Kepler’s Laws

I. The equation of motion

We consider the motion of a point mass under the influence of a gravitational field created by point mass that is fixed at the origin.

Newton’s laws give the basic equation of motion for such a system. We denote by \( q(t) \) the position of the movable point mass at time \( t \), by \( m \) the mass of the movable point mass, and by \( M \) the mass of the point mass fixed at the origin. By Newton’s law on gravition, the force exerted by the fixed mass on the movable mass is in the direction of the vector \(-q(t)\). It is proportional to the product of the two masses and the inverse of the square of the distances between the two masses. The proportionality constant is the gravitational constant \( G \). In formulæ

\[
\text{force} = -G mM \frac{q}{\|q\|^3}
\]

Newton’s second law states that

\[
\text{force} = \text{mass} \times \text{acceleration} = m \ddot{q}
\]

From these two equation one gets

\[
m \ddot{q} = -G mM \frac{q}{\|q\|^3}
\]

Dividing by \( m \),

\[
\ddot{q} = -\mu \frac{q}{\|q\|^3}
\]  

(1)

with \( \mu = GM \). This is the basic equation of motion. It is an ordinary differential equation, so the motion is uniquely determined by initial point and initial velocity. In particular \( q(t) \) will always lie in the linear subspace of \( \mathbb{R}^3 \) spanned by the initial position and initial velocity. This linear subspace will in general have dimension two, and in any case has dimension at most two.

II. Statement of Kepler’s Laws

Johannes Kepler (1571-1630) had stated three laws about planetary motion. We state these laws below. Isaac Newton (1643-1727) showed that these laws are consequences of Newton’s laws.

\[\text{(1)}\] For a discussion of the amount of rigour in Newton’s Principia, see [Pourciau]. In fact, Newton seems to have been more interested in the “direct problem” of finding basic laws consistent with the Kepler laws; see [Brackenridge].

1
of the basic equation of motion (1). We first state Kepler’s laws in an informal way, then discuss a more rigid formulation, and then give various proofs of the fact that (1) implies these laws.

Kepler’s first law: Let \( q(t) \) be a maximal solution of the basic equation of motion. Its orbit is either an ellipse which has one focal point at the origin, a branch of a hyperbola which has one focal point at the origin, a parabola whose focal point is the origin, or an open ray emanating from the origin.

Kepler’s second law (Equal areas in equal times): The area swept out by the vector joining the origin to the point \( q(t) \) in a given time is proportional to the time.

Kepler’s third law: The squares of the periods of the planets are proportional to the cubes of their semimajor axes.

We now comment on the terms used in the formulation of Kepler’s laws.

The Picard Lindelöf theorem about existence and uniqueness of solutions applies to the differential equation (1). It implies that for any solution \( q(t) \) defined on an interval \( I \), there is a unique maximal open interval \( I' \) containing \( I \) such that the solution can be extended to \( I' \). Such a solution we call maximal. \( I' \) may be the whole real axis or have boundary points. If \( t_0 \) is a boundary point then \( \lim_{t \in I', t \to t_0} |q(t)| = 0 \) or \( \lim_{t \in I', t \to t_0} |q(t)| = \infty \). The orbit of the solution is by definition the set \( \{ q(t) \mid t \in I' \} \) in \( \mathbb{R}^2 \).

One of the standard definitions of ellipses is the following: Let \( F \) and \( F' \) be two points in the plane. Fix a length, say \( 2a \), which is bigger or equal to the distance between \( F \) and \( F' \). Then the curve consisting of all points \( P \) for which the sum of the distances from \( P \) to \( F \) and from \( P \) to \( F' \) is equal to \( 2a \) is called an ellipse with focal points \( F, F' \). \( a \) is called the semimajor axis of the ellipse.

Similarly, fix a length \( 2a > 0 \), which is smaller than the distance between \( F \) and \( F' \). Then the curve consisting of all points \( P \) for which the absolute value of the difference of the distances from \( P \) to \( F \) and from \( P \) to \( F' \) is equal to \( 2a \) is called a hyperbola with focal points \( F, F' \).

Finally, fix a point \( F \) and a line \( g \) that does not contain \( F \). Then the curve consisting of all points \( P \) for which the distance from \( P \) to the point \( F \) is equal to the distance from \( P \) to the line \( g \) is called a parabola with focal point \( F \).
Ellipses, parabolas and hyperbolas make up the conic sections. There are many other ways to describe conic sections. see the appendix below and the references cited therein.

For two vectors \( v = (v_1, v_2) \) and \( w = (w_1, w_2) \) in the plane, we denote, by abuse of notation, by

\[ v \times w = v_1 w_2 - v_2 w_1 \]

the third component of the cross product of \( v \) and \( w \). Its absolute value is the double of the area of the triangle spanned by the points 0, \( v \) and \( v + w \).

Now let \( q(t), t \in (a, b) \) be a differentiable curve in the plane. Using Riemann sums, one sees that the area swept out by the vector joining the origin 0 to the point \( q(t) \) in the time between between \( t_1 \) and \( t_2, a < t_1 \leq t_2 < b \), is equal to \( \frac{1}{2} \int_{t_1}^{t_2} q(t) \times \dot{q}(t) \, dt \).

Kepler's second law states that there is a proportionality constant \( L \) such that \( \int_{t_1}^{t_2} q(t) \times \dot{q}(t) \, dt = L(t_2 - t_1) \). By the fundamental theorem of calculus, this equivalent to saying that

\[ q(t) \times \dot{q}(t) = L \quad (2) \]

is constant. Up to a constant depending on the mass of the particle, the quantity \( q(t) \times \dot{q}(t) \) is the angular momentum vector of the particle with respect to the origin. Thus, Kepler's second law is a consequence of the principle of conservation of angular momentum.

When the orbit of the solution is an ellipse, we talk of planetary motion. In this case it follows from Kepler's second law that the motion is periodic. The period is the minimal \( T > 0 \) such that \( q(t + T) = q(t) \) for all \( t \in \mathbb{R} \). The precise form of the third law is, that

\[ \frac{T^2}{a^3} = \frac{4\pi^2}{\mu} \]

where \( a \) is the major semiaxis of the ellipse.

III. Proofs of Kepler’s Laws

The proof of Kepler’s second law is straightforward. By (1)

\[ \frac{d}{dt} q(t) \times \dot{q}(t) = \dot{q}(t) \times \dot{q}(t) + q(t) \times \ddot{q}(t) = 0 - \frac{\mu}{|q(t)|^3} q(t) \times q(t) = 0 \]

so that \( q(t) \times \dot{q}(t) \) is constant. As discussed in the previous section, this is the content of Kepler’s second law.
Another conserved quantity is the total energy

\[ E = \frac{1}{2} \| \dot{q} \|^2 - \frac{\mu}{|q|} \]  

Indeed,

\[ \frac{d}{dt} \left( \frac{1}{2} \| \dot{q} \|^2 - \frac{\mu}{|q|} \right) = \frac{d}{dt} \left( \frac{1}{2} \dot{q} \cdot \dot{q} - \frac{\mu}{|q|^3} q \cdot \dot{q} \right) = \dot{q} \cdot \ddot{q} + \frac{\mu}{|q|^3} q = 0 \]

by Kepler’s equation (1).

The proof of the first law is not as obvious as that of the second law. We give several proofs, always using Kepler’s second law and conservation of energy. Let

\[ L = q \times \dot{q} \]

be the constant of the second law (the angular momentum). We assume from now on that \( L \neq 0 \); otherwise one has motion on a ray emanating from the origin.

**Polar coordinates**

Kepler’s equation is rotation symmetric. Therefore it is a natural idea to use polar coordinates in the plane where the motion of the particle takes place. Without loss of generality we may assume that this is the \((q_1, q_2)\)-plane. The polar coordinates \( r, \varphi \) of a point are defined by

\[ q_1 = r \cos \varphi, \quad r = \sqrt{q_1^2 + q_2^2} \]
\[ q_2 = r \sin \varphi, \quad \tan \varphi = \frac{q_2}{q_1} \]

We first express the absolute value of the angular momentum \( L \) and the energy in polar coordinates. Observe that

\[ \| q \times \dot{q} \| = |q_1 \dot{q}_2 - q_2 \dot{q_1}| = |r \dot{r} \cos \varphi \sin \varphi + r^2 \dot{\varphi} \cos^2 \varphi - r \dot{r} \sin \varphi \cos \varphi + r^2 \dot{\varphi} \sin^2 \varphi| = |r^2 \dot{\varphi}| \]

Without loss of generality we assume from now on that \( \dot{\varphi} > 0 \). So

\[ \| L \| = r^2 \dot{\varphi}, \quad \text{or, equivalently,} \quad \dot{\varphi} = \frac{\| L \|}{r^2} \]  

Also observe that

\[ \| \dot{q} \|^2 = \left( \frac{d}{dt} r \cos \varphi \right)^2 + \left( \frac{d}{dt} r \sin \varphi \right)^2 \]
\[ = (\dot{r} \cos \varphi - r \dot{\varphi} \sin \varphi)^2 + (\dot{r} \sin \varphi + r \dot{\varphi} \cos \varphi)^2 \]
\[ = r^2 \cos^2 \varphi + r^2 \dot{\varphi}^2 \sin^2 \varphi + \dot{r}^2 \sin^2 \varphi + r^2 \dot{\varphi}^2 \cos^2 \varphi \]
\[ = \dot{r}^2 + r^2 \dot{\varphi}^2 \]  

4
By (5) and (4) the total energy (3) is

\[ E = \frac{1}{2} \| \dot{q} \|^2 - \frac{\mu}{r} = \frac{1}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2) - \frac{\mu}{r} = \frac{1}{2} (\dot{r}^2 + \| L \|^2) - \frac{\mu}{r} \]

by (4) so that

\[ \dot{r}^2 = 2E + \frac{2\mu}{r} - \frac{\| L \|^2}{r^2} \]  

\[(6)\]

is a differential equation for \( r \) as a function of \( t \). It is difficult to solve explicitly and we do not do this here. Instead we derive a differential equation for the angle \( \varphi \) as a function of the radius \( r \). Observe that

\[ \frac{d\varphi}{dr} = \frac{d\varphi}{dt} \frac{dt}{dr} = \frac{\dot{\varphi}}{r} = \frac{\| L \|}{r \sqrt{r^2 - \| L \|^2}} \]

by (4). Inserting (6) gives

\[ \frac{d\varphi}{dr} = \frac{\| L \|}{r^2 \sqrt{2E + \frac{2\mu}{r} - \frac{\| L \|^2}{r^2}}} = \frac{\| L \|}{r^2 \sqrt{2E + \frac{2\mu}{r} - (\| L \| - \frac{\mu}{r})^2}} = \frac{1}{r^2 \sqrt{\frac{2\mu}{r^2} - (\frac{\mu}{r} - \frac{\| L \|^2}{r^2})}} \]

with

\[ e = \sqrt{1 + \frac{2E\| L \|^2}{\mu^2}}, \quad l = \frac{\| L \|^2}{\mu} \]  

\[(7)\]

Therefore

\[ \varphi = \int \frac{dr}{r^2 \sqrt{\left(\frac{\mu}{r^2} - 1\right)^2}} = \int \frac{1}{r^2 \sqrt{1 - \left(\frac{l}{\mu} - 1\right)^2}} \frac{l}{r^2} \frac{1}{\sqrt{\left(\frac{l}{\mu} - 1\right)^2}} \left( \frac{d}{dr} \left( \frac{l}{\mu} - 1 \right) \right) dr \]

Thus, with an integration constant \( \varphi_0 \)

\[ -(\varphi - \varphi_0) = \arccos \left( \frac{l}{e} - 1 \right) \]

or

\[ r = \frac{l}{1 + e \cos(\varphi - \varphi_0)} \]  

\[(8)\]

This is the equation of a conic section with eccentricity \( e \). See Appendix A below. If the energy \( E \) is negative, the eccentricity \( e \) is smaller than one and we have an ellipse. Similarly, we for \( E = 0 \) we have \( e = 1 \) and the conic section is a parabola. Finally, if \( E > 0 \), we have \( e > 1 \) and the conic section is a hyperbola.

This ends the proof of Kepler’s first law using polar coordinates. Observe that we used only the equations (4) and (6) of conservation of angular momentum and of energy.

