DELFT UNIVERSITY OF TECHNOLOGY

BACHELOR THESIS

APPLIED MATHEMATICS and APPLIED PHYSICS

The Runge-Lenz vector and symmetries of the Kepler problem

Author:

Dirkzwager, F.J.M.

June 25, 2025

Committee:

Dr. S.E. Zegers

Prof. dr. Y.M. Blanter

Dr. A.R. Akhmerov

Dr. P.M. Visser

TU Delft Delft University of Technology

Abstract

This thesis investigates the conserved Runge–Lenz vector in systems governed by inverse-square central forces. By analyzing its associated symmetries through the framework of Lie groups and Lie algebras, we explore its role in both classical and quantum mechanical settings. In each case a hidden $\mathfrak{so}(4)$ symmetry is revealed. In the classical regime, this is mapped to a $SO(4)$ group action. Whilst in the quantum regime, this symmetry is used to calculate the energy levels of the hydrogen atom. This thesis was written as part of the Bachelor's programs in Applied Physics and Applied Mathematics at Delft University of Technology.

Contents

1 Introduction

Symmetries, not only can they be visually beautiful, they can also be incredibly useful in mathematics and physics. In physics, when certain conditions are met, symmetries directly lead to conservation laws through Noether's theorem [1]. An example of this is the Runge-Lenz vector, a conserved quantity in the Kepler problem, which also plays an interesting role in understanding the hydrogen atom in quantum mechanics.
The history of the Runge-Lenz vector is surprisingly convoluted. Although it is named after Carl Runge and Wilhelm Lenz, its origins can be traced back much earlier. A formulation of a similar conserved vector appeared in correspondence between Jakob Hermann and Johann Bernoulli in 1710 [2], and later in Laplace's *Traité de mécanique céleste* (1799) [3], which is why the vector is sometimes referred to as the Laplace-Runge-Lenz vector. William Rowan Hamilton independently rediscovered the vector in 1845[4]. Runge introduced it as an example in a book on vector analysis, and Lenz used it when studying the hydrogen atom. Despite these layered origins, the name "Runge-Lenz vector" persisted.
In 1926, Wolfgang Pauli, a former research assistant of Lenz, used the Runge-Lenz vector to compute the energy levels of the hydrogen atom without solving Schrödinger's equation directly [5]. Later, in 1935, Vladimir Fock offered a complementary approach using Lie groups [6]. More recently, regularization methods such as those of Moser and Ligon-Schaaf [7, 8], have deepened our understanding of the symmetries in the Kepler problem. These approaches have found their way into more recent work as well, including studies connected to the Birkhoff conjecture [9].
The aim of this thesis is to provide an overview of the symmetry behind the Runge-Lenz vector in both classical and quantum mechanical settings. In particular, we will show how certain mathematical tools can be used to describe this symmetry and to compute the energy levels of the hydrogen atom.
This thesis is structured as follows:

- **Chapter 2** begins with a Newtonian formulation of the Kepler problem and introduce the classical Runge-Lenz vector.
- **Chapter 3** reviews the necessary geometric tools, focusing on the Hamiltonian formalism and elements of differential geometry need for describing the space where the Kepler problem lives.
- **Chapter 4** develops a general theory of symmetry via Lie groups and Lie algebras. For classical systems, we also study Poisson algebras. We conclude with an overview of representation theory and Casimir operators, which we will be using to compute the energy levels of the hydrogen atom.
- **Chapter 5** revisits the classical Kepler problem and describe its symmetry using the formalism developed in previous chapters, including a discussion of the Moser and Ligon-Schaaf regularization maps.
- **Chapter 6** examines the quantum Kepler problem and compute the hydrogen atom energy levels using the Casimir operator, analogous to how Pauli did it. At the end we will also glance at Fock's method.

2 The Classical Kepler Problem

The Kepler problem is one of the cornerstones of classical mechanics, describing the motion of two bodies under the influence of an inverse-square central force. Historically, it emerged from efforts to understand the motion of planets in the solar system.
In the early 17^{th} century, the German astronomer Johannes Kepler, using the precise observational data of Tycho Brahe, formulated three empirical laws of planetary motion. These laws, which describe how planets orbit the Sun, were published between 1609 and 1619[10]:

1. The orbit of a planet is an ellipse with the Sun at one of the foci.
2. A line segment joining a planet and the Sun sweeps out equal areas during equal intervals of time.
3. The square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit.

Kepler derived these laws without a fundamental theory of dynamics; they were found empirically. A deeper physical understanding came with the work of Isaac Newton.
Later in the 17^{th} century, Newton formulated the **law of universal gravitation**, which quantitatively explained the forces underlying Kepler's laws. According to Newton, every two masses in the universe attract each other with a force proportional to the product of their masses and inversely proportional to the square of the distance between them:

$$\vec{F}_{\text{gravity}} = -G\frac{m_1 m_2}{r^2}\hat{r} \tag{1}$$

Here, G is the gravitational constant, m_1 and m_2 are the masses of the two interacting bodies, r is the distance between them, and $\hat{r}$ is the unit vector pointing from one mass to the other. The negative sign indicates that the force is attractive.
This gravitational force is an example of an **inverse-square law**, meaning the force decreases with the square of the distance between the objects. Another famous inverse-square law is **Coulomb's law**, which describes the electrostatic force between two electric charges:

$$\vec{F}_{\text{electrostatic}} = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}\hat{r} \tag{2}$$

In this expression, q_1 and q_2 are the electric charges, ε_0 is the vacuum permittivity, and $\hat{r}$ again denotes the direction of the force. The electrostatic force can be attractive or repulsive depending on the signs of the charges.
In the remainder of this chapter, we will study these inverse-square force laws in detail, by looking at some conserved quantities and finding out the shapes of the orbits of the 2 particle system.

2.1 Inverse Square Laws

In physics, we often encounter forces governed by the so-called **inverse square law**. Under such a law, the force $\mathbf{F}$ between two bodies is given by

$$\mathbf{F} = -\frac{k}{r^2}\hat{\mathbf{r}} \tag{3}$$

where k is a constant, $r = |\mathbf{r}|$ is the distance between the bodies, and $\hat{\mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|}$ is the unit vector in the direction of $\mathbf{r}$. Two of the most prominent examples of inverse square laws are Newton's law of gravitation and Coulomb's law as discussed earlier. For gravity, we have

$k = Gm_1m_2$, see (1). In the case of the electrostatic (Coulomb) force, we have $k = \frac{q_1q_2}{4\pi\varepsilon_0}$, as in (2). Using Newton's second law in conjunction with (3), the equation of motion becomes:

$$m\frac{d^2\mathbf{r}}{dt^2} = -k\frac{\mathbf{r}}{r^3}.$$

Taking the scalar product of both sides with $\frac{d\mathbf{r}}{dt}$ and integrating with respect to time yields to conservation of energy:

$$m\frac{d^2\mathbf{r}}{dt^2}\cdot\frac{d\mathbf{r}}{dt} = -k\frac{\mathbf{r}}{r^3}\cdot\frac{d\mathbf{r}}{dt},$$

$$\frac{m}{2}\left|\frac{d\mathbf{r}}{dt}\right|^2 - \frac{k}{r} = E,$$

$$E = \frac{m}{2}\left|\dot{r}^2 + r^2\dot{\theta}^2\right|^2 - \frac{k}{r}. \tag{4}$$

Where the integration constant E represents the total energy of the system. So since $\frac{dE}{dt} = 0$ we say that E is conserved, in physics conserved quantities relate to symmetries, in the case of energy it is related to time invariance of a system. How these symmetries appear will be discussed in Chapter 5. There are two other conserved quantities that we are interested in, angular momentum and the Runge-Lenz vector which we will now study.

2.2 Angular Momentum

Angular momentum plays a central role in rotational dynamics and is especially important in systems with an inverse square force like (3). We begin by defining angular momentum.

Definition 2.1. The **angular momentum L** of a particle of mass m moving with velocity $\mathbf{v}$ and located at position vector $\mathbf{r}$ relative to some origin is defined as:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \tag{5}$$

where $\mathbf{p} = m\mathbf{v}$ is the linear momentum of the particle.

We can rewrite $L = |\mathbf{L}|$ in polar coordinates:

$$L = mr\dot{\theta}, \tag{6}$$

and use this to rewrite (4) to:

$$E = \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} - \frac{k}{r}.r \tag{7}$$

Where we have thus created an effective potential given by:

$$V_{\text{eff}}(r) = \frac{L^2}{2mr^2} - \frac{k}{r}.$$

We will use this later on when classifying different orbits, but now we will look at a very important property of angular momentum.

Proposition 2.2. *If a particle is subject to the inverse square force such as* (3)*then its angular momentum is conserved:*[1]

$$\frac{d\mathbf{L}}{dt} = 0. \tag{8}$$

[1] Angular momentum is conserved for any central force and the proof of this is effectively the same as the one below, but we do not need this generalization of this result.

Proof. We compute the time derivative of the angular momentum:

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}).$$

Using the product rule for derivatives:

$$\frac{d\mathbf{L}}{dt} = \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt}.$$

Since $\mathbf{p} = m\frac{d\mathbf{r}}{dt}$, we have:

$$\frac{d\mathbf{r}}{dt} \times \mathbf{p} = \frac{d\mathbf{r}}{dt} \times m\frac{d\mathbf{r}}{dt} = \mathbf{0}.$$

Thus,

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \frac{d\mathbf{p}}{dt}.$$

By Newton's second law, $\frac{d\mathbf{p}}{dt} = \mathbf{F}$. Therefore:

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \mathbf{F}.$$

If the force is inverse square force, it has the form:

$$\mathbf{F} = -\frac{k}{r^2}\hat{\mathbf{r}},$$

where $\hat{\mathbf{r}} = \frac{\mathbf{r}}{r}$ is the unit vector in the direction of $\mathbf{r}$. Since $\mathbf{r}$ and $\hat{\mathbf{r}}$ are parallel, their cross product is zero:

$$\mathbf{r} \times \mathbf{F} = \mathbf{r} \times \left(-\frac{k}{r^2}\hat{\mathbf{r}}\right) = \mathbf{0}.$$

Therefore,

$$\frac{d\mathbf{L}}{dt} = \mathbf{0}.$$

□

2.3 Runge–Lenz Vector

In addition to angular momentum, another conserved quantity arises in the case of the inverse-square central force: the **Runge–Lenz vector**. This vector characterizes the orientation and shape of planetary orbits in the Kepler problem.

Definition 2.3. The Runge–Lenz vector $\mathbf{A}$ for a two body problem subject to an inverse-square force is defined as:

$$\mathbf{A} = \mathbf{p} \times \mathbf{L} - mk\frac{\mathbf{r}}{r}, \tag{9}$$

where $\mathbf{p}$ is the linear momentum, $\mathbf{L}$ is the angular momentum, $\mathbf{r}$ is the position vector, k is a constant determining the strength of the force and m the reduced mass calculated by $m = \frac{m_1 m_2}{m_1+m_2}$.

Proposition 2.4 ([11])**.** *For a particle moving under an inverse-square central force, the Runge–Lenz vector* $\mathbf{A}$ *is conserved:*

$$\frac{d\mathbf{A}}{dt} = \mathbf{0}. \tag{10}$$

Proof. Starting from Newton's second law with the inverse-square force:

$$m\ddot{\mathbf{r}} = -\frac{k}{r^2}\hat{\mathbf{r}},$$

we take the cross product of both sides with the angular momentum $\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}}$:

$$m\ddot{\mathbf{r}} \times \mathbf{L} = -\frac{k}{r^2}\hat{\mathbf{r}} \times \mathbf{L}.$$

Using the vector triple product identity:

$$\hat{\mathbf{r}} \times (\mathbf{r} \times \dot{\mathbf{r}}) = (\hat{\mathbf{r}} \cdot \dot{\mathbf{r}})\mathbf{r} - (\hat{\mathbf{r}} \cdot \mathbf{r})\dot{\mathbf{r}},$$

and the identity $\mathbf{r} = r\hat{\mathbf{r}}$, so $\dot{\mathbf{r}} = \dot{r}\hat{\mathbf{r}} + r\dot{\hat{\mathbf{r}}}$, it follows after simplification that:

$$\ddot{\mathbf{r}} \times \mathbf{L} = k\dot{\hat{\mathbf{r}}}.$$

Now taking the time derivative of $\dot{\mathbf{r}} \times \mathbf{L} - k\hat{\mathbf{r}}$:

$$\frac{d}{dt}(\dot{\mathbf{r}} \times \mathbf{L}) = \ddot{\mathbf{r}} \times \mathbf{L} = k\dot{\hat{\mathbf{r}}}, \quad \text{and} \quad \frac{d}{dt}(k\hat{\mathbf{r}}) = k\dot{\hat{\mathbf{r}}},$$

so:

$$\frac{d}{dt}(\dot{\mathbf{r}} \times \mathbf{L} - k\hat{\mathbf{r}}) = 0.$$

Multiplying by m, obtains:

$$\frac{d}{dt}(\mathbf{p} \times \mathbf{L} - mk\frac{\mathbf{r}}{r}) = 0,$$

and thus:

$$\frac{d\mathbf{A}}{dt} = 0.$$

□

The conservation of the Runge–Lenz vector can actually be used to derive the shape of the orbits.

Proposition 2.5 ([11]). *Under an inverse-square central force, the orbit of a particle is a conic section and the trajectory satisfies:*

$$r = \frac{\mathbf{L} \cdot \mathbf{L}}{km} \cdot \frac{1}{1 + A\cos\theta}, \tag{11}$$

where $A = |\mathbf{A}|$ is the magnitude of the Runge–Lenz vector and θ is the angle between $\mathbf{A}$ and $\mathbf{r}$.

Proof. Starting from the definition of the Runge–Lenz vector:

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

we take the dot product of both sides with $\mathbf{r}$:

$$\mathbf{A} \cdot \mathbf{r} = (\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} - mkr.$$

We now compute $(\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r}$. Using $\mathbf{p} = m\dot{\mathbf{r}}$ and $\mathbf{L} = \mathbf{r} \times m\dot{\mathbf{r}}$, we have:

$$(\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} = (m\dot{\mathbf{r}} \times (\mathbf{r} \times m\dot{\mathbf{r}})) \cdot \mathbf{r} = m^2(\dot{\mathbf{r}} \times (\mathbf{r} \times \dot{\mathbf{r}})) \cdot \mathbf{r}.$$

Using the vector triple product identity:

$$\dot{\mathbf{r}} \times (\mathbf{r} \times \dot{\mathbf{r}}) = (\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})\mathbf{r} - (\dot{\mathbf{r}} \cdot \mathbf{r})\dot{\mathbf{r}},$$

and taking the dot product with $\mathbf{r}$:

$$[(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})\mathbf{r} - (\dot{\mathbf{r}} \cdot \mathbf{r})\dot{\mathbf{r}}] \cdot \mathbf{r} = \dot{\mathbf{r}}^2 r^2 - (\dot{\mathbf{r}} \cdot \mathbf{r})^2.$$

Therefore:

$$(\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} = m^2 \left(r^2 \dot{\mathbf{r}}^2 - (\dot{\mathbf{r}} \cdot \mathbf{r})^2 \right).$$

On the other hand, the magnitude squared of the angular momentum is:

$$|\mathbf{L}|^2 = m^2 |\mathbf{r} \times \dot{\mathbf{r}}|^2 = m^2 \left(r^2 \dot{\mathbf{r}}^2 - (\dot{\mathbf{r}} \cdot \mathbf{r})^2 \right).$$

Thus:

$$(\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} = |\mathbf{L}|^2,$$

and plugging this back into our earlier expression:

$$\mathbf{A} \cdot \mathbf{r} = \frac{|\mathbf{L}|^2}{m} - mkr.$$

Solving for r we obtain:

$$r = \frac{|\mathbf{L}|^2}{km} \cdot \frac{1}{1 + \frac{\mathbf{A} \cdot \mathbf{r}}{kr}}.$$

Now observe that $\mathbf{A} \cdot \mathbf{r} = Ar\cos\theta$, where θ is the angle between $\mathbf{A}$ and $\mathbf{r}$. Substituting:

$$r = \frac{|\mathbf{L}|^2}{km} \cdot \frac{1}{1 + A\cos\theta}. \tag{12}$$

This is the polar form of a conic section. For $0 < A < 1$, the orbit is an ellipse, $A = 1$ a parabola and $A > 1$ a hyperbola. □

From (12) we see that the magnitude of the Runge-Lenz vector is the eccentricity of the elliptical orbit. And it points to the perihelium, where $p = \frac{|\mathbf{L}|^2}{km}$. We can actually relate E to A by looking at the minimal value of r, at r_{min} $\dot{r} = 0$. We can use this with (4), to lead to :

$$E = \frac{L^2}{2mr_{min}^2} - \frac{k}{r_{min}}$$

and we can derive:

$$r_{min} = \frac{|\mathbf{L}|^2}{km} \cdot \frac{1}{1 + A},$$

as $\cos\theta$ can be at most 1. Hence substituting and rewriting gives:

$$E = \frac{k^2 m}{2L^2}(A^2 - 1)$$

So we conclude that:

$$\begin{aligned} E < 0 &\quad : \text{bound elliptical orbit} \\ E = 0 &\quad : \text{parabolic escape orbit} \\ E > 0 &\quad : \text{hyperbolic unbound orbit} \end{aligned}$$

This is visualized in Figure 1.

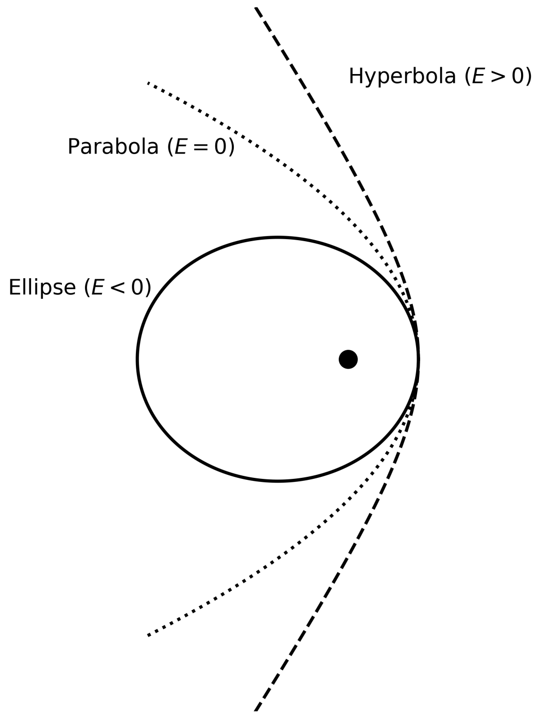

Figure 1: Orbital trajectories for different total energy levels are shown above. The solid ellipse represents a bound elliptical orbit for $E < 0$. The dashed line shows a parabolic escape trajectory where the total energy $E = 0$. The dotted line corresponds to a hyperbolic trajectory representing an unbound orbit with $E > 0$. All orbits are centered around a central mass at the origin (black dot).

3 Geometry

In this section we will discuss a couple of geometry-related mathematical concepts that we need to study the symmetries of the Kepler problem and the hydrogen atom. First we will take a short look at the Hamiltonian formalisms in which we are going to rewrite the Kepler problem in Chapter 5. Then we will look into the basic concepts of smooth manifolds which we will need for understanding Lie groups. Finally, we will apply this knowledge in an example of the unit sphere which we will expand upon in Chapter 4.

3.1 Hamiltonian Mechanics

Besides Newtons formulation of physics there are also other formulations such as the Lagrangian and Hamiltonian formulations. In this section we take a look at the latter. We will do this using [1] and [12] as our main sources. From the Lagrangian formalism we remember that for most systems we have $\mathcal{L} = T - V$ i.e. the difference of kinetic and potential energy. The Lagrangian is a function of n generalized coordinates q_i and n generalized velocities $\dot{q}_i$. sometimes time is also a parameter, hence

$$\mathcal{L}(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t) = T - V$$

In Hamiltonian mechanics instead of using generalized velocities we use generalized momenta defined by:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

Definition 3.1. The **Hamiltonian** $\mathcal{H}$ is defined as

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L}, \tag{13}$$

where q_i and p_i are the generalized coordinates and momenta. These coordinates together will be referred to as canonical coordinates.

The 2n-dimensional space these canonical coordinates live in is called the **phase space**. From the Hamiltonian we can derive equations of motions, the so called Hamilton's equations. These equations are defined as follows:

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \quad \text{for } i \in 1, \ldots, n \tag{14}$$

One can change from coordinate system to describe your problem, in Hamiltonian mechanics one has a specific format for coordinate changes in the form of a canonical transformation.

