to be isotropic, meaning any direction is equally likely. We then consider all the vector directions that give the same inclination. In Figure 6.2.1, we see the angle i and some new inclination $i + \mathrm{d}i$. The surface area between i and $i + \mathrm{d}i$ represents the amount of vectors that have a direction between i and $i + \mathrm{d}i$ and is given by $2\pi \sin i \ \mathrm{d}i$. We can rewrite $2\pi \sin i \mathrm{d}i = -2\pi \mathrm{d} \cos i$. So, inclination is uniform in cosine [2].
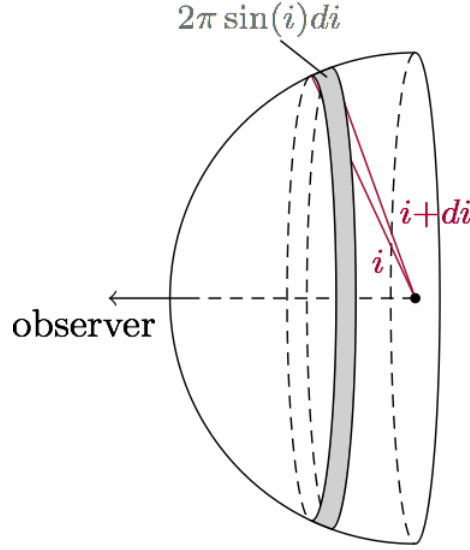
Figure 6.2.1: An variable that is uniform in space, or isotropic, is uniform in cosine of the polar angle. The surface area (and thus the amount of vectors) is larger on the great circle than towards the edge of the sphere [2].

6.2.2   Deriving the Probability Density for the Circular Case

So, our priors on i and ν are $f_i(\cos i) = 1$ and $f_\nu(\nu) = \frac{2}{\pi}$, where the latter value is adopted because the orbital symmetry makes integration over a quarter orbit most convenient.

Now, we want to find the posterior distribution given some measurement of the projected radius, r_p. Since we are not measuring one of our variables a, i or ν directly, we have to consider the relationship between these quantities and the measured quantity. In particular,

$$r_p = a\sqrt{1 - \sin\nu^2 \sin^2 i}. \tag{6.2.1}$$

We then want to change variables from ν to r_p for which we need the Jacobian factor,

$$\left|\frac{\mathrm{d}\nu}{\mathrm{d}r_p}\right| = \frac{\sqrt{1 - \sin i^2 \sin\nu^2}}{a\sin\nu \cos\nu \sin i^2} = \frac{r_p/a}{\sqrt{a^2 - r_p^2}\sqrt{(r_p/a)^2 - \cos^2 i}},$$

where the second equality elliminates ν in favor of r_p using Eqn 6.2.1. Notice that in the circular case $0 \le r_p \le a$, so the value inside both square roots are positive, justifying neglecting the absolute values.

Substituting our priors and implementing the change of variables, the joint prior distribution becomes

$$\frac{2}{\pi} f_a(a) \left| \frac{\mathrm{d}\nu}{\mathrm{d}r_p} \right| \mathrm{d}r_p \, \mathrm{d}\cos i \, \mathrm{d}a = \frac{2}{\pi} f_a(a) \frac{r_p/a}{\sqrt{a^2 - r_p^2}\sqrt{(r_p/a)^2 - \cos^2 i}} \mathrm{d}r_p \, \mathrm{d}\cos i \, \mathrm{d}a$$

Marginalizing over $\cos i$ correlates r_p and a, leaving us with the joint distribution

$$f(a, r_p) = \frac{2}{\pi} f_a(a) \int_0^{\frac{r_p}{a}} \frac{r_p/a}{\sqrt{a^2 - r_p^2}\sqrt{(r_p/a)^2 - \cos^2 i}} \mathrm{d}\cos i = \frac{f_a(a)}{(a/r_p)^2 \sqrt{(a/r_p)^2 - 1}}.$$
$$(6.2.2)$$

The posterior after taking a measurement $r_p = r_p^*$, is just a section of this joint distribution,

$$f(a|r_p^*) = \frac{f_a(a)}{(a/r_p^*)^2 \sqrt{(a/r_p^*)^2 - 1}},$$

keeping in mind that by simply replacing $r_p = r_p*$ we assume an infinitely precise measurement of the projected radius. We can then plot the distribution for two different choices for priors on semi-major axis as in Figure 6.2.2.
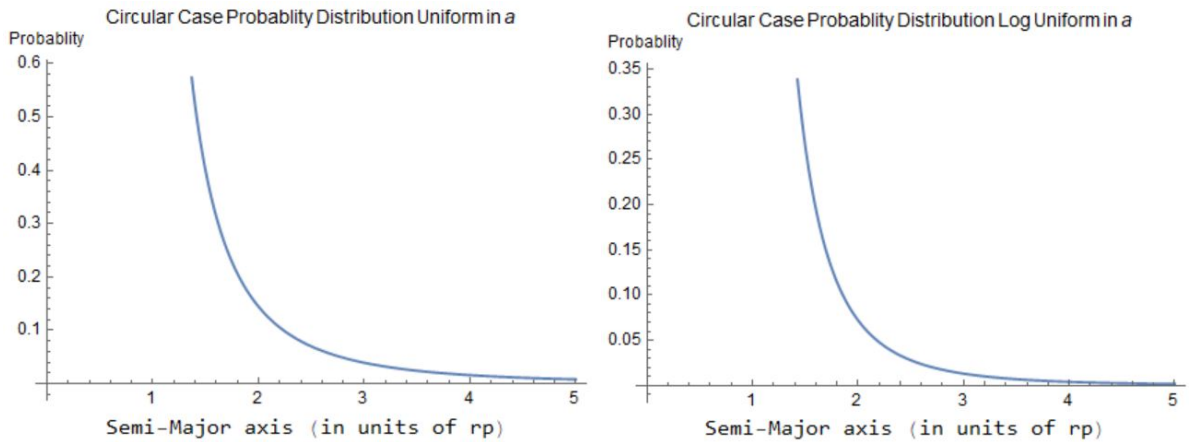
Figure 6.2.2: On the left is the probability distribution for a uniform prior on semi-major axis and on the right is the distribution for a log-uniform prior on semi-major axis. Notice that both favor smaller semi-major axes, but the log-uniform drops off more quickly.

We notice that in both cases the peaks are at $a = r_p$ and the probability falls of very quickly after that. Since we know that we have a circular orbit the planet is at a constant distance from its star. So, we know that if $a \neq r_p$, then we have an inclined orbit. The

larger the a/r_p ratio, the more inclined the orbit must have been. Since near edge-on
orbits are less likely, it makes sense that the probability of larger a/r_p ratio is smaller.

6.3   Elliptical Case

We proceed along similar lines for the elliptical case, but with a few additional parameters.
An elliptical orbit can be described by five orbital parameters: the semi-major axis a, the
eccentricity e, the orbital inclination i, the true anomaly ν, and the argument of periapsis
ω. These parameters were described in detail in Chapter 2. As in the circular case we
would like to simplify the general joint prior $f(a, \nu, \cos i, e, \omega)dad\nu d\cos id ed\omega$. However,
the parameters ν and e are not independent. We know that the planet moves faster closer
to the star, so highly eccentric orbits will decrease the probability that a planet is observed
at positions nearer to periapsis. Before we consider the full joint prior distribution, we
derive the joint prior on nu and e.