Using the formula from Appendix A, we see that in the case of ellipses \( E < 0 \) the major axis is

\[ a = \frac{l}{1-e^2} = \frac{\mu}{2|E|} \]  

\[(9)\]
since \(1 - e^2 = \frac{2|E|\|L\|^2}{\mu^2}\). The minor axis is

\[
b = a\sqrt{1 - e^2} = \frac{\|r\|}{2|E|} \cdot \sqrt{\frac{2|E|\|L\|^2}{\mu^2}} = \frac{\|L\|}{\sqrt{2|E|}} = \sqrt{\frac{a\|L\|}{\sqrt{\mu}}}
\]

(10)

Therefore the area of the ellipse is \(\pi ab = \pi a^3/2\|L\|/\sqrt{\mu}\).

Now let \(T\) be the period of the orbit. By Kepler’s second law the the area swept out after time \(T\) is

\[
\frac{1}{2} \int_0^T \|q(t) \times \dot{q}(t)\| \, dt = \frac{1}{2}\|L\|T
\]

On the other hand this is equal to the area of the ellipse. Therefore we get

\[
\frac{1}{2}\|L\|T = \pi a^3/2\|L\|/\sqrt{\mu}
\]

or

\[
T = \frac{2\pi}{\sqrt{\mu}}a^{3/2}
\]

This is Kepler’s third law.

The previous proof of Kepler’s first law is based on the conservation laws for energy and angular momentum and on the idea of writing the angle \(\varphi\) as a function of the radius \(r\). In a variant of this proof one writes \(\sigma = \frac{1}{r}\) as a function of \(\varphi\). By the change of variables formula, (4) and (5)

\[
\|\dot{q}\|^2 = \dot{r}^2 + r^2 \dot{\varphi}^2 = \dot{\varphi}^2 \left(\left(\frac{\dot{r}}{r}\right)^2 + r^2\right) = \|\dot{L}\|^2 \left(\frac{(\frac{d\sigma}{d\varphi})^2 + (\frac{r}{r^2})^2}{r^2}\right) = \|\dot{L}\|^2 \left(\frac{(\frac{d\sigma}{d\varphi})^2}{r^2}\right) + \frac{1}{r^2}
\]

Then by (3)

\[
E = \frac{1}{2}\|\dot{q}\|^2 - \frac{\mu}{\|q\|} = \frac{\|\dot{L}\|^2}{2} \left((\frac{d\sigma}{d\varphi})^2 + \sigma^2\right) - \mu\sigma
\]

Differentiating with respect to \(\varphi\) and dividing by \(\|\dot{L}\|^2\) gives

\[
0 = \frac{d\sigma}{d\varphi} \left(\frac{d^2\sigma}{d\varphi^2} + \sigma - \frac{\mu}{\|\dot{L}\|^2}\right)
\]

Since \(\frac{d\sigma}{d\varphi} \neq 0\) this implies

\[
\frac{d^2\sigma}{d\varphi^2} + \sigma - \frac{\mu}{\|\dot{L}\|^2} = 0
\]

* It is remarkable that \(a\) depends only on the energy. So all bounded orbits with the same energy are ellipses with the same length of the major axis, but of course their position in space is different because the eccentricity then varies with \(\|L\|\)
The general solution of this differential equation is

\[ \sigma = \frac{\mu}{\|L\|^2}(1 + e \cos(\varphi - \varphi_0)) \]

with integration constants \( e, \varphi_0 \). Remembering that \( r = \frac{1}{\sigma} \) and setting \( l = \frac{\|L\|^2}{\mu} \) as in (7) we get again the equation

\[ r = \frac{l}{1 + e \cos(\varphi - \varphi_0)} \]

for the orbit.
In this approach \( e \) just comes as an integration constant, but now it can easily be identified with the quantity of (7).

The Laplace Lenz Runge vector

The basic equation of motion is a second order differential equation. Its solution is completely determined by the position \( q \) and velocity \( \dot{q} \) at any given time \( t \). Therefore one expects that the quantities characterizing the orbits can be expressed in terms of \( q \) and \( \dot{q} \).

For the eccentricity of the conic section this is done in (7), since \( E \) and \( L \) are expressed in terms of \( q \) and \( \dot{q} \) in (3) and (2). Similarly, in the case of ellipses, the length of the major and minor axis are described in terms of \( q \) and \( \dot{q} \) by (9) and (10). If one considers the Kepler problem as a problem in three dimension, the plane of the orbit is determined as the plane through the origin perpendicular to the angular momentum vector \( L \).

These data determine the Kepler ellipse up to rotation around the origin (the integration constant \( \varphi_0 \) of the previous subsection. To fix this ambiguity we would like to find a vector in the direction of the major axis that can be expressed purely in terms of \( q \) and \( \dot{q} \).

The standard choice is the \textit{Laplace Lenz Runge vector}

\[ A = -\frac{q}{\|q\|} + \frac{1}{\mu} \dot{q} \times L = \frac{1}{\mu} (\|\dot{q}\|^2 - \frac{\mu}{\|q\|^3}) q - \frac{1}{\mu} (q \cdot \dot{q}) \dot{q} = \frac{1}{\mu} (2E + \frac{\mu}{\|q\|^3}) q - \frac{1}{\mu} (q \cdot \dot{q}) \dot{q} \quad (11) \]

Here we used the identity \( x \times (y \times z) = -(x \cdot y) z + (x \cdot z) y \) for vectors \( x, y, z \in \mathbb{R}^3 \) to see that \( \dot{q} \times L = \dot{q} \times (q \times \dot{q}) = -(q \cdot \dot{q})\dot{q} + \|\dot{q}\|^2 q \), and, in the second step, the definition (3) of the energy. Observe that, for a vector \( v \) in the “invariant plane” orthogonal to \( L \), \( v \times \frac{L}{\|L\|} \) is the vector in this plane obtained from \( v \) by rotating by 90°. So \( \frac{1}{\|q\|} \dot{q} \times L \) is the vector obtained from \( \dot{q} \) by rotating by 90°.

First we verify that \( A \) really is constant during a Kepler motion. Using the basic equation of motion (1) and the definition of \( L \)

\[ \frac{dA}{dt} = -\frac{1}{\|q\|} \dot{q} + \frac{q \cdot \dot{q}}{\|q\|^3} q + \frac{1}{\mu} \dot{q} \times L = -\frac{1}{\|q\|} \dot{q} + \frac{q \cdot \dot{q}}{\|q\|^3} q - \frac{1}{\|q\|^3} q \times (q \times \dot{q}) = 0 \]

We also used the fact that \( \frac{d}{dt} \frac{1}{\|q\|} = \frac{d}{dt} (q \cdot q)^{-1/2} = -(q \cdot \dot{q})(q \cdot q)^{-3/2} \) and again the identity \( x \times (y \times z) = -(x \cdot y) z + (x \cdot z) y \).
Next we claim that the length of the Laplace Lenz Runge vector is equal to the eccentricity \( e \). For this reason \( A \) is called the eccentricity vector in [Cushman]. To prove the statement about the length of \( A \) observe that
\[
A \cdot A = 1 - \frac{2\mu}{\mu^2 \| q \|^2} \| \dot{q} \times L \|^2
\]
Since \( \ddot{q} \) and \( L = q \times \dot{q} \) are perpendicular, \( \| \ddot{q} \times L \|^2 = \| \dot{q} \|^2 \| L \|^2 \). By the standard vector identity \( x \cdot (y \times z) = z \cdot (x \times y) \)
\[
q \cdot (\dot{q} \times L) = L \cdot (q \times \dot{q}) = L \cdot L = \| L \|^2
\]
Therefore
\[
A \cdot A = 1 - \frac{2\mu}{\mu^2 \| q \|^2} \| L \|^2 + \frac{1}{\mu^2} \| \dot{q} \|^2 \| L \|^2
\]
\[
= 1 + \frac{\| L \|^2}{\mu^2} \left( \| \dot{q} \|^2 - \frac{2\mu}{\mu^2 \| q \|^2} \right)
\]
\[
= 1 + 2 \frac{\| L \|^2}{\mu^2} E = e^2
\]
by (3).

Before using the Laplace Lenz Runge vector, we describe how one could get the idea to consider it, once one already knows Kepler’s laws*. Let us consider the case of negative energy, so that the orbits are ellipses. A natural invariant of the ellipse is the vector joining to consider it, once one already knows Kepler’s laws*. Let us consider the case of negative energy, so that the orbits are ellipses. A natural invariant of the ellipse is the vector joining to the two foci. The origin is one focus of the ellipse, call the other one the energy, so that the orbits are ellipses. A natural invariant of the ellipse is the vector joining to the other focus form opposite equal angles with the tangent line of the ellipse. See Appendix A. Therefore a vector from \( q \) in the direction of \( f \) is obtained by adding to \(-q\) twice the orthogonal projection of \( q \) to the tangent direction of the ellipse at \( q \).

A unit tangent vector to the ellipse at the point \( q \) is \( \frac{\dot{q}}{\| \dot{q} \|} \). Therefore \( f - q \) has the direction \(-q + 2(q \cdot \frac{\dot{q}}{\| \dot{q} \|}) \frac{\dot{q}}{\| \dot{q} \|} \). This means that there is a scalar function \( \alpha(t) \) such that
\[
f = q(t) + \alpha(t) \left( -q(t) + \frac{2q(t) \cdot \dot{q}(t)}{\| q(t) \|^2} \dot{q}(t) \right)
\]
for all \( t \). Differentiating and using Kepler’s law (1) gives
\[
0 = \dot{f} = \dot{q} + \ddot{\alpha}(q) \left( -q + 2 \frac{q \cdot \dot{q}}{\| \dot{q} \|^2} \dot{q} \right) + \alpha \left[ -q + 2 \left( \frac{\dot{q} \cdot \dot{q}}{\| \dot{q} \|^2} - \frac{q \cdot \ddot{q}}{\| \dot{q} \|^2} \right) \dot{q} + \frac{2q \cdot \dddot{q}}{\| \dot{q} \|^2} \right]
\]
\[
= -\dddot{\alpha} q + \frac{1}{\| q \|^2} \left\{ \| \dot{q} \|^2 + 2\alpha (q \cdot \dot{q}) + \alpha \left[ -\| \dot{q} \|^2 + 2\| \ddot{q} \|^2 + 2q \cdot ( -\frac{q}{\| q \|^2} ) - 2 \frac{q \cdot \dddot{q}}{\| q \|^2} ( -\gamma ) \right] \right\} \dot{q}
\]
\[
+ 2\alpha \frac{q \cdot \dddot{q}}{\| \dot{q} \|^2} \gamma
\]
\[
= -\left[ \dddot{\alpha} + 2\alpha \mu \frac{q \cdot \dddot{q}}{\| q \|^2 \| q \|^2} \right] q + \frac{1}{\| q \|^2} \left\{ \| \dot{q} \|^2 + 2(q \cdot \dot{q}) \left( \dddot{\alpha} + 2\alpha \mu \frac{q \cdot \dddot{q}}{\| q \|^2 \| q \|^2} \right) \right\} \dot{q}
\]
\[
= -\dddot{\alpha} \left( \frac{1}{\| q \|^2} \right) q + \frac{1}{\| q \|^2} \left\{ \| \dot{q} \|^2 + 2(q \cdot \dot{q}) \left( \dddot{\alpha} + 2\alpha \mu \frac{q \cdot \dddot{q}}{\| q \|^2 \| q \|^2} \right) \right\} \dot{q}
\]
for all \( t \).

* For other arguments, see [Heintz] and [Kaplan]. Remarks on the history of this vector can be found in [Goldstein].
Since \( q \) and \( \dot{q} \) are linearly independent, this implies that the coefficients of \( q \) and \( \dot{q} \) are both zero. From the coefficient of \( q \) we get

\[
\dot{\alpha} + 2\alpha \mu \frac{\dot{q} \cdot \dot{q}}{\|q\|^2} = 0
\]

Inserting this, and the fact that \( \frac{1}{2} \|\dot{q}\|^2 - \frac{\mu}{\|q\|} = E \) into the coefficient of \( \dot{q} \) gives

\[
\|\dot{q}\|^2 + 2\alpha E = 0
\]

So \( \alpha = -\frac{\|\dot{q}\|^2}{2E} \). Inserting this into (13) gives

\[
f = q - \frac{\|\dot{q}\|^2}{2E} \left( -q + \frac{2q \cdot \dot{q}}{\|q\|^2} \dot{q} \right)
\]

\[
= \frac{1}{E} \left( (E + \frac{\|\dot{q}\|^2}{2}) q - (q \cdot \dot{q}) \dot{q} \right)
\]

\[
= \frac{1}{E} \left( (\|\dot{q}\|^2 - \frac{\mu}{\|q\|}) q - (q \cdot \dot{q}) \dot{q} \right)
\]

\[
= \frac{\mu}{E} A \tag{14}
\]

The argument started with the observation that \( f \) should be a conserved quantity. As \( \frac{\mu}{E} \) is conserved, this suggests that \( A \) is a conserved quantity. Above, we have proven this directly. Observe from (9), that in the case of ellipses, \( \frac{\mu}{|E|} \) is twice the major semiaxis \( a \) of the ellipse. This is consistent with the fact that the distance between the two foci is \( 2ea \).