Definition 3.2 ([13])**.** A **canonical transformation** is a change of variables from (q_i, p_i) to (Q_i, P_i) that preserves the canonical form of Hamilton's equations of motion. That is, the new variables also satisfy Hamilton's equations:

$$\dot{Q}_i = \frac{\partial \mathcal{H}(Q, P, t)}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial \mathcal{H}(Q, P, t)}{\partial Q_i}, \tag{15}$$

for some Hamiltonian function $\mathcal{H}(Q, P, t)$.

An example of a canonical transformation would be the geometric Fourier transform, given by $(q, p) \mapsto (x, y) = (p, -q)$. We are now going to take a look at some of the basics of differential geometry that we will use to study phase spaces.

3.2 Manifolds

In this section we take a look at the basics of differential geometry, here we will define smooth manifolds, geodesics and the cotangent bundle. These objects will be used to look at the geometry of phase spaces with the help of symplectic geometry and we need it to define Lie groups which are needed to study symmetries. For our study of differential geometry we follow the book [14] and the reader [15]. The main object we study in differential geometry is a smooth manifold. To define what a smooth manifold is, we first need to lay some groundwork. The idea behind a smooth manifold is that we want to define a space that looks locally like $\mathbb{R}^n$, we will first define a topological manifold which forms the topological backbone we need for a smooth manifold.

Definition 3.3 ([14])**.** A **topological manifold** of dimension n is a topological space M that satisfies the following conditions:

1. M is a **Hausdorff Space**, meaning that for every pair of distinct points $p, q \in M$, there exist disjoint open subsets $U, V \subseteq M$ such that $p \in U$ and $q \in V$.

2. M is **Second-Countable**, meaning there exists a countable basis for the topology of M.

3. M is **locally Euclidean** of dimension n, meaning each point of M has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

Now we are going to build the structure needed to elevate this to a smooth manifold, for this we need to define charts.

Definition 3.4. Let M be a topological n-manifold. A **chart** on M is a pair (U, φ), where U is an open subset of M and $\varphi\colon U \to \widehat{U}$ is a homeomorphism from U to an open subset $\widehat{U} = \varphi(U) \subseteq \mathbb{R}^n$.

A chart will only map an open subset of a topological manifold to $\mathbb{R}^n$. To check if a topological manifold locally looks like $\mathbb{R}^n$ we are going to need multiple charts.

Definition 3.5. Let M be a topological set and $\mathcal{A} = \{(U_\alpha, \varphi_\alpha) : \alpha \in A\}$, with A an index set. We call $\mathcal{A}$ a topological atlas of M if $M = \bigcup_{\alpha \in A} U_\alpha$. We call $\mathcal{A}$ a smooth atlas if all charts are compatible (as defined below).

For defining compatibility, we look at intersections of charts. Suppose we have two charts $(U_\alpha, \varphi_\alpha)$ and (U_β, φ_β) that have a nonempty intersection $U_\alpha \cap U_\beta$ we wish for the points in this intersection to have a smooth connection between the charts. We do this by defining the transition functions $\kappa_{\alpha\beta} = \varphi_\beta \varphi_\alpha^{-1}$ and $\kappa_{\beta\alpha} = \varphi_\alpha \varphi_\beta^{-1}$ on $U_\alpha \cap U_\beta$. We call charts compatible when these transition functions are smooth.

Definition 3.6. A **smooth manifold** is a topological manifold of dimension n with a smooth atlas $\mathcal{A}$

Obviously $\mathbb{R}^n$ is a smooth manifold, but there are many others. In the next section, we will take a look at the unit sphere as an example. One commonly used method to show that something is a manifold, is by proving it is an embedded submanifold.

Definition 3.7 ([15], Definition 2.37)**.** Let M be a smooth manifold, then the subset $\Sigma \subseteq M$ is called a k-dimensional embedded submanifold of M if for every point $p \in \Sigma$, there exists a coordinate chart $(U_\alpha, \varphi_\alpha)$ around p in M such that

$$\varphi_\alpha(U_\alpha \cap \Sigma) = \varphi_\alpha(U_\alpha) \cap (\mathbb{R}^k \oplus \{0\}),$$

where $\mathbb{R}^k \oplus \{0\} \subseteq \mathbb{R}^k \oplus \mathbb{R}^{n-k}$ is a k-dimensional hyperplane in $\mathbb{R}^n$. A chart with this property is called a **slice chart**.

It is sometimes interesting to study if two smooth manifolds are "the same" we call this diffeomorphic.

Definition 3.8. Let M,N be smooth manifolds then a **diffeomorphism** $\phi : M \to N$ is a smooth bijection with a smooth inverse, if such a diffeomorphism exists between M and N then we call M and N **diffeomorphic**.

Another really important attribute of a smooth manifold are tangent spaces. For this we will define the following equivalence class:

Definition 3.9 ([15], Definition 4.2)**.** A tangent vector at $p \in M$ is an equivalence class $v_p = [\gamma]$ of curves through p with respect to the relation $\sim_p$. Two curves γ and λ are equivalent, $\gamma \sim_p \lambda$, if there exists a chart (U_α, ϕ_α) around p such that their derivatives at $t = 0$ in coordinates coincide, i.e.,

$$\frac{d}{dt}(\phi_\alpha \circ \gamma)(t)\bigg|_{t=0} = \frac{d}{dt}(\phi_\alpha \circ \lambda)(t)\bigg|_{t=0}.$$

This condition is independent of the chosen chart. The set T_pM of tangent vectors is called the *tangent space* of M at p.

If we take the union of all tangent spaces of a manifold we get the **tangent bundle**. A commonly used type of smooth manifold in physics is a Riemannian manifold, this is a smooth manifold on which we define a metric, a Riemannian metric, which gives us a way to measure distances.

Definition 3.10 ([15], Definition 9.1)**.** A **Riemannian metric** g on M is a covariant tensor field of rank 2 such that for every point $p \in M$, the bilinear form $g_p : T_pM \times T_pM \to \mathbb{R}$ is an inner product

To gain a proper mathematical understanding of a covariant tensor field of rank 2, where 'rank 2' specifically refers to the presence of two inputs, one is encouraged to read chapter 12 and 13 of [14] for a more in-depth treatment. If we equip a Riemannian metric on a smooth manifold M we get a Riemannian manifold denoted (M,g). The metric we will be using is the **Euclidean metric** which is the metric on $\mathbb{R}^n$ that is defined as the inner product. Using this metric we can find a lot Riemannian manifolds, namely using embedded submanifolds. When we have that (M, g), a Riemannian manifold, and $\Sigma \subseteq M$ is an embedded submanifold, then Σ is itself a Riemannian manifold in a natural way. To define a metric from this we will be using a pullback, which is defined as follows.

Definition 3.11 ([15], Definition 4.20)**.** The *pullback* along $F : M \to N$ of a function $f \in C^\infty(N)$ is the function $F^*f \in C^\infty(M)$ defined by $F^*f := f \circ F$.

The metric on Σ is then simply the pullback ι^*g of g along the canonical inclusion $\iota : \Sigma \hookrightarrow M$. Since ι_*: $T_\sigma\Sigma \to T_\sigma M$ is injective, this is indeed a Riemannian metric, for a detailed proof see proposition 9.6 of [15]. With a metric we can define a **geodesic** which is a curve of unit speed that is locally the shortest path between the points it connects. Finally we will define the covector and the cotangent space.

Definition 3.12. A **covector** at a point $p \in M$ is a linear functional $\alpha_p : T_pM \to \mathbb{R}$, meaning it maps tangent vectors at p to real numbers in a linear fashion. The **cotangent space** at p, denoted T_p^*M, is the vector space consisting of all covectors at p. It has the same dimension n as the tangent space T_pM.

Again the union of all contangent spaces of a manifold is called the **cotangent bundle** and is denoted by T^*M. We can link this cotangent bundle back to the phase space from Hamiltonian mechanics, because the phase space is the cotangent bundle of the configuration space, here the configuration space is the space spanned by the generalized coordinate q_i. There is actually a whole subfield in differential geometry called symplectic geometry which uses the language of differential geometry to formulate an even more general form of Hamiltonian mechanics. Here the central role is taken by the symplectic 2-form which plays the role of the Hamiltonian form and instead of canonical transformations their exist symplecteomorphisms. Sadly we do not have the time to delve deeper into the symplectic geometry, but a good source for further reading is [16]. We will now apply our gained knowledge to the unit sphere.

3.3 The Unit Sphere

To apply the concepts discussed in the previous section we will look at the n-dimensional unit sphere and proof it is a smooth manifold, as well as an embedded submanifold of $\mathbb{R}^{n+1}$. In the next chapter we will come back to the unit sphere and define a group action on it, which we end up using in our final analysis of the Kepler problem.

Definition 3.13. The **unit sphere** in n-dimensions, denoted by $\mathbb{S}^n$, is the surface of a n-dimensional ball. We can define this as follows.

$$\mathbb{S}^n = \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\| = 1 \right\}$$

To create charts we will use the stereographic projection. to explain this projection we will first define it for $\mathbb{S}^2$. You effectively choose a North Pole, in this case we choose to work in Cartesian coordinates and take the North Pole as $N = (0, 0, 1)$.
If we then take (x, y, z) on the sphere in Cartesian coordinates and (X, Y) on the plane, the projection and its inverse are given by the formula [17]:

$$(X, Y) = \left(\frac{x}{1-z}, \frac{y}{1-z} \right),$$

$$(x, y, z) = \left(\frac{2X}{1 + X^2 + Y^2}, \frac{2Y}{1 + X^2 + Y^2}, \frac{-1 + X^2 + Y^2}{1 + X^2 + Y^2} \right).$$

Geometrically what we are doing is for any point $\mathbf{x} = (x, y, z) \in \mathbb{S}^2$ we draw a line through N and $\mathbf{x}$ and the projected coordinates (X, Y) will be the intersection with the plane of $z = 0$, for illustration see Figure 2.

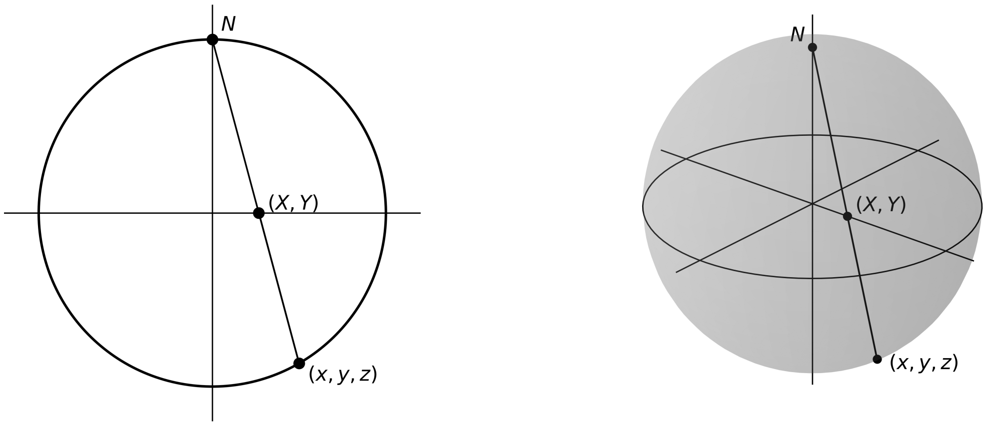

Figure 2: The geometric process behind the stereographic projection is depicted above. A line is drawn from the North Pole (N) through a point (x,y,z) on the unit sphere. The intersection of this line with the projection plane (z=0) defines the projected coordinates (X,Y). The left diagram serves as a cross-sectional view of the 3D projection depicted on the right.

We can generalize this projection to $\mathbb{S}^n$ as follows. We again choose to work with Cartesian coordinates and define the North Pole as $N = (0, 0, \dots, 0, 1)$ then we can define the projection as follows:

$$X_i = \frac{x_i}{1 - x_n} \quad \forall i \in \{1, 2, \dots, n, n+1\}. \tag{16}$$

For the inverse we first define:

$$s^2 = \sum_{j=1}^{n} X_j^2 = \frac{1 + x_{n+1}}{1 - x_{n+1}},$$

the inverse is given by

$$x_{n+1} = \frac{s^2 - 1}{s^2 + 1} \quad \text{and} \quad x_i = \frac{2X_i}{s^2 + 1} \quad (i = 1, \dots, n).$$

Proposition 3.14. *$\mathbb{S}^n$ is a smooth manifold of dimension n.*

Proof. We will leave the proof that $\mathbb{S}^n$ is a topological manifold to the reader and instead focus on constructing a smooth atlas. We will do this by constructing two charts (U_1, φ_1) and (U_2, φ_2). For this we define both a North Pole and a South Pole, with $N = (0, 0, \dots, 0, 1)$ and $S = (0, 0, \dots, 0, -1)$. Then we will take $U_1 = \mathbb{S}^n \backslash \{N\}$ and $U_2 = \mathbb{S}^n \backslash \{S\}$ We see that $\mathbb{S}^n = U_1 \cup U_2$. Then we take φ_1 as the stereographic projection defined above and take φ_2 as a stereographic projection from the South Pole. To show that $\mathbb{S}^n$ is indeed a smooth manifold it remains to show that the transition functions are smooth. We trivially see that both the stereographic projection and its inverse as defined above are smooth, hence $\kappa_{12} = \varphi_2 \varphi_1^{-1}$ and $\kappa_{21} = \varphi_1 \varphi_2^{-1}$ are smooth. □

We can actually show that $\mathbb{S}^n$ is an embedded submanifold of $\mathbb{R}^{n+1}$. This can be done by constructing a slice chart from (U_1, φ_1). We can also use this to find a metric for the unit sphere by restricting the Euclidean metric, hence we get.

$$\iota^* g_\sigma^E(v, w) = v^1 w^1 + \cdots + v^{n+1} w^{n+1}$$

With this metric we can now take a look at the geodesics of $\mathbb{S}^n$.These geodesics are also called great circles. A great circle can be parametrized as follows

$$X(t) = a \cos t + b \sin t,$$

where $a, b \in S^n$ and they are orthogonal:

$$\langle a, b \rangle = 0.$$

Later we will be specifically interested in projecting great circles. If we use (16) we get

$$X_i(t) = \frac{a_i \cos t + b_i \sin t}{1 - (a_n \cos t + b_n \sin t)}$$

Specifically for the great circles of $\mathbb{S}^3$ this becomes:

$$X_i(t) = \frac{a_i \cos t + b_i \sin t}{1 - (a_4 \cos t + b_4 \sin t)} \tag{17}$$

4 Lie Groups, Lie Algebras and Representations

In this chapter, the mathematical framework of Lie groups, Lie algebras and their representations is developed. First, Lie groups and Lie algebras are introduced and specifically the groups $SO(n)$ and $SU(n)$ are examined. These groups, and their associated algebras, will later be used to describe the symmetries in the Kepler problem. In the last section we examine representations and the Casimir operator, tools that will be used to calculate the energy levels of the hydrogen atom.

4.1 Lie Groups

In this section the definition of a Lie Group and some general information about them is discussed, then we are going to look at some important examples including a study on how $SO(n)$ acts on the units sphere.

Definition 4.1 ([18], Definition 1.20)**.** A **Lie group** is a smooth manifold G which is also a group such that the group product

$$G \times G \to G$$

and the inverse map $G \to G$ are smooth.

For our purposes we will be interested in a specific subset of Lie groups, the so called matrix Lie Groups. These are Lie groups that are subgroups of the general linear group.

Definition 4.2 ([18], Definition 1.1)**.** The **general linear group** over the real numbers, denoted $\mathrm{GL}(n;\mathbb{R})$, is the group of all $n \times n$ invertible matrices with real entries. The general linear group over the complex numbers, denoted $\mathrm{GL}(n;\mathbb{C})$, is the group of all $n \times n$ invertible matrices with complex entries.

It is useful to observe that $\mathrm{GL}(n,\mathbb{K})$ is an open subset of $\mathbb{K}^{n\times n}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. This follows from the fact that the determinant map

$$\det : \mathbb{K}^{n\times n} \to \mathbb{K}$$

is a continuous function, and $\mathrm{GL}(n,\mathbb{K}) = \det^{-1}(\mathbb{K} \setminus \{0\})$. Since $\mathbb{K} \setminus \{0\}$ is an open set in $\mathbb{K}$, the preimage under a continuous map is also open. Therefore, $\mathrm{GL}(n,\mathbb{K})$ is open in $\mathbb{K}^{n\times n}$. We will use this topological property in our proofs later.

Definition 4.3 ([18], Definition 1.4)**.** A **matrix Lie group** is a subgroup G of $\mathrm{GL}(n;\mathbb{C})$ with the following property: if A_m is any sequences of matrices in G, and A_m converges [2] to some matrix A, then either A is in G or A is not invertible.

This Definition is saying that G is a closed subgroup of $\mathrm{GL}(n;\mathbb{C})$. We also see that $\mathrm{GL}(n;\mathbb{C})$ itself is a matrix Lie group. One very important property of matrix Lie groups is that every matrix Lie Group is a Lie group as can be seen in Corollary 3.45 in [18]. Finally, just like with groups it is really useful to compare Lie groups, to do this we need to define homomorphisms between them.

Definition 4.4 ([18], Definition 1.18)**.** Let G and H be matrix Lie groups. A map Φ from G to H is called a **Lie group homomorphism** if (1) Φ is a group homomorphism and (2) Φ is continuous. If, in addition, Φ is one-to-one and onto and the inverse map Φ^{-1} is continuous, then Φ is called a **Lie group isomorphism**.

[2] Here convergence is defined with the Hilbert-Schmidt norm explained in Appendix B

4.1.1 Important Lie Groups

We will now discuss a few important Lie groups. The first Lie group we will study is the special orthogonal group.

Definition 4.5. The **special orthogonal group** $SO(n)$ is defined as:

$$SO(n) = \left\{ A \in \mathbb{R}^{n \times n} \mid AA^T = I, \det(A) = 1 \right\}$$

where $AA^T = I$ implies that

$$\langle a, b \rangle = \langle Aa, Ab \rangle \quad \forall A \in SO(n) \text{ and } a, b \in \mathbb{R}^n.$$

Geometrically speaking $SO(n)$ resembles rotation matrices. We will look more at the geometric interpretation in the next section where we let elements of $SO(n)$ act on great circles of the unit sphere.

Proposition 4.6. *$SO(n)$ is a matrix Lie Group.*

Proof. We need to show that $SO(n)$ is a **closed subgroup** of $\mathrm{GL}(n; \mathbb{R})$, since it consists of real matrices. We will first show that $SO(n)$ is a subgroup of $\mathrm{GL}(n; \mathbb{R})$. Let $A, B \in SO(n)$. Then:

$$(AB)(AB)^T = ABB^T A^T = AIA^T = AA^T = I,$$

so AB is orthogonal. Also,

$$\det(AB) = \det(A)\det(B) = 1 \cdot 1 = 1.$$

Hence, $AB \in SO(n)$. The identity matrix $I \in SO(n)$, and for any $A \in SO(n)$, its inverse is A^T, and $\det(A^T) = \det(A) = 1$, so $A^{-1} \in SO(n)$. Thus, $SO(n)$ is a subgroup of $\mathrm{GL}(n; \mathbb{R})$. We will now show that $SO(n)$ is closed in $\mathbb{R}^{n \times n}$, since $\mathrm{GL}(n; \mathbb{R})$ is an open subset of $\mathbb{R}^{n \times n}$, and $SO(n) \subset \mathrm{GL}(n; \mathbb{R})$, it follows that $SO(n)$ is closed in $\mathrm{GL}(n; \mathbb{R})$ if it is closed in $\mathbb{R}^{n \times n}$. Define the functions:

$$f : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}, \quad f(A) = AA^T, \quad \text{and} \quad g : \mathbb{R}^{n \times n} \to \mathbb{R}, \quad g(A) = \det(A).$$

Both f and g are continuous. Then:

$$SO(n) = f^{-1}(\{I\}) \cap g^{-1}(\{1\}),$$

which is the intersection of preimages of closed sets under continuous functions, hence closed in $\mathbb{R}^{n \times n}$. To conclude $SO(n)$ is a closed subgroup of $\mathrm{GL}(n; \mathbb{R})$, so by definition, it is a matrix Lie group.

□

In a similar fashion for the complex matrices we have the **special unitary group**.