6.3.1   Prior on ν and e

To derive the joint prior on e and ν, we begin with Kepler's second law, that the planet
traces out equal area in equal time. For an ellipse, $dA = \frac{1}{2}r^2 d\nu$, where r is the distance to
a point on the ellipse and one focus. In our case, $r = a(1-e^2)/(1+e\cos\nu)$, is the distance
between the planet and the star. By dividing by dt on both sides we get $\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\nu}{dt}$.

From here we need one intermediate argument to show that $\frac{dA}{dt}$ is constant. First, lets
consider the angular momentum of the system,

$$\vec{L} = \vec{r} \times \vec{p} = r\hat{r} \times m\vec{v} = r\hat{r} \times m(v_r\hat{r} + v_\nu\hat{\nu}), \tag{6.3.1}$$

$$= mrv_r(\hat{r} \times \hat{r}) + mrv_\nu(\hat{r} \times \hat{\nu}), \tag{6.3.2}$$

$$= mrv_\nu\hat{z} = mr^2\frac{d\nu}{dt}\hat{z}. \tag{6.3.3}$$

Thus, $L = mr^2\frac{d\nu}{dt}$ and it follows that, $\frac{L}{2m} = \frac{1}{2}r^2\frac{d\nu}{dt} = \frac{dA}{dt}$. Since L and m are both constants
so is $\frac{dA}{dt}$. Over one whole period the area traced by a planet is equal to the area of the

ellipse that forms its orbit which implies that $\frac{dA}{dt} = \frac{\pi ab}{T}$ where a is the semi major axis of the ellipse, b is the semi-minor axis, and T is the period of the orbit.

It follows that,

$$\frac{\pi ab}{T} = \frac{1}{2}r^2\frac{d\nu}{dt}, \tag{6.3.4}$$

$$\frac{\pi a(a\sqrt{1-e^2})}{T} = \frac{1}{2}\frac{a(1-e^2)}{1+e\cos\nu}r\frac{d\nu}{dt}, \tag{6.3.5}$$

$$\frac{dt}{T} = \frac{1}{2\pi a}\frac{(1-e^2)^{1/2}}{(1+e\cos\nu)}r d\nu, \tag{6.3.6}$$

$$= \frac{1}{2\pi}\frac{(1-e^2)^{3/2}}{(1+e\cos\nu)^2}d\nu. \tag{6.3.7}$$

Note, $dM = \frac{dt}{T}$ where M is the mean anomaly. Since the prior on M is uniform, the prior true anomaly, $d\nu$, (which also depends on e) is described by $\frac{1}{2\pi}\frac{(1-e^2)^{3/2}}{(1+e\cos\nu)^2}$

6.3.2 Deriving the Probability Density for the Elliptic Case

As in the circular case on physical grounds, we can argue for some simplifications of the general prior; the parameters ν and e are excepted as we saw in the previous section. We can rewrite the prior distribution as

$$f_a(a)f_{\nu,e}(\nu,e)f_i(\cos i)f_\omega(\omega)da\ d\nu\ d\cos i\ de\ d\omega.$$

As before, $f_i(\cos i) = 1$ and $f_\omega(\omega) = \frac{1}{2\pi}$ are uniform priors, while Kepler's second law implies $f_{\nu,e}(\nu,e) = \frac{1}{2\pi}\frac{(1-e^2)^{3/2}}{(1+e\cos\nu)^2}$. Substituting these priors, we have

$$\frac{1}{(2\pi)^2}f_a(a)\frac{(1-e^2)^{3/2}}{(1+e\cos\nu)^2}da\ d\nu\ d\cos i\ de\ d\omega.$$

Again, we need to change variables this time from ω to r_p. The choice of which variable to exchange for r_p was made for convenience of the calculations. For elliptical orbits the relationship between r_p and the other parameters is:

$$r_p = r\sqrt{\cos^2(\omega+\nu) + \sin^2(\omega+\nu)\cos^2 i},$$

where r_p is the projected semi major axis and $r = r(a,\nu,e) = a(1-e^2)/(1+e\cos\nu)$, describes the distance from the star to the location of planet on the ellipse.

Then the Jacobian for this change of variables is:

$$\left|\frac{\mathrm{d}\omega}{\mathrm{d}r_p}\right| = \left|\frac{1}{r_p}\frac{1}{\sqrt{[1 - (\frac{r\cos i}{r_p})^2][(\frac{r}{r_p})^2 - 1]}}\right|.$$

Here we notice that in the elliptical case, we no longer have the condition that $0 \leq r_p \leq a$. The reason is that the semi-major axis is measured form the center of the ellipse, whereas r_p is measured as a distance from the star. The star is located at a focus, so at periapsis the planet would have a physical distance of $r = a(1 - e)$ and at apoapsis (the farthest point) the planet would be $r = a(1 + e)$ away from the star. So, we now have to consider cases, as the Jacobian takes a different form for different relationships between r and r_p.

Case 1: In the first case we take $r/r_p > 1$ and $(r\cos i)/r_p < 1$. Given these assumptions we have a joint probability density:

$$\frac{1}{(2\pi)^2}f_a(a)\frac{(1 - e^2)^{3/2}}{(1 + e\cos\nu)^2}\frac{1}{r_p}\frac{1}{\sqrt{[1 - (\frac{r\cos i}{r_p})^2][(\frac{r}{r_p})^2 - 1]}}\mathrm{d}a\ \mathrm{d}\nu\ \mathrm{d}\cos i\ \mathrm{d}e\ \mathrm{d}r_p.$$

Case 2: In the second case we make the assumption that $r/r_p < 1$ and $(r\cos i)/r_p < 1$ we get the probability density:

$$\frac{1}{(2\pi)^2}f_a(a)\frac{(1 - e^2)^{3/2}}{(1 + e\cos\nu)^2}\frac{1}{r_p}\frac{1}{\sqrt{[(\frac{r\cos i}{r_p})^2 - 1][(\frac{r}{r_p})^2 - 1]}}\mathrm{d}a\ \mathrm{d}\nu\ \mathrm{d}\cos i\ \mathrm{d}e\ \mathrm{d}r_p.$$

Case 3: In the third case we have $r/r_p < 1$ and $(r\cos i)/r_p < 1$ giving the probability density:

$$\frac{1}{(2\pi)^2}f_a(a)\frac{(1 - e^2)^{3/2}}{(1 + e\cos\nu)^2}\frac{1}{r_p}\frac{1}{\sqrt{[1 - (\frac{r\cos i}{r_p})^2][1 - (\frac{r}{r_p})^2]}}\mathrm{d}a\ \mathrm{d}\nu\ \mathrm{d}\cos i\ \mathrm{d}e\ \mathrm{d}r_p.$$

Just as in the circular case, our next step is to marginalize over all other variables to obtain a single variable probability distribution in terms of a.

6.4   Marginalizing: Case 1

We now want to marginalize over the three variables, $\cos i$, ν and e, to obtain the joint probability density $f(a, r_p)$. In the first case we can integrate analytically over $\cos i$ and over ν leaving only the e integral to do numerically.