Once one knows that the Laplace Lenz Runge vector and the angular momentum vector are constants of the Kepler motion, the proof of Kepler’s first law is relatively fast. As

\[
q \cdot A = q \cdot \left( -\frac{q}{\|q\|} + \frac{1}{\mu} \dot{q} \times L \right) = -\|q\| + \frac{1}{\mu} q \cdot (\dot{q} \times L) = -\|q\| + \frac{1}{\mu} L \cdot (q \times \dot{q}) = -\|q\| + \frac{1}{\mu} \|L\|^2
\]

we have, setting again \( e = \|A\| \)

\[
\|q\| = e \left( \frac{1}{\mu} \|L\|^2 - q \cdot \frac{A}{\|A\|} \right)
\]

The expression in brackets is the distance from \( q \) to the line perpendicular to \( A \) through the point \( \frac{\|L\|^2}{e^2 \mu} A \). If we call this line the directrix, then the equation above states that the ratio of the distance between \( q \) and the origin and the distance between \( q \) and the directrix is equal to \( e \). As pointed out in Appendix A, this is one of the characterizing properties of conic sections.

**Exercise:** In the case of negative energy, show that \( \|q - f\| + \|q\| \) is constant! Here, \( f \) is the vector of (14).
Taking the cross product of the Laplace Lenz Runge vector $A$ with the angular momentum vector $L$ gives

$$L \times A = \frac{1}{||q||} L \times (-q) + \frac{1}{\mu} L \times (\dot{q} \times L) = \frac{1}{||q||} q \times L + \frac{1}{\mu} ||L||^2 \dot{q}$$

(15)

using the vector identity $x \times (y \times z) = -(x \cdot y) z + (x \cdot z) y$ and the fact that $\dot{q}$ and $L$ are perpendicular. Therefore

$$\dot{q} = \frac{\mu}{||L||^2} \left( L \times A - \frac{q}{||q||} \times L \right) = \frac{\mu}{||L||^2} L \times A + \frac{\mu}{||L||^2} L \times \frac{q}{||q||}$$

This equation determines the velocity vector $\dot{q}$ of the Kepler motion as a function of the position $q$. Observe that $\frac{\mu}{||L||^2} L \times A$ is independent of $t$ and that $\frac{q}{||q||}$ is always a vector of length one perpendicular to $L$. So $\frac{\mu}{||L||^2} L \times \frac{q}{||q||}$ is always a vector of length $\frac{\mu}{||L||}$ in the plane perpendicular to $L$. Consequently, for a fixed Kepler orbit, the velocity vectors all lie on the circle around the point $\frac{\mu}{||L||^2} L \times A$ of radius $\frac{\mu}{||L||}$. This circle is called the momentum hodograph.

**A momentum space argument**

In this subsection, we prove Kepler’s first law, starting with the analysis of the hodograph*, that is the curve traced out by the momentum vector

$$p(t) = \dot{q}(t)$$

The equation of motion (1) is

$$\ddot{p} = -\frac{\mu q}{||q||^3}$$

so that

$$\ddot{p} = -\frac{\mu}{||q||^3} \dot{q} + \frac{3\mu q \cdot \dot{q}}{||q||^5} q$$

Consequently

$$\dot{p} \times \ddot{p} = \frac{\mu^2}{||q||^6} q \times \dot{q} = \frac{\mu^2}{||q||^6} L$$

The standard formula for the curvature of a plane curve shows that the curvature of the hodograph at the point $q$ is

$$\kappa = \frac{||\dot{p} \times \ddot{p}||}{||p||^3} = \frac{\mu^2 ||L|| ||q||^6}{||q||^6} = \frac{||L||}{\mu}$$

* Another derivation of the equation of the hodograph, attributed to Hamilton, can be found in [Hankins], ch.24.
This proves that the curvature of the hodograph is constant. So it is a circle. Its radius is \( \frac{1}{\kappa} = \frac{\mu}{\|L\|} \). At each of its points \( p \), the vector pointing to the center \( u \) of the circle is perpendicular to \( L \) and to the tangent vector \( \dot{p} \). Its length is the radius \( \frac{\mu}{\|L\|} \). So

\[
p - u = \frac{\mu}{\|L\|} \frac{\dot{p}}{\|\dot{p}\|} \times \frac{L}{\|L\|} = -\frac{\mu}{\|q\|\|L\|} q \times L \tag{16}
\]

and our argument shows that

\[
u = p + \frac{\mu}{\|q\|\|L\|} q \times L \tag{17}
\]

is a conserved quantity\(^(*)\). The hodograph has the equation

\[
\|p - u\| = \frac{\mu}{\|L\|}
\]

So the hodograph is the circle around \( u \) of radius \( \frac{\mu}{\|L\|} \). Observe that \( \|u\| = \frac{\mu e}{\|L\|} \) where \( e = \sqrt{1 + \frac{2E\|L\|^2}{\mu^2}} \) as in (7), since

\[
\|u\|^2 = \|p\|^2 + \frac{\mu^2}{\|q\|^2\|L\|^2} \|q \times L\|^2 + 2 \frac{\mu}{\|q\|\|L\|} p \cdot (q \times p)
\]

\[
= \|p\|^2 + \frac{\mu^2}{\|L\|^2} - 2 \frac{\mu}{\|q\|\|L\|} \cdot L \cdot (p \times q)
\]

\[
= \|p\|^2 + \frac{\mu^2}{\|L\|^2} - 2 \frac{\mu}{\|q\|}
\]

\[
= 2E + \frac{\mu^2}{\|L\|^2} = \left( \frac{\mu e}{\|L\|} \right)^2
\]

Here we used, as in (12), that \( p \cdot (q \times L) = -L \cdot (q \times p) = -\|L\|^2 R \).

We now describe the position \( q \) in terms of the momentum \( p \). By (16) and the fact that \( q \) is perpendicular to \( q \times L \), we have \( q \cdot (p - u) = 0 \). So \( q \) is perpendicular to \( (p - u) \) and therefore it is a multiple of \( (p - u) \times L \). Write

\[
q = r \frac{(p-u) \times L}{\|(p-u) \times L\|} = \frac{r}{\mu} (p - u) \times L \quad \text{with} \quad r = \pm \|q\| \tag{18}
\]

As pointed out in (12), \( q \cdot (p \times L) = \|L\|^2 \). If we insert the representation of \( q \) above, we get

\[
r \frac{1}{\mu} ((p - u) \times L) \cdot (p \times L) = \|L\|^2
\]

Cross product by \( \frac{L}{\|L\|} \) corresponds to rotation by 90\(^o\) around the axis through the origin in direction of \( L \). Therefore this implies

\[
\mu = r (p - u) \cdot p
\]

\[
= r \left( (p - u) \cdot (p - u) + (p - u) \cdot u \right) \tag{19}
\]

\[
= r \left( \frac{\mu^2}{\|L\|^2} + \frac{\mu^2 e}{\|L\|^2} \frac{(p-u) \cdot u}{\|u\|^2} \right)
\]

\(^(*)\) From (15) one sees that \( u = -\frac{\mu}{\|L\|^2} L \times A \), where \( A \) is the Laplace Lenz Runge vector of the previous subsection. Conversely, \( A = -\frac{1}{\mu} u \times L \). One can prove directly that \( u \) is a conserved quantity and then derive the Laplace Lenz Runge vector using this formula.
Here we used that $\|p - u\| = \frac{\mu}{\|L\|}$ and that $\|u\| = \frac{\mu e}{\|L\|}$.

Denote the angle between $q$ and the (constant) vector $u \times L$ by $\varphi$. By (18) this is the angle between $(p - u) \times L$ and $u \times L$. Again, as cross product by a fixed vector does not change angles, $\phi$ also is the angle between $p - u$ and $u$. Now (19) gives

$$r = \frac{l}{1 + e \cos \varphi}$$

with $l = \frac{\|L\|^2}{\mu}$ as in (7). So we obtain again the equation of the conic as in Appendix A.

A purely geometric proof of Kepler’s first law using the hodograph was given by Hamilton, Kelvin and Tait, Maxwell, Fano and Feynman independently; see Feynman’s lost lecture [Goodstein]. It first describes the hodograph in geometric terms and then deduces the Kepler orbit as the enveloppe of its tangent lines. This argument is in parts very close to Newton’s original argument. There is a debate whether Newton’s original argument meets the 21\textsuperscript{st} century standards of rigour. For references see the introduction of [Derbes]. [Brackenridge] provides a guide and historical perspective on Newton’s treatment of the Kepler problem.

The eccentric anomaly

In the proofs of Kepler’s first law given above we derived the shape of the orbit, but did not actually write down a parametrization by the time $t$. Such a parametrization would correspond to a solution of the basic equation of motion (1). However, (8) is a parametrization of the orbit in terms of the angle $\varphi - \varphi_0$ (also called the true anomaly, see Appendix A), and conservation of angular momentum (4) implicitly gives the dependence of $\varphi$ on $t$ by the differential equation $\frac{d\varphi}{dt} = \frac{\|L\|}{r^2}$. To simplify the discussion we put $\varphi_0 = 0$.

We consider the case of negative energy, that is, the case of bounded orbits. Write

$$E = -\frac{\varepsilon^2}{2}$$

In this case the orbit is described by the equation

$$\frac{(q_1 + ea)^2}{a^2} + \frac{q_2^2}{b^2} = 1$$

where the principal axes $a, b$ of the ellipse and the eccentricity $e$ are determined by (9), (10) and (7) respectively. We use the parametrization

$$q_1 = a \cos s - ea, \quad q_2 = b \sin s$$

by the eccentric anomaly $s$ (see Appendix A). To get the dependence of $s$ on the time $t$, observe that by formula (30) of the appendix and (7)

$$\frac{ds}{d\varphi} = \frac{\sqrt{1 - e^2}}{l} r = \frac{\sqrt{2\varepsilon^2}}{\|L\|} r = \frac{\varepsilon}{\|L\|} r$$
As observed in (4), \( \frac{d\phi}{dt} = \frac{\|E\|}{r} \). Therefore, by the chain rule

\[
\frac{ds}{dt} = \frac{\varepsilon}{r}
\]

(20)

By (9), \( \varepsilon^2 = 2|E| = \frac{\sqrt{\mu}}{a} \). In formula (31) of the appendix, we show that \( r = a(1 - \cos s) \). Hence (20) gives

\[
\frac{ds}{dt} = \frac{\sqrt{\mu}}{a^{3/2}(1 - \cos s)}
\]

or, equivalently

\[
\frac{dt}{ds} = \frac{a^{3/2}(1 - \cos s)}{\sqrt{\mu}}
\]

Integrating both sides with respect to \( s \) gives Kepler’s equation

\[
t - t_0 = \frac{a^{3/2}}{\sqrt{\mu}} (s - e \sin s)
\]

(21)

Kepler’s equation is an implicit equation for \( s \) as a function of \( t \). It is not a differential equation anymore, but just an equation that involves the inversion of the “elementary function” \( s \mapsto s - e \sin s \). This inversion cannot be performed by elementary functions. See, however, Appendix E.