Definition 4.7. The **special unitary group** $SU(n)$ is defined as:

$$SU(n) = \left\{ A \in \mathbb{R}^{n \times n} \mid AA^\dagger = I, \det(A) = 1 \right\}$$

where $AA^\dagger = I$ implies that

$$\langle a, b \rangle = \langle Aa, Ab \rangle \quad \forall A \in SU(n) \text{ and } a, b \in \mathbb{R}^n.$$

Here the † means the Hermitian of a matrix i.e. $A^\dagger = \overline{A^T}$. $SU(n)$ is also matrix Lie group, the proof of this is very similar to the proof of $SO(n)$
Interestingly enough, there is a relation between $SO(3)$ and $SU(2)$, $SU(2)$ is a double cover of $SO(3)$, so the exists a 2-to-1 Lie group homomorphism from $SU(2)$ to $SO(3)$, this also leads to the Pauli spin matrices, for more information see [19]. We have only discussed the "special" groups, here "special" refers to the $det(A) = 1$ conditions, if we drop this conditions we would get the definition of the orthogonal and unitary groups, these are however not of our interest.

4.1.2 SO(n) and the Unit Sphere

We will now look into what happens if we let elements of the Lie group $SO(n)$ act on great circles. For this we will first revise what it means for a group to act on a set.

Definition 4.8 ([20], Definition 8.1)**.** Let G be a group, and let X be a set. We say that G acts on X if for every $g \in G$ and every $x \in X$ an element $g \circ x \in X$ is given such that

- $e \circ x = x$ for all $x \in X$.
- $(gh) \circ x = g \circ (h \circ x)$ for all $g, h \in G$ and $x \in X$.

If G Acts on X, then the map $G \times X \to X$, given by $(g, x) \mapsto g \circ x$, is an action of G on X

Note that this definition is technically a left action, one can also have $x \circ g$ instead which would quantify as a right action, we will only consider left actions and drop the left quantifier. If we now take the set X as the set of great circles defined as

$$X = \{X(t) = a \cos t + b \sin t \mid a, b \in \mathbb{S}^n, \langle a, b \rangle = 0\}$$

and let $SO(n)$ act on it as follows,

$$Q \circ x = Qa \cos t + Qb \sin t \quad \forall Q \in SO(n), \forall x \in X.$$

Since:

$$\langle Qa, Qb \rangle = \langle a, b \rangle = 0,$$

we have that $Q \circ x \in X$ and by the properties of the matrix vector product we have that $SO(n)$ acts on X as in the definition. So geometrically speaking the action of an element of $SO(n)$ acting on an element of X is a rotation of the great circle. We will use see this group action again at the end of 5.3.

4.2 Lie Algebras

Besides Lie Groups there is another algebraic cornerstone in Lie theory, namely the Lie algebra. To define Lie algebras we will first define the matrix exponential. After we define Lie algebras we will also define the Poisson algebra with which we can define the symmetries in the phase space. We will also take a look at a couple of important Lie algebras and the relations between them.

4.2.1 The Matrix Exponential

Before we can study Lie algebras it is necessary to understand the matrix exponential.

Definition 4.9 ([18])**.** If X is an $n \times n$ matrix, we define the *exponential* of X, denoted e^X or $\exp X$, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}, \tag{18}$$

where X^0 is defined to be the identity matrix I and where X^m is the repeated matrix product of X with itself.

Proposition 4.10 ([18], Proposition 2.1)**.** *The series* (18) *converges for all* $X \in M_n(\mathbb{C})$ *and* e^X *is a continuous function of* X*.*

Proof. For the proof of this proposition see Appendix B □

Proposition 4.11 ([18], Proposition 2.3)**.** *Let* X *and* Y *be arbitrary* $n \times n$ *matrices. Then we have the following:*

1. $e^0 = I$.
2. $\left(e^X\right)^* = e^{X^*}$.
3. *The matrix exponential* e^X *is invertible, and its inverse is given by*

$$(e^X)^{-1} = e^{-X}.$$

Theorem 4.12 ([18], Theorem 2.12)**.** *For any* $X \in M_n(\mathbb{C})$*, we have*

$$\det\left(e^X\right) = e^{tr(X)}.$$

Proof. If X is diagonalizable with eigenvalues $\lambda_1, \dots, \lambda_n$, then e^X is diagonalizable with eigenvalues $e^{\lambda_1}, \dots, e^{\lambda_n}$. Thus, $\mathrm{tr}(X) = \sum_j \lambda_j$ and

$$\det(e^X) = e^{\lambda_1} \cdots e^{\lambda_n} = e^{\lambda_1 + \cdots + \lambda_n} = e^{\mathrm{tr}(X)}.$$

If X is not diagonalizable, we can approximate it by matrices that are diagonalizable, see exercise 4 chapter 2 of [18].

□

4.2.2 Lie Algebra

Definition 4.13 ([18], Definition 3.1)**.** A finite-dimensional real or complex Lie algebra is a finite-dimensional real or complex vector space $\mathfrak{g}$, together with a map $[\cdot,\cdot]$ from $\mathfrak{g} \times \mathfrak{g}$ into $\mathfrak{g}$, with the following properties:

1. $[\cdot,\cdot]$ is bilinear.
2. $[\cdot,\cdot]$ is skew symmetric: $[X,Y] = -[Y,X]$ for all $X,Y \in \mathfrak{g}$.
3. The Jacobi identity holds:
$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$
for all $X,Y,Z \in \mathfrak{g}$.

Two elements X and Y of a Lie algebra $\mathfrak{g}$ commute if $[X,Y] = 0$. A Lie algebra $\mathfrak{g}$ is **commutative** if $[X,Y] = 0$ for all $X,Y \in \mathfrak{g}$. The map $[\cdot,\cdot]$ is referred to as the **bracket operation** on $\mathfrak{g}$. Note also that Condition 2 implies that $[X,X] = 0$ for all $X \in \mathfrak{g}$.

Theorem 4.14 ([18], Example 3.3)**.** *Let $\mathcal{A}$ be an associative algebra and let $\mathfrak{g}$ be a subspace of $\mathcal{A}$ such that $XY - YX \in \mathfrak{g}$ for all $X,Y \in \mathfrak{g}$. Then $\mathfrak{g}$ is a Lie algebra with bracket operation given by*
$$[X,Y] = XY - YX.$$
This bracket is also commonly called a commutator.

Proof. Let $X,Y,Z \in \mathfrak{g}$, and $\lambda \in \mathbb{K}$. Then:

1. Bilinear:

   Linearity in the first argument:
$$[X+Z,Y] = (X+Z)Y - Y(X+Z) = XY - YX + ZY - YZ = [X,Y] + [Z,Y]$$
$$[\lambda X,Y] = (\lambda X)Y - Y(\lambda X) = \lambda XY - \lambda YX = \lambda[X,Y]$$
   Linearity in the second argument:
$$[X,Y+Z] = X(Y+Z) - (Y+Z)X = XY - YX + XZ - ZX = [X,Y] + [X,Z]$$
$$[X,\lambda Y] = X(\lambda Y) - (\lambda Y)X = \lambda XY - \lambda YX = \lambda[X,Y]$$
   Thus, the commutator is bilinear.

2. Skew symmetric:
$$[X,Y] = XY - YX = -(YX - XY) = -[Y,X].$$

3. Jacobi identity:
   To prove:
$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0.$$
   We start working out each nested term:
$$[X,[Y,Z]] = X(YZ - ZY) - (YZ - ZY)X = XYZ - XZY - YZX + ZYX.$$

$$[Y,[Z,X]] = Y(ZX - XZ) - (ZX - XZ)Y = YZX - YXZ - ZXY + XZY.$$

$$[Z,[X,Y]] = Z(XY - YX) - (XY - YX)Z = ZXY - ZYX - XYZ + YXZ.$$

Adding the three terms:

$$\begin{aligned}&(XYZ - XZY - YZX + ZYX) + (YZX - YXZ - ZXY + XZY)\\&+ (ZXY - ZYX - XYZ + YXZ) = 0\end{aligned}$$

Hence the Jacobi identity holds.

□

To see how different or similar Lie algebras are we can use Lie algebra homomorphisms.

Definition 4.15 ([18], Definition 3.6)**.** If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then a linear map $\varphi : \mathfrak{g} \to \mathfrak{h}$ is called a *Lie algebra homomorphism* if

$$\varphi([X,Y]) = [\varphi(X), \varphi(Y)]$$

for all $X, Y \in \mathfrak{g}$. If, in addition, φ is one-to-one and onto, then φ is called a *Lie algebra isomorphism.*

Now for the Lie algebra of a matrix Lie group we will be needing the matrix exponential.

Definition 4.16 ([18], Definition 3.18)**.** Let G be a matrix Lie group. The Lie algebra of G, denoted $\mathfrak{g}$, is the set of all matrices X such that e^{tX} is in G for all real numbers t.

Theorem 4.17 ([18], Theorem 3.20)**.** *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. If X and Y are elements of $\mathfrak{g}$, the following results hold.*

1. $AXA^{-1} \in \mathfrak{g}$ *for all* $A \in G$.
2. $sX \in \mathfrak{g}$ *for all real numbers* s.
3. $X + Y \in \mathfrak{g}$.
4. $XY - YX \in \mathfrak{g}$.

It follows from Theorem 4.14 and Theorem 4.17 that the Lie algebra of a matrix Lie group is a real Lie algebra, with bracket given by $[X,Y] = XY - YX$. For X and Y in $\mathfrak{g}$, i.e. the commutator bracket.
In Hamiltonian mechanics we are particularly interested in Poisson algebras, the algebras can be used to study the symmetries of the phase space.

Definition 4.18 ([21])**.** A **Poisson algebra** is a vector space over a field $\mathbb{K}$ equipped with two bilinear products, $\cdot$ and $\{,\}$, satisfying the following properties:

1. The product $\cdot$ forms an associative $\mathbb{K}$-algebra.

2. The product $\{,\}$, called the *Poisson bracket*, forms a Lie algebra.

3. The Poisson bracket acts as a derivation of the associative product $\cdot$, so that for any three elements x, y, and z in the algebra:

$$\{x, y \cdot z\} = \{x, y\} \cdot z + y \cdot \{x, z\}$$

also called the Leibnitz rule

In canonical coordinates (q_i, p_i) in the phase space the Poisson bracket is defined as:

$$\{f, g\} = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \tag{19}$$

Given two functions $f(p_i, q_i, t)$ and $g(p_i, q_i, t)$. We will now calculate the **fundamental Poisson brackets**. They are defined as:

$$\{q_k, q_l\}_{qp} = \sum_i \left(\frac{\partial q_k}{\partial q_i} \frac{\partial q_l}{\partial p_i} - \frac{\partial q_l}{\partial q_i} \frac{\partial q_k}{\partial p_i} \right) = \sum_i (\delta_{ki} \cdot 0 - 0 \cdot \delta_{li}) = 0$$

$$\{p_k, p_l\}_{qp} = \sum_i \left(\frac{\partial p_k}{\partial q_i} \frac{\partial p_l}{\partial p_i} - \frac{\partial p_l}{\partial q_i} \frac{\partial p_k}{\partial p_i} \right) = \sum_i (0 \cdot \delta_{li} - \delta_{ki} \cdot 0) = 0$$

$$\{q_k, p_l\}_{qp} = \sum_i \left(\frac{\partial q_k}{\partial q_i} \frac{\partial p_l}{\partial p_i} - \frac{\partial p_l}{\partial q_i} \frac{\partial q_k}{\partial p_i} \right) = \sum_i (\delta_{ki} \cdot \delta_{li} - 0 \cdot 0) = \delta_{kl}$$

With this Poisson bracket we can actually build Hamiltonian mechanics up, we will partly do this with the following useful identity $\frac{dG}{dt} = \frac{\partial G}{\partial t} + \{G, H\}$. We can for example formulate the Hamiltonian's equations using this identity as follows.

$$\dot{q}_k = \{q_k, H\} = \frac{\partial H}{\partial p_k}$$
$$\dot{p}_k = \{p_k, H\} = -\frac{\partial H}{\partial q_k}$$

The Possion bracket is also invariant under canonical transformations, and finally we can give an alternative condition for constants of motion. Since $\frac{dG}{dt} = \frac{\partial G}{\partial t} + \{G, H\}$, we have that a observable $G(p, q, t)$ will be a constant of motion if $\frac{dG}{dt} = 0$, thus

$$\frac{\partial G}{\partial t} + \{G, H\} = 0$$

That is

$$\frac{\partial G}{\partial t} = \{H, G\}$$

Additionally, this leads to the conclusion that when the constant of motion G does not depend explicitly on time then

$$\{G, H\} = 0.$$

4.2.3 Examples

In the sections about Lie Groups we discussed the Lie Groups $SO(n)$ and $SU(n)$, Now we will look at their associated Lie algebras $\mathfrak{so}(n)$ and $\mathfrak{su}(n)$.

Proposition 4.19. *The Lie Algebra of $SO(n)$ is defined as follows:*

$$\mathfrak{so}(n) = \left\{X \in M_n(\mathbb{R}) \mid X^T = -X\right\}$$

Proof. $A \in SO(n)$ if $A^T = A^{-1}$ and $\det(A) = 1$ so $e^{tX} \in SO(n)$ if and only if

$$(e^{tX})^T = (e^{tX})^{-1} = e^{-tX}$$

then by Proposition 4.11 we have that $(e^{tX})^T = e^{tX^T}$ so

$$e^{tX^T} = e^{-tX}$$

hence $X^T = -X$. Also because of this $trace(X) = 0$, we have with Theorem 4.12

$$det(e^{tX}) = e^{t \cdot trace(X)} = e^0 = 1$$

□

Proposition 4.20. *The Lie algebra of $SU(n)$ is defined as follows:*

$$\mathfrak{su}(n) = \left\{X \in M_n(\mathbb{C}) \mid X^\dagger = -X\right\}$$

The proof is very similar to the proposition 4.19. To better understand the Lie algebra $\mathfrak{so}(n)$ we will find a basis, this is useful as we can then define Lie algebra homomorphisms by solely defining where the basis elements are mapped to.

Proposition 4.21 ([22], CH 12.1)**.** *Let $n \in \mathbb{N}$. For $1 \leq a, b \leq n$, define the $n \times n$ matrix $\mathcal{M}_{ab}$ by*

$$(\mathcal{M}_{ab})_{k\ell} \;=\; -\,\delta_{ak}\,\delta_{b\ell} \;+\; \delta_{a\ell}\,\delta_{bk} \quad (k, \ell = 1, \dots, n).$$

Then:

1. *Each $\mathcal{M}_{ab}$ is real and antisymmetric*
2. *$\mathcal{M}_{ba} = -\,\mathcal{M}_{ab}$, and $\mathcal{M}_{aa} = 0$.*
3. *The set $\{\mathcal{M}_{ab} : 1 \leq a < b \leq n\}$ is a basis of $\mathfrak{so}(n)$, of dimension $\frac{1}{2}(n^2 - n)$.*
4. *Every $A \in \mathfrak{so}(n)$ can be written as*

$$A = \sum_{1 \leq i < j \leq n} \alpha_{ij}\,\mathcal{M}_{ij},$$

where the real coefficients satisfy $\alpha_{ji} = -\alpha_{ij}$ and $\alpha_{ii} = 0$.

Proof.

1. By construction,

$$(\mathcal{M}_{ab})_{\ell k} = -\delta_{a\ell}\,\delta_{bk} + \delta_{ak}\,\delta_{b\ell} = -(\mathcal{M}_{ab})_{k\ell},$$

so $\mathcal{M}_{ab}^\top = -\mathcal{M}_{ab}$, and all entries are real.

2. If $a = b$, then
$$(\mathcal{M}_{aa})_{k\ell} = -\delta_{ak}\,\delta_{a\ell} + \delta_{a\ell}\,\delta_{ak} = 0.$$
Swapping $a \leftrightarrow b$ gives $\mathcal{M}_{ba} = -\mathcal{M}_{ab}$.

3. Since $\mathcal{M}_{aa} = 0$ and $\mathcal{M}_{ba} = -\mathcal{M}_{ab}$, the independent generators are those with $a < b$, of which there are $\binom{n}{2} = \frac{1}{2}(n^2 - n)$, That this set is a basis follows from 4.

4. Take α_{ij} as the ij^{th} component of A then $A = \sum_{i<j} \alpha_{ij}\, \mathcal{M}_{ij}$ and using $\alpha_{ji} = -\alpha_{ij},\ \alpha_{ii} = 0$ yields the claimed form.

□

The Lie Algebras $\mathfrak{su}(2)$, $\mathfrak{so}(3)$ and $\mathfrak{so}(4)$

Now we will describe the Lie Algebras $\mathfrak{su}(2)$, $\mathfrak{so}(3)$ and $\mathfrak{so}(4)$ in detail.
First for $\mathfrak{su}(2)$ we have by example 3.27 of [18] that the following elements for a basis for $\mathfrak{su}(2)$

$$E_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad E_2 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad E_3 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{20}$$

If we calculate the commutation relations we see

$$[E_i, E_j] = \epsilon_{ijk} E_k. \tag{21}$$

Now, for $\mathfrak{so}(3)$, we see from Proposition 4.21 that the following matrices form a basis:

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These matrices satisfy the following commutation relations:

$$[A_i, A_j] = \epsilon_{ijk} A_k, \tag{22}$$

where ϵ_{ijk} is the Levi-Civita symbol (see Appendix A for details). We see from the matching commutation relations that $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic. Finally, for $\mathfrak{so}(4)$, we again use Proposition 4.21 to construct a basis, given by the following six antisymmetric matrices:

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

These six generators satisfy the following commutation relations:

$$[A_i, A_j] = \epsilon_{ijk} A_k, \quad [B_i, B_j] = \epsilon_{ijk} B_k, \quad [A_i, B_j] = \epsilon_{ijk} B_k. \tag{23}$$

Following [23], we introduce a new basis by defining:

$$X_i = \frac{1}{2}(A_i + B_i), \quad Y_i = \frac{1}{2}(A_i - B_i), \quad \text{for } i = 1, 2, 3. \tag{24}$$

Using the relations in equation (23), we compute the new commutation relations:

$$[X_i, X_j] = \epsilon_{ijk} X_k, \quad [Y_i, Y_j] = \epsilon_{ijk} Y_k, \quad [X_i, Y_j] = 0. \tag{25}$$

This shows that the sets $\{X_i\}$ and $\{Y_i\}$ each satisfy the $\mathfrak{so}(3)$ commutation relations independently, and they commute with each other. Hence, we conclude that the Lie algebra $\mathfrak{so}(4)$ decomposes as a direct sum:

$$\mathfrak{so}(4) \cong \mathfrak{so}(3) \oplus \mathfrak{so}(3). \tag{26}$$

And since $\mathfrak{so}(3) \cong \mathfrak{su}(2)$

$$\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2). \tag{27}$$

4.3 Representation Theory

Lie groups and Lie algebras can both be studied through their **representations**, linear actions on vector spaces that reflect the structure of the original object. Before we define what it means to represent a Lie group or Lie algebra, we need to recall some key spaces and operations. If we take V as a finite complex or real vector space, we can denote its general linear group with $GL(V)$ as the group of invertible linear transformations of V. Once a basis for V is chosen, this group can be identified with the matrix Lie groups $\mathrm{GL}(n;\mathbb{C})$ or $\mathrm{GL}(n;\mathbb{R})$, depending on the field over which V is defined. The other space we will define is the space of all linear operators from V to itself as $\mathfrak{gl}(V) = End(V)$.

$$[A, B] := AB - BA, \quad \text{for } A, B \in \mathfrak{gl}(V).$$

4.3.1 Representations of Lie Groups

A representation of a group G provides a way to understand the group in terms of linear operators. We choose a vector space V and associate each group element with a linear transformation of V, ensuring that the group's multiplication corresponds to the composition of these transformations.

Definition 4.22 ([24, 18])**.** Let G be a Lie group and V a finite-dimensional complex vector space. A **representation** of G on V is a Lie group homomorphism

$$\rho : G \to \mathrm{GL}(V).$$

We refer to the triple (G, V, ρ) as a **Lie group representation**.

4.3.2 Representations of Lie Algebras

Just as groups can act on vector spaces through linear maps, so too can Lie algebras. A representation of a Lie algebra is a homomorphism into the space of endomorphisms of a vector space that respects the Lie bracket.

Definition 4.23 ([24])**.** Let $\mathfrak{g}$ be a Lie algebra and V a complex vector space. A **representation** of $\mathfrak{g}$ on V is a Lie algebra homomorphism

$$\rho : \mathfrak{g} \to \mathfrak{gl}(V).$$

We will often refer to a Lie algebra representation by the triple $(\mathfrak{g}, V, \rho)$, or simply by ρ or V when the context is clear. Just as with group representations, we can define morphisms between Lie algebra representations.