We integrate over all allowed values for $\cos i$ giving a lower bound of 0 and an upper bound of $\frac{r_p}{a} \frac{1+e\cos(\nu)}{(e^2-1)}$. For ν we find a lower bound of $\arccos\left(\frac{a}{r_p}((1-e^2)-\frac{r_p}{a})\right)$ and an upper bound of π. After the $\cos i$ and ν integrals we obtain the probability density for **Case 1** in terms of e, a and r_p. In the equation below K and F represent the elliptic functions of the first and second kind respectively.

$$\int \frac{(1-e^2)^{3/2}}{(1+e\cos\nu)^2} \frac{1}{r_p} \frac{1}{\sqrt{[1-(\frac{r\cos i}{r_p})^2][(\frac{r}{r_p})^2-1]}} \mathrm{d}\nu \ \mathrm{d}\cos i =$$

$$\frac{r_p}{\pi a} \sqrt{\frac{1}{(a(1-e)-r_p)(a(e+1)+r_p)}} \left[K\left(1-\frac{(a(e-1)-r_p)(ea+a-r_p)}{(a(e-1)+r_p)(ea+a+r_p)}\right) \right.$$

$$\left. - F\left(\arctan\left(\sqrt{-\frac{(a(e-1)+r_p)(ea+a+r_p)}{(e^2-1)\,a^2-2ear_p+r_p}}\right) \middle| 1 - \frac{(a(e-1)-r_p)(ea+a-r_p)}{(a(e-1)+r_p)(ea+a+r_p)}\right) \right].$$

We can then make a contour plot of the above equation as shown in Figure 6.4.1. Note we drop constant factors since they can be found by normalization and we drop the prior on a as we can take it into account after integration. I also only include the differentials for the integrals we have actually executed. We can see that the contour plot covers accurately some but not the whole distribution we would expect. The one constraint we have is that the projected radius r_p can never be measured greater then the largest physical distance for some eccentricity, which implies that $r_p \leq a(1+e)$. The curve between regions D and C is given by the equation $1 = a(1+e)$ as we have set $r_p = 1$. It therefore makes sense that all combinations of a and e in region D are not physical. They would give orbits where the projected radius is greater then the maximum possible distance between the planet and the star. We might similarly try to argue that the projected radius cannot be smaller then the minimum physical distance, however by inclining the orbits, we can achieve an arbitrarily small projected radius. We should therefore still have expected to find probability in all three regions A, B, and C.

We would need to solve the integrals for case two and case three in order to fill in the rest of the plot. One strategy would be to use the symmetries that we have studied here to find more appropriate variables that make integration easier. While we have yet
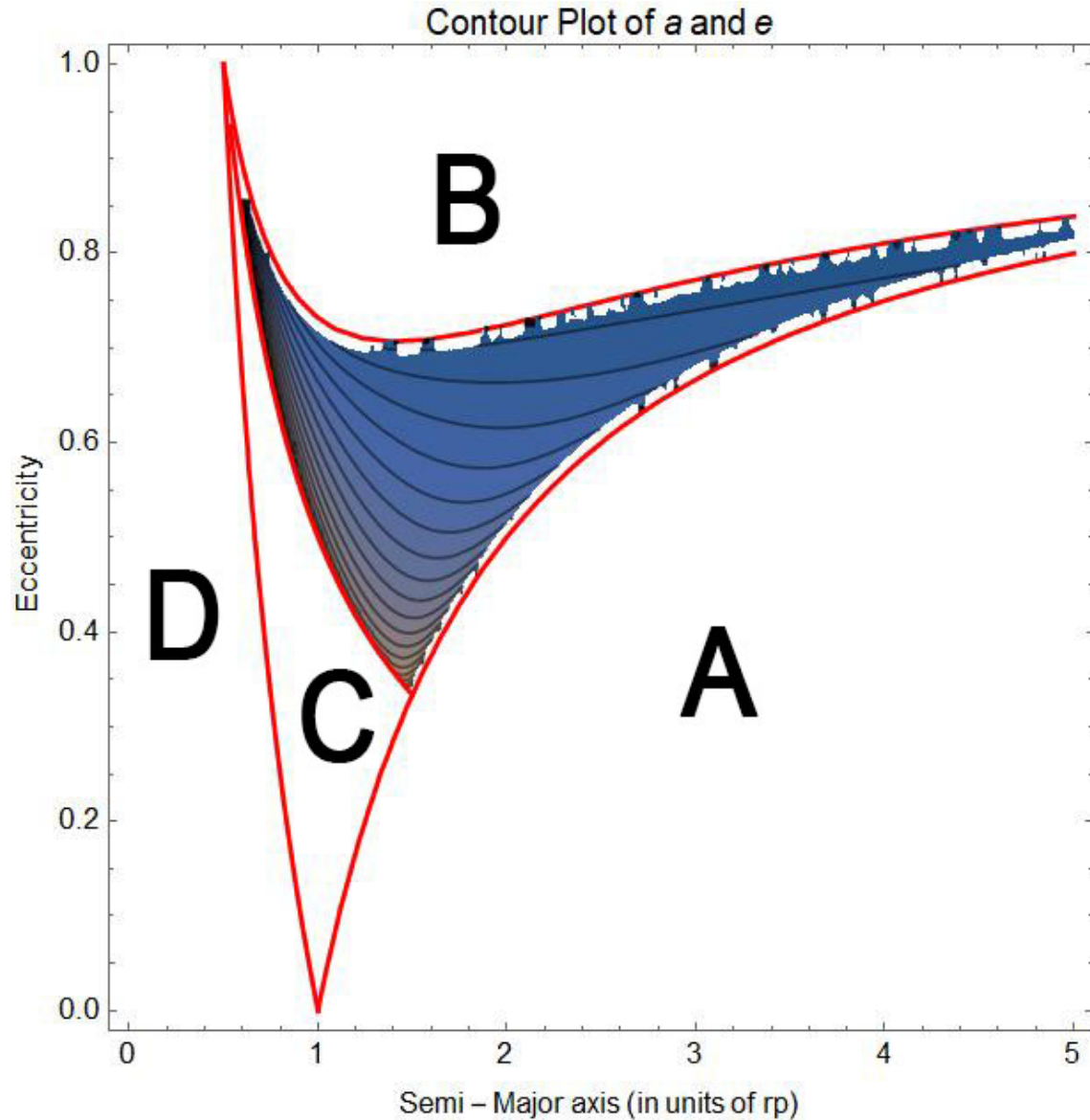
Figure 6.4.1: A contour plot of the joint probability distribution, found by marginalizing the **Case 1** probability density. The y-axis is eccentricity and the x-axis is semi-major axis and r_p is fixed to $r_p = 1$. We notice that the full distribution should include the region all of regions A, B and C only excluding region D.