One of the advantages of the eccentric anomaly is, that it is well suited for a description of the Kepler motion in position space. Therefore we derive the equation of motion with respect to this parameter. We make the change of variables (20) in the basic equation of motion (1). Recall that \( r = \|q\| \). Then

\[
\frac{dq}{ds} = \hat{q} \frac{dt}{ds} = \frac{\|q\|}{\varepsilon} \hat{q} \quad \text{or, equivalently} \quad \hat{q} = \frac{\varepsilon}{\|q\|} \frac{dq}{ds}
\]

(22)

Therefore, using (1)

\[
\frac{d^2q}{ds^2} = \frac{1}{\varepsilon} \frac{d\|q\|}{ds} \hat{q} + \frac{\|q\|}{\varepsilon} \frac{dq}{ds} = \frac{1}{\|q\|} \frac{d\|q\|}{ds} \frac{dq}{ds} + \frac{\|q\|^2}{\varepsilon^2} \hat{q} = \frac{1}{\|q\|} \frac{d\|q\|}{ds} \frac{dq}{ds} - \frac{\mu}{\varepsilon^2 \|q\|} q
\]

(23)

Once one knows that (20) is a good change of variables for the Kepler problem and that the Laplace Lenz Runge vector is a conserved quantity, one can give a quick proof of Kepler’s first law:

By the definition of \( \varepsilon = \sqrt{-2E} \), (11) and (22), the Laplace Lenz Runge vector is

\[
A = \frac{1}{\mu} \left( -\varepsilon^2 + \frac{\mu}{\|q\|} \right) q - \frac{1}{\mu} (q \cdot \hat{q}) \hat{q} = -\frac{\varepsilon^2}{\mu} \left( \frac{1}{\|q\|} \|q\| \frac{dq}{ds} - \frac{\mu}{\varepsilon^2 \|q\|} q \right)
\]

since \( \varepsilon^2 = -2E \). Therefore (23) gives

\[
\frac{d^2q}{ds^2} + q = -\frac{\mu}{\varepsilon^2} A
\]

13
The general solution of this second order inhomogeneous linear differential equation with constant coefficients is

$$q(s) = C_1 \cos s + C_2 \sin s - aA$$

(24)

with constant vectors $C_1, C_2$ and $a = \frac{d\mu}{ds}$. This is the parametrization of an ellipse. We set $\|A\| = e$. To identify $C_1$ and $C_2$ we use the fact that, by (22),

$$\|q\|L = \varepsilon q \times \frac{dq}{ds} = \varepsilon(C_1 \times C_2 + aA \times C_1 \sin s - aA \times C_2 \cos s)$$

Since the functions $1, \sin s$ and $\cos s$ are linearly independent over $\mathbb{R}$, this implies that the vectors $C_1 \times C_2, A \times C_1$ and $A \times C_2$ are all proportional to $L$. Consequently $C_1, C_2$ and $A$ lie in one plane. We may assume (by replacing $s$ by $s - s_0$ and modifying $C_1, C_2$) that $q(0)$ points in the direction of $A$, in other words, that $C_1$ and $A$ are collinear. Then the equation above gives

$$\|q\|L = \varepsilon (C_1 \times C_2 - aA \times C_2 \cos s)$$

(25)

By (24)

$$\|q\|^2 = \|C_1\|^2 \cos^2 s + \|C_2\|^2 \sin^2 s + e^2a^2$$

$$+ 2C_1 \cdot C_2 \sin s \cos s - 2aA \cdot C_1 \cos s - 2aA \cdot C_2 \sin s$$

$$= (\|C_1\|^2 - \|C_2\|^2) \cos^2 s + (\|C_2\|^2 + e^2a^2)$$

$$+ 2C_1 \cdot C_2 \sin s \cos s - 2aA \cdot C_1 \cos s - 2aA \cdot C_2 \sin s$$

(26)

The square of the norm of the right hand side of (25) is a linear combination of $1, \cos s$ and $\cos^2 s$. As the functions $\cos^2 s, \sin s \cos s, \sin s, \cos s$ and $1$ are linearly independent over $\mathbb{R}$, the fact that this square of the norm of the right hand side of (25) is equal to $\|q\|^2\|L\|^2$ implies that the coefficients of $\sin s \cos s$ and $\sin s$ in (26) are zero. That is, $C_2$ is perpendicular to $C_1$ and $A$. So

$$\|q\|^2 = (\|C_1\|^2 - \|C_2\|^2) \cos^2 s + (\|C_2\|^2 + e^2a^2) - 2ae\|C_1\| \cos s$$

Taking the square of the absolute values of both sides of (25) now gives

$$\|L\|^2 ((\|C_1\|^2 - \|C_2\|^2) \cos^2 s + (\|C_2\|^2 + e^2a^2) - 2ea\|C_1\| \cos s)$$

$$= \varepsilon^2 (\|C_1\|^2\|C_2\|^2 - 2ae\|C_1\|\|C_2\|^2 \cos s + a^2e^2\|C_2\|^2 \cos^2 s)$$

Equating the coefficients of $\cos s$ gives $\|C_2\| = \frac{|L\|}{\varepsilon}$ and our identity gives

$$(\|C_1\|^2 - \|C_2\|^2) \cos^2 s + (\|C_2\|^2 + e^2a^2) = \|C_1\|^2 + a^2e^2 \cos^2 s$$

and hence $\|C_1\|^2 = \|C_2\|^2 + e^2a^2$ and $\|q\| = \|C_1\| - ea \cos s$. The equation for the energy, combined with (22) gives

$$-\frac{\varepsilon^2}{2} = \frac{1}{2} \frac{\varepsilon^2}{\|q\|^2} \frac{dq}{ds}^2 - \frac{\mu}{\|q\|}$$
or equivalently
\[ \varepsilon^2 \left( \frac{\| \dot{q} \|^2}{\| q \|^2} + \| \dot{q} \|^2 \right) = 2\mu \| q \| \]
Inserting we get
\[ \varepsilon^2 \left( \| C_1 \|^2 \sin^2 s + \| C_2 \|^2 \cos^2 s + \| C_1 \|^2 - 2ea\| C_1 \| \cos s + \varepsilon^2 a^2 \cos^2 s \right) = 2\mu (\| C_1 \| - ea \cos s) \]
or
\[ 2\| C_1 \|^2 - 2ea\| C_1 \| \cos s = 2 \frac{\mu}{\varepsilon^2} (\| C_1 \| - ea \cos s) \]
Therefore \( \| C_1 \| = \frac{\mu}{\varepsilon^2} = a \).

In the parametrization (24) all orbits have period \( 2\pi \). This is even true for the orbits with zero angular momentum (in this case \( C_1 \) and \( C_2 \) are linearly dependent), in contrast to the true Kepler flow where the point mass crashes into the origin. One says that the eccentric anomaly ”regularizes the collisions”.

**IV. The Hamiltonian point of view**

**Hamiltonian vectorfields**

**Definition.** Let \( H(q, p) \) be a differentiable function on \( \mathbb{R}^n \times \mathbb{R}^n \). The system of first order differential equations
\[ \dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \text{for } i = 1, \ldots, n \]
is called the *Hamiltonian system* associated to the Hamiltonian function \( H \).

Observe that the Hamiltonian vectorfield does not change when one adds a constant to the Hamiltonian.

**Example 1** (Kepler Hamiltonian)

The Hamiltonian system of the Hamiltonian
\[ K(q, p) = \frac{1}{2} \| p \|^2 - \frac{\mu}{\| q \|} \]
is
\[ \dot{q}_i = p_i, \quad \dot{p}_i = -\mu \frac{q_i}{\| q \|^3} \]
Differentiating the first equations once and inserting the second gives
\[ \ddot{q}_i = -\mu \frac{q_i}{\| q \|^3} \]
This is the basic equation (1) for the Kepler flow.
Example 2 (Harmonic Oscillator)

\[ H(q, p) = \frac{1}{2} (||p||^2 + ||q||^2) \]

The corresponding Hamiltonian system

\[ \dot{q}_i = p_i , \quad \dot{p}_i = -q_i \]

gives the second order system

\[ \ddot{q}_i = -q_i \quad \text{for } i = 1 \cdots n \]

This is the harmonic oscillator.

More generally, whenever one has the motion of a particle on \( \mathbb{R}^n \) under the influence of a potential \( V(q) \), the corresponding flow is described by the Hamiltonian

\[ H(q, p) = \frac{1}{2} ||p||^2 + V(q) \]

Indeed, the corresponding Hamiltonian system is

\[ \dot{q} = p , \quad \dot{p} = -\nabla V(q) \]

so that, as above, \( \ddot{q} = -\nabla V(q) \). In this description, \( q \) is the position variable and \( p \) is the momentum variable.

The relation between the Hamiltonian formalism and the Lagrange formalism is given by the Legendre transform, see for example [Arnold] ch.15.

Example 3 (Regularized Kepler Hamiltonian)

Let \( \varepsilon \neq 0 \). The system of the Hamiltonian

\[ \tilde{K}(q, p) = \frac{1}{2\varepsilon} ||q|| (||p||^2 + \varepsilon^2) - \frac{\mu}{||q||} \]

is

\[ \dot{q} = \frac{1}{\varepsilon} \frac{q}{||q||} p \]

\[ \dot{p} = -\frac{1}{2\varepsilon} \frac{q}{||q||} (||p||^2 + \varepsilon^2) = -\frac{1}{||q||} (\tilde{K}(q, p) + \frac{\mu}{\varepsilon}) q \]  \hspace{1cm} (27)

By the first equation, \( p = \frac{\varepsilon}{||q||} \dot{q} \) and

\[ \ddot{q} = \frac{1}{\varepsilon} \frac{d}{dt} ||q|| p + \frac{1}{\varepsilon} ||q|| \dot{p} = \frac{1}{\varepsilon} q \cdot \dot{q} \frac{\varepsilon}{||q||} \dot{q} - \frac{1}{\varepsilon} ||q|| \frac{1}{||q||} (\tilde{K}(q, p) + \frac{\mu}{\varepsilon}) q \]

\[ = \frac{1}{||q||} (q \cdot \dot{q}) \dot{q} - \frac{1}{\varepsilon ||q||} (\varepsilon \tilde{K}(q, p) + \mu) q \]

Its restriction to the level set \( \{ (q, p) \mid \tilde{K}(q, p) = 0 \} \) is the flow (23) of the Kepler problem parametrized by the eccentric anomaly.
Example 4 (Geodesics)

As in Appendix C, let $G(q) = (g_{ab}(q))_{a,b=1,...,n}$ be a Riemannian metric on an open subset $U$ of $\mathbb{R}^n$. We claim that the geodesics are described by the Hamiltonian

$$H(q,p) = \frac{1}{2} p^\top G^{-1}(q) p = \frac{1}{2} \sum_{a,b=1}^n g_{ab}(q) p_a p_b$$

Indeed, the associated Hamiltonian system is

$$\dot{q} = G^{-1} p$$

$$\dot{p}_a = -\frac{1}{2} p^\top \frac{\partial G^{-1}}{\partial q_a} p$$

for $a = 1, \cdots, n$

Since $G G^{-1} = 1$, we have $\frac{\partial G}{\partial q_a} G^{-1} + G \frac{\partial G^{-1}}{\partial q_a} = 0$ so that $\frac{\partial G^{-1}}{\partial q_a} = -G^{-1} \frac{\partial G}{\partial q_a} G^{-1}$. We insert this into the equation above and get the equations

$$G \dot{q} = p$$

$$\dot{p}_a = \frac{1}{2} (G p)^\top \frac{\partial G}{\partial q_a} (G p) = \frac{1}{2} q^\top \frac{\partial G}{\partial q_a} \dot{q}$$

for $a = 1, \cdots, n$

Differentiating the first equation gives $G \ddot{q} = \dot{p} - \dot{G} \dot{q}$ and therefore, as in ???

$$(G \ddot{q})_a = \dot{p}_a - \sum_{b=1}^n \dot{g}_{ab} \dot{q}_b$$

$$= \frac{1}{2} \sum_{b,c=1}^n \frac{\partial g_{bc}}{\partial q_a} \dot{q}_b \dot{q}_c - \sum_{b,c=1}^n \frac{\partial g_{ac}}{\partial q_b} \dot{q}_b \dot{q}_c$$

$$= \frac{1}{2} \left( \sum_{b,c=1}^n \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ac}}{\partial q_b} \right) \right) \dot{q}_b \dot{q}_c$$

and therefore the equation

$$\ddot{q}_a + \sum_{b,c=1}^n \Gamma_{bc}^a \dot{q}_b \dot{q}_c = 0$$

for geodesics; see equation (37) below.