Definition 4.24 ([24])**.** Let $(\mathfrak{g}, V, \rho)$ and $(\mathfrak{g}, \widetilde{V}, \widetilde{\rho})$ be two representations of the same Lie algebra $\mathfrak{g}$. A linear map $T : V \to \widetilde{V}$ is a *homomorphism of representations* if

$$T \circ \rho(X) = \widetilde{\rho}(X) \circ T, \quad \text{for all } X \in \mathfrak{g}.$$

If T is bijective, it is an **isomorphism of representations**, and the two representations are said to be **isomorphic**.

Now we will look at irreducible representations that are specifically interesting to us.

Definition 4.25 ([24], Definition 8.6)**.** Suppose $\mathfrak{g}$ is an arbitrary Lie algebra and $(\mathfrak{g}, V, \rho)$ is a Lie algebra representation. A subspace W of V is an invariant subspace for ρ if

$$\rho(A)w \in W \quad \text{for every } A \in \mathfrak{g} \text{ and every } w \in W.$$

If W is an invariant subspace for ρ, then the representation $\rho_W : \mathfrak{g} \to \mathfrak{gl}(W)$ defined by

$$\rho_W(A) := \rho(A)\big|_W$$

is called a subrepresentation of ρ. If V and $\{0\}$ are the only invariant subspaces of V, then we say that $(\mathfrak{g}, V, \rho)$ is an irreducible representation.

For irreducible representation we have an important Lemma, Shur's Lemma that discusses the uniqueness of these representations.

Proposition 4.26 ([24] Proposition 8.4 (**Schur's Lemma**)). *Suppose $(\mathfrak{g}, V_1, \rho_1)$ and $(\mathfrak{g}, V_2, \rho_2)$ are irreducible representations of the Lie algebra $\mathfrak{g}$. Suppose that $T : V_1 \to V_2$ is a homomorphism of representations. Then there are only two possible cases:*

- *The function T is the zero function.*
- *The representations $(\mathfrak{g}, V_1, \rho_1)$ and $(\mathfrak{g}, V_2, \rho_2)$ are isomorphic (and T is an isomorphism).*

4.3.3 Studying the Representation of $\mathfrak{su}(2)$

We will start by constructing a family of irreducible representations of the Lie algebra $\mathfrak{su}(2)$ by considering subrepresentations of a single representation on $\mathcal{P}$, the space of complex-coefficient polynomials in two variables. Recall from Recall from 20 that $\mathfrak{su}(2)$ can be viewed as a real vector space with basis $\{E_1, E_2, E_3\}$ and bracket relations:

$$[E_i, E_j] = \epsilon_{ijk} E_k.$$

For real numbers $c_1, c_2, c_3 \in \mathbb{R}$, define a map $\mathsf{U} : \mathfrak{su}(2) \to \mathfrak{gl}(\mathcal{P})$ by:

$$\mathsf{U}(c_1 E_1 + c_2 E_2 + c_3 E_3) := c_1 \mathsf{U}_1 + c_2 \mathsf{U}_2 + c_3 \mathsf{U}_3. \tag{28}$$

Define the operators:

$$\begin{aligned} \mathsf{U}_1 &= \frac{i}{2}(x\partial_x - y\partial_y), \\ \mathsf{U}_2 &= \frac{1}{2}(x\partial_y - y\partial_x), \\ \mathsf{U}_3 &= \frac{i}{2}(x\partial_y + y\partial_x). \end{aligned}$$

Each operator preserves the degree of homogeneous polynomials. For example, acting with U_1 on a degree-n monomial $x^k y^{n-k}$ yields:

$$\mathsf{U}_1(x^k y^{n-k}) = \frac{i}{2}\left(k x^{k-1} y^{n-k} - (n-k) x^k y^{n-k-1}\right) = \frac{i(2k-n)}{2} x^k y^{n-k}.$$

Similarly,

$$\begin{aligned} \mathsf{U}_2(x^k y^{n-k}) &= \frac{1}{2}\left((n-k) x^{k+1} y^{n-k-1} - k x^{k-1} y^{n-k+1}\right), \\ \mathsf{U}_3(x^k y^{n-k}) &= \frac{i}{2}\left((n-k) x^{k+1} y^{n-k-1} + k x^{k-1} y^{n-k+1}\right). \end{aligned}$$

Thus, the operators $\mathsf{U}_1, \mathsf{U}_2, \mathsf{U}_3$ preserve the degree of any monomial. Henceforth These operators U_i preserve the degree of a polynomial and maps any homogeneous polynomial to another of the same degree. Thus the subspace $\mathcal{P}^n$ of homogeneous polynomials of degree n forms an irreducible representation of $\mathfrak{su}(2)$. These representations are effectively the only one as captured in the following Proposition

Proposition 4.27 ([24], Proposition 8.9). *Suppose* $(\mathfrak{su}(2), V, \rho)$ *is a finite-dimensional irreducible Lie algebra representation. Set*

$$n := \dim V - 1.$$

Then $(\mathfrak{su}(2), V, \rho)$ *is isomorphic to the representation* $(\mathfrak{su}(2), \mathcal{P}^n, U)$.

In other words, the representations U of $\mathfrak{su}(2)$ as differential operators on homogeneous polynomials in two variables are essentially the only finite-dimensional irreducible representations, and they are classified by their dimensions. For the proof see [24].

Casimir Operator

We will now introduce the Casimir operator, we will later use this operator directly to calculate the energy levels of the hydrogen atom.

Definition 4.28 ([24], Definition 8.9). Suppose $(\mathfrak{su}(2), V, \rho)$ is a Lie algebra representation. The Casimir operator for ρ is the linear transformation $C : V \to V$ defined by

$$C := \rho(E_1)^2 + \rho(E_2)^2 + \rho(E_3)^2.$$

The Casimir operator is not derived from any specific element of the Lie algebra $\mathfrak{su}(2)$. Nonetheless, in the algebra $\mathfrak{gl}(V)$ of linear transformations on a vector space V, operations like squaring and adding linear maps are well-defined. This allows us to define the Casimir operator once a representation is specified.

A key property of the Casimir operator is that it commutes with all operators in the image of the representation.

Proposition 4.29 ([24], Proposition 8.10). *Suppose* $(\mathfrak{su}(2), V, \rho)$ *is a representation and* C *is its Casimir operator. Then* C *commutes with* ρ.

Proof. First note that C commutes with $\rho(E_1)$:

$$\begin{aligned}
[C, \rho(E_1)] &= [\rho(E_1)^2 + \rho(E_2)^2 + \rho(E_3)^2, \rho(E_1)] \\
&= \rho(E_2)^2\rho(E_1) - \rho(E_1)\rho(E_2)^2 + \rho(E_3)^2\rho(E_1) - \rho(E_1)\rho(E_3)^2 \\
&= \rho(E_2)[\rho(E_2), \rho(E_1)] - [\rho(E_1), \rho(E_2)]\rho(E_2) \\
&\quad + \rho(E_3)[\rho(E_3), \rho(E_1)] - [\rho(E_1), \rho(E_3)]\rho(E_3) \\
&= -\rho(E_2)\rho(E_3) - \rho(E_3)\rho(E_2) + \rho(E_3)\rho(E_2) + \rho(E_2)\rho(E_3) \\
&= 0.
\end{aligned}$$

Similarly $[C, \rho(E_2)] = [C, \rho(E_3)] = 0$. Because $\{E_1, E_2, E_3\}$ is a basis for $\mathfrak{su}(2)$, it follows that $[C, \rho(q)] = 0$ for any element $q \in \mathfrak{su}(2)$. □

For $(\mathfrak{su}(2), V, \rho)$ we can actually restrict the values of the Casimir operator as follows:

Proposition 4.30 ([24], Proposition 8.11). *Suppose* $(su(2), V, \rho)$ *is a finite-dimensional irreducible Lie algebra representation. Then the Casimir operator is a scalar multiple of the identity on* V*, specifically we can write the Casimir operator as*

$$C = -\frac{1}{4}(n^2 + 2n)I \quad \text{for some } n \in \mathbb{N}$$

Proof. First we will proof that the Casimir operator is a scalar multiple of the identity.Since V is a finite-dimensional complex vector space, C must have at least one eigenvalue λ. Define

$$W := \{v \in V : Cv = \lambda v\};$$

i.e, W is the eigenspace corresponding to λ. By Proposition 8.5, this subspace is invariant under ρ because C commutes with the representation by Proposition 8.10. But because λ is an eigenvalue for C, the subspace W is not equal to $\{0\}$. Hence, since ρ is irreducible, we conclude by Schur's Lemma 4.26 that $W = V$. So $Cv = \lambda I v$ for every $v \in V$. In other words, C is a scalar multiple of the identity.
Now we can narrow down this scalar multiple by evaluating the Casimir operator C restricted to P^n for arbitrary n. It suffices to evaluate C on any one element of P^n, say, x^n. We find that

$$Cx^n = -\frac{1}{4}(n^2 + 2n).$$

Hence on P^n we have

$$C = -\frac{1}{4}(n^2 + 2n)I.$$

Thus by Proposition 4.27, each finite-dimensional irreducible representation of $su(2)$ is isomorphic to P^n for some n, it follows that the only possible eigenvalue of the Casimir operator on a finite-dimensional representation is $-\frac{1}{4}(n^2 + 2n)$ for some n. □

In physics $-\frac{1}{4}(n^2+2n)$ is often written as $-l(l+1)$ for l an half integer i.e. $l = \frac{n}{2}$ for $n \in \mathbb{N}$. We are now going to extend the notion of the Casimir operator to $\mathfrak{so}(4)$. Given that $\mathfrak{so}(4)$ is isomorphic to $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, the representations of these two Lie algebras must be identical. Hence, classifying the finite-dimensional irreducible representations of $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ is sufficient for understanding those of $\mathfrak{so}(4)$.

Proposition 4.31 ([24], Proposition 8.13)**.** *Suppose $(\mathfrak{su}(2) \oplus \mathfrak{su}(2), V, \rho)$ is a finite-dimensional irreducible representation. Then there exist irreducible representations*

$$(\mathfrak{su}(2), W_1, \rho_1) \quad \textit{and} \quad (\mathfrak{su}(2), W_2, \rho_2)$$

such that the representation $(\mathfrak{su}(2) \oplus \mathfrak{su}(2), V, \rho)$ is isomorphic to the Lie algebra representation

$$(\mathfrak{su}(2) \oplus \mathfrak{su}(2), W_1 \otimes W_2, \rho_1 \otimes I + I \otimes \rho_2).$$

The proof of this proposition can be found in [24], from this we can see that we have two Casimir operators corresponding to the two $\mathfrak{su}(2)$ algebras.

5 Classical Kepler Problem revisited

Now that we have seen the mathematical methods needed to describe symmetries and to find symmetries in the phase space, we will apply these methods to the Kepler problem, for this we will first describe the Kepler problem in the Hamiltonian formalism, after which we will use Poisson bracket to understand the symmetries of the phase space and finally we will use a canonical transformation to describe the symmetry as a group action of a symmetry group.

5.1 Hamiltonian of the Kepler Problem

The Hamiltonian of the Kepler problem is the sum of the kinetic energy and the potential energy resulting in (29).

$$H = \frac{p^2}{2m} - \frac{k}{r} \tag{29}$$

Here we have that $p = \sqrt{p_1^2 + p_2^2 + p_3^2}$ and $r = \sqrt{q_1^2 + q_2^2 + q_3^2}$ with q_i, p_i our canonical coordinates. When we calculate the Hamilton's equations we get

$$\dot{q}_i = \frac{p_i}{m} \quad \dot{p}_i = -\frac{kq_i}{r^3}$$

5.2 Poisson Algebra of the Kepler Hamiltonian

We will now look at applying the possion bracket to our phase-space, remember from 2 that we defined the angular momentum vector $\mathbf{L}$, and the Runge-Lenz vector $\mathbf{A}$ for the Kepler problem as follows:

$$\begin{aligned} \mathbf{L} &= \mathbf{r} \times \mathbf{p}, \\ \mathbf{A} &= \mathbf{p} \times \mathbf{L} - mk\frac{\mathbf{r}}{r}. \end{aligned}$$

We find that the Poisson brackets of the Hamiltonian with $\mathbf{L}$ and $\mathbf{A}$ vanish:

$$\begin{aligned} \{H, L_i\} &= 0, \\ \{H, A_i\} &= 0. \end{aligned}$$

This indicates that both $\mathbf{L}$ and $\mathbf{A}$ are conserved quantities under the Kepler Hamiltonian. We now compute the remaining Poisson brackets [25]:

$$\begin{aligned} \{L_i, L_j\} &= \epsilon_{ijk} L_k, \\ \{L_i, A_j\} &= \epsilon_{ijk} A_k, \\ \{A_i, A_j\} &= -2mH\epsilon_{ijk} L_k. \end{aligned}$$

These relations reveal the underlying symmetry algebra. For bound states (i.e., when $H < 0$), we define a rescaled vector:

$$\tilde{\mathbf{A}} = \frac{\mathbf{A}}{\sqrt{-2mH}}, \tag{30}$$

which satisfies the Poisson brackets:

$$\begin{aligned} \{L_i, L_j\} &= \epsilon_{ijk} L_k, \\ \{L_i, \tilde{A}_j\} &= \epsilon_{ijk} \tilde{A}_k, \\ \{\tilde{A}_i, \tilde{A}_j\} &= \epsilon_{ijk} L_k. \end{aligned}$$

Thus, the set $\{L_i, \tilde{A}_i\}$ satisfies the commutation relations of the Lie algebra $\mathfrak{so}(4)$. And L_i on its own satisfies the commutation relations of the Lie Algebra $\mathfrak{so}(4)$. The latter is true for all spherically symmetric Hamiltonians, but the extra symmetry of the Runge-Lenz vector is unique to this problem and creates an extra hidden symmetry, we will now try to understand this symmetry as group action.

5.3 Ligon-Schaaf Regularization

Now that we found that for $H < 0$ the phase space admits to a $\mathfrak{so}(4)$ lie algebra caused by the conservation of the Runge-Lenz vector and Angular momentum. We would like to see if we can describe the symmetry with a Lie group action of $SO(4)$ on $\mathbb{S}^4$ to better our understanding of the symmetry. We will do this by using a regularization map from $T^*\mathbb{R}^3 \to T^*(\mathbb{S}^4 \setminus \{N\})$. We will be using [26] as our main source and will only give a global sketch of how to get to this regularization map and how to use it, at the end will give some sources required for a better understanding. The first map that was found that does this was the Moser-Regularization.

Proposition 5.1 ([26], Corollary 2.4)**.** *The Moser regularization map $\Phi_M : T^*\mathbb{R}^n \to T^*(S^n - \{N\})$ is defined as the composition of stereographic projection with geometric Fourier transform. It is a symplectomorphism and explicitly given by the formula*

$$(\mathbf{q}, \mathbf{p}) \mapsto \Phi_M(\mathbf{q}, \mathbf{p}) = (\mathbf{u}, \mathbf{v})$$

with

$$\mathbf{u} = \left(\frac{2\mathbf{p}}{\mathbf{p}^2 + 1}, \frac{2\mathbf{p}^2}{\mathbf{p}^2 + 1} - 1 \right), \quad \mathbf{v} = \left(-\frac{(\mathbf{p}^2 + 1)\mathbf{q}}{2} + (\mathbf{q} \cdot \mathbf{p})\mathbf{p}, -\mathbf{q} \cdot \mathbf{p} \right)$$

Based on this map a more general regularization map was developed by Ligon and Schaaf, we will only look at the negative energy part of this map, therefore we define:

$$P_- = \{(\mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^n : \mathbf{q} \neq \mathbf{0}, H(\mathbf{q}, \mathbf{p}) < 0\}$$

$$T_- = \{(\mathbf{u}, \mathbf{v}) \in T^*\mathbb{S}^n; \mathbf{u} \neq \mathbf{n}, v \neq 0\}$$

where $T^*\mathbb{S}^n = \{(\mathbf{u}, \mathbf{v}) \in T^*\mathbb{R}^{n+1}; u = 1, \mathbf{u} \cdot \mathbf{v} = 0\}$.
Then The Ligon–Schaaf regularization map $\Phi_{LS} : P_- \to T_-$ is an adaptation of the Moser regularization and can be defined as follows in the form of a rotation and scaling. That has the useful property for us that

$$\Phi_{LS}^* L_{ij} = \tilde{A}_{ij}$$

And with this property one can prove that there is a $SO(4)$ action on the great circles of $\mathbb{S}^3$ as discussed in Section 3.3. Thus, the orbit of a body is the stereographic projection of a great circle of $\mathbb{S}^3$. For a full understanding why this is true, one is referred to the full paper of [26], for this one needs a bit more background on Delaunay variables, which is explained in the Appendix C, Hamiltonian Flow explained in [27] and symplectic geometry.

6 Symmetries of the Hydrogen Atom

In the previous chapter we looked into the symmetries of the classical Kepler problem, now we will delve into the quantum case. We will study the symmetries of the hydrogen atom. There are two ways one could do this, the Lie group approach using Fock's method from 1936, or the Lie algebra approach based on a modern version of Pauli's method with the use of the Casimir operator. The one method is not necessarily a lot better than the other, but we choose to go with Lie algebraic method as it shares some similarities with how we choose to attack the Kepler problem. After we have seen how this method works we will also take a short look at Fock's method. However we will first discuss some quantum mechanical preliminaries.

6.1 Introduction to Quantum Mechanics

In this section we will give an overview of the basics of quantum mechanics, this overview is mostly based on [28] and if anything is unclear this is a good source for further reading. In quantum mechanics we cannot describe a particle as a point particle with a parameterization of its position, speed, and other properties like in classical mechanics, instead we use the wavefunction $\Psi(\mathbf{r}, t)$ to describe the state of the system. These wave functions live in a complex Hilbert space; physicists usually call it the Hilbert space.

Definition 6.1 ([29], p. 357)**.** A **complex Hilbert space** is a complete normed complex vector space where the norm is induced by an inner product, i.e., $\|u\| = \langle u, u\rangle^{1/2}$.

In three-dimensional space, it's customary to take $\mathcal{H} = L^2(\mathbb{R}^3)$. The way we calculate the wave function is with the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r}, t) + V(\mathbf{r}, t)\Psi(\mathbf{r}, t) \tag{31}$$

Here, $\hbar$ is the reduced Planck constant, ∇ is the gradient, Ψ is the wave function, and V is the potential energy.
The other important tool in quantum mechanics are operators, we use operators to represent observables. Operators act on wave functions like linear transformations, below we find a mathematical definition.

Definition 6.2. Let $\mathcal{H}$ be a vector space. An operator in $\mathcal{H}$ is a map $\hat{Q} : \mathcal{H} \to \mathcal{H}$ such that

$$\hat{Q}(\lambda|\psi\rangle + \mu|\phi\rangle) = \lambda\hat{Q}(|\psi\rangle) + \mu\hat{Q}(|\phi\rangle),$$

for all $\lambda, \mu \in \mathbb{K}$ and $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, which means that the map is linear. Here, $\mathbb{K}$ denotes an arbitrary field.

We usually denote operators with a hat. If an operator is Hermitian, the operator is observable, that is to say for operator $\hat{Q}$ we have that $\forall f, g \in \mathcal{H}, \quad \langle f|\hat{Q}g\rangle = \langle\hat{Q}f|g\rangle$. Two of the most important operators in quantum mechanics are the position and momentum operators. They are defined as:

$$\hat{R}_x = x, \qquad \hat{P}_x = -i\hbar\frac{\partial}{\partial x}$$

We see that both the position and momentum operators are Hermitian. For operators we can also use the commutator in a very similar fashion as the Poisson bracket is used. For example we can calculate commutators between the position and momentum operators.

$$\begin{aligned}\left[\hat{R}_x,\hat{P}_x\right] &= \left(\hat{R}_x\hat{P}_x - \hat{P}_x\hat{R}_x\right)\\ &= \hat{R}_x\hat{P}_x - \hat{P}_x\hat{R}_x\\ &= x(-i\hbar)\partial_x - (-i\hbar)\partial_x x\\ &= (-i\hbar)\left(x\partial_x - \partial_x x\right)\\ &= (-i\hbar)\left(x\partial_x - x\partial_x - 1\right)\\ &= (-i\hbar)(-1)\\ &= i\hbar.\end{aligned}$$

We also see that:

$$\left[\hat{R}_y,\hat{P}_x\right] = 0 \quad \text{and} \quad \left[\hat{R}_z,\hat{P}_x\right] = 0$$

Hence we see that

$$\left[\hat{R}_i,\hat{P}_j\right] = i\hbar\delta_{ij}, \tag{32}$$

where δ_{ij} is the Kronecker delta. This commutation relation is also called the **canonical commutation relation**. Finally we will take another look at the Schrödinger equation. First we will write the Schrödinger as follows

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \hat{H}\Psi(\mathbf{r},t) \tag{33}$$

Here the $\hat{H}$ is the Hamiltonian operator and is quite similar to the classical Hamiltonian. We write it as follows:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)$$

We see that it is the sum of the kinetic and potential energy. There is also a time-independent version of the Schrödinger equation.

$$\hat{H}\psi = E\psi$$

This is sometimes called the Schrödinger eigenfunction equation, where E the total energy is the eigenvalue. The collection of all eigenvalues is called a spectrum, which can both be discrete as continuous. We will know take a deeper look into the Hydrogen atom.