to find coordinates that would allow us to solve the other integrals we have found a coordinate system that simplifies the boundaries in the contour plot. In particular we let $u = a(1 - e) - r_p$ and $v = a(1 + e) + r_p$. The new coordinates u and v are nicer as they represent the curves when $r_p = r_{min} = a(1 - e)$ and $r_p = r_{max} = a(1 + e)$, respectively. Using these variables we get a new contour plot.
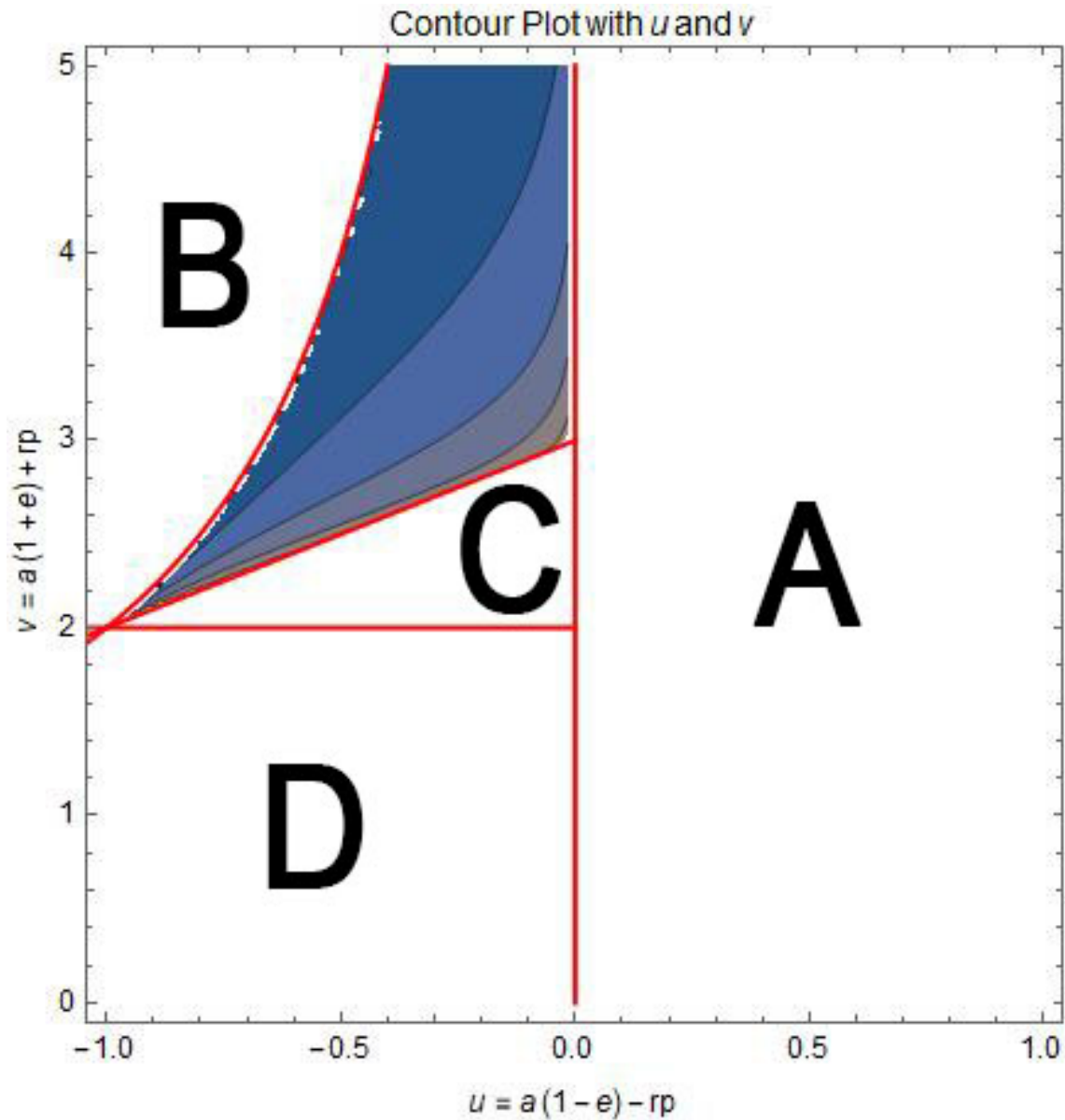
Figure 6.4.2: A contour plot of the joint probability distribution, found by marginalizing the **Case 1** probability density, plotted with the variable $u = a(1 - e) - r_p$ on the x-axis and $v = a(1 + e) + r_p$ on the y-axis and $r_p = 1$. We notice that the full distribution should include the region all of regions A, B and C only excluding region D.

We are hopeful that we will be able to use analytic continuation to extend the validity of the the integral we can solve into regions A, B and C.

7
Conclusion

In preparation for direct image data, I have been working to derive probability distributions for the semi-major axis given a single direct image. I have been able to derive the probability in some regions of the semi-major axis-eccentricity plane, i.e the ea-plane, but so far in other regions the integrals have been intractable. To address these difficulties I have been working on two strategies, one involving analytic continuation, and the other to use the symmetries of the Kepler problem to find more a more appropriate coordinate system. I have focused in this paper on understanding the relationship between conserved quantities and symmetries in an effort to understand the subtle symmetry associated with the Laplace-Runge-Lenz Vector. The plots in Figure 5.1.1 demonstrate the phase space rotational symmetry of the LRL vector. The orbits change shape as a result of a projection from phase space into configuration space.

I intended to make the project accessible to physics students in roughly their junior year. In order to do so, I explained the Kepler problem, the conserved quantities and symmetries of Keplerian orbits in-depth, introduced the Poisson bracket and flow variables, and provided a graphical representation of the symmetry of the LRL Vector. I have also included the derivation of the partial results for the probability distribution of a given r_p.

Going forward, I would love to see the connection between the symmetry of the LRL vector and a more appropriate coordinate system strengthened. For instance by considering the various coordinate systems I have already used, including the action angle variables and the linear combinations of action angle variables, as well as those outlined in Gořansson ([9]). Gořansson recasts the energy of the Kepler problem as a three-sphere via a nonstandard parameterization. It could also be fruitful to dive deeper into the connection between group theory and the LRL and angular momentum vectors; I scratch the surface of the connection in Appendix C.

Appendix A
Action Angle Variables

The Hamilton-Jacobi equation is a formulation of classical mechanics equivalent to New-
tonian, Lagrangian, and Hamiltonian Mechanics, but that allows us to think of particles
as waves. We use the Hamilton-Jacobi equation to derive the action angle variables. As
a result the action-angle variables are well adapted for a particular system. The angle
variable represents moving along a wave front, and the action variable perpendicularly to
the wave front. The Hamilton-Jacobi Equation in one dimension is given by

$$H(p, x, t) = H\left(\frac{\partial S}{\partial x}, x, t\right) = -\frac{\partial S}{\partial t}. \qquad \text{(A.0.1)}$$

We can see that once we substitute in the Hamiltonian for a given system, this becomes a
partial differential equation that we may or may not be able to solve. The action variable
is defined as $J = \oint p\, dq$ where we use the Hamilton-Jacobi equation to find p in terms of
q. We also use the Hamilton Jacobi Equation to get the generating function that allows
us to find the angle variables that correspond to the action variables.