A vectorfield $X$ on an open subset $U$ of $\mathbb{R}^m$ is a map that associates to point $x \in U$ a tangent vector $X(x)$. The tangent vector lies in $\mathbb{R}^m$, viewed as tangent space to $U$ in the point $x$. An integral curve to the $X$ is a differentiable curve $t \mapsto x(t)$ whose derivative at each point is given by the vectorfield, that is $\dot{x}(t) = X(x(t))$ for all $t$. In this sense, ordinary differential equations on $U$ are the same as vectorfields on $U$. Assume that the vectorfield has the components $X_1, \cdots, X_m$. That is

$$X(x) = (X_1(x), \cdots X_m(x))$$
Then we also write

\[ X = X_1 \frac{\partial}{\partial x_1} + \cdots + X_m \frac{\partial}{\partial x_m} \]

The reason for this notation is the following. Let \( \varphi(x) \) be any function on \( U \), and \( t \mapsto x(t) \) an integral curve. Then, by the chain rule

\[ \frac{d}{dt} \varphi(x(t)) = X_1 \frac{\partial \varphi}{\partial x_1} + \cdots + X_m \frac{\partial \varphi}{\partial x_m} \]

The resulting function is the directional derivative of \( \varphi \) in direction \( X \) and is denoted by \( X(\varphi) \) or \( L_X \varphi \).

If \( X \) is a vectorfield with continuous coefficients, then for each point \( x \in U \) there exists a unique integral curve \( t \mapsto X^t(x) \) with \( X^0(x) = x \). The map \( (t, x) \mapsto X^t(x) \) is called the flow of the vectorfield. Clearly, \( X^{s+t}(x) = X^s(X^t(x)) \). So, if all integral curves are defined for all times, the flow defines an action of the additive group of real numbers on \( U \).

For a Hamiltonian function \( H(q, p) \), the vectorfield associated to its Hamiltonian system is the Hamiltonian vectorfield

\[ X_H = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \]

The Hamiltonian system is the flow of this vectorfield.

**The Poisson Bracket**

**Definition.** The Poisson bracket of two differentiable functions \( F(q, p), G(q, p) \) on \( \mathbb{R}^n \times \mathbb{R}^n \) is

\[ \{F, G\} = \sum_{i=1}^{n} \left( \frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i} \frac{\partial F}{\partial p_i} \right) \]

Obviously In particular, \( \{F, G\} = -\{G, F\} \). We say that \( F \) and \( G \) are in involution if \( \{F, G\} = 0 \).

By the remarks of the previous section, \( \{F, G\} = X_G(F) = -X_F(G) \). In particular, \( F \) and \( G \) are in involution if and only if \( G \) is a conserved quantity for the Hamiltonian system with Hamiltonian \( F \). Observe that, for any Hamiltonian \( H \), the Poisson bracket \( \{H, H\} \) vanishes. So the Hamiltonian is always a conserved quantity for its flow.

**Example 5** As above, let \( K(q, p) = \frac{1}{2} ||p||^2 - \frac{\mu}{||q||} \) be the Kepler Hamiltonian. We know that the the components of the angular momentum vector \( L(q, p) = q \times p \) are conserved quantities for the associated flow. Therefore they must be in involution with the Hamiltonian.
Example 6  The components of the angular momentum are not in involution among each other. In fact

\[ \{L_1, L_2\} = L_3, \quad \{L_2, L_3\} = L_1, \quad \{L_3, L_1\} = L_2 \quad (28) \]

We verify the first equation:

\[
\begin{align*}
\{L_1, L_2\} &= \{q_2p_3 - q_3p_2, q_3p_1 - q_1p_3\} \\
&= \{q_2p_3, q_3p_1\} + \{q_3p_2, q_1p_3\} - \{q_3p_2, q_3p_1\} - \{q_2p_3, q_1p_3\} \\
&= -q_2p_1 + p_2q_1 = q_1p_2 - q_2p_1 = L_3
\end{align*}
\]

Example 7  Since the Laplace Lenz Runge vector

\[
A(q, p) = -\frac{q}{|q|} + \frac{1}{\mu} \cdot p \times L = \frac{1}{\mu} (||p||^2 - \frac{\mu}{||q||}) q - \frac{1}{\mu} (q \cdot p) p
\]

is a conserved quantity for the Kepler flow, its components are in involution with the Hamiltonian \(K\). On the other hand

\[
\begin{align*}
\{L_1, A_1\} &= \{L_2, A_2\} = \{L_3, A_3\} = 0 \\
\{L_1, A_2\} &= \{A_1, L_2\} = A_3 \\
\{L_2, A_3\} &= \{A_2, L_3\} = A_1 \\
\{L_3, A_1\} &= \{A_3, L_1\} = A_2
\end{align*}
\]

and

\[
\begin{align*}
\{A_1, A_2\} &= -\frac{2E}{\mu^2} L_3, \quad \{A_2, A_3\} = -\frac{2E}{\mu^2} L_1, \quad \{A_3, A_1\} = -\frac{2E}{\mu^2} L_2
\end{align*}
\]

Proof:  In the proof we shall use the general identity

\[
\{F, G_1G_2\} = G_1\{F, G_2\} + G_2\{F, G_1\}
\]

and the preliminary calculations that

\[
\begin{align*}
\{L_1, q_1\} &= \{L_2, q_2\} = \{L_3, q_3\} = 0 \\
\{L_1, q_2\} &= \{q_1, L_2\} = q_3, \quad \{L_2, q_3\} = \{q_2, L_3\} = q_1, \quad \{L_3, q_1\} = \{q_3, L_1\} = q_2 \\
\{L_1, p_1\} &= \{L_2, p_2\} = \{L_3, p_3\} = 0 \\
\{L_1, p_2\} &= \{p_1, L_2\} = p_3, \quad \{L_2, p_3\} = \{p_2, L_3\} = p_1, \quad \{L_3, p_1\} = \{p_3, L_1\} = p_2
\end{align*}
\]

that for \(i = 1, 2, 3\)

\[
\{L_i, \frac{1}{||q||}\} = 0, \quad \{(q \cdot p), q_i\} = -q_i, \quad \{(q \cdot p), p_i\} = p_i
\]
and that 
\[ \{(q \cdot p), \|p\|^2\} = 2\|p\|^2, \quad \{(q \cdot p), \frac{1}{\|q\|}\} = \frac{1}{\|q\|} \]

Using these identities we get

\[ \{L_1, A_1\} = \left\{ L_1, \frac{q_1}{\|q\|} + \frac{1}{\mu} (p_2 L_3 - p_3 L_2) \right\} \]
\[ = \frac{1}{\mu} \left[ p_2 \{L_1, L_3\} + L_3 \{L_1, p_2\} - p_3 \{L_1, L_2\} - L_2 \{L_1, p_3\} \right] \]
\[ = \frac{1}{\mu} \left[ -p_2 L_2 + p_3 L_3 - p_3 L_3 + p_2 L_2 \right] = 0 \]

\[ \{L_1, A_2\} = \left\{ L_1, \frac{q_2}{\|q\|} + \frac{1}{\mu} (p_3 L_1 - p_1 L_3) \right\} \]
\[ = \frac{q_2}{\|q\|} + \frac{1}{\mu} \left[ p_3 \{L_1, L_1\} + L_1 \{L_1, p_3\} - p_1 \{L_1, L_3\} - L_3 \{L_1, p_1\} \right] \]
\[ = \frac{q_2}{\|q\|} + \frac{1}{\mu} \left[ -p_2 L_1 + p_1 L_2 \right] = A_3 \]

\[ \{L_1, A_3\} = \left\{ L_1, \frac{q_3}{\|q\|} + \frac{1}{\mu} (p_1 L_2 - p_2 L_1) \right\} \]
\[ = -\frac{q_3}{\|q\|} + \frac{1}{\mu} \left[ p_1 \{L_1, L_2\} + L_2 \{L_1, p_1\} - p_2 \{L_1, L_1\} - L_1 \{L_1, p_2\} \right] \]
\[ = -\frac{q_3}{\|q\|} + \frac{1}{\mu} \left[ p_1 L_3 - p_3 L_1 \right] = -A_2 \]

The other Poisson brackets \(\{L_i, A_j\}\) are obtained by cyclic permutation. Using the second representation of the Laplace Lenz Runge vector given above, one calculates

\[ \mu^2 \{A_1, A_2\} = \left\{ (\|p\|^2 - \frac{\mu}{\|q\|}) q_1 - (q \cdot p) p_1, \ (\|p\|^2 - \frac{\mu}{\|q\|}) q_2 - (q \cdot p) p_2 \right\} \]
\[ = \left\{ (\|p\|^2 - \frac{\mu}{\|q\|}) q_1, \ (\|p\|^2 - \frac{\mu}{\|q\|}) q_2 \right\} + \{(q \cdot p) p_1, \ (q \cdot p) p_2\} \]
\[ - \left\{ (\|p\|^2 - \frac{\mu}{\|q\|}) q_1, \ (q \cdot p) p_2 \right\} - \{(q \cdot p) p_1, \ (\|p\|^2 - \frac{\mu}{\|q\|}) q_2\} \]
\[ = (\|p\|^2 - \frac{\mu}{\|q\|}) \left[ q_1 \{(\|p\|^2 - \frac{\mu}{\|q\|}), q_2\} + q_2 \{(\|p\|^2 - \frac{\mu}{\|q\|}), q_1\} \right] \]
\[ + (q \cdot p) \left[ p_1 \{(q \cdot p), p_2\} + p_2 \{(q \cdot p), p_1\} \right] \]
\[ - q_1 p_2 \{(\|p\|^2 - \frac{\mu}{\|q\|}), (q \cdot p)\} - (\|p\|^2 - \frac{\mu}{\|q\|}) p_2 \left\{ q_1, \ (q \cdot p) \right\} \]
\[ - (q \cdot p) q_1 \{(\|p\|^2 - \frac{\mu}{\|q\|}), p_2\} \]
\[ - p_1 q_2 \{(q \cdot p), (\|p\|^2 - \frac{\mu}{\|q\|})\} - (\|p\|^2 - \frac{\mu}{\|q\|}) p_1 \left\{ (q \cdot p), q_2 \right\} \]
\[ - (q \cdot p) q_2 \left\{ p_1, (\|p\|^2 - \frac{\mu}{\|q\|}) \right\} \]
\[ = (\|p\|^2 - \frac{\mu}{\|q\|}) \left[ q_1 \{(\|p\|^2, q_2\} + q_2 \{(q_1, \|p\|^2)\} \right] + (q \cdot p) [p_1 p_2 - p_2 p_1] \]
\[ + q_1 p_2 \left( 2\|p\|^2 - \frac{\mu}{\|q\|} \right) - (\|p\|^2 - \frac{\mu}{\|q\|}) p_2 q_1 + (q \cdot p) q_1 \left\{ \frac{\mu}{\|q\|}, p_2 \right\} \]
\[ - p_1 q_2 \left( 2\|p\|^2 - \frac{\mu}{\|q\|} \right) + (\|p\|^2 - \frac{\mu}{\|q\|}) p_1 q_2 + (q \cdot p) q_2 \left\{ p_1, \frac{\mu}{\|q\|} \right\} \]
\[ = (\|p\|^2 - \frac{\mu}{\|q\|}) \left[ -2 q_1 p_2 + 2 q_2 p_1 - p_2 q_1 + q_2 p_1 \right] \]
\[ + (2 \|p\|^2 - \frac{\mu}{\|q\|}) (q_1 p_2 - p_1 q_2) - (q \cdot p) q_1 \frac{\mu q_2}{\|q\|^2} + (q \cdot p) q_1 \frac{\mu p_2}{\|q\|^2} \]
\[ = - \left\{ (\|p\|^2 - 2 \frac{\mu}{\|q\|}) q_1 p_2 - q_2 p_1 \right\} = -2 E L_3 \]

The other brackets \(\{A_2, A_3\}\) and \(\{A_3, A_1\}\) follow by cyclic permutation.
The Lie bracket of two vectorfields \( X = \sum_{i=1}^{m} X_i \frac{\partial}{\partial x_i} \) and \( Y = \sum_{i=1}^{m} Y_i \frac{\partial}{\partial x_i} \) is defined as

\[
[X, Y] = \sum_{i=1}^{m} \sum_{j=1}^{m} (Y_j \frac{\partial X_i}{\partial x_j} - X_j \frac{\partial Y_i}{\partial x_j}) \frac{\partial}{\partial x_i}
\]

For any function \( \varphi \)

\[
[X, Y](\varphi)(x) = \frac{\partial^2}{\partial s \partial t} \left[ \varphi(X^t(Y^s(x))) - \varphi(Y^s(X^t(x))) \right] \bigg|_{s=t=0}
\]