6.2 Hydrogen Atom

In this section we will take a look at the Hamiltonian of the hydrogen atom and its energy eigenvalues. The hydrogen consists of an essentially motionless proton of charge e, which we will choose as our origin and an orbiting electron of charge $-e$. For this we find that the force between these two bodies is given by Coulomb's law.

$$\mathbf{F} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r^2}\hat{r}$$

Hence with $V = -\nabla F$ we get the coulomb potential

$$V(r) = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}$$

Hence we can write the Hamiltonian operator for the hydrogen atom as

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r} \tag{34}$$

One can solve Schrödinger equation for this Hamiltonian operator in a rather tedious calculation using separation of variables, for more details one can look at chapter 4 of [28]. From this one finds for negative energies a discrete spectra of eigenvalues governed by the equation

$$E_n = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{n^2} \approx -\frac{13.6 \text{ eV}}{n^2} \tag{35}$$

Here $n \in \mathbb{Z}_{>0}$.This spectrum of energy levels can be measured from the emission spectra, where the electron falls from one energy level to another, releasing a photon containing the difference of energy creating an emission spectrum which is also visible below. Rydberg derived this equation empirically before it was theoretically derived and expanded on by Bohr.

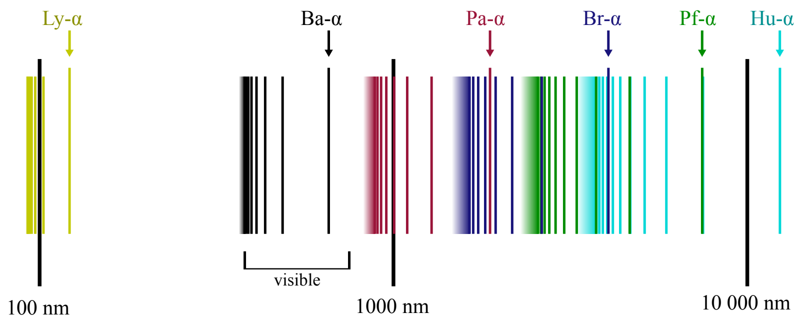

Figure 3: The spectral lines of hydrogen, divided into series, displayed on a logarithmic scale [30].

6.3 Angular Momentum

Classically we have seen that the angular momentum is defined as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, see equation (5). We can still use this formula to derive the angular momentum operators. Using the operators $\hat{R}_n = x_n$ and $\hat{P}_n = -i\hbar\partial_n$ we get that:

$$\hat{L}_x = y\hat{P}_z - z\hat{P}_y$$
$$\hat{L}_y = z\hat{P}_x - x\hat{P}_z$$
$$\hat{L}_z = x\hat{P}_y - y\hat{P}_x$$

This is allowed since no symmetrization is needed due to the fact that every pair in each term commute. Substituting $\hat{P}_n = -i\hbar\partial_n$, the explicit differential form of the angular momentum operators is:

$$\hat{L}_x = -i\hbar(y\partial_z - z\partial_y)$$
$$\hat{L}_y = -i\hbar(z\partial_x - x\partial_z)$$
$$\hat{L}_z = -i\hbar(x\partial_y - y\partial_x)$$

These represent the fundamental angular momentum operators in quantum mechanics. They obey the commutation relations:

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y.$$

Hence we see that the angular momentum operators form a representation of $\mathfrak{so}(3)$ with the map $\rho(A_n) = i\hbar A_n$, where A_n a basis element of $\mathfrak{so}(3)$.
Since the Hamiltonian of the hydrogen atom is spherically symmetric one would inspect that the Hamiltonian would commute with the angular momentum operators and this is indeed the case. We have:

$$\hat{H} = -\frac{\hbar^2}{2m}(\partial_x^2 + \partial_y^2 + \partial_z^2) - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{\sqrt{x^2+y^2+z^2}}$$
$$\hat{L}_x = -i\hbar(y\partial_z - z\partial_y).$$

We will split H in the kinetic energy operator:

$$\hat{T} = -\frac{\hbar^2}{2m}(\partial_x^2 + \partial_y^2 + \partial_z^2).$$

and the potential energy:

$$\hat{V} = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{\sqrt{x^2+y^2+z^2}}, \quad r = \sqrt{x^2+y^2+z^2}.$$

Then starting with the kinetic energy:

$$[\hat{T}, \hat{L}_x] = -\frac{\hbar^2}{2m}(-i\hbar)[\partial_x^2 + \partial_y^2 + \partial_z^2, y\partial_z - z\partial_y].$$

Computing individual commutators:

$$[\partial_x^2, y\partial_z - z\partial_y] = 0,$$
$$[\partial_y^2, y\partial_z - z\partial_y] = 2\partial_y\partial_z,$$
$$[\partial_z^2, y\partial_z - z\partial_y] = -2\partial_y\partial_z.$$

Summing these:

$$[\partial_x^2 + \partial_y^2 + \partial_z^2, y\partial_z - z\partial_y] = 0.$$

Thus:

$$[\hat{T}, \hat{L}_x] = 0.$$

Now for the potential, we have the commutator:

$$[\hat{V}, \hat{L}_x] = -i\hbar[\hat{V}, y\partial_z - z\partial_y].$$

By linearity:

$$[\hat{V}, \hat{L}_x] = -i\hbar([V, y\partial_z] - [\hat{V}, z\partial_y]).$$

Evaluating:

$$[\hat{V}, y\partial_z] = -y\partial_z\hat{V}, \quad [\hat{V}, z\partial_y] = -z\partial_y\hat{V}.$$

Since:

$$\partial_z\hat{V} = \frac{1}{4\pi\varepsilon_0}\frac{e^2 z}{r^3}, \quad \partial_y\hat{V} = \frac{1}{4\pi\varepsilon_0}\frac{e^2 y}{r^3},$$

we obtain:

$$[\hat{V}, \hat{L}_x] = \frac{-i\hbar}{4\pi\varepsilon_0}\left(-y\frac{e^2 z}{r^3} + z\frac{e^2 y}{r^3}\right) = \frac{-i\hbar}{4\pi\varepsilon_0}(0) = 0.$$

Both the kinetic energy and potential energy terms commute with $\hat{L}_x$, so:

$$[\hat{H}, \hat{L}_x] = [\hat{T}, \hat{L}_x] + [\hat{V}, \hat{L}_x] = 0 + 0 = 0.$$

Same holds for $[\hat{H}, \hat{L}_y]$ and $[\hat{H}, \hat{L}_z]$. Thus we indeed see that the angular momentum operators commute with $\hat{H}$.

6.4 The Runge-Lenz Vector and the Hydrogen Atom

Similarly to the classical regime there is more then just the $\mathfrak{so}(3)$ symmetry of the angular momentum. In this section we will look into the quantization of the Runge-Lenz vector and study its related symmetry, we will base this on Chapter 8 of [24] and fill in the missing details. Classically we had the Runge-Lenz vector $\mathbf{A} = \mathbf{p} \times \mathbf{L} - mk\hat{\mathbf{r}}$ as in (9). Sadly, because $[\hat{L}_i, \hat{P}_j] \neq 0$, we need to apply symmetrization for the associated quantum mechanical operator $\hat{A}$, we get $\frac{1}{2}(\hat{P} \times \hat{L} - \hat{L} \times \hat{P})$, hence we get:

$$\hat{A}_n = \frac{1}{2}\left[(\hat{P} \times \hat{L})_n - (\hat{L} \times \hat{P})_n\right] - \frac{mk}{r}\hat{R}_n.$$

So with $k = \frac{1}{4\pi\varepsilon_0}e^2$ and a normalization factor of $\frac{1}{\sqrt{-2mE}}$ we have:

$$\begin{aligned}
\hat{A}_x &= \frac{-i\hbar}{2\sqrt{-2mE}}\left(\partial_y L_z + L_z\partial_y - \partial_z L_y - L_y\partial_z + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 x}{i\hbar r}\right), \\
\hat{A}_y &= \frac{-i\hbar}{2\sqrt{-2mE}}\left(\partial_z L_x + L_x\partial_z - \partial_x L_z - L_z\partial_x + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 y}{i\hbar r}\right), \\
\hat{A}_z &= \frac{-i\hbar}{2\sqrt{-2mE}}\left(\partial_x L_y + L_y\partial_x - \partial_y L_x - L_x\partial_y + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 z}{i\hbar r}\right).
\end{aligned}$$

If we now calculate all the relevant commutation relation, a very tedious task, as is done in Appendix D and the previous section, one gets:

$$\begin{aligned}
[\hat{H}, \hat{L}_i] &= 0 \\
[\hat{H}, \hat{A}_i] &= 0 \\
[\hat{L}_i, \hat{L}_j] &= i\hbar\epsilon_{ijk}\hat{L}_k, \\
[\hat{L}_i, \hat{A}_j] &= i\hbar\epsilon_{ijk}\hat{A}_k, \\
[\hat{A}_i, \hat{A}_j] &= i\hbar\epsilon_{ijk}\hat{L}_k.
\end{aligned}$$

We will now define a Lie Algebra isomorphism as follows:

$$\begin{aligned}
J_{+x} &= \frac{\hat{L}_x + \hat{A}_x}{2i\hbar}, \qquad & J_{-x} &= \frac{\hat{L}_x - \hat{A}_x}{2i\hbar}, \\
J_{+y} &= \frac{\hat{L}_y + \hat{A}_y}{2i\hbar}, \qquad & J_{-y} &= \frac{\hat{L}_y - \hat{A}_y}{2i\hbar}, \\
J_{+z} &= \frac{\hat{L}_z + \hat{A}_z}{2i\hbar}, \qquad & J_{-z} &= \frac{\hat{L}_z - \hat{A}_z}{2i\hbar}.
\end{aligned}$$

Then we can calculate all the commutators assuming that $E < 0$.

$$
\begin{aligned}
[J_{+x}, J_{+y}] &= \frac{-1}{4\hbar^2}\big([\hat{L}_x, \hat{L}_y] + [\hat{A}_x, \hat{A}_y] + [\hat{L}_x, \hat{A}_y] + [\hat{A}_x, \hat{L}_y]\big) \\
&= \frac{1}{4i\hbar}(2\hat{L}_z + 2\hat{A}_z) = J_{+z}, \\
[J_{-x}, J_{-y}] &= \frac{-1}{4\hbar^2}\big([\hat{L}_x, \hat{L}_y] + [\hat{A}_x, \hat{A}_y] - [\hat{L}_x, \hat{A}_y] - [\hat{A}_x, \hat{L}_y]\big) \\
&= \frac{1}{4i\hbar}(2\hat{L}_z - 2\hat{A}_z) = J_{-z}, \\
[J_{+x}, J_{-x}] &= \frac{-1}{4\hbar^2}\big([\hat{L}_x, \hat{L}_x] - [\hat{A}_x, \hat{A}_x] + [\hat{A}_x, \hat{L}_x] - [\hat{L}_x, \hat{A}_x]\big) = 0, \\
[J_{+x}, J_{-y}] &= \frac{-1}{4\hbar^2}\big([\hat{L}_x, \hat{L}_y] - [\hat{A}_x, \hat{A}_y] - [\hat{A}_x, \hat{L}_y] + [\hat{L}_x, \hat{A}_y]\big) = \frac{1}{4i\hbar}(\hat{L}_z - \hat{L}_z - \hat{A}_z + \hat{A}_z) = 0.
\end{aligned}
$$

Similarly, we also have:

$$
\begin{aligned}
[J_{+y}, J_{+z}] &= J_{+x}, \qquad [J_{+z}, J_{+x}] = J_{+y}, \\
[J_{-y}, J_{-z}] &= J_{-x}, \qquad [J_{-z}, J_{-x}] = J_{-y}, \\
[J_{+y}, J_{-y}] &= 0, \qquad [J_{+z}, J_{-z}] = 0, \\
[J_{+y}, J_{-z}] &= 0, \qquad [J_{+z}, J_{-x}] = 0.
\end{aligned}
$$

Hence we see that both J_+ and J_- form a representation of $\mathfrak{su}(2)$ which are independent of each other. We also see that the Runge-Lenz operator and angular momentum operator form a $\mathfrak{so}(4)$ algebra by the direct sum of the algebras from J_+ and J_-. Now when we calculate the Casimir operator of both $\mathfrak{su}(2)$ algebras we get:

$$
J_+^2 = J_-^2 = \frac{-1}{4\hbar^2}\left(\hat{L}^2 + \hat{A}^2\right) = \frac{1}{4}\left(1 + \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \frac{2me^4}{E\hbar^2}\right).
$$

Since $\hat{A} \cdot \hat{L} = 0$ and $\hat{L} \cdot \hat{A} = 0$, we will check this and that $\frac{-1}{\hbar^2}(\hat{L}^2 + \hat{A}^2) = 1 + \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \frac{2me^4}{E\hbar^2}$ in Appendix D. Now, if we recall from Proposition 4.30 we can write:

$$
(i^2 + 2i) = 1 + \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \frac{2me^4}{E\hbar^2}.
$$

For i a non-negative integer, we can rewrite this to

$$
E = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{1}{n^2}, \tag{36}
$$

where $n = i + 1$, which is the energy level formula we know and love.

6.5 Fock's Method

In 1935 Fock showed that the hydrogen atom also adheres the $SO(4)$ symmetry group and solved the energy level formula for the hydrogen atom that way. Result wise this method is no better then the Lie algebraic approach but mathematically the $\mathfrak{so}(4)$ symmetry follows from Fock's method, but this does not hold the other way around. An English translation of his paper is found in chapter 9 of [24], but the main idea is that solution is mapped from $L^2(\mathbb{R}^3)$ to $L^2(\mathbb{S}^3)$ using the stereographic projection, this is done in the momentum space using a Fourier transform.

7 Conclusions and Outlook

In this thesis, we have explored the inherent symmetries of the Kepler problem, examining both its classical and quantum mechanical formulations. Revealing an $\mathfrak{so}(4)$ Lie algebra symmetry in both cases, illustrated through the use of Poisson brackets in the classical domain and commutators in the quantum domain. This parallel between the classical and quantum descriptions of the system highlights a shared mathematical framework that bridges both domains.
Classically, we have shown that the Kepler problem exhibits a rotational symmetry via a $SO(4)$ Lie group acting on the great circles of a four-dimensional sphere, where its great circles can be stereographically projected into three dimensions forming elliptical orbits in the bounded case. Quantum mechanically we derived the energy levels of the hydrogen atom through the use of Casimir operators. Additionally, we briefly explored Fock's method where we have shown that there is also a $SO(4)$ Lie group symmetry.
These insights naturally lead to further inquiry. A notable area for future investigation is whether a quantized version of the Ligon-Schaaf map exists to bridge the gap from algebra to group in our quantum mechanical analysis.
A significant omission in our current study are symmetries related to when $E \geq 0$ as we have exclusively focused on the bounded case where $E < 0$. For systems where $E \geq 0$, there also exists a Lie algebraic symmetry, namely the Lorentz algebra, which we did not cover.
Finally, one could dive deeper into the relationship between the Poisson bracket and the commutator to better understand the relations between the quantum mechanical Kepler problem and its classical counterpart.

Acknowledgments

Finally, I would like to thank my supervisors for providing guidance, feedback, and support throughout my study of the Runge-Lenz vector and the symmetries of the Kepler problem.

References

[1] John R. Taylor. *Classical Mechanics.* University Science Books, Sausalito, CA, 2005.

[2] Johann I. Bernoulli. *Histoires de L'Académie Royale des Sciences avec les Mémoires de Mathématique et Physique.* Académie Royale des Sciences, Paris, 1712. Also pp. 685–703. Amsterdam reprint of 1713. Includes letter dated July 12, 1710, by Jakob Hermann to Johann Bernoulli.

[3] P. S. de Laplace. *Traité de mécanique céleste*, volume I. Paris, 1799. Also in Oeuvres (Gauthier-Villars, Paris, 1843), Tome I, p. 187 ff.

[4] Herbert Goldstein. More on the prehistory of the laplace or runge–lenz vector. *American Journal of Physics*, 44(11):1123–1124, 1976.

[5] Wolfgang Pauli. Über das wasserstoffspektrum vom standpunkt der neuen quantenmechanik. *Zeitschrift für Physik*, 36:336–363, 1926.

[6] V. Fock. Bemerkung zur quantelung des harmonischen oszillators im magnetfeld. *Zeitschrift für Physik*, 98:145–154, 1935.

[7] Jürgen Moser. Regularization of kepler's problem and the averaging method on a manifold. *Communications on Pure and Applied Mathematics*, 23(4):609–636, 1970.

[8] Thomas Ligon and Martin Schaaf. On the global symmetry of the classical kepler problem. *Reports on Mathematical Physics*, 9(3):281–300, 1976.

[9] Thomas S. Ligon and William B. Ligon. The symmetry of the kepler problem, the inverse ligon-schaaf mapping and the birkhoff conjecture. *arXiv preprint arXiv:1804.03844*, 2018. Also published in PLoS ONE, 13(9): e0203821, 2018. DOI: 10.1371/journal.pone.0203821.

[10] J. L. Russell. Kepler's laws of planetary motion: 1609–1666. *The British Journal for the History of Science*, 2(1):1–24, 1964.

[11] John Baez. The kepler problem revisited: The runge–lenz vector. `https://math.ucr.edu/home/baez/classical/gravitational2.pdf`, 2023. Math 241 Homework.

[12] Herbert Goldstein, Charles Poole, and John Safko. *Classical Mechanics.* Addison-Wesley, San Francisco, 3rd edition, 2002. See Chapter 9: Canonical Transformations and Chapter 10: Action-Angle Variables.

[13] Physics LibreTexts. 15.3: Canonical transformations in hamiltonian mechanics, 2023. Accessed: 2025-06-09.

[14] John M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics.* Springer, 2012.

[15] Bas Janssens. *Introduction to Differential Geometry.* TU Delft, Faculty of Electrical Engineering, Mathematics and Computer Science, 2024. Course AM3580, Academic Year 2024/2025, Artikelnummer 06917510011.

[16] Gert Heckman. Symplectic geometry. `https://www.math.ru.nl/~heckman/symplgeom.pdf`, 2010. Dedicated to the memory of Hans Duistermaat (1942–2010).

[17] H. S. M. Coxeter. *Introduction to Geometry.* Wiley, New York, 2nd edition, 1969. See pp. 93 and 289–290.

[18] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Springer Cham, 2015.

[19] Jeremy Atkins. Su(2)'s double-covering of so(3). *University of Rochester Undergraduate Research*, 2019.

[20] D.C. Gijswijt. Algebra 1. Course document, 2023. Course AM1060, Academic Year 2023/2024.

[21] Douglas Cline. 15.2: Poisson bracket representation of hamiltonian mechanics, 2023. Accessed: 2025-06-17.

[22] Eef van Beveren. Some notes on group theory, 2012. Accessed: 2025-05-14.

[23] Howard Haber. Physics 251 solution set 4, 2013. Accessed: 2025-05-14.

[24] Stephanie Frank Singer. *Linearity, Symmetry, and Prediction in the Hydrogen Atom.* Undergraduate Texts in Mathematics. Springer New York, NY, 1 edition, 2005.

[25] Guowu Meng. The universal kepler problem. *Journal of Geometry and Symmetry in Physics*, 36:47–57, 2014.

[26] Gert Heckman and Tim de Laat. On the regularization of the kepler problem. 2010. Dedicated to the memory of Hans Duistermaat (1942-2010).

[27] Makrina Agaoglou, Broncio Aguilar-Sanjuan, Víctor José García Garrido, Francisco González-Montoya, Matthaios Katsanikas, Vladimír Krajňák, Shibabrat Naik, and Stephen Wiggins. Lagrangian descriptors: Discovery and quantification of phase space structure and transport. `https://champsproject.github.io/lagrangian_descriptors/content/chapter1_2.html`, 2023. Chapter 1.2 – Hamiltonian Systems.