A.1   Action Angle Variables of the Kepler problem

In the case of the Kepler problem, we need the multivariate version of the Hamilton-Jacobi equation given by

$$H(q_i, p_i, t) = -\frac{\partial S}{\partial t}.$$

where $p_i = \partial S/\partial q_i$. We can then substitute in the Hamiltonian for the Kepler problem

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2 \sin(\theta)^2} - \frac{k}{r}$$

which gives the partial differential equation,

$$\frac{1}{2m}\frac{\partial S^2}{\partial r} + \frac{1}{2mr^2}\frac{\partial S^2}{\partial \theta} + \frac{1}{2mr^2 \sin(\theta)^2}\frac{\partial S^2}{\partial \phi} - \frac{k}{r} = -\frac{\partial S}{\partial t}.$$

Now we separate variables with the separation ansatz $S(r, \theta, \phi, t) = T(t) + W(r, \theta, \phi)$ giving

$$\frac{1}{2m}\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{2mr^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{1}{2mr^2 \sin(\theta)^2}\frac{\partial W^2}{\partial \phi} - \frac{k}{r} = -\frac{dT}{dt} \qquad (A.1.1)$$

Now we see that the left side does not depend on t nor does the right side depend on r or ϕ. So, each side is equal to a constant, and in particular the energy, which gives us a solution for T(t),

$$-\frac{dT}{dt} = E \implies T(t) = -Et. \qquad (A.1.2)$$

We then get another partial differential equation:

$$\frac{1}{2m}\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{2mr^2}\left(\frac{\partial W}{\partial \phi}\right)^2 + \frac{1}{2mr^2 \sin(\theta)^2}\frac{\partial W^2}{\partial \phi} - \frac{k}{r} = E \qquad (A.1.3)$$

We can now separate variables again with the ansatz, $W(r, \theta, \phi) = W_1(r, \theta) + W_\phi(\phi)$ giving the PDE

$$\frac{1}{2m}\frac{\partial W_1^2}{\partial r} + \frac{1}{2mr^2}\frac{\partial W_1^2}{\partial \theta} + \frac{1}{2mr^2 \sin(\theta)^2}\frac{dW_\phi^2}{d\phi} - \frac{k}{r} = E.$$

Here we multiply both sides by $2mr^2 \sin(\theta)^2$ and bringing all terms with r, θ to one side and those with ϕ to the other gives the equation

$$\frac{dW_\phi}{d\phi} = \sqrt{2mr^2 \sin(\theta)^2 E - r^2 \sin(\theta)^2 \frac{\partial W_1}{\partial r}^2 - \sin(\theta)^2 \frac{\partial W_1}{\partial \theta}^2 + 2mr \sin(\theta)^2 k}.$$

A.1.1 Solving for J_ϕ

Again, since the left side only depends on ϕ and the right side only depends on r and θ they are both equal to some constant α_ϕ. Then, we have

$$\frac{dW_\phi}{d\phi} = \alpha_\phi \implies W_\phi(\phi) = \alpha_\phi \phi.$$

Since $p_\phi = \frac{\partial S}{\partial \phi} = \frac{dW}{d\phi} = \alpha_\phi$, we can solve for the action variable J_ϕ.

$$J_\phi = \oint p \, dq = \oint \alpha_\phi d\phi = 2\pi \alpha_\phi$$

We can thus solve for $\alpha_\phi = J_\phi/2\pi$, which gives $W_\phi(\phi) = J_\phi \phi/2\pi$ taking ϕ to have a period of 2π. Notice that $p_\phi = 2\pi J_\phi$.

A.1.2 Solving for J_θ

Returning to the PDE we now have that

$$\sqrt{2mr^2 \sin(\theta)^2 E - r^2 \sin(\theta)^2 \frac{\partial W_1}{\partial r}^2 - \sin(\theta)^2 \frac{\partial W_1}{\partial \theta}^2 + 2mr\sin(\theta)^2 k} = \alpha_\phi.$$

Squaring both sides gives us,

$$2mr^2 \sin(\theta)^2 E - r^2 \sin(\theta)^2 \frac{\partial W_1}{\partial r}^2 - \sin(\theta)^2 \frac{\partial W_1}{\partial \theta}^2 + 2mr\sin(\theta)^2 k = \alpha_\phi^2$$

Taking the ansatz $W_1(r, \theta) = W_r(r) + W_\theta(\theta)$ the equation becomes

$$2mr^2 \sin(\theta)^2 E - r^2 \sin(\theta)^2 \frac{dW_r}{dr}^2 - \sin(\theta)^2 \frac{dW_\theta}{d\theta}^2 + 2mr\sin(\theta)^2 k = \alpha_\phi^2$$

Dividing by $\sin(\theta)^2$ and moving all r dependence to one side and all θ dependence to the other, we have

$$2mr^2 E - r^2 \frac{dW_r}{dr}^2 + 2mrk = \frac{\alpha_\phi^2}{\sin(\theta)^2} + \frac{dW_\theta}{d\theta}^2.$$

Again, since the left side only depends on r and the right only on θ both sides are equal to a constant, α_θ^2, giving

$$\frac{\alpha_\phi^2}{\sin(\theta)^2} + \frac{dW_\theta}{d\theta}^2 = \alpha_\theta^2.$$

Since by definition $p_\theta = \partial S/\partial \theta = dW_\theta/d\theta$, we solve for $dW_\theta/d\theta$, to find that

$$p_\theta = \frac{dW_\theta}{d\theta} = \pm\sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin(\theta)^2}}.$$

Then, we use the definition of J_θ,

$$J_\theta = \oint p_\theta d\theta = \pm \oint \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{(\sin\theta)^2}} d\theta.$$

To evaluate the integral, following [8], we first define $\cos\gamma = \alpha_\phi/\alpha_\theta$ which allows us to write,

$$J_\theta = \pm \oint \sqrt{1 - \frac{(\cos\gamma)^2}{(\sin\theta)^2}} d\theta.$$

We can then let $\sin\theta_0 = \cos\gamma$ which ensures that $p_\theta = 0$ at θ_0 and $\pi - \theta_0$. Since $(\sin\theta)^2$ is even about $\theta = \pi/2$, we integrate from $\pi/2$ to θ_0 giving,

$$J_\theta = 4\alpha_\theta \int_{\frac{\pi}{2}}^{\theta_0} \sqrt{1 - \frac{(\cos\gamma)^2}{(\sin\theta)^2}} d\theta.$$

After two more substitutions:

$$\cos\theta \equiv \sin\gamma\sin\psi, \quad \text{and} \quad u \equiv \tan\psi,$$

and some work we find that,

$$J_\theta = 4\alpha_\theta \int_0^\infty \left(\frac{1}{1-u^2} - \frac{(\cos\gamma)^2}{1 + u^2(\cos\gamma)^2} \right) du = 2\pi\alpha_\theta(1 - \cos\gamma) = 2\pi(\alpha_\theta - \alpha_\phi).$$

We also notice that the sum $J_\theta + J_\phi = 2\pi\alpha_\theta$.

