Kepler's Laws

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.

FIGURE

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 formuli

$$\text{force} = -G\,mM\,\frac{q}{|q|^3}$$

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,

$$\ddot{q} = -\mu\,\frac{q}{|q|^3} \tag{1}$$

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.

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[(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 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 theorem about existence and uniqueness of solutions (see e.g. .???) 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 $\{\ q(t)\ \big|\ t \in I'\ \}$ 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

---

(1) In developping the basic laws of mechanics, the problem for Newton was more to show, that his second law and Kepler's laws imply the universal law on gravitation. See ???

the the semimajor axis of the ellipse.

FIGURE

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'.

FIGURE

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.

FIGURE

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 $v + w$

FIGURE

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 $\frac{1}{2} \int_{t_1}^{t_2} q(t) \times \dot{q}(t)\, dt$.

FIGURE

Keplers 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

$$q(t) \times \dot{q}(t) = L \tag{2}$$

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.

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

$$E = \frac{1}{2}\|\dot{q}\|^2 - \frac{\mu}{\|q\|} \tag{3}$$

Indeed,

$$\frac{d}{dt}\Big(\frac{1}{2}\|\dot{q}\|^2 - \frac{\mu}{\|q\|}\Big) = \frac{d}{dt}\Big(\frac{1}{2}\dot{q} \cdot \dot{q} - \frac{\mu}{(q \cdot q)^{1/2}}\Big) = \dot{q} \cdot \ddot{q} + \frac{\mu}{(q \cdot q)^{3/2}}\, q \cdot \dot{q} = \dot{q}\Big(\ddot{q} + \frac{\mu}{\|q\|^3}\, q\Big) = 0$$

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 $(q_1, q_1)-$plane. The polar coordinates $r, \varphi)$ of a point are defined by

$$q_1 = r\cos\varphi, \qquad r = \sqrt{q_1^2 + q_2^2}$$

$$q_2 = r\sin\varphi, \qquad \tan\varphi = \tfrac{q_2}{q_1}$$

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

$$\begin{aligned}\|q \times \dot q\| &= |q_1\dot q_2 - q_2\dot q_1| \\ &= \left|r\dot r\,\cos\varphi\,\sin\varphi + r^2\dot\varphi\,\cos^2\varphi - r\dot r\,\sin\varphi\,\cos\varphi + r^2\dot\varphi\,\sin^2\varphi\right| \\ &= |r^2\dot\varphi|\end{aligned}$$

Without loss of generality we assume from now on that $\dot\varphi > 0$. So

$$\|L\| = r^2\dot\varphi, \qquad \text{or, equivalently,} \qquad \dot\varphi = \tfrac{\|L\|}{r^2} \tag{4}$$

Also observe that

$$\begin{aligned}\|\dot q\|^2 &= \left(\tfrac{d}{dt}\,r\cos\varphi\right)^2 + \left(\tfrac{d}{dt}\,r\sin\varphi\right)^2 \\ &= \dot r^2\cos^2\varphi + r^2\dot\varphi^2\sin^2\varphi + \dot r^2\sin^2\varphi + r^2\dot\varphi^2\cos^2\varphi \\ &= \dot r^2 + r^2\dot\varphi^2\end{aligned} \tag{5}$$

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

$$E \;= \tfrac12\|\dot q\|^2 - \tfrac{\mu}{r} \;=\; \tfrac12\left(\dot r^2 + r^2\dot\varphi^2\right) - \tfrac{\mu}{r} \;=\; \tfrac12\left(\dot r^2 + \tfrac{\|L\|^2}{r^2}\right) - \tfrac{\mu}{r}$$

by (4) so that

$$\dot r^2 = 2E + \tfrac{2\mu}{r} - \tfrac{\|L\|^2}{r^2} \tag{6}$$

(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 φ as a function of the radius r. Observe that

$$\tfrac{d\varphi}{dr} = \tfrac{d\varphi}{dt}\,\tfrac{dt}{dr} = \tfrac{\dot\varphi}{\dot r} = \tfrac{\|L\|}{r^2\,\dot r}$$

by (4). Inserting (6) gives

$$\frac{d\varphi}{dr} = \frac{\|L\|}{r^2\sqrt{2E+\frac{2\mu}{r}-\frac{\|L\|^2}{r^2}}} = \frac{\|L\|}{r^2\sqrt{2E+\frac{\mu^2}{\|L\|^2}-\left(\frac{\|L\|}{r}-\frac{\mu}{\|L\|}\right)^2}} = \frac{1}{r^2\sqrt{\frac{e^2}{l^2}-\left(\frac{1}{r}-\frac{1}{l}\right)^2}}$$

with

$$e = \sqrt{1+\frac{2E\|L\|^2}{\mu^2}} \qquad , \qquad l = \frac{\|L\|^2}{\mu} \tag{7}$$

Therefore

$$\varphi = \int \frac{dr}{r^2\sqrt{\frac{e^2}{l^2}-\left(\frac{1}{r}-\frac{1}{l}\right)^2}} = \int \frac{1}{\sqrt{1-\left(\frac{l}{er}-\frac{1}{e}\right)^2}}\,\frac{l}{er^2}\,dr = \int \frac{-1}{\sqrt{1-\left(\frac{l}{er}-\frac{1}{e}\right)^2}}\left(\frac{d}{dr}\left(\frac{l}{er}-\frac{1}{e}\right)\right)dr$$

Thus, with an integration constant φ_0

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

or

$$r = \frac{l}{1+e\cos(\varphi-\varphi_0)} \tag{8}$$

This is the equation of a conic section with eccentricity e. see Appendix ???? 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 formuli from Appendix ???, we see that in the case of ellipses ($E<0$) the major axis is

$$a = \frac{l}{1-e^2} = \frac{\mu}{2|E|} \tag{9}$$

since $1-e^2 = \frac{2|E|\|L\|^2}{\mu^2}$ *. The minor axis is

$$b = a\sqrt{1-e^2} = \frac{\mu}{2|E|}\cdot\sqrt{\frac{2|E|\|L\|^2}{\mu^2}} = \frac{\|L\|}{\sqrt{2|E|}} = \sqrt{a}\,\frac{\|L\|}{\sqrt{\mu}} \tag{10}$$

---

* 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 ddifferent because the eccentricity then varies with $\|L\|$

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

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

$$\tfrac{1}{2}\int_0^T q(t) \times \dot{q}(t)\, dt