[28] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics.* Cambridge University Press, 3rd edition, 2018.

[29] N. L. Carothers. *Real Analysis.* Cambridge University Press, Cambridge, 2000.

[30] OrangeDog. Hydrogen spectrum, 2009. Creative Commons Attribution-Share Alike 3.0 Unported.

Appendix

A Einstein Summation and the Levi-Civita Symbol

The Einstein summation convention is a notational convention which when an index variable appears twice in a single term and is not otherwise defined, it implies a summation over the term over all values of the index. So for example

$$v^\mu \partial_\mu = \sum_{\mu=1}^{4} v^\mu \partial_\mu$$

Here $\partial_\mu \in T_pM$ is a partial differential operator for some manifold M and forms a basis for the tangent space of M at p. Note that when one uses Greek letters for indices one usually sums over 4 indices and for Latin letters it is usually 3.
To write determinants and cross products in Einstein notation or to write the components of, the Levi-Civita symbol was created. In 3 dimensions, the Levi-Civita symbol is defined by:

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

And generally for the N dimensional case we have:

$$\varepsilon_{a_1 a_2 \dots a_n} = \begin{cases} +1 & \text{if } (a_1,a_2,\dots,a_n) \text{ is an even permutation of } (1,2,\dots,n) \\ -1 & \text{if } (a_1,a_2,\dots,a_n) \text{ is an odd permutation of } (1,2,\dots,n) \\ 0 & \text{otherwise} \end{cases}$$

For more information about how permutations work see Chapter 4 of [20]. Now with this Levi-Civita symbol we can write:

$$(\mathbf{a} \times \mathbf{b})^i = \epsilon_{ijk} a^j b^k.$$

B Hilbert-Schmidt Norm and Convergence of the Matrix Exponential

For the definition of a matrix Lie group, we require a notion of convergence for sequences of matrices. Similarly, for the definition of the matrix exponential, we require convergence of a series of matrices. To rigorously define these convergences, we equip the space $M_n(\mathbb{C})$ of $n \times n$ complex matrices with a norm. In what follows, we use the **Hilbert-Schmidt norm**.

Definition 7.1 ([18], Definition 2.2)**.** For any $X \in M_n(\mathbb{C})$, the **Hilbert-Schmidt norm** is defined by

$$\|X\| = \left(\sum_{j,k=1}^{n} |X_{jk}|^2 \right)^{1/2}.$$

This may also be expressed in a basis-independent way as

$$\|X\| = (\mathrm{tr}(X^* X))^{1/2}.$$

This norm is induced by the **Hilbert-Schmidt inner product** on $M_n(\mathbb{C})$, defined by

$$\langle A, B \rangle = \mathrm{tr}(A^* B),$$

which is conjugate symmetric and linear in the second factor. It satisfies the usual inner product properties:

$$\|A\|^2 = \langle A, A\rangle = \sum_{k,l=1}^{n} |A_{kl}|^2.$$

The Hilbert-Schmidt norm satisfies the following important inequalities:

$$\|X + Y\| \le \|X\| + \|Y\|, \tag{37}$$

$$\|XY\| \le \|X\| \cdot \|Y\|, \tag{38}$$

where the first is the triangle inequality and the second follows from the Cauchy-Schwarz inequality for the inner product:

$$|\langle A, B\rangle| \le \|A\| \cdot \|B\|.$$

Convergence and continuity of the matrix exponential

The matrix exponential is defined by the power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}.$$

To show that this series converges in the Hilbert–Schmidt norm, we observe using (38) that

$$\|X^m\| \le \|X\|^m.$$

Hence,

$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \le \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} < \infty.$$

This shows that the series converges absolutely with respect to the Hilbert–Schmidt norm. To show continuity, note that each X^m is a continuous function of X, and the partial sums

$$S_N(X) = \sum_{m=0}^{N} \frac{X^m}{m!}$$

are continuous functions. By the Weierstrass M-test, the convergence is uniform on any norm-bounded subset of $M_n(\mathbb{C})$. Thus, the exponential map $X \mapsto e^X$ is continuous on all of $M_n(\mathbb{C})$.

C Action Angle Coordinates and the Delaunay Variables

One of the more common canonical transformations is the transform to action-angle coordinates. These types of coordinates have certain properties that are useful for perturbative methods. To transform to these coordinates we use

$$J_i = \frac{1}{2\pi} \oint p_i dq^i \quad \text{and} \quad \theta^i = \frac{\partial}{\partial J_i} \oint p_i dq^i \tag{39}$$

Here J_i refers to the action coordinates and θ^i to the angle coordinates. These integrals are over energy contours where $\mathcal{H} = E$, hence we can write the Hamiltonian as a function just depending on the action coordinate. So $\mathcal{H}(J_1, \ldots, J_n)$, hence we have that Hamilton's equations become

$$\dot{\theta}^i = \frac{\partial \mathcal{H}(J_1, \ldots, J_n)}{\partial J_i} = \omega^i(J_1, \ldots, J_n), \quad \dot{J}_i = -\frac{\partial \mathcal{H}(J_1, \ldots, J_n)}{\partial \theta^i} = 0, \tag{40}$$

where ω^i is called the angular frequency. For more information about the theory behind action-angle coordinates see chapter 10 of [12]. We will now calculate the action angle coordinates of the Kepler problem for this we will rewrite (29) to spherical coordinates

$$H(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\right) - \frac{k}{r} \tag{41}$$

The action-angle coordinates specifically for the Kepler problem are called **Delaunay** variables. We will be mostly interested in the actions. The easiest action is J_ϕ, we have

$$J_\phi = \frac{1}{2\pi}\oint p_\phi\, d\phi = \frac{1}{2\pi}p_\phi \cdot \oint d\phi = p_\phi$$

For J_θ it is useful to remember that we can write:

$$L^2 = p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta}$$

Hence, solving for p_θ:

$$p_\theta = \pm\sqrt{L^2 - \frac{p_\phi^2}{\sin^2\theta}}$$

Then, the action variable J_θ is defined as:

$$J_\theta = \frac{1}{2\pi}\oint p_\theta\, d\theta$$

Substituting the expression for p_θ:

$$J_\theta = \frac{1}{2\pi}\oint \sqrt{L^2 - \frac{p_\phi^2}{\sin^2\theta}}\, d\theta$$

Upon evaluation, this integral yields:

$$J_\theta = L - |p_\phi|$$

Finally, we will calculate J_r. We start by expressing the radial momentum p_r from the Hamiltonian of the Kepler problem:

$$H = \frac{1}{2m}(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}) - \frac{k}{r} = E$$

From the angular momentum relations, we know that $p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta} = L^2$, so we can write:

$$E = \frac{1}{2m}(p_r^2 + \frac{L^2}{r^2}) - \frac{k}{r}$$

Solving for p_r^2:

$$p_r^2 = 2mE + \frac{2mk}{r} - \frac{L^2}{r^2}$$

Thus, p_r is given by:

$$p_r = \pm\sqrt{2mE + \frac{2mk}{r} - \frac{L^2}{r^2}}$$

The action variable J_r is defined as the integral of p_r over a full radial period:

$$J_r = \frac{1}{2\pi} \oint p_r \, dr = \frac{1}{2\pi} \oint \sqrt{2mE + \frac{2mk}{r} - \frac{L^2}{r^2}} \, dr$$

The integral evaluates to:

$$J_r = \frac{mk}{\sqrt{-2mE}} - L$$

This expression is valid for bound elliptical orbits, where the total energy E is negative. solving for the Hamiltonian.

$$H \equiv E = -\frac{mk^2}{2(J_r + J_\theta + J_\phi)^2}$$

Since the Hamiltonian is symmetric w.r.t. the actions, all the frequencies will be the same, hence we calculate

$$v = \frac{\partial H}{\partial J_r} = \frac{\partial H}{\partial J_\theta} = \frac{\partial H}{\partial J_\phi} = \frac{mk^2}{(J_r + J_\theta + J_\phi)^3}$$

We can now perform another canonical transformation to get rid of the degeneracy in the frequencies:

$$\begin{aligned} v_1 &= v_\phi - v_\theta \quad J_1 = J_\phi, \\ v_2 &= v_\theta - v_r \quad J_2 = J_\phi + J_\theta, \\ v_3 &= v_r \quad J_3 = J_\phi + J_\theta + J_r. \end{aligned}$$

With these actions we can rewrite the Hamiltonian to the so called Delaunay Hamiltonian.

$$H = -\frac{mk^2}{2J_3^2} \tag{42}$$

Where $v_3 = v$ is the only nonzero frequency. For a full derivation and discussion with all the detailed integrals, see Section 10.8 of [12].

D The Tedious Mathematical Details of the Hydrogen Atom

In this appendix we will look at the full mathematical detail of the calculations of the commutators of 6.4. To calculate the relevant commutators we will split the problem such that:

$$\hat{A}_n = \frac{-i\hbar}{2\sqrt{-2mE}} (\hat{M}_n + \hat{K}_n),$$

where:

$$\begin{aligned} \hat{M}_x &= \partial_y \hat{L}_z + \hat{L}_z \partial_y - \partial_z \hat{L}_y - \hat{L}_y \partial_z, \quad \hat{K}_x = \frac{1}{4\pi\varepsilon_0} \frac{2me^2 x}{i\hbar r}, \\ \hat{M}_y &= \partial_z \hat{L}_x + \hat{L}_x \partial_z - \partial_x \hat{L}_z - \hat{L}_z \partial_x, \quad \hat{K}_y = \frac{1}{4\pi\varepsilon_0} \frac{2me^2 y}{i\hbar r}, \\ \hat{M}_z &= \partial_x \hat{L}_y + \hat{L}_y \partial_x - \partial_y \hat{L}_x - \hat{L}_x \partial_y, \quad \hat{K}_z = \frac{1}{4\pi\varepsilon_0} \frac{2me^2 z}{i\hbar r}. \end{aligned}$$

We start with the commutator $[\hat{H}, \hat{A}_x]$.

$$
\begin{aligned}
[\hat{H}, \hat{A}_x] &= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \hat{M}_x] + \frac{2me^2}{4\pi\varepsilon_0 i\hbar}[\hat{H}, \frac{x}{r}]\right) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \hat{M}_x] + \frac{2me^2}{4\pi\varepsilon_0 i\hbar}[-\frac{\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}, \frac{x}{r}]\right) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \hat{M}_x] - \frac{\hbar^2}{2m}\frac{2me^2}{4\pi\varepsilon_0 i\hbar}[\nabla^2, \frac{x}{r}]\right) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \partial_y\hat{L}_z + \hat{L}_z\partial_y - \partial_z\hat{L}_y - \hat{L}_y\partial_z] - \frac{\hbar e^2}{4\pi\varepsilon_0 i}[\nabla^2, \frac{x}{r}]\right) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \partial_y\hat{L}_z] + [\hat{H}, \hat{L}_z\partial_y] - [\hat{H}, \partial_z\hat{L}_y] - [\hat{H}, \hat{L}_y\partial_z] - \frac{\hbar e^2}{4\pi\varepsilon_0 i}[\nabla^2, \frac{x}{r}]\right) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\Big([\hat{H}, \partial_y]\hat{L}_z + \partial_y[\hat{H}, \hat{L}_z] + [\hat{H}, \hat{L}_z]\partial_y + \hat{L}_z[\hat{H}, \partial_y] \\
&\quad -[\hat{H}, \partial_z]\hat{L}_y - \partial_z[\hat{H}, \hat{L}_y] - [\hat{H}, \hat{L}_y]\partial_z - \hat{L}_y[\hat{H}, \partial_z] - \frac{\hbar e^2}{4\pi\varepsilon_0 i}[\nabla^2, \frac{x}{r}]\Big) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{H}, \partial_y]\hat{L}_z + \hat{L}_z[\hat{H}, \partial_y] - [\hat{H}, \partial_z]\hat{L}_y - \hat{L}_y[\hat{H}, \partial_z] - \frac{\hbar e^2}{4\pi\varepsilon_0 i}[\nabla^2, \frac{x}{r}]\right).
\end{aligned}
$$

We have $[\hat{H}, \partial_y] = [-\frac{\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}, \partial_y] = [-\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}, \partial_y] = -\frac{e^2}{4\pi\varepsilon_0}[\frac{1}{r}, \partial_y] = -\frac{e^2}{4\pi\varepsilon_0}\frac{y}{r^3}$, simmilary $[\hat{H}, \partial_z] = -\frac{e^2}{4\pi\varepsilon_0}[\frac{1}{r}, \partial_z] = -\frac{e^2}{4\pi\varepsilon_0}\frac{z}{r^3}$, hence:

$$
\begin{aligned}
[\hat{H}, \hat{A}_x] &= \frac{-i\hbar}{2\sqrt{-2mE}}\left(-\frac{e^2}{4\pi\varepsilon_0}\left([\frac{1}{r}, \partial_y]\hat{L}_z + \hat{L}_z[\frac{1}{r}, \partial_y] - [\frac{1}{r}, \partial_z]\hat{L}_y - \hat{L}_y[\frac{1}{r}, \partial_z]\right) - \frac{\hbar e^2}{4\pi\varepsilon_0 i}[\nabla^2, \frac{x}{r}]\right) \\
&= \frac{i\hbar}{2\sqrt{-2mE}}\frac{e^2}{4\pi\varepsilon_0}\left([\frac{1}{r}, \partial_y]\hat{L}_z + \hat{L}_z[\frac{1}{r}, \partial_y] - [\frac{1}{r}, \partial_z]\hat{L}_y - \hat{L}_y[\frac{1}{r}, \partial_z] - i\hbar[\nabla^2, \frac{x}{r}]\right).
\end{aligned}
$$

Now focusing on the term:

$$
\left[\frac{1}{r}, \partial_y\right]\hat{L}_z + \hat{L}_z\left[\frac{1}{r}, \partial_y\right] = -i\hbar\left(\left[\frac{1}{r}, \partial_y\right](x\partial_y - y\partial_x) + (x\partial_y - y\partial_x)\left[\frac{1}{r}, \partial_y\right]\right).
$$

Using the Leibniz rule:

$$
\left[\frac{x}{r}, \partial_y^2\right] = \left[\frac{x}{r}, \partial_y\right]\partial_y + \partial_y\left[\frac{x}{r}, \partial_y\right].
$$

We obtain:

$$
-i\hbar\left(\left[\frac{1}{r}, \partial_y\right](x\partial_y - y\partial_x) + (x\partial_y - y\partial_x)\left[\frac{1}{r}, \partial_y\right]\right) = -i\hbar\left(\left[\frac{x}{r}, \partial_y^2\right] - \left[\frac{1}{r}, \partial_y\right]y\partial_x - y\partial_x\left[\frac{1}{r}, \partial_y\right]\right).
$$

Similarly, for the z-component:

$$
-\left[\frac{1}{r}, \partial_z\right]\hat{L}_y - \hat{L}_y\left[\frac{1}{r}, \partial_z\right] = -i\hbar\left(\left[\frac{x}{r}, \partial_z^2\right] - \left[\frac{1}{r}, \partial_z\right]z\partial_x - z\partial_x\left[\frac{1}{r}, \partial_z\right]\right).
$$

Substituting back into the full expression gives:

$$
\begin{aligned}
[\hat{H}, \hat{A}_x] &= \frac{i\hbar}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \Bigg(-i\hbar \left([\frac{x}{r}, \partial_y^2] - [\frac{1}{r}, \partial_y] y\partial_x - y\partial_x [\frac{1}{r}, \partial_y] \right) \\
&\quad - i\hbar \left([\frac{x}{r}, \partial_z^2] - [\frac{1}{r}, \partial_z] z\partial_x - z\partial_x [\frac{1}{r}, \partial_z] \right) - i\hbar [\nabla^2, \frac{x}{r}] \Bigg) \\
&= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \Bigg([\frac{x}{r}, \partial_y^2] - [\frac{1}{r}, \partial_y] y\partial_x - y\partial_x [\frac{1}{r}, \partial_y] \\
&\quad + [\frac{x}{r}, \partial_z^2] - [\frac{1}{r}, \partial_z] z\partial_x - z\partial_x [\frac{1}{r}, \partial_z] + [\nabla^2, \frac{x}{r}] \Bigg) \\
&= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \left([\partial_x^2, \frac{x}{r}] - [\frac{1}{r}, \partial_y] y\partial_x - y\partial_x [\frac{1}{r}, \partial_y] - [\frac{1}{r}, \partial_z] z\partial_x - z\partial_x [\frac{1}{r}, \partial_z] \right) \\
&= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \left([\partial_x^2, \frac{x}{r}] - \frac{y}{r^3} y\partial_x - y\partial_x \frac{y}{r^3} - \frac{z}{r^3} z\partial_x - z\partial_x \frac{z}{r^3} \right) \\
&= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \left([\partial_x^2, \frac{x}{r}] + \frac{y^2}{r^5}(3x - 2r^2\partial_x) + \frac{z^2}{r^5}(3x - 2r^2\partial_x) \right).
\end{aligned}
$$

Focusing on $[\partial_x^2, \frac{x}{r}]$.

$$
\begin{aligned}
[\partial_x^2, \frac{x}{r}] &= \partial_x^2 \frac{x}{r} - \frac{x}{r}\partial_x^2 = \partial_x \left(\frac{1}{r} - \frac{x^2}{r^3} + \frac{x}{r}\partial_x \right) - \frac{x}{r}\partial_x^2 \\
&= \frac{-x}{r^3} + \frac{1}{r}\partial_x - \frac{2x}{r^3} + \frac{3x^3}{r^5} - \frac{x^2}{r^3}\partial_x + \frac{1}{r}\partial_x - \frac{x^2}{r^3}\partial_x + \frac{x}{r}\partial_x^2 - \frac{x}{r}\partial_x^2 \\
&= -\frac{3x}{r^3} + \frac{3x^3}{r^5} + \frac{2}{r}\partial_x - \frac{2x^2}{r^3}\partial_x = 3x \left(\frac{x^2}{r^5} - \frac{1}{r^3} \right) + 2 \left(\frac{1}{r} - \frac{x^2}{r^3} \right) \partial_x.
\end{aligned}
$$

Hence:

$$
\begin{aligned}
[\hat{H}, \hat{A}_x] &= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \left(3x \left(\frac{x^2}{r^5} - \frac{1}{r^3} \right) + 2 \left(\frac{1}{r} - \frac{x^2}{r^3} \right) \partial_x + \frac{y^2}{r^5}(3x - 2r^2\partial_x) + \frac{z^2}{r^5}(3x - 2r^2\partial_x) \right) \\
&= \frac{\hbar^2}{2\sqrt{-2mE}} \frac{e^2}{4\pi\varepsilon_0} \left(3x \left(\frac{x^2}{r^5} + \frac{y^2}{r^5} + \frac{z^2}{r^5} - \frac{1}{r^3} \right) + 2 \left(\frac{1}{r} - \frac{x^2}{r^3} - \frac{y^2}{r^3} - \frac{z^2}{r^3} \right) \partial_x \right) = 0.
\end{aligned}
$$

Similarly we get $[\hat{H}, \hat{A}_y] = 0$ and $[\hat{H}, \hat{A}_z] = 0$.
We will now look at the commutator of $\hat{A}_x$ and $\hat{A}_y$, this becomes:

$$
[\hat{A}_x, \hat{A}_y] = \frac{\hbar^2}{8mE} \left([\hat{M}_x, \hat{M}_y] + [\hat{M}_x, \hat{K}_y] + [\hat{K}_x, \hat{M}_y] + [\hat{K}_x, \hat{K}_y] \right). \tag{43}
$$

Before we are going to attack this problem term by term we will try to simplify our definition of $\hat{M}_n$, starting with $\hat{M}_x$:

$$
\hat{M}_x = \partial_y \hat{L}_z + \hat{L}_z \partial_y - \partial_z \hat{L}_y - \hat{L}_y \partial_z.
$$

Computing each term:

$$
\partial_y \hat{L}_z = \partial_y \left(-i\hbar(x\partial_y - y\partial_x) \right) = -i\hbar \left(x\partial_y^2 - \partial_y(y\partial_x) \right) = -i\hbar \left(x\partial_y^2 - \partial_x - y\partial_y\partial_x \right),
$$

$$\hat{L}_z\partial_y = -i\hbar(x\partial_y - y\partial_x)\partial_y = -i\hbar\left(x\partial_y^2 - y\partial_x\partial_y\right),$$
$$\partial_z\hat{L}_y = \partial_z\left(-i\hbar(z\partial_x - x\partial_z)\right) = -i\hbar\left(\partial_x + z\partial_z\partial_x - x\partial_z^2\right),$$
$$\hat{L}_y\partial_z = -i\hbar(z\partial_x - x\partial_z)\partial_z = -i\hbar\left(z\partial_x\partial_z - x\partial_z^2\right).$$