A.1.3   Solving for J_r

We have one more differential equation to solve.

$$2mr^2E - r^2\frac{dW_r}{dr}^2 + 2mrk = \alpha_\theta^2.$$

Then, we have that

$$p_r = \frac{dW_r}{dr} = \sqrt{2mE + 2m\frac{k}{r} - \frac{\alpha_\theta^2}{r^2}},$$

and so, we have

$$J_r = \oint \sqrt{2mE + 2m\frac{k}{r} - \frac{\alpha_\theta^2}{r^2}} dr.$$

The integral can be preformed by contour integration as stated by [8], to obtain,

$$J_r = -(J_\theta + J_\phi) + \pi k \sqrt{\frac{2m}{-E}},$$

for $E < 0$. Conveniently we can invert to find $E = E(J_r, J_\theta + J_\phi)$ gives

$$E = -\frac{2\pi^2 k^2 m}{(J_r + J_\theta + J_\phi)^2}.$$

A.1.4   Finding the Angle Variables

Now, that we have found all the action variables J_r, J_θ and, J_ϕ we need to find their corresponding angle variables. We do so by treating $W = W_r + W_\theta + W_\phi$ as the generating function. I will discuss generating functions in more depth in the next section. Returning to our formulas for W_r, W_θ, and W_ϕ from the previous sections, we find that

$$W = \frac{\phi J_\phi}{2\pi} \pm \oint \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{(\sin\theta)^2}} d\theta \pm \oint \sqrt{-\frac{4\pi^2 k^2 m^2}{(J_r + J_\theta + J_\phi)^2} + 2m\frac{k}{r} - \frac{\alpha_\theta^2}{r^2}} dr$$

Substituting for α_i we have

$$W = \frac{\phi J_\phi}{2\pi} \pm \oint \frac{1}{2\pi} \sqrt{(J_r + J_\theta + J_\phi)^2 - \frac{J_\phi^2}{(\sin\theta)^2}} d\theta$$

$$\pm \oint \sqrt{-\frac{(2\pi k m)^2}{(J_r + J_\theta + J_\phi)^2} + 2m\frac{k}{r} - \frac{(J_\theta + J_\phi)^2}{(2\pi r)^2}} dr.$$

The angle variables, Θ_r, Θ_θ and Θ_ϕ, are found by taking derivatives of the generating function with respect to each action variable. Mathematically,

$$\Theta_r = \frac{\partial W}{\partial J_r}, \quad \Theta_\theta = \frac{\partial W}{\partial J_\theta} \quad \text{and,} \quad \Theta_\phi = \frac{\partial W}{\partial J_{\partial\phi}}.$$

I have left the simpler case, using the two-dimensional Hamiltonian for the Kepler problem, as an exercise for the reader. The results for the two-dimensional case, as quoted in Chapter 5 with $I_i = J_i/2\pi$, are

$$I_\phi = p_\phi,$$

$$H = \frac{-mk^2}{2(I_r + I_\phi)^2},$$

$$\Theta_r = \sqrt{2mkr - \frac{m^2k^2r^2}{2(I_r + I_\phi)^2} - I_\phi^2} - \arcsin \frac{1 - \frac{mkr}{2(I_r+I_\phi)^2}}{\sqrt{1 - \frac{I_\phi^2}{2(I_r+I_\phi)^2}}}, \quad \text{and}$$

$$\Theta_\phi = \phi - \arcsin \frac{1 - \frac{I_\phi^2}{mkr}}{\sqrt{1 - \frac{I_\phi^2}{2(I_r+I_\phi)^2}}} + \Theta_r.$$

Appendix B
Generating Functions

In this chapter, I give the Hamiltonian understanding of generating functions, as well as what I call the "big phase space" picture of generating functions. The purpose is to give some justification for our derivation of the angle variables and to give some intuition for the connection between generating functions and our phase space coordinate transformations.

B.1 Hamiltonian Mechanics and generating functions

Since we have already briefly described the structure of Hamiltonian Mechanics in Chapter 4, I will begin here with generating functions as they relate directly to Hamiltonian Mechanics.

In Hamiltonian Mechanics we know that for a Hamiltonian $H = H(q, p, t)$

$$\dot{q} = \frac{\partial H}{\partial p} \text{ and } \dot{p} = -\frac{\partial H}{\partial q}$$

regardless of the coordinate system we choose. So, if we want to transform to a new coordinate system $Q(q, p, t)$ and $P(q, p, t)$ and new Hamiltonian $K(Q, P, t)$ the same relationships must hold. In particular,

$$\dot{Q} = \frac{\partial K}{\partial P} \quad \text{and} \quad \dot{P} = -\frac{\partial K}{\partial Q}.$$

It is now useful to point out that Hamilton's equations are derived from Hamilton's principle that states that objects trace the path that minimizes the action. Mathematically that the variation of the action is 0, or ,

$$\delta S = 0.$$

Hamilton's Principle is equivalent to Hamilton's equations. For now, we will just take Hamilton's principle and the equivalence of these formulations as given.

The action which we also defined in Chapter 4 is given by

$$S = \int_{t_1}^{t_2} \mathcal{L} dt,$$

where the Lagrangian ($\mathcal{L}$) is defined by $\mathcal{L} = \mathcal{L}(q, \dot{q}, t) = T - V$ and where T is the kinetic energy and V is the potential energy. We can rewrite the Lagrangian as $\mathcal{L} = T - V = 2T - H$. We can also rewrite $T = m\dot{q}^2/2 = p\dot{q}/2$, allowing us to conclude that

$$S = \int_{t_1}^{t_2} (p\dot{q} - H(q, p, t)) dt, \quad \text{and} \quad S = \int_{t_1}^{t_2} (P\dot{Q} - K(Q, P, t)) dt.$$

Since we know that Hamilton's equations and Hamilton's principle are equivalent and that both coordinate systems satisfy Hamilton's equations,

$$\delta \int_{t_1}^{t_2} (p\dot{q} - H(q, p, t)) dt = 0, \quad \text{and} \quad \delta \int_{t_1}^{t_2} (P\dot{Q} - K(Q, P, t)) dt = 0$$

These equations do not allow us to set the integrands equal. However, we can say that they are related by the addition of the time derivative of some function F,

$$\lambda(p\dot{q} - H) = P\dot{Q} - K + \frac{dF}{dt}.$$

To see why this is true, notice that

$$\delta \int_{t_1}^{t_2} \frac{dF}{dt} dt = \delta(F(t_1) - F(t_2)).$$

Since, our end points t_1 and t_2 are not changing,

$$\delta \int_{t_1}^{t_2} \frac{dF}{dt} dt = \delta(F(t_1) - F(t_2)) = 0.$$

We call F the generating function for the transformation and can choose one of four versions of F which each depend on a different pair of coordinates: $F_1(q, Q, t), F_2(q, P, t), F_3(p, Q, t)$ and, $F_4(p, P, t)$. I will explain why these are the possible generating functions in the next section. Notice that each of these depends on one new coordinate and one old coordinate. We choose the one that is simplest in a given context.