See [Arnold], ch. 39. The Hamiltonian vectorfield associated to the Poisson bracket of two functions \( F \) and \( G \) is the Lie bracket of the Hamiltonian vectorfields associated to \( F \) and \( G \) respectively. That is

\[
X_{\{F,G\}} = [X_F, X_G]
\]

Indeed,

\[
[X_F, X_G] = \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \frac{\partial G}{\partial p_j} \frac{\partial^2 F}{\partial p_i \partial q_j} - \frac{\partial F}{\partial q_j} \frac{\partial^2 G}{\partial p_i \partial p_j} - \frac{\partial G}{\partial q_j} \frac{\partial^2 F}{\partial q_i \partial p_j} + \frac{\partial F}{\partial q_j} \frac{\partial^2 G}{\partial q_i \partial p_j} \right) \frac{\partial}{\partial q_i} 
\]

\[
+ \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \frac{\partial G}{\partial q_j} \frac{\partial^2 F}{\partial q_i \partial p_j} - \frac{\partial F}{\partial p_j} \frac{\partial^2 G}{\partial q_i \partial q_j} - \frac{\partial G}{\partial p_j} \frac{\partial^2 F}{\partial p_i \partial q_j} + \frac{\partial F}{\partial p_j} \frac{\partial^2 G}{\partial p_i \partial q_j} \right) \frac{\partial}{\partial p_i}
\]

\[
= \sum_{i=1}^{n} \frac{\partial}{\partial p_i} \sum_{j=1}^{n} \left( \frac{\partial G}{\partial q_j} \frac{\partial^2 F}{\partial q_i \partial p_j} - \frac{\partial F}{\partial p_j} \frac{\partial^2 G}{\partial q_i \partial q_j} \right) \frac{\partial}{\partial q_i} - \sum_{i=1}^{n} \frac{\partial}{\partial q_i} \sum_{j=1}^{n} \left( \frac{\partial G}{\partial q_j} \frac{\partial^2 F}{\partial q_i \partial p_j} - \frac{\partial F}{\partial p_j} \frac{\partial^2 G}{\partial q_i \partial q_j} \right) \frac{\partial}{\partial p_i}
\]

\[
= X_{\{F,G\}}
\]

Finally, we remark that both the Poisson bracket and the Lie bracket fulfil the Jacobi identity, that is

\[
\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0
\]

\[
[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0
\]

For \( C^2 \) functions \( F, G, H \) and \( C^2 \) vector fields \( X, Y, Z \)

**Infinitesimal Symmetries**

Whenever \( q(t) \) is a solution of the basic equation (1) for the Kepler problem and \( R \in SO(3) \) is a rotation around an axis through the origin, then \( R q(t) \) is again a solution of (1). In other words, the Kepler problem has \( SO(3) \)–symmetry. This is reflected by the fact that the Kepler Hamiltonian is \( SO(3) \)–invariant:

\[
K(Rq, Rp) = K(q, p) \quad \text{for all} \ R \in SO(3)
\]

In particular, the Kepler Hamiltonian \( K \) is a conserved quantity for the flows

\[
(t, (q, p)) \mapsto (R_t(t)q, R_t(t)p)
\]
associated to the rotation $R_i(t)$ around the $i^{th}$ axis ($i = 1, 2, 3$) with angle $t$. The associated vectorfields $X_i$ are given by

$$\dot{q} = \frac{d}{dt} R_i(t) q \bigg|_{t=0} = E_i q$$

$$\dot{p} = \frac{d}{dt} R_i(t) p \bigg|_{t=0} = E_i p$$

where $E_i = \frac{dR_i(t)}{dt} \bigg|_{t=0}$. See Appendix F. This construction is compatible with the Lie brackets defined on the Lie algebra $so(3)$ (see appendix F) and for vectorfields. For example, $[X_1, X_2]$ is the vectorfield, which, at the point $(q, p)$, takes the value

$$\frac{\partial^2}{\partial s \partial t} \left[ (R_1(s)R_2(t) q, R_1(s)R_2(t) p) - (R_2(t)R_1(s) q, R_2(t)R_1(s) p) \right]_{s,t=0}$$

$$= \frac{\partial}{\partial t} \left[ (E_1 R_2(t) q, E_1 R_2(t) p) - (E_2 R_1(t) q, E_2 R_1(t) p) \right]_{t=0}$$

$$= ([E_1, E_2] q, [E_1, E_2] p)$$

$$= X_3(q, p)$$

Similarly, $[X_2, X_3] = X_1$ and $[X_3, X_1] = X_2$. Since the Kepler Hamiltonian $K$ is a conserved quantity for the flows of the vectorfields $X_1, X_2, X_3$, we also have

$$[X_i, X_K] = 0 \text{ for } i = 1, 2, 3$$

In fact, the vectorfields $X_i$ are already known to us. They are the vectorfields associated to the components $L_i$ of the angular momentum:

$$X_i = X_{L_i} \text{ for } i = 1, 2, 3$$

This is trivial to verify.

In the case of negative energy $E = -\frac{\mu^2}{\varepsilon^2}$, by Example 7

$$\{L_1 \pm \frac{\mu}{\varepsilon} A_1, L_2 \pm \frac{\mu}{\varepsilon} A_2\} = \{L_1, L_2\} \pm \frac{\mu}{\varepsilon^2} \left(\{L_1, A_2\} + \{A_1, L_2\}\right) + \frac{\mu^2}{\varepsilon^2} \{A_1, A_2\}$$

$$= \left(1 + \frac{\mu^2}{\varepsilon^2}(-\frac{2\mu}{\mu^2})\right) L_3 \pm 2\frac{\mu}{\varepsilon} A_3 = 2[L_3 \pm \frac{2\mu}{\varepsilon} A_3]$$

Thus, for each fixed choice of $\pm$, the Poisson brackets of the functions

$$B_i^\pm = \frac{1}{\sqrt{2}} [L_i \pm \frac{\mu}{\varepsilon} A_i]$$

fulfil the same relation as the generators $E_1, E_2, E_3$ of the Lie algebra $so(3)$. Precisely,

$$\{B_1^+, B_2^+\} = B_3^+, \quad \{B_2^+, B_3^+\} = B_1^+, \quad \{B_3^+, B_1^+\} = B_2^+$$
Also, they are in involution with the Kepler Hamiltonian $K$. Furthermore
\[
\{B_i^+, B_j^-\} = 0 \quad \text{for } i, j = 1, 2, 3
\]

For the planar problem the relations
\[
\{\frac{\mu}{\varepsilon} A_1, \frac{\mu}{\varepsilon} A_2\} = L_3 \quad \{\frac{\mu}{\varepsilon} A_2, L_3\} = \frac{\mu}{\varepsilon} A_1 \quad \{L_3, \frac{\mu}{\varepsilon} A_1\} = \frac{\mu}{\varepsilon} A_2
\]
suggest an $SO(3)$ symmetry.

**Explicit Symmetries of the two dimensional Kepler problem**

Let $\tilde{K}$ be the Hamiltonian of the regularized Kepler problem of Example 3. We consider the case of dimension two and think of $q$ and $p$ as complex variables. We restrict ourselves to the case of negative energy. By scaling, we can assume that $\varepsilon = 1$. Then
\[
\tilde{K}(q, p) = \frac{1}{2} |q|(|p|^2 + 1) - \mu
\]

For an invertible complex $2 \times 2$ matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $(q, p) \in \mathbb{C} \times \mathbb{C}$ set
\[
A \cdot (q, p) = \left( (\bar{c}p + \bar{d})^2 \cdot q, \frac{ap + b}{cp + d} \right)
\]
whenever it is defined. Observe that the second component is the standard action of $GL(2, \mathbb{C})$ on the complex plane by fractional linear transformations. If $\det A = 1$, the factor $(\bar{c}p + \bar{d})^2$ is the complex conjugate of the inverse of the derivative of the fractional linear transformation $p \mapsto \frac{ap + b}{cp + d}$. Therefore, for $A_1, A_2 \in SL(2, \mathbb{C})$
\[
(A_1 A_2) \cdot (q, p) = A_1 \cdot (A_2 \cdot (q, p))
\]
whenever defined.

**Theorem 8** Let $(q, p) \in \mathbb{C} \times \mathbb{C}$. Then for all $A \in SU(2)$

(i) $\tilde{K}(A \cdot (q, p)) = \tilde{K}(q, p)$, whenever defined.

(ii) If $(q(t), p(t))$ solves the Kepler equations (27) with respect to the eccentric anomaly, that is if
\[
\dot{q} = |q| p \quad , \quad \dot{p} = -\frac{1}{2} \frac{\mu}{|q|^4}(|p|^2 + 1) = -\frac{1}{|q|^2} (\tilde{K}(q, p) + \mu) q
\]
then $(P(t), Q(t)) = A \cdot (p(t), q(t))$ also fulfills (27).
Proof: Let \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2) \).

(i)  
\[ \tilde{K}(A \cdot (q, p)) + \mu = \frac{1}{2}|c\bar{p} + \bar{d}|^2 |q| (|\frac{ap + b}{c\bar{p} + d}|^2 + 1) = \frac{1}{2}|q| (|ap + b|^2 + |cp + d|^2) \]
\[ = |q| (|p_1|^2 + |p_2|^2) \]
where
\[ \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = A \cdot \begin{pmatrix} p \\ 1 \end{pmatrix} \]

As \( A \in SU(2) \), \( |p_1|^2 + |p_2|^2 = |p|^2 + 1 \).

(ii) By part (i)
\[ \dot{P} + \frac{1}{|Q|^2} (\tilde{K}(Q, P) + \mu)Q = \frac{1}{(cp + d)^2} \dot{p} + \frac{1}{|cp + d|^2 |q|^2} (\tilde{K}(q, p) + \mu)(c\bar{p} + \bar{d})^2 q \]
\[ = \frac{1}{(cp + d)^2} (\dot{p} + \frac{1}{|q|^2} (\tilde{K}(q, p) + \mu) q) = 0 \]
\[ \dot{Q} - |Q|P = 2c(c\bar{p} + \bar{d})\dot{p} q + (c\bar{p} + \bar{d})^2 \dot{q} - |c\bar{p} + \bar{d}|^2 |q| \frac{ap + b}{cp + d} \]
\[ = (c\bar{p} + \bar{d}) (2c\dot{p} q + (c\bar{p} + \bar{d}) \dot{q} - |q|(ap + b)) \]
\[ = (c\bar{p} + \bar{d}) (b\frac{1}{|q|} (|p|^2 + 1) q + (-b\bar{p} + a)|q|p - |q|(ap + b)) \]
\[ = 0 \]
since \( c = -\bar{b} \) and \( d = \bar{a} \).

The \( SU(2) \) action described above preserves the symplectic form \( \operatorname{Re} d\bar{p} \wedge dq \). This can be used to deduce the second part of the theorem from its first part.

\[ \text{24} \]
Appendix A: Conic sections

Conic sections are the nonsingular curves that are obtained by intersecting a quadratic cone with a plane. The relation with the focal description given after the statement of Kepler’s laws can be seen using the ”Dandelin spheres”, see [Knörrer] 4.7.3. Also, conic sections are the zero sets of quadratic polynomials in two variables that do not contain a line. We discuss the most relevant properties of the conic sections

Ellipses

Consider the ellipse consisting of all points $P$ for which the sum of the distances $\|P - F\|$ and $\|P - F'\|$ is equal to $2a$. The eccentricity $e$ of the ellipse is defined by

$$\|F - F'\| = 2 e a$$

Clearly, $0 \leq e < 1$. Let $M$ be the midpoint between the two foci (the center of the ellipse). The line through the two foci is called the major axis of the ellipse. The two points of the ellipse on the major axis have each distance $a$ from the center. The foci both have distance $e a$ from the center. The line through $M$ perpendicular to the major axis is called the minor axis. It is the perpendicular bisector of the two foci. The points of the ellipse on the minor axis have distance $a$ from both foci. It follows from Pythagoras’ theorem that the distance of these points from the center is

$$b = a \sqrt{1 - e^2}$$

If one introduces Cartesian coordinates $x, y$ centered at $M$ with the major axis as $x$–axis and the minor axis as $y$–axis then the equation of the ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

(29)

For a proof see [Knörrer], Satz 4.3. The area enclosed by the ellipsoid is equal to $\pi a b$; see for example [Zorich], 6.4.3, Example 5.