Combining these:

$$\partial_y\hat{L}_z + \hat{L}_z\partial_y = -i\hbar\left(2x\partial_y^2 - \partial_x - 2y\partial_x\partial_y\right),$$
$$\partial_z\hat{L}_y + \hat{L}_y\partial_z = -i\hbar\left(\partial_x + 2z\partial_x\partial_z - 2x\partial_z^2\right).$$

Hence:

$$\begin{aligned}\hat{M}_x &= (\partial_y\hat{L}_z + \hat{L}_z\partial_y) - (\partial_z\hat{L}_y + \hat{L}_y\partial_z),\\ &= -i\hbar\left[2x(\partial_y^2 + \partial_z^2) - 2y\partial_x\partial_y - 2z\partial_x\partial_z - 2\partial_x\right],\\ &= -2i\hbar\left[x(\partial_y^2 + \partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x\right].\end{aligned}$$

Similarly:

$$\hat{M}_y = -2i\hbar\left[y(\partial_z^2 + \partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y\right],$$
$$\hat{M}_z = -2i\hbar\left[z(\partial_x^2 + \partial_y^2) - x\partial_z\partial_x - y\partial_z\partial_y - \partial_z\right].$$

Now attacking (43) term by term:

$$[M_x, M_y] = -4\hbar^2\left[x(\partial_y^2 + \partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x, y(\partial_z^2 + \partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y\right],$$

by linearly splitting:

$$\begin{aligned}[M_x, M_y] = -4\hbar^2\Big(&[x(\partial_y^2 + \partial_z^2), y(\partial_z^2 + \partial_x^2)] + [x(\partial_y^2 + \partial_z^2), -z\partial_y\partial_z]\\ &+ [x(\partial_y^2 + \partial_z^2), -x\partial_y\partial_x] + [x(\partial_y^2 + \partial_z^2), -\partial_y]\\ &+ [-y\partial_x\partial_y, y(\partial_z^2 + \partial_x^2)] + [-y\partial_x\partial_y, -z\partial_y\partial_z]\\ &+ [-y\partial_x\partial_y, -x\partial_y\partial_x] + [-y\partial_x\partial_y, -\partial_y]\\ &+ [-z\partial_x\partial_z, y(\partial_z^2 + \partial_x^2)] + [-z\partial_x\partial_z, -z\partial_y\partial_z]\\ &+ [-z\partial_x\partial_z, -x\partial_y\partial_x] + [-z\partial_x\partial_z, -\partial_y]\\ &+ [-\partial_x, y(\partial_z^2 + \partial_x^2)] + [-\partial_x, -z\partial_y\partial_z]\\ &+ [-\partial_x, -x\partial_y\partial_x] + [-\partial_x, -\partial_y]\Big).\end{aligned}$$

Calculating each term:

$$
\begin{aligned}
[x(\partial_y^2+\partial_z^2), y(\partial_z^2+\partial_x^2)] &= x(\partial_y^2+\partial_z^2)y(\partial_x^2+\partial_y^2) - y(\partial_x^2+\partial_z^2)x(\partial_y^2+\partial_z^2) = \\
&= x\partial_y^2 y(\partial_x^2+\partial_z^2) - y\partial_x^2 x(\partial_y^2+\partial_z^2) + xy\partial_z^2(\partial_x^2-\partial_y^2) \\
&= x(2\partial_y + y\partial_y^2)(\partial_x^2+\partial_z^2) - y(2\partial_x + x\partial_x^2)(\partial_y^2+\partial_z^2) + xy\partial_z^2(\partial_x^2-\partial_y^2) \\
&= 2x\partial_y(\partial_x^2+\partial_z^2) - 2y\partial_x(\partial_y^2+\partial_z^2), \\
[x(\partial_y^2+\partial_z^2), -z\partial_y\partial_z] &= -x\partial_y(\partial_y^2+\partial_z^2)z\partial_z + xz\partial_y\partial_z(\partial_y^2+\partial_z^2) = -x\partial_y\partial_z^2 z\partial_z + xz\partial_y\partial_z^3 \\
&= -x\partial_y(2\partial_z^2 + z\partial_z^3) + xz\partial_y\partial_z^3 = -2x\partial_y\partial_z^2, \\
[x(\partial_y^2+\partial_z^2), -x\partial_y\partial_x] &= -x^2\partial_x\partial_y(\partial_y^2+\partial_z^2) + x\partial_x x\partial_y(\partial_y^2+\partial_z^2) = x\partial_y, \\
[x(\partial_y^2+\partial_z^2), -\partial_y] &= 0, \\
[-y\partial_x\partial_y, y(\partial_z^2+\partial_x^2)] &= -y\partial_x\partial_y y(\partial_x^2+\partial_z^2) + y^2\partial_x\partial_y(\partial_x^2+\partial_z^2) = -y\partial_x(\partial_x^2+\partial_z^2), \\
[-y\partial_x\partial_y, -z\partial_y\partial_z] &= yz\partial_x\partial_y^2\partial_z - z\partial_x\partial_y y\partial_y\partial_z = -z\partial_x\partial_y\partial_z, \\
[-y\partial_x\partial_y, -x\partial_y\partial_x] &= y\partial_x x\partial_x\partial_y^2 - x\partial_y y\partial_x^2\partial_y = y\partial_x\partial_y^2 - x\partial_x^2\partial_y, \\
[-y\partial_x\partial_y, -\partial_y] &= y\partial_x\partial_y^2 - \partial_y y\partial_x\partial_y = y\partial_x\partial_y^2 - \partial_x\partial_y - y\partial_x\partial_y^2 = -\partial_x\partial_y, \\
[-z\partial_x\partial_z, y(\partial_z^2+\partial_x^2)] &= -yz\partial_x\partial_z(\partial_x^2+\partial_z^2) + y\partial_x(\partial_x^2+\partial_z^2)z\partial_z = -yz\partial_x\partial_z^3 + y\partial_x\partial_z^2 z\partial_z = 2y\partial_x\partial_z^2, \\
[-z\partial_x\partial_z, -z\partial_y\partial_z] &= 0, \\
[-z\partial_x\partial_z, -x\partial_y\partial_x] &= z\partial_x x\partial_x\partial_y\partial_z - xz\partial_x^2\partial_y\partial_z = z\partial_x\partial_y\partial_z + xz\partial_x^2\partial_y\partial_z - xz\partial_x^2\partial_y\partial_z = z\partial_x\partial_y\partial_z, \\
[-z\partial_x\partial_z, -\partial_y] &= 0, \\
[-\partial_x, y(\partial_z^2+\partial_x^2)] &= -\partial_x y(\partial_x^2+\partial_z^2) + y(\partial_x^2+\partial_z^2)\partial_x = y\partial_x(\partial_x^2+\partial_z^2) - y\partial_x(\partial_x^2+\partial_z^2) = 0, \\
[-\partial_x, -z\partial_y\partial_z] &= 0, \\
[-\partial_x, -x\partial_y\partial_x] &= \partial_x x\partial_x\partial_y - x\partial_x^2\partial_y = \partial_x\partial_y + x\partial_x^2\partial_y - x\partial_x^2\partial_y = \partial_x\partial_y, \\
[-\partial_x, -\partial_y] &= 0.
\end{aligned}
$$

Then filling everything in again:

$$
\begin{aligned}
[\hat{M}_x, \hat{M}_y] = -4\hbar^2\Big(&2x\partial_y(\partial_x^2+\partial_z^2) - 2y\partial_x(\partial_y^2+\partial_z^2) - 2x\partial_y\partial_z^2 + x\partial_y(\partial_y^2+\partial_z^2) \\
&+ 0 - y\partial_x(\partial_x^2+\partial_z^2) - z\partial_x\partial_y\partial_z + y\partial_x\partial_y^2 - x\partial_x^2\partial_y \\
&- \partial_x\partial_y + 2y\partial_x\partial_z^2 + 0 + z\partial_x\partial_y\partial_z + 0 + 0 + \partial_x\partial_y + 0\Big).
\end{aligned}
$$

Simplifying:

$$
\begin{aligned}
[\hat{M}_x, \hat{M}_y] &= -4\hbar^2\Big(2x\partial_y(\partial_x^2+\partial_z^2) - 2y\partial_x(\partial_y^2+\partial_z^2) - 2x\partial_y\partial_z^2 + x\partial_y(\partial_y^2+\partial_z^2) - y\partial_x(\partial_x^2+\partial_z^2) \\
&\qquad + y\partial_x\partial_y^2 - x\partial_x^2\partial_y + 2y\partial_x\partial_z^2\Big) \\
&= -4\hbar^2\Big(x\partial_y(\partial_x^2+\partial_y^2+\partial_z^2) - y\partial_x(\partial_x^2+\partial_y^2+\partial_z^2)\Big) \\
&= -4i\hbar\Big(\hat{L}_z(\partial_x^2+\partial_y^2+\partial_z^2)\Big) = -4i\hbar\hat{L}_z\nabla^2.
\end{aligned}
$$

Due to the circular nature we also have that $[\hat{M}_y, \hat{M}_z] = -4i\hbar\hat{L}_x\nabla^2$ and $[\hat{M}_z, \hat{M}_x] = -4i\hbar\hat{L}_y\nabla^2$. Now we will take a look at $[\hat{M}_x, \hat{K}_y] + [\hat{K}_x, \hat{M}_y]$, we remind ourself that we defined:

$$
\hat{M}_x = -2i\hbar\big(x(\partial_y^2+\partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x\big), \quad \hat{K}_x = \frac{1}{4\pi\varepsilon_0}\frac{2me^2x}{i\hbar r},
$$

$$
\hat{M}_y = -2i\hbar\big(y(\partial_z^2+\partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y\big), \quad \hat{K}_y = \frac{1}{4\pi\varepsilon_0}\frac{2me^2y}{i\hbar r},
$$

$$\hat{M}_z = -2i\hbar\big(z(\partial_x^2+\partial_y^2) - x\partial_z\partial_x - y\partial_z\partial_y - \partial_z\big), \quad \hat{K}_z = \frac{1}{4\pi\varepsilon_0}\frac{2me^2 z}{i\hbar r}.$$

Hence:

$$\begin{aligned}
[\hat{M}_x,\hat{K}_y] + [\hat{K}_x,\hat{M}_y] &= [-2i\hbar\big(x(\partial_y^2+\partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x\big), \frac{1}{4\pi\varepsilon_0}\frac{2me^2 y}{i\hbar r}] \\
&\quad + [\frac{1}{4\pi\varepsilon_0}\frac{2me^2 x}{i\hbar r}, -2i\hbar\big(y(\partial_z^2+\partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y\big)] \\
&= \frac{-me^2}{\pi\varepsilon_0}\Big([x(\partial_y^2+\partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x, \frac{y}{\sqrt{x^2+y^2+z^2}}] \\
&\quad + [\frac{x}{\sqrt{x^2+y^2+z^2}}, y(\partial_z^2+\partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y]\Big)
\end{aligned}$$

To calculate these commutators we will use that with $r=\sqrt{x^2+y^2+z^2}$ we have $\partial_n(r) = \frac{x_n}{r}$ and $\partial_n(\frac{1}{r}) = \frac{-x_n}{r^3}$. We will know attack each commutator separately. By linearity:

$$[x(\partial_y^2+\partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x, \frac{y}{r}] = [x\partial_y^2, \frac{y}{r}] + [x\partial_z^2, \frac{y}{r}] - [y\partial_x\partial_y, \frac{y}{r}] - [z\partial_x\partial_z, \frac{y}{r}] - [\partial_x, \frac{y}{r}].$$

Computing each component separately:

$$\begin{aligned}
[x\partial_y^2, \frac{y}{r}] &= x\partial_y(\frac{y}{r}\partial_y + \frac{1}{r} - \frac{y^2}{r^3}) - \frac{xy}{r}\partial_y^2 = x(2(\frac{1}{r} - \frac{y^2}{r^3})\partial_y + 3(\frac{y^3}{r^5} - \frac{y}{r^3})), \\
[x\partial_z^2, \frac{y}{r}] &= xy\partial_z^2\frac{1}{r} - \frac{xy}{r}\partial_z^2 = xy\partial_z(\frac{1}{r}\partial_z - \frac{z}{r^3}) - \frac{xy}{r}\partial_z^2 = xy(\frac{3z^2}{r^5} - \frac{1}{r^3}(1+2z\partial_z)), \\
[y\partial_x\partial_y, \frac{y}{r}] &= y\partial_x(\frac{1}{r} - \frac{y^2}{r^3} + \frac{y}{r}\partial_y) - \frac{y^2}{r}\partial_x\partial_y = y(\frac{3xy^2}{r^5} - \frac{1}{r^3}(x + xy\partial_y + y^2\partial_x) + \frac{1}{r}\partial_x), \\
[z\partial_x\partial_z, \frac{y}{r}] &= yz\partial_x\partial_z\frac{1}{r} - \frac{yz}{r}\partial_x\partial_z = yz(\partial_x(\frac{1}{r}\partial_z - \frac{z}{r^3}) - \frac{1}{r}\partial_x\partial_z) = yz(\frac{3xz}{r^5} - \frac{x\partial_z + z\partial_x}{r^3}), \\
[\partial_x, \frac{y}{r}] &= y(\partial_x\frac{1}{r} - \frac{1}{r}\partial_x) = -\frac{xy}{r^3}.
\end{aligned}$$

Now the other commutator:

$$\left[\frac{x}{r}, y(\partial_z^2+\partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y\right] = \left[\frac{x}{r}, y\partial_z^2\right] + \left[\frac{x}{r}, y\partial_x^2\right] - \left[\frac{x}{r}, z\partial_y\partial_z\right] - \left[\frac{x}{r}, x\partial_y\partial_x\right] - \left[\frac{x}{r}, \partial_y\right].$$

Computing these separately:

$$\begin{aligned}
\left[\frac{x}{r}, y\partial_z^2\right] &= \frac{xy}{r}\partial_z^2 - y\partial_z^2\frac{x}{r} = \frac{xy}{r}\partial_z^2 - y\partial_z(\frac{x}{r}\partial_z - \frac{xz}{r^3}) = -xy(\frac{3z^2}{r^5} - \frac{1}{r^3}(1+2z\partial_z)), \\
\left[\frac{x}{r}, y\partial_x^2\right] &= \frac{xy}{r}\partial_x^2 - y\partial_x^2\frac{x}{r} = \frac{xy}{r}\partial_x^2 - y\partial_x(\frac{x}{r}\partial_x + \frac{1}{r} - \frac{x^2}{r^3}) = -y(2(\frac{1}{r} - \frac{x^2}{r^3})\partial_x + 3(\frac{x^3}{r^5} - \frac{x}{r^3})), \\
\left[\frac{x}{r}, z\partial_y\partial_z\right] &= \frac{xz}{r}\partial_y\partial_z - xz\partial_y\partial_z\frac{1}{r} = \frac{xz}{r}\partial_y\partial_z - xz\partial_y(\frac{-z}{r^3} + \frac{1}{r}\partial_z) = -xz(\frac{3yz}{r^5} - \frac{y\partial_z + z\partial_y}{r^3}), \\
\left[\frac{x}{r}, x\partial_y\partial_x\right] &= \frac{x^2}{r}\partial_y\partial_x - x\partial_y\partial_x\frac{x}{r} = \frac{x^2}{r}\partial_y\partial_x - x\partial_y(\frac{x}{r}\partial_x + \frac{1}{r} - \frac{x^2}{r^3}) = -x(\frac{3x^2y}{r^5} - \frac{1}{r^3}(y + xy\partial_x + x^2\partial_y) + \frac{1}{r}\partial_y), \\
\left[\frac{x}{r}, \partial_y\right] &= x(\partial_y\frac{1}{r} - \frac{1}{r}\partial_y) = \frac{xy}{r^3}.
\end{aligned}$$

Now we will sum the related pairs of commutators:

$$[x\partial_y^2, \tfrac{y}{r}] + [\tfrac{x}{r}, y\partial_x^2] = x\left(2\left(\tfrac{1}{r} - \tfrac{y^2}{r^3}\right)\partial_y + 3\tfrac{y^3}{r^5}\right) - y\left(2\left(\tfrac{1}{r} - \tfrac{x^2}{r^3}\right)\partial_x + 3\tfrac{x^3}{r^5}\right),$$

$$[x\partial_z^2, \tfrac{y}{r}] + [\tfrac{x}{r}, y\partial_z^2] = 0,$$

$$[y\partial_x\partial_y, \tfrac{y}{r}] + [\tfrac{x}{r}, x\partial_y\partial_x] = y\left(\tfrac{3xy^2}{r^5} - \tfrac{1}{r^3}(xy\partial_y + y^2\partial_x) + \tfrac{1}{r}\partial_x\right) - x\left(\tfrac{3x^2y}{r^5} - \tfrac{1}{r^3}(xy\partial_x + x^2\partial_y) + \tfrac{1}{r}\partial_y\right),$$

$$[z\partial_x\partial_z, \tfrac{y}{r}] + [\tfrac{x}{r}, z\partial_y\partial_z] = -\tfrac{1}{r^3}\left(yz^2\partial_x - xz^2\partial_y\right),$$

$$[\partial_x, \tfrac{y}{r}] + [\tfrac{x}{r}, \partial_y] = 0.$$

Filling these sums in gives:

$$\begin{aligned}
[\hat{M}_x, \hat{K}_y] + [\hat{K}_x, \hat{M}_y] &= \frac{-me^2}{\pi\varepsilon_0}\big([x(\partial_y^2 + \partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x, \tfrac{y}{r}] \\
&\quad + [\tfrac{x}{r}, y(\partial_z^2 + \partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y]\big) \\
&= \frac{-me^2}{\pi\varepsilon_0}\Bigg(x\left(2\left(\tfrac{1}{r} - \tfrac{y^2}{r^3}\right)\partial_y + 3\tfrac{y^3}{r^5}\right) - y\left(2\left(\tfrac{1}{r} - \tfrac{x^2}{r^3}\right)\partial_x + 3\tfrac{x^3}{r^5}\right) \\
&\quad - y\left(\tfrac{3xy^2}{r^5} - \tfrac{1}{r^3}(xy\partial_y + y^2\partial_x) + \tfrac{1}{r}\partial_x\right) + x\left(\tfrac{3x^2y}{r^5} - \tfrac{1}{r^3}(xy\partial_x + x^2\partial_y) + \tfrac{1}{r}\partial_y\right) \\
&\quad + \tfrac{1}{r^3}\left(yz^2\partial_x - xz^2\partial_y\right) \\
&= \frac{-me^2}{\pi\varepsilon_0}\Bigg(\frac{2}{r}\left(x\partial_y - y\partial_x\right) + \frac{2}{r^3}\left(-xy^2\partial_y + x^2y\partial_x\right) \\
&\quad + \frac{1}{r}\left(-y\partial_x + x\partial_y\right) - y\left(-\tfrac{1}{r^3}(xy\partial_y + y^2\partial_x)\right) + x\left(-\tfrac{1}{r^3}(xy\partial_x + x^2\partial_y)\right) \\
&\quad + \tfrac{1}{r^3}\left(yz^2\partial_x - xz^2\partial_y\right) \\
&= \frac{1}{4\pi\varepsilon_0}\frac{-8me^2}{r}\left(x\partial_y - y\partial_x\right) \\
&= \frac{1}{4\pi\varepsilon_0}\frac{8me^2}{i\hbar r}\hat{L}_z,
\end{aligned}$$

via the circularity of the system, we also have $[\hat{M}_y, \hat{K}_z] + [\hat{K}_y, \hat{M}_z] = \frac{1}{4\pi\varepsilon_0}\frac{8me^2}{i\hbar r}\hat{L}_x$ and $[\hat{M}_z, \hat{K}_x] + [\hat{K}_z, \hat{M}_x] = \frac{1}{4\pi\varepsilon_0}\frac{8me^2}{i\hbar r}\hat{L}_y$. Now it rests us to calculate $[\hat{K}_x, \hat{K}_y]$, remember we defined:

$$\hat{K}_x = \frac{1}{4\pi\varepsilon_0}\frac{2me^2x}{i\hbar r}, \quad \hat{K}_y = \frac{1}{4\pi\varepsilon_0}\frac{2me^2y}{i\hbar r}, \quad \hat{K}_z = \frac{1}{4\pi\varepsilon_0}\frac{2me^2z}{i\hbar r}.$$