Once we have a generating function we can find relationships between the two coordinate systems. For instance, suppose we had an F_1 generating function. We see that

$$p\dot{q} - H = P\dot{Q} - K + \frac{dF_1(q, Q, t)}{dt}.$$

Then, via chain rule we find that

$$p\dot{q} - H = P\dot{Q} - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q}\frac{\partial q}{\partial t} + \frac{\partial F_1}{\partial Q}\frac{\partial Q}{\partial t},$$

which simplifies to

$$p\dot{q} - H = P\dot{Q} - K + \frac{\partial F_1}{\partial t} + \frac{\partial F_1}{\partial q}\dot{q} + \frac{\partial F_1}{\partial Q}\dot{Q}.$$

Since our coordinates are independent, we can compare the $\dot{q}$, $\dot{Q}$, and constant terms separately giving,

$$p = \frac{\partial F_1}{\partial q}, \quad P = \frac{\partial F_1}{\partial Q} \quad \text{and,} \quad K = H + \frac{\partial F_1}{\partial t}.$$

While this explicitly only tells us p and P, we can invert equations to find P, Q and K in terms of p and q. We can follow a similar process for each F_2, F_3 and, F_4 to find equations for the new coordinates in terms of the old. Notice that by choosing an generating function that does not depend on time, we would have that $H = K$. Therefore by choosing various F we can generate any number of canonical coordinate transformations.

B.2 Big Phase Space

In the previous section we used the fact that both our initial and new coordinate systems must satisfy Hamilton's principle to find the appropriate transformations. In terms of the

phase space, the condition that both systems must satisfy Hamilton's principle is related
to the preservation of the "area" in phase space called symplectic area. Notice that sym-
plectic area has different units of momentum times distance rather then distance squared.
Transformations that preserve symplectic area are called canonical transformations.

At this point I begin to use the language of differential forms. For readers not familiar
with differential forms, it's most convenient to think of an n-form as a function that takes
n vectors as input and outputs a real number. In the case of a two-form, $\omega = dp \wedge dq$, we
can think of it as the function,

$$\omega(\vec{v_1}, \vec{v_2}) = dp(\vec{v_1})dq(\vec{v_2}) - dp(\vec{v_2})dq(\vec{v_1}).$$

In phase space, vectors have q and p components i.e. $\vec{v} = q\hat{q} + p\hat{p}$. We can translate this
to the language of differential forms as $\vec{v} = q\frac{\partial}{\partial q} + p\frac{\partial}{\partial p}$. If we take $\vec{v_1} = q_1\frac{\partial}{\partial q} + p_1\frac{\partial}{\partial p}$ and
$\vec{v_2} = q_2\frac{\partial}{\partial q} + p_2\frac{\partial}{\partial p}$, then the output of the function is given by

$$\omega(\vec{v_1}, \vec{v_2}) = p_1q_2 - p_2q_1.$$

The equation above has a geometric interpretation as the symplectic area of the parallel-
ogram formed by $\vec{v_1}$ and $\vec{v_2}$.

Now, we construct a "big phase space" by considering our initial coordinate system
described by p and q and our new coordinate system described by P and Q as the basis
for the "big phase space". Then, every point (p, q, P, Q) in the big phase space corresponds
to a transformation that takes p to P and q to Q. As mentioned before, we only want to
consider transformations that preserve the phase space area. So, we impose the condition
that the symplectic areas are not changed by the transformation. In other words the
condition that $(dp \wedge dq) = (dP \wedge dQ)$. So, we can define a two-form $\Omega = (dp \wedge dq) - (dP \wedge dQ)$.

Imposing the condition $\Omega = (dp \wedge dq) - (dP \wedge dQ) = 0$, we constrain our transformations
to a two-dimensional surface. We can describe the surface via a function F that will depend
on two of the four phase space variables, since it describes a two dimensional surface. The
function F is again called the generating function. We also know that if the surface only

depended on the initial coordinates or only depended on the new coordinates, it would live entirely in the pq-plane or PQ-plane and not serve as transformations between the two. We therefore have four versions of the generating function $F_1(q, Q, t)$, $F_2(q, P, t)$, $F_3(p, Q, t)$ and, $F_4(p, P, t)$. We can use the process described in the previous section to find the formulas for P and Q in terms of p and q.

Appendix C
LRL symmetry as SO(4)

In this chapter we look more closely at the mathematical relationships between $E, \mathbf{L}$ and, $\mathbf{A}$, and see how we can derive the symmetry of the LRL vectors via group theory. I give a brief introduction to group theory, first through the formal definitions and then with some familiar examples. Those already familiar with group theory may want skip to section 2 where I describe the Lie algebra created by the Poisson brackets and the components of angular momentum and the LRL vector. Section 2 does not give as much explanatory exposition as my previous chapters. It's purpose is to give direction for a more mathematical justification for the claim that the symmetry corresponding to the LRL vector is the set of four-dimensional rotations. It also provides a jumping point for those searching for other interesting connections between my work and other areas of research.

C.1 Group Theory

C.1.1 Definition of a Group

A group G is a set imbued with an operation $*$. For understanding the following definition, the operation $*$ f intuitively can be understood as a product, though it does not have to be. The set and operation must satisfy the following properties:

- **Identity:** There exists an identity element $e \in G$ such that $e * g = g$ and $g * e = g$ for all $g \in G$.

- **Inverses:** For all $g \in G$ there exists an inverse g^{-1} such that $g * g^{-1} = e$ and $g^{-1} * g = e$.

- **Closure:** For all $g, h \in G$, the quantity $g * h \in G$.

- **Associative:** For all $f, g, h \in G$, we have that $f * (g * h) = (f * g) * h$.

Note that while associativity is required for G to be a group, commutativity is not. A group that is commutative is called an abelian group and has the additional property that for all $g, h \in G$, the quantity $g * h = h * g$.

C.1.2 *Examples of Groups*

I will now give a few examples of groups and give brief discussions of how they satisfy the properties above.

Real Numbers: The set of all real numbers, $\mathbb{R}$, under addition is a group and more specifically an abelian group. First, we know that there exists an identity element namely 0. If I add 0 to any number in $\mathbb{R}$, I get the same number back. We also see that g and $-g$ are inverses since $g + (-g) = 0$ and $(-g) + g = 0$. For closure, we note that no matter what two numbers I start with in the real numbers if I add them I still get back a real number. We also know that addition is associative and therefore $G = (\mathbb{R}, +)$ is a group.

I won't go through the properties but you should check that the set $\mathbb{R} - \{0\}$ referring to all the real numbers except 0 is a group under multiplication. Also consider why we must exclude 0 in order to have a group.

Rotations on points in two dimensions: The set of all points in two dimensions under rotations also forms a group. We know that the identity rotation is the transformation that simply does nothing. Inverses are rotations in the opposite direction. we cannot rotate a point in anyway such that it leaves the two-dimensional plane. It's harder to

check but you can also show that rotations are associative. It's also possible to show that rotations in any dimension also form groups.

C.1.3 Non-examples of Groups

I will now give a few examples of sets that are not groups.

The set of natural numbers under addition: Notice that the set of natural numbers, $\mathbb{N}$ does not contain 0 and therefore does not satisfy the identity property. Even if we consider the set $\mathbb{N} \cup \{0\}$, there are no negative numbers and therefore no inverses in $\mathbb{N} \cup \{0\}$.

The set $A = \{-3, -2, -1, 0, 1, 2, 3\}$ under addition: Notice that the set A satisfies the first two properties. There is an identity, and inverses in A. However, the set is not closed. If we take the elements $2 + 3$ we get 5 which is outside the set.