An important consequence of the focal description of the ellipse is the following: For each point $q$ of the ellipse, the lines joining $q$ to the origin and $q$ to the other focus $f$ form opposite equal angles with the tangent line of the ellipse. For a proof, see [Knörrer], Satz 4.6.

Another useful description of the ellipse is the following. The line perpendicular to the major axis that has distance $\frac{1}{e} a$ to the center and lies on the same side of the center as the focus $F$ is called the directrix with respect to $F$. One can show that the ellipse is the set of points $P$ for which the ratio of the distance from $P$ to $F$ to the distance from
$P$ to the directrix is equal to the eccentricity $e$. To see this, note that a point $(x, y)$ fulfils the directrix condition with focus $(ea, 0)$ if and only if

$$(x - ea)^2 + y^2 = e^2(x - \frac{a}{e})^2 \iff \frac{x^2}{a^2} + \frac{y^2}{(1-e^2)a^2} = 1$$

This description directly gives the equation of the ellipse in polar coordinates\(^*\) $(r, \varphi)$ centered at the focus $F$. Choose the angular variable $\varphi$ such that $\varphi = 0$ corresponds to the ray starting at $F$ in the direction away from the center\(^**\). If a point $P$ has polar coordinates $(r, \varphi)$ then its distance to $F$ is $r$. Since the distance from $F$ to the directrix is $(\frac{1}{e} - e)a$, the distance from $P$ to the directrix is $(\frac{1}{e} - e)a - r \cos \varphi$. Thus the equation of the ellipse is

$$r = e \left((\frac{1}{e} - e)a - r \cos \varphi\right)$$

or, equivalently

$$r = \frac{l}{1 + e \cos \varphi} \quad (30)$$

where $l = (1 - e^2)a$. The quantity $l$ is called the parameter of the ellipse. The angle $\varphi$ is called the true anomaly of a point on the ellipse with respect to the focus $F$. Observe that $\varphi = 0$ corresponds to the point of the ellipse closest to $F$. In celestial mechanics, this point is called the perihelion.

Formula (30) is a parametrization of the ellipse, giving the distance $r$ to the focus as a function of the true anomaly $\varphi$. Equation (29) suggests the parametrization

$$x = a \cos s, \quad y = b \sin s$$

of the ellipse. The parameter $s$ is called the eccentric anomaly. To get the relation between the eccentric anomaly and the true anomaly with respect to the focus $F = (ea, 0)$, let $(x, y)$ be a point of the ellipse with true anomaly $\varphi$ and eccentric anomaly $s$. Then

$$x = a \cos s$$
$$y = b \sin s = a \sqrt{1 - e^2} \sin s$$

and

$$x = r \cos \varphi + ea = \frac{l}{1 + e \cos \varphi} \cos \varphi + ea = \frac{(1-e^2)a}{1+e \cos \varphi} \cos \varphi + ea$$
$$y = r \sin \varphi = \frac{(1-e^2)a}{1+e \cos \varphi} \sin \varphi$$

Comparing these two representation, we obtain

$$\cos s = e + \frac{1-e^2}{1+e \cos \varphi} \cos \varphi, \quad \sin s = \frac{\sqrt{1-e^2}}{1+e \cos \varphi} \sin \varphi$$

\(^*\) For another derivation, see [Knörrer] 4.7.1
\(^**\) This is the direction from the focus $F$ to the closest point on the ellipse.
We differentiate the first equation with respect to $\varphi$ and insert the second to get

$$-\frac{ds}{d\varphi} \sin s = -\frac{1-e^2}{1+e \cos \varphi} \sin \varphi + \frac{1-e^2}{(1+e \cos \varphi)^2} e \sin \varphi \cos \varphi$$

$$= -\sqrt{1-e^2} \sin s + \sqrt{1-e^2} \frac{e \cos \varphi}{1+e \cos \varphi} \sin s$$

$$= -\sqrt{1-e^2} \sin s \cdot \frac{1}{1+e \cos \varphi}$$

Dividing by $\sin s$ and using (30) gives

$$\frac{ds}{d\varphi} = \sqrt{1-e^2} \frac{r}{1} \quad (31)$$

We also need the expression of $r$ in terms of $s$. Since

$$r^2 = (x-ea)^2 + y^2 = a^2(\cos s - e)^2 + (1-e^2)a^2 \sin^2 s$$

$$= a^2(\cos^2 s - 2e \cos s + e^2 + (1-e^2) - \cos^2 s + 2e \cos s)$$

$$= a^2(1-e \cos s)^2$$

we have

$$r = a (1 - e \cos s) \quad (32)$$

**Hyperboli**

Consider now the hyperbola consisting of all points $P$ for which the difference of the distances $\|P - F\|$ and $\|P - F'\|$ has absolute value equal to $2a$. The *eccentricity* $e$ of the hyperbola is defined by

$$\|F - F'\| = 2e a$$

Clearly, $e > 1$. As before, let $M$ be the midpoint between the two foci (called the center). The line through the two foci is called the major axis of the hyperbola. The two points of the hyperbola on the major axis have each distance $a$ from the center. The foci both have distance $ea$ from the center. If one introduces Cartesian coordinates $x, y$ centered at $M$ with the major axis as $x$–axis then the equation of the ellipse is

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

where $b = a \sqrt{e^2 - 1}$.

The description using a directrix is almost identical to the one for ellipses. The line perpendicular to the major axis that has distance $\frac{1}{e} a$ to the center and lies on the same side of the center as the focus $F$ is called the *directrix* with respect to $F$. The hyperbola is the set of points $P$ for which the ratio of the distance from $P$ to $F$ to the distance from $P$
to the directrix is equal to the eccentricity \( e \). In polar coordinates centered at \( F \) one gets as equation for the hyperbola

\[
r = \frac{(e^2 - 1) a}{1 + e \cos \varphi}
\]

(33)

Here, the angular variable \( \varphi \) has been chosen such that \( \varphi = 0 \) corresponds to the ray starting at \( F \) in the direction to the center\(^*\).

**Parabola**

Finally consider the parabola consisting of all points that have equal distance from the focus \( F \) and the line \( g \) (which we call the *directrix* of the parabola). Let \( l \) be the distance from \( F \) to \( g \). Furthermore let \( \ell \) be the line perpendicular to \( g \) through \( F \) and \( M \) the midpoint of its segment between \( F \) and \( g \). It has distance \( l/2 \) both from \( F \) and \( g \).

**FIGURE**

If one introduces Cartesian coordinates \( x, y \) centered at \( M \) with the \( x \)--axis being \( \ell \) oriented in direction of \( M \) then the equation of the parabola is

\[
y^2 + 2lx = 0
\]

Again we choose polar coordinates centered at \( F \) such that \( \varphi = 0 \) corresponds to the ray from \( F \) to \( M \). As before one sees that the equation of the parabola is

\[
r = \frac{l}{1 + \cos \varphi}
\]

(34)

**Appendix B: The two body problem**

The basic equation of motion (1) also governs the general two body problem. Here we consider two point masses with masses \( m_1, m_2 \) and time dependent positions \( r_1(t), r_2(t) \). In this situation, Newton’s laws give

\[
m_1 \ddot{r}_1(t) = G m_1 m_2 \frac{r_2(t) - r_1(t)}{|r_2(t) - r_1(t)|^3}
\]

\[
m_2 \ddot{r}_2(t) = G m_1 m_2 \frac{r_1(t) - r_2(t)}{|r_1(t) - r_2(t)|^3}
\]

(35)

Denote by \( R(t) = \frac{1}{m_1 + m_2} (m_1 r_1(t) + m_2 r_2(t)) \) the center of gravity. Adding the two equations of (35) gives \( \ddot{R}(t) = 0 \). That is, the center of gravity moves with constant speed. Also, set \( q(t) = r_2(t) - r_1(t) \). (35) gives

\[
\ddot{r}_1(t) = G m_2 \frac{q(t)}{|q(t)|^3}, \quad \ddot{r}_2(t) = -G m_1 \frac{q(t)}{|q(t)|^3}
\]

\(^*\) Again, this is the direction from the focus \( F \) to the closest point on the ellipse.
so that
\[ \ddot{q}(t) = -\mu \frac{q(t)}{|q(t)|^3} \]
with \( \mu = G(m_1 + m_2) \). This is the basic equation (1).

**Appendix C: The Lagrangian formalism**

The Euler Lagrange equations associated to a function \( \mathcal{L}(t, q, p) \) is
\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \bigg|_{p=\dot{q}} \right) - \frac{\partial \mathcal{L}}{\partial q_i} \bigg|_{p=\dot{q}} = 0 \tag{36}
\]
Conventionally one views \( \mathcal{L} \) as a function of the variables \( t, q, \dot{q} \) and uses the shorthand formulation
\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0
\]
The Lagrange function for the Kepler problem is the difference between the kinetic and the potential energy
\[
\mathcal{L}_K(q, p) = \frac{1}{2} \|p\|^2 + \frac{\mu}{\|q\|}
\]
It is independent of \( t \). Then
\[
\frac{\partial \mathcal{L}_K}{\partial p_i} = p_i, \quad \frac{d}{dt} \left( \frac{\partial \mathcal{L}_K}{\partial p_i} \bigg|_{p=\dot{q}} \right) = \ddot{q}_i, \quad \frac{\partial \mathcal{L}_K}{\partial q_i} = -\mu \frac{q_i}{\|q\|^3}
\]
so that the associated Euler Lagrange equation are
\[ \ddot{q}_i + \mu \frac{q_i}{\|q\|^3} = 0 \]
This is the basic equation of motion (1).

The Euler Lagrange are related to the following variational problem. Fix points \( q', q'' \in \mathbb{R}^3 \) and times \( t' < t'' \). For each twice differentiable curve
\[
\gamma : [t', t''] \to \mathbb{R}^3 \quad t \mapsto q(t) \quad \text{with} \quad q(t') = p', \quad q(t'') = q''
\]
define the action
\[
\Phi(\gamma) = \int_{t'}^{t''} \mathcal{L}(t, q(t), \dot{q}(t)) \, dt
\]
If \( \gamma \) minimizes the action \( \Phi(\gamma) \) among all curves as above, then the Euler Lagrange equations hold. See [Arnold, Section 13]. The fact that the equations of motion of an autonomous mechanical system are minimizers of the action “(kinetic energy) - (potential energy)” is called the principle of least action (or principle of Maupertuis).
Another instance of a variational problem is the construction of geodesics. Let $U$ be an open subset of $\mathbb{R}^n$. Assume that for each point $q \in U$ one is given a positive definite symmetric bilinear form on $\mathbb{R}^n$ (a Riemannian metric). It is represented by a positive symmetric $n \times n$ matrix $G(q) = (g_{ab}(q))_{a,b=1,\ldots,n}$. For simplicity we assume that the coefficients $g_{ab}(q)$ are $C^\infty$ functions of $q$. The length of a curve $\gamma : [t', t''] \to U, \ t \mapsto q(t)$ with respect to the Riemannian metric is by definition

$$\text{length}(\gamma) = \int_{t'}^{t''} \left( \dot{q}(t)^\top G(q(t)) \dot{q}(t) \right)^{1/2} \, dt$$

By definition geodesics are “locally shortest connections”. That is, a curve $t \mapsto q(t)$ is a geodesic if and only if, for each $t$, there is $\epsilon > 0$ such that for all $s$ with $t < s < t + \epsilon$, the curve is a shortest connection between the points $q(t)$ and $q(s)$. It follows that geodesics fulfil the Euler Lagrange equations (36) with

$$\mathcal{L}(q, p) = \sqrt{p^\top G(q) p} = \left( \sum_{a,b=1}^{n} g_{ab}(q) p_a p_b \right)^{1/2}$$

Obviously the length of a curve does not change under reparametrization. In particular, the critical point for the variational problem is degenerate. To normalize the situation, we look for minimizing curves that are parametrized with constant speed (different from zero). That is, curves which are parametrized in such a way that $\dot{q}(t)^\top G(q(t)) \dot{q}(t) = \text{const}$ for all $t$. If one has a minimizer for the variational problem that is parametrized by arclength, then it also fulfills the equation

$$\frac{d}{dt} \left( \sum_{b=1}^{n} g_{ab}(q) \dot{q}_b \right) - \sum_{b,c=1}^{n} \frac{\partial^2 \mathcal{L}}{\partial q_b \partial q_c} \left. \right|_{p=\dot{q}} = 0$$

Indeed, multiplying the Euler Lagrange equation (36) by $\text{const} = \mathcal{L}(q, \dot{q})$, we get

$$\text{const} \left[ \frac{d}{dt} \left( \mathcal{L}(q, \dot{q}) \left. \frac{\partial \mathcal{L}}{\partial p_i} \right|_{p=\dot{q}} \right) - \mathcal{L}(q, \dot{q}) \left. \frac{\partial \mathcal{L}}{\partial q_i} \right|_{p=\dot{q}} \right] = 0$$

which implies (37).