We instantly see that $[\hat{K}_n, \hat{K}_m] = 0 \quad \forall n, m \in \{x, y, z\}$, hence we can now finally compute the commutator of $\hat{A}_x$ and $\hat{A}_y$:

$$\begin{aligned}
[\hat{A}_x, \hat{A}_y] &= \frac{\hbar^2}{8mE}\Big([\hat{M}_x, \hat{M}_y] + [\hat{M}_x, \hat{K}_y] + [\hat{K}_x, \hat{M}_y] + [\hat{K}_x, \hat{K}_y]\Big) \\
&= \frac{\hbar^2}{8mE}(-4i\hbar\hat{L}_z\nabla^2 + \frac{1}{4\pi\varepsilon_0}\frac{8me^2}{i\hbar r}\hat{L}_z) \\
&= \frac{-i\hbar}{E}(\frac{\hbar^2}{2m}\hat{L}_z\nabla^2 + \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}\hat{L}_z) = i\hbar\hat{L}_z.
\end{aligned}$$

Due to the circularity of the system we conclude that:

$$[\hat{A}_x, \hat{A}_y] = i\hbar \hat{L}_z,$$
$$[\hat{A}_y, \hat{A}_z] = i\hbar \hat{L}_x,$$
$$[\hat{A}_z, \hat{A}_x] = i\hbar \hat{L}_y.$$

Now we will look into the commutator of $\hat{L}_x$ and $\hat{A}_y$. Before we do this we will verify two identities we will need, namely $[\hat{L}_x, \hat{M}_y] = -i\hbar \hat{M}_z$ en $[\hat{L}_n, \frac{1}{r}] = 0$. starting with the latter, we will prove it for $n = x$ and due to the cyclic nature of the angular momentum operator the same proof also holds for the other coordinates.

$$[\hat{L}_x, \tfrac{1}{r}] = [-i\hbar(y\partial_z - z\partial_y), \tfrac{1}{r}] = -i\hbar((y\partial_z - z\partial_y)\tfrac{1}{r} - \tfrac{1}{r}(y\partial_z - z\partial_y)) = -i\hbar(\tfrac{1}{r^3}(yz - zy)) = 0$$

We will use this in combination with the product rule to show:

$$\begin{aligned}\left[\hat{L}_x, \frac{y}{r}\right] &= \left[\hat{L}_x, \frac{1}{r}\right] y + \frac{1}{r}\left[\hat{L}_i, y\right] \\ &= \frac{1}{r}\left[\hat{L}_x, y\right] \\ &= \frac{i\hbar z}{r},\end{aligned}$$

and thus also $\left[\hat{L}_y, \frac{z}{r}\right] = \frac{i\hbar x}{r}$ and $\left[\hat{L}_z, \frac{x}{r}\right] = \frac{i\hbar y}{r}$. Now for the first identity:

$$\begin{aligned}[\hat{L}_x, \hat{M}_y] &= -2\hbar^2[y\partial_z - z\partial_y, y(\partial_z^2 + \partial_x^2) - z\partial_y\partial_z - x\partial_y\partial_x - \partial_y] \\ &= -2\hbar^2([y\partial_z, -z\partial_y\partial_z - x\partial_y\partial_x - \partial_y] - [z\partial_y, y(\partial_z^2 + \partial_x^2) - z\partial_y\partial_z]) \\ &= -2\hbar^2(-[y\partial_z, z\partial_y\partial_z] - x[y, \partial_y]\partial_x\partial_z - [y, \partial_y]\partial_z - [z\partial_y, y\partial_z^2] - z[\partial_y, y]\partial_x^2 + [z, z\partial_z]\partial_y^2) \\ &= -2\hbar^2(-y\partial_y\partial_z + z\partial_z^2 + x\partial_x\partial_z + \partial_z + 2y\partial_y\partial_z - z\partial_z^2 - z\partial_x^2 - z\partial_y^2) \\ &= -2\hbar^2(x\partial_x\partial_z + \partial_z + y\partial_y\partial_z - z(\partial_x^2 + \partial_y^2)) \\ &= i\hbar(-2i\hbar\big(z(\partial_x^2 + \partial_y^2) - x\partial_z\partial_x - y\partial_z\partial_y - \partial_z\big)) \\ &= i\hbar\hat{M}_z.\end{aligned}$$

By the circular nature of the problem we then also have $[\hat{L}_y, \hat{M}_z] = i\hbar\hat{M}_x$ and $[\hat{L}_z, \hat{M}_x] = i\hbar\hat{M}_y$. Now for the final commutator:

$$\begin{aligned}[\hat{L}_x, \hat{A}_y] &= [\hat{L}_x, \frac{-i\hbar}{2\sqrt{-2mE}}\left(\hat{M}_y + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 y}{i\hbar r}\right)] \\ &= \frac{-i\hbar}{2\sqrt{-2mE}}[\hat{L}_x, \hat{M}_y + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 y}{i\hbar r}] \\ &= \frac{-i\hbar}{2\sqrt{-2mE}}\left([\hat{L}_x, \hat{M}_y] + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}[\hat{L}_x, \tfrac{y}{r}]\right) \\ &= \frac{-i\hbar}{2\sqrt{-2mE}}\left((i\hbar\hat{M}_z) + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}(i\hbar\tfrac{z}{r})\right) \\ &= i\hbar\frac{-i\hbar}{2\sqrt{-2mE}}\left(\hat{M}_z + \frac{1}{4\pi\varepsilon_0}\frac{2me^2 z}{i\hbar r}\right) \\ &= i\hbar\hat{A}_z.\end{aligned}$$

Now due to the circularity we also have $[\hat{L}_y, \hat{A}_z] = -i\hbar\hat{A}_x$ and $[\hat{L}_z, \hat{A}_x] = -i\hbar\hat{A}_y$. With a very similar computation we also get $[\hat{A}_x, \hat{L}_y] = i\hbar\hat{A}_z$, $[\hat{A}_y, \hat{L}_z] = i\hbar\hat{A}_x$ and $[\hat{A}_z, \hat{L}_x] = i\hbar\hat{A}_y$. Now it rests us to show that $\hat{A}_n$ and $\hat{L}_n$ commute to see that the commutation relations match that of $\mathfrak{so}(4)$.

$$\begin{aligned}
[\hat{L}_x, \hat{A}_x] &= [\hat{L}_x, \frac{-i\hbar}{2\sqrt{-2mE}}(M_x + K_x)] \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}([\hat{L}_x, M_x] + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}[\hat{L}_x, \frac{x}{r}]) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}([\hat{L}_x, M_x] + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}([\hat{L}_x, \frac{1}{r}]x + \frac{1}{r}[\hat{L}_x, x]) \\
&= \frac{-i\hbar}{2\sqrt{-2mE}}([\hat{L}_x, M_x]) \\
&= \frac{i\hbar^3}{\sqrt{-2mE}}([y\partial_z - z\partial_y, x(\partial_y^2 + \partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z - \partial_x]) \\
&= \frac{i\hbar^3}{\sqrt{-2mE}}([y\partial_z - z\partial_y, x(\partial_y^2 + \partial_z^2) - y\partial_x\partial_y - z\partial_x\partial_z]) \\
&= \frac{i\hbar^3}{\sqrt{-2mE}}([y\partial_z, x\partial_y^2 - y\partial_x\partial_y - z\partial_x\partial_z] - [z\partial_y, x\partial_z^2 - y\partial_x\partial_y - z\partial_x\partial_z]) \\
&= \frac{i\hbar^3}{\sqrt{-2mE}}(x[y, \partial_y^2]\partial_z - [y, y\partial_y]\partial_x\partial_z - y[\partial_z, z\partial_z]\partial_x - x[z, \partial_z^2]\partial_y + z[\partial_y, y\partial_y]\partial_x + [z, z\partial_z]\partial_x\partial_y) \\
&= \frac{i\hbar^3}{\sqrt{-2mE}}(-2x\partial_y\partial_z - y\partial_x\partial_z + y\partial_x\partial_z + 2x\partial_z\partial_y + z\partial_x\partial_y - z\partial_x\partial_y) = 0
\end{aligned}$$

Again due to the circularity of the system we also have that $[\hat{L}_y, \hat{A}_y] = 0$ and $[\hat{L}_z, \hat{A}_z] = 0$. Now we have finally calculated all commutation relations.
Now we still need to check some details. First $\hat{A} \cdot \hat{L} = 0, \quad \hat{L} \cdot \hat{A} = 0$. With $[\hat{L}_n, \hat{A}_n] = 0$ showing either is true is sufficient. We will now check $\hat{A} \cdot \hat{L} = 0$, we will do this by looking at $\hat{M} \cdot \hat{L}, \quad \frac{\hat{R}}{r} \cdot \hat{L}$.

$$\begin{aligned}
\frac{-1}{\hbar^2}\hat{M}_x\hat{L}_x &= \left(2x\partial_y^2 - 2y\partial_x\partial_y - 2z\partial_x\partial_z + 2x\partial_z^2 - 2a\partial_x\right)\left(y\partial_z - z\partial_y\right) \\
&= 2\left(xy\partial_y^2\partial_z - yz\partial_x\partial_z^2\right) + 2\left(yz\partial_x\partial_y^2 - xz\partial_y\partial_z^2\right) \\
&\quad + 4\left(z\partial_x\partial_y - y\partial_x\partial_z\right) + 2\left(xy\partial_z^2 - xz\partial_y^3\right) \\
&\quad + \left(2z^2\partial_x\partial_y\partial_z - 2y^2\partial_x\partial_y\partial_z\right).
\end{aligned}$$

If we calculate the other two products, then the sum of the nth term of each product sums to zero, for the first term we get for example:

$$\left(xy\partial_y^2\partial_z - yz\partial_x\partial_z^2\right) + \left(yz\partial_z^2\partial_x - zx\partial_y\partial_x^2\right) + \left(zx\partial_x^2\partial_y - xy\partial_z\partial_y^2\right) = 0.$$

Hence:

$$\hat{M} \cdot \hat{L} = \hat{M}_x\hat{L}_x + \hat{M}_y\hat{L}_y + \hat{M}_z\hat{L}_z = 0.$$

Now for the second term:

$$\frac{\hat{R}}{r}\cdot\hat{L} = \frac{x}{\sqrt{x^2+y^2+z^2}}\hat{L}_x + \frac{y}{\sqrt{x^2+y^2+z^2}}\hat{L}_y + \frac{z}{\sqrt{x^2+y^2+z^2}}\hat{L}_z$$
$$= \frac{-i\hbar}{\sqrt{x^2+y^2+z^2}}\left(xz\partial_y - xy\partial_z + yx\partial_z - yz\partial_x + zy\partial_x - zx\partial_y\right) = 0.$$

Thus by linearity we can conclude that $\hat{A}\cdot\hat{L} = 0$. Now we still need to check that indeed $\hat{L}^2 + \hat{A}^2 = 1 + \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \frac{me^4}{2E\hbar^2}$.First we look at $\hat{L}^2$,

$$\begin{aligned}
\hat{L}^2 &:= \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \\
&= -\hbar^2((z\partial_y - y\partial_z)^2 + (x\partial_z - z\partial_x)^2 + (y\partial_x - x\partial_y)^2) \\
&= -\hbar^2((y^2+z^2)\partial_x^2 + (x^2+y^2)\partial_z^2 + (x^2+z^2)\partial_y^2) \\
&\quad - 2yz\partial_y\partial_z - 2xz\partial_x\partial_z - 2xy\partial_x\partial_y - 2x\partial_x - 2y\partial_y - 2z\partial_z) \\
&= -\hbar^2((x^2+y^2+z^2)\nabla^2 - (x\partial_x + y\partial_y + z\partial_z)^2 - (x\partial_x + y\partial_y + z\partial_z)).
\end{aligned}$$

Now we will look at $\hat{A}^2$ for this we remind ourself that

$$\hat{A}_x = \frac{-i\hbar}{2\sqrt{-2mE}}\left(\hat{M}_x + \frac{1}{4\pi\varepsilon_0}\frac{2me^2x}{i\hbar r}\right)$$

so

$$\hat{A}^2 = \frac{\hbar^2}{8mE}\left(\hat{M} + \frac{1}{4\pi\varepsilon_0}\frac{2me^2\hat{R}}{i\hbar r}\right)^2$$

Hence we will first look at $\hat{M}^2$

$$\begin{aligned}
\frac{-1}{4\hbar^2}\hat{M}_x^2 &= \left(x\partial_y^2 - y\partial_x\partial_y - z\partial_x\partial_z + x\partial_z^2 - \partial_x\right)^2 \\
&= x^2\partial_y^4 + y^2\partial_x^2\partial_y^2 + y\partial_x^2\partial_y + z^2\partial_x^2\partial_z^2 + x\partial_z^2\partial_x + x^2\partial_z^4 + \partial_x^2 \\
&\quad - 2xy\partial_y^3 - 2x\partial_x\partial_y^2 - y\partial_y^3 - 2xz\partial_x\partial_y^2\partial_z - z\partial_y^2\partial_z + 2x^2\partial_y^2\partial_z^2 \\
&\quad - 2x\partial_x\partial_y^2 - \partial_y^2 + 2yz\partial_x^2\partial_y\partial_z - 2xy\partial_x\partial_y\partial_z^2 - y\partial_y\partial_z^2 + 2y\partial_x^2\partial_y \\
&\quad - 2xz\partial_x\partial_z^3 - z\partial_z^3 - 2x\partial_x\partial_z^2 + 2z\partial_x^2\partial_z - 2x\partial_x\partial_z^2 - \partial_z^2 \\
&= \left(\partial_x^2 - \partial_y^2 - \partial_z^2\right) \\
&\quad + \left(x^2\partial_y^4 + x^2\partial_z^4 + y^2\partial_x^2\partial_y^2 + z^2\partial_x^2\partial_z^2 + 2x^2\partial_y^2\partial_x^2 + x^2\partial_x^4 + y^2\partial_x^2\partial_y^2 + z^2\partial_x^2\partial_z^2\right) \\
&\quad - \left(x^2\partial_x^4 + y^2\partial_x^2\partial_y^2 + z^2\partial_x^2\partial_z^2 + 2xy\partial_x\partial_y^3 + 2xy\partial_x\partial_y\partial_z^2\right) \\
&\quad + 2xz\partial_x\partial_z^3 + z\partial_z^3 + y\partial_y\partial_z^2 + x\partial_x\partial_z^2 \\
&\quad - \left(y\partial_y^3 + z\partial_y^2\partial_z + 2x\partial_x\partial_y^2\right) \\
&\quad + \left(2yz\partial_x\partial_y\partial_z - 2xz\partial_x\partial_y^2\partial_z - 2x\partial_y\partial_z^2 + 2y\partial_x^2\partial_y\right) \\
&\quad - 3x\partial_x\partial_z^2 + 3z\partial_x^2\partial_z + x\partial_y^2\partial_y.
\end{aligned}$$

Now in a very similar manner one can calculate $\hat{M}_y^2$ and $\hat{M}_z^2$ and when we sum these we get.

$$\frac{-1}{4\hbar^2}\hat{M}^2 = \left(-1 - \frac{\hat{L}^2}{\hbar^2}\right)\nabla$$

The second square term is just

$$-\left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2$$

since the $\hat{R}$ cancels with the r^2 term. Thus we are left with the cross terms. For the x cross terms we get:

$$\begin{aligned}
&\left(\hat{M}_x\frac{x}{\sqrt{x^2+y^2+z^2}}+\frac{x}{\sqrt{x^2+y^2+z^2}}\hat{M}_x\right)\\
&\qquad=\left[\hat{M}_x,\frac{x}{\sqrt{x^2+y^2+z^2}}\right]+2\frac{x}{\sqrt{x^2+y^2+z^2}}\hat{M}_x\\
&\qquad=\frac{1}{\sqrt{x^2+y^2+z^2}}[\hat{M}_x,x]\\
&\qquad\quad+x\left[\hat{M}_x,\frac{1}{\sqrt{x^2+y^2+z^2}}\right]+2\frac{x}{\sqrt{x^2+y^2+z^2}}\hat{M}_x,
\end{aligned}$$

similar results hold for y and z. We will now calculate the individual terms:

$$\begin{aligned}
&\frac{1}{\sqrt{x^2+y^2+z^2}}\left([\hat{M}_x,x]+[\hat{M}_y,y]+[\hat{M}_z,z]\right)\\
&=\frac{2i\hbar}{\sqrt{x^2+y^2+z^2}}(3+2x\partial_x+2y\partial_y+2z\partial_z),
\end{aligned}$$

$$\begin{aligned}
\left[x\hat{M}_x+y\hat{M}_y+z\hat{M}_z,\frac{1}{\sqrt{x^2+y^2+z^2}}\right]&=\frac{2i\hbar}{\left(\sqrt{x^2+y^2+z^2}\right)^3}\Big(x^2+y^2+z^2+x^2y\partial_y+y^2\partial_z+z^2x\partial_x+x^2z\partial_z\\
&\quad+y^2x\partial_x+z^2y\partial_y-x^2\partial_x-xy\partial_y-x^2\partial_y-xy^2\partial_y-y^2z\partial_z-xz^2\partial_x\Big)\\
&=\frac{2i\hbar}{\sqrt{x^2+y^2+z^2}},
\end{aligned}$$

$$\begin{aligned}
\frac{2}{\sqrt{x^2+y^2+z^2}}\left(x\hat{M}_x+y\hat{M}_y+z\hat{M}_z\right)&=\frac{4i\hbar}{\sqrt{x^2+y^2+z^2}}\big(2xy\partial_x\partial_y+2yz\partial_y\partial_z+2xz\partial_z\partial_x\\
&\quad-x^2\partial_y^2-x^2\partial_z^2-y^2\partial_x^2-y^2\partial_z^2-z^2\partial_x^2-z^2\partial_y^2+x\partial_x+y\partial_y+z\partial_z\big).
\end{aligned}$$

summing gives:

$$\begin{aligned}
&\frac{4i\hbar}{\sqrt{x^2+y^2+z^2}}\big(1+2x\partial_x+2y\partial_y+2z\partial_z+2xy\partial_x\partial_y+2yz\partial_y\partial_z\\
&\quad+2xz\partial_x\partial_z-x^2\partial_y^2-x^2\partial_z^2-y^2\partial_x^2-y^2\partial_z^2-z^2\partial_x^2-z^2\partial_y^2\big).\\
&=\frac{4i\hbar(\frac{\hat{L}^2}{\hbar^2}+1)}{\sqrt{x^2+y^2+z^2}}.
\end{aligned}$$

So in total:

$$
\begin{aligned}
\hat{L}^2 + \hat{A}^2 &= \hat{L}^2 + \frac{\hbar^2}{8mE}\left(4\hbar^2\left(1+\frac{\hat{L}^2}{\hbar^2}\right)\nabla + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}\frac{4i\hbar(\frac{\hat{L}^2}{\hbar^2}+1)}{\sqrt{x^2+y^2+z^2}} - \left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2\right) \\
&= \hat{L}^2\left(1+\frac{\hbar^2}{2mE}\nabla + \frac{1}{E}\frac{e^2}{4\pi\varepsilon_0}\frac{1}{\sqrt{x^2+y^2+z^2}}\right) \\
&\quad + \frac{\hbar^2}{8mE}\left(4\hbar^2\nabla + \frac{1}{4\pi\varepsilon_0}\frac{2me^2}{i\hbar}\frac{4i\hbar}{\sqrt{x^2+y^2+z^2}} - \left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2\right) \\
&= \frac{\hbar^2}{8mE}\left(4\hbar^2\nabla + \frac{2me^2}{4\pi\varepsilon_0}\frac{4}{\sqrt{x^2+y^2+z^2}} - \left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2\right) \\
&= \frac{\hbar^2}{E}\left(\frac{\hbar^2}{2m}\nabla + \frac{e^2}{4\pi\varepsilon_0}\frac{1}{\sqrt{x^2+y^2+z^2}}\right) - \frac{\hbar^2}{8mE}\left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2 \\
&= -\hbar^2 - \frac{\hbar^2}{8mE}\left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2.
\end{aligned}
$$

Hence:

$$
\frac{-1}{\hbar^2}(\hat{L}^2 + \hat{A}^2) = 1 + \frac{\hbar^2}{2mE}\left(\frac{2me^2}{4\pi\varepsilon_0\hbar}\right)^2 = 1 + \left(\frac{1}{4\pi\varepsilon_0}\right)^2\frac{2me^4}{E\hbar^2}.
$$