C.2   Group Algebra via Poisson bracket

From Rogers, [14] we know the components of the angular momentum form a Lie algebra via the Poisson bracket. Lie algebras describe the tangent space of the identity of a continuous smooth group. In particular, we have the relationship,

$$\{L_i, L_j\} = \epsilon_{ijk} L_k,$$

where ϵ_{ijk} is the Levi-Civita symbol defined by

$$\begin{cases} -1 & \text{if } (i, j, k) \text{ is } (1, 2, 3), (2, 3, 1) \text{ or } (3, 1, 2) \\ 0 & \text{if } i = j, j = k, \text{ or, } k = i \\ -1 & \text{if } (i, j, k) \text{ is } (3, 2, 1), (2, 1, 3) \text{ or } (1, 3, 2). \end{cases}$$

Combining both angular momentum and the LRL vector we have that the components together form another Lie algebra, via the Poisson bracket. Using the normalization $\widetilde{\boldsymbol{A}} =$

$\frac{1}{\sqrt{-2mE}}\boldsymbol{A} = \frac{1}{\sqrt{-2mE}}(\boldsymbol{p} \times \boldsymbol{L} - mk\hat{\boldsymbol{r}})$ we get the following relationships,

$$\{L_i, L_j\} = \epsilon_{ijk}L_k$$

$$\{\tilde{A}_i, L_j\} = \epsilon_{ijk}\tilde{A}_k$$

$$\{\tilde{A}_i, \tilde{A}_j\} = \epsilon_{ijk}L_k.$$

These relationships define a Lie algebra related to the group $SO(4)$, which is the group of all four-dimensional rotations.

From Rogers, [14], we also have the relationship between the symmetry group, $SO(4)$ and the two $SO(3)$ groups, by defining two new quantities $\boldsymbol{M}, \boldsymbol{N}$ that are linear combinations of $\boldsymbol{L}$ and $\boldsymbol{A}$. Then, we get three new relationships,

$$\boldsymbol{M} = \frac{1}{2}(\boldsymbol{L} + \boldsymbol{A})$$

$$\boldsymbol{N} = \frac{1}{2}(\boldsymbol{L} - \boldsymbol{A})$$

These linear combinations satisfy the Poisson bracket relations:

$$\{N_i, N_j\} = \epsilon_{ijk}N_k$$

$$\{M_i, M_j\} = \epsilon_{ijk}M_k$$

$$\{M_i, N_j\} = 0.$$

Here we see that $SO(4)$ can indeed be written instead as two $SO(3)$ groups as $SO(4) = SO(3) \bigotimes SO(3)$. We are able to make many simplifications by fixing $\boldsymbol{L}$. In doing so we break some the $SO(4)$ symmetry into $SO(3)$. For a more in-depth discussion see [14].

Bibliography

[1] Astronomical Returns, *Orbital Position and Velocity*, `https://www.astronomicalreturns.com/p/section-45-orbital-position-and-velocity.html`. Accessed 30 April 2022.

[2] Keaton Bell, *Inclination angles and a uniform distribution for isotropy* (19 Feb, 2018), `http://keatonb.github.io/archivers/uniforminclination`. Accessed 30 April 2022.

[3] Pat Brennan, *Observing Exoplanets: What Can We Really See?* (28 October 2019), `https://exoplanets.nasa.gov/news/1605/observing-exoplanets-what-can-we-really-see/`. Accessed 30 April 2022.

[4] Margret Bruna, Nicolas Cowan, Audrey Bourdon, Hal M. Haggard, Mathilde Mâlin, and Julia Sheffler, *Combining Photometry and Astrometry to Improve Orbit Retrieval of Directly Imaged Exoplanets*, in preparation.

[5] CheCheDaWaff, *File:Eccentric and True Anomaly.svg* (26 April 2016), `https://commons.wikimedia.org/w/index.php?curid=48384905`. Accessed 30 April 2022.

[6] Ben Farr, Will M. Farr, Nicolas B. Cowan, Hal M. Haggard, and Tyler Robinson, *exocartographer: A Bayesian Framework for Mapping Exoplanets in Reflected Light* **156** (19 Feb 2018), no. 4, 146, DOI arXiv:1802.06805 [astro-ph.EP].

[7] Herbert Goldstein, *Classical Mechanics*, second edition, Addison-Wesley series in physics, Addison-Wesley Longman, Incorporated, 1980.

[8] Joanna Gonera, Pitor Kosiński, and Patryk Michel, *Global Symmetries of the Kepler Problem*, posted on 28 Apr 2021, DOI arXiv:2104.14416 [physics.class-ph].

[9] Jesper Goŕansson, *Symmetries of the Kepler Problem* (8 Mar 2015).

[10] Lasunncty, *File:Orbit1.svg* (10 October 2007), `https://commons.wikimedia.org/w/index.php?curid=8971052`. Accessed 30 April 2022.

[11] Carl D. Murray and Alexandre C.M. Correia, *Keplerian Orbits and Dynamics of Exoplanets*, posted on 25 Feb 2011, 1-4, DOI arXiv:1009.1738 [astro-ph.EP].

[12] NASA, *Exoplanet Discoveries* (14 December 2017), `https://www.nasa.gov/sites/default/files/thumbnails/image/fig10-exoplanetdisc-dec14.jpg`. Accessed 30 April 2022.

[13] _____, *Transiting Exoplanet Survey Satellite (TESS)* (22 March 2021), `https://exoplanets.nasa.gov/tess/`. Accessed 30 April 2022.

[14] Harold H. Rogers, Journal of Mathematical Physics **14** (1973), 1125, DOI https://doi.org/10.1063/1.1666448.

[15] Ross C. O'Connell and Kannan Jagannathan, *Illustrating dynamical symmetries in classical mechanics: The Laplace–Runge–Lenz vector revisited*, American Journal of Physics **71** (13 February 2003).

[16] Tom Schulte, *Kepler's Laws of Planetary Motion* (2017), `https://www.aplusphysics.com/courses/honors/ucm/Kepler.html`. Accessed 30 April 2022.

[17] Ian Stewart, *Lecture Notes 8.09(F14) CHAPTER 4: CANONICAL TRANSFORMATIONS, HAMILTON-JACOBI EQUATIONS, AND ACTION-ANGLE VARIABLES* (Fall 2014), `https://ocw.mit.edu/courses/8-09-classical-mechanics-iii-fall-2014/resources/mit8_09f14_chapter_4/`.

[18] user121799, *"Nice" 2D ellipse in 3D space* (12 April 2018), `https://tex.stackexchange.com/questions/426241/nice-2d-ellipse-in-3d-space`. Accessed 30 April 2022.

[19] Vinod Wadhawan, *The Vinod Wadhawan Blog* (21 April 2012), `http://vinodwadhawan.blogspot.com/2012/04/24-attractors-in-phase-space.html`. Accessed 30 April 2022.

[20] Robert Wilson, *Astronomy through the Ages*, Princeton University Press, Princeton, New Jersey, 1997.

[21] Roland Winkler, *Un packing the proposed exo-planet imaging telescope HabEx* (19 September 2019), `https://www.nasaspaceflight.com/2019/09/unpacking-proposed-exo-planet-imaging-telescope-habex/`. Accessed 30 April 2022.