Equation (37) is simpler than the Euler Lagrange equations associated to the original Lagrange function. We now derive the equation of motion associated to (37). To see that we really get geodesics, we then have to verify that the resulting curves are indeed parametrized with constant speed. Equation (37) gives

$$2 \frac{d}{dt} \left( \sum_{b=1}^{n} g_{ab}(q) \dot{q}_b \right) - \sum_{b,c=1}^{n} \frac{\partial g_{ab}}{\partial q_c} \dot{q}_b \dot{q}_c = 0 \quad \text{for } a = 1, \ldots, n$$
Since \( \frac{d}{dt} g_{ab}(q) = \sum_{c=1}^{n} \frac{\partial g_{ab}}{\partial q_c} \dot{q}_c \), (37) is equivalent to

\[
\sum_{b=1}^{n} g_{ab} \ddot{q}_b = \frac{1}{2} \sum_{b,c=1}^{n} \frac{\partial g_{bc}}{\partial q_a} \dot{q}_b \dot{q}_c - \sum_{b,c=1}^{n} \frac{\partial g_{ab}}{\partial q_c} \dot{q}_b \dot{q}_c
\]

\[
= \frac{1}{2} \left( \sum_{b,c=1}^{n} \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \right) \right) \dot{q}_b \dot{q}_c
\]

for \( a = 1, \ldots, n \). Here we used that \( \frac{\partial g_{ab}}{\partial q_c} = \frac{\partial g_{ba}}{\partial q_c} \) and exchanged the summation indices \( b \) and \( c \). The equations above state that the \( a \)-component of the vector \( \dot{G} \ddot{q} \) is equal to the vector with entries \( \frac{1}{2} \left( \sum_{b,c=1}^{n} \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \right) \right) \dot{q}_b \dot{q}_c \). We denote by \( g^{ab} \) the entries of the inverse matrix \( G^{-1} \). Then the equations are equivalent to

\[
\ddot{q}_a + \sum_{b,c=1}^{n} \Gamma^a_{bc} \dot{q}_b \dot{q}_c = 0 \tag{38}
\]

where \( \Gamma^a_{bc} \) are the Christoffel symbols

\[
\Gamma^a_{bc} = \frac{1}{2} \sum_{d=1}^{n} g^{ad} \sum_{b,c=1}^{n} \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \right)
\]

To verify that (38) really describes geodesics, we still have to verify that solutions of (38) are curves that are parametrized by constant speed. So let \( t \mapsto q(t) \) be a solution of (38). Reversing the calculation above, we see that shows \( 2 G \ddot{q} \) is the vector with entries \( \sum_{b,c=1}^{n} \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \right) \dot{q}_b \dot{q}_c \). Therefore

\[
\frac{d}{dt} \dot{G}(q) \dot{q} = 2 \dot{q} G(q) \ddot{q} + \dot{q} \ddot{G} \dot{q}
\]

\[
= \sum_{a,b,c=1}^{n} \left( \frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \right) \dot{q}_a \dot{q}_b \dot{q}_c + \sum_{a,b,c=1}^{n} \frac{\partial g_{bc}}{\partial q_a} \dot{q}_a \dot{q}_b \dot{q}_c = 0
\]

and \( \dot{q} G(q) \dot{q} \) is indeed constant.

**Appendix E: The Kepler equation**

The Kepler equation (21)

\[
\tau = s - e \sin s \tag{39}
\]

where \( \tau = \frac{\sqrt{\mu}}{a^{3/2}} (t - t_0) \) is the “mean anomaly” can be solved for \( s \) as a function of \( \tau \) (and thus of the time \( t \)) using Fourier series and Bessel functions. Since \( e < 1 \), the right hand side is a strictly monotonically increasing function of \( s \), and for \( s = n \pi, n \in \mathbb{Z} \) one has
\[ \tau = s. \text{ Therefore we can write } s \text{ as a } C^\infty \text{ function } s(\tau) \text{ of } \tau, \text{ and } \sin s(\tau) \text{ is an odd, } 2\pi \text{ periodic function of } \tau. \text{ It has a Fourier series} \]
\[ \sin s(\tau) = \sum_{n=1}^{\infty} a_n \sin n\tau \]

with Fourier coefficients
\[ a_n = \frac{2}{\pi} \int_0^\pi (\sin s(\tau))(\sin n\tau) \, d\tau \]

Partial integration gives
\[ a_n = -\frac{2}{n\pi} (\sin s(\tau))(\cos n\tau) \bigg|_{\tau=\pi}^{\tau=0} + \frac{2}{n\pi} \int_0^\pi (\frac{d}{d\tau} \sin s(\tau))(\cos n\tau) \, d\tau \]

The first term vanishes. For the second term, observe that by (39)
\[ \frac{d}{d\tau} \sin s(\tau) = \frac{1}{e} \frac{d}{d\tau} (s(\tau) - \tau) = \frac{1}{e} \left( \frac{ds}{d\tau} - 1 \right) \]

Since \( \int_0^\pi (\cos n\tau) \, d\tau = 0 \), we get by the substitution rule and (39)
\[ a_n = \frac{2}{n\pi} \int_0^\pi (\cos n\tau) \frac{ds}{d\tau} \, d\tau = \frac{2}{n\pi} \int_0^\pi (\cos n\tau(s)) \, ds = \frac{2}{n\pi} \int_0^\pi \cos (n(s - e \sin s)) \, ds \quad (40) \]

The \( n \text{th} \) Bessel function is defined as
\[ J_n(x) = \frac{1}{\pi} \int_0^\pi \cos (x \sin t - nt) \, dt \]

See [Watson]. So
\[ a_n = \frac{2}{n\pi} \int_0^\pi \cos \left( (ne) \sin s - ns \right) \, ds = \frac{2}{ne} J_n(ne) \]

Consequently we get from (39) the “Kapteyn” series
\[ s = \tau + \sum_{n=1}^{\infty} \frac{2}{n} J_n(ne) \sin n\tau \]

**Appendix F: The group \( SO(3) \) and its Lie algebra \( sO(3) \)**

By definition, \( SO(3) \) is the group of all real \( 3 \times 3 \) matrices \( R \) with determinant 1 for which \( R^T R = I \). One can show that it consists of all rotations around an axis in \( \mathbb{R}^3 \) around an axis through the origin. It is generated by the one parameter subgroups

\[ R_1(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}, \quad R_2(t) = \begin{pmatrix} \cos t & 0 & \sin t \\ 0 & 1 & 0 \\ -\sin t & 0 & \cos t \end{pmatrix}, \quad R_3(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

32
of rotations around the $x_1$, $x_2$ and $x_3$ axis, respectively.

The Lie algebra $so(3)$ is the set of all derivatives $\left.\frac{d}{dt}R(t)\right|_{t=0}$ of differentiable curves in $SO(3)$ with $R(0) = 1$. One can show that it consists of all real skew symmetric $3 \times 3$ matrices. A basis for the vectorspace $so(3)$ consists of the matrices

\[
E_1 = \left.\frac{d}{dt}R_1(t)\right|_{t=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},
\]
\[
E_2 = \left.\frac{d}{dt}R_2(t)\right|_{t=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},
\]
\[
E_3 = \left.\frac{d}{dt}R_3(t)\right|_{t=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

Whenever $R \in SO(3)$ and $Y \in so(3)$ then $RYR^{-1}$ lies again in $so(3)$. This defines an action of the group $SO(3)$ on its Lie algebra $so(3)$, called the adjoint representation. If $X = \left.\frac{d}{dt}R(t)\right|_{t=0}$ is an element of the Lie algebra and $Y$ another element of the Lie algebra then $[X,Y] = \left.\frac{d}{dt} R(t) Y R(t)^{-1}\right|_{t=0} = XY - YX$ lies again in the Lie algebra. $[X,Y]$ is called the Lie bracket in the Lie algebra $so(3)$. It fulfils $[X,Y] = -[Y,X]$ and the Jacobi identity

\[
[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0
\]

One easily verifies that

\[
[E_1,E_2] = E_3, \quad [E_2,E_3] = E_1, \quad [E_2,E_1] = E_2
\]

**Appendix G: Reparametrization**

In Example 3, we reparametrized the Kepler flow using the elliptic anomaly as a new independent parameter. This reparametrization depends on the energy $E$ of the system. In general, let $H(q,p)$ be any Hamiltonian and fix an energy $E$. Furthermore, let $f(q,p)$ be any differentiable function. As observed above, the Hamiltonian vectorfields associated to the Hamiltonians $H(q,p)$ and $H(q,p) - E$ agree. Set

\[
\tilde{H}(q,p) = f(q,p)(H(q,p) - E)
\]
On the level set of $H$ to the energy $E$

$$H^{-1}(E) = \{ (q, p) \mid H(q, p) = E \}$$

the Hamiltonian vectorfield for $\tilde{H}$ is

$$\frac{dq_i}{ds} = f(q, p) \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{ds} = -f(q, p) \frac{\partial H}{\partial q_i},$$

So, on $H^{-1}(E)$, the Hamiltonian vectorfield for $\tilde{H}$ is obtained from the Hamiltonian vectorfield for $H$ by the reparametrization

$$\frac{dt}{ds} = f(q, p), \quad \text{or, equivalently} \quad ds = \frac{1}{f(q, p)}.$$

In the case of negative energy $E = -\frac{\epsilon^2}{2}$, the regularized Kepler flow is obtained from the standard Kepler flow by the reparametrization $\frac{dt}{ds} = \|q\| \epsilon$. So it is described by the Hamiltonian

$$\frac{\|q\|}{\epsilon} (K(q, p) + \frac{\epsilon^2}{2}) = \frac{\|q\|}{\epsilon} (\frac{1}{2} \|p\|^2 - \|x\| + \frac{\epsilon^2}{2}) = K(q, p)$$

as in Example 3.

### Appendix H: Hamiltonian flows on the sphere

We denote by

$$S^n = \{ x = (x_0, x_1, \cdots, x_n) \in \mathbb{R}^{n+1} \mid \|x\| = (x_0^2 + x_1^2 + + \cdots + x_n^2)^{1/2} = 1 \}$$

the $n$-dimensional sphere. The tangent hyperplane to $S^n$ in $x \in S^n$ is

$$T_xS^n = \{ y \in \mathbb{R}^{n+1} \mid x \cdot y = 0 \}$$

and the total tangent space of the sphere is

$$TS^n = \{ (x, y) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \mid \|x\| = 1, x \cdot y = 0 \}$$

Given a differentiable function $H(x, y)$ on $TS^n$, the associated Hamiltonian system on $TS^n$ is

$$\dot{x}_i = \frac{\partial H}{\partial y_i} - \left( \sum_{j=1}^n x_j \frac{\partial H}{\partial y_j} \right) x_i, \quad \dot{y}_i = -\frac{\partial H}{\partial x_i} + \left( \sum_{j=1}^n (x_j \frac{\partial H}{\partial x_j} - y_j \frac{\partial H}{\partial y_j}) \right) x_i$$

The terms $\left( \sum_{j=1}^n x_j \frac{\partial H}{\partial y_j} \right) x_i$ and $\left( \sum_{j=1}^n (x_j \frac{\partial H}{\partial x_j} - y_j \frac{\partial H}{\partial y_j}) \right) x_i$ are chosen such that the flow stays on the tangent space of the sphere, that is, that for an integral curve $t \mapsto (x(t), y(t))$

$$\frac{d}{dt} \|x(t)\| = 0 \quad \text{and} \quad \frac{d}{dt} x(t) \cdot y(t) = 0$$

34
Example. The flow associated to the Hamiltonian $H(x, y) = \frac{1}{2}\|y\|$ is

$$\dot{x}_i = \frac{1}{\|y\|} y_i, \quad \dot{y}_i = -\|y\| x_i$$

It describes the uniform flow of the point $x$ on big circles, that is the geodesic flow on the sphere.

Hamiltonian systems can be defined on much more general spaces, for example on ”symplectic manifolds” or ”Poisson manifolds”. See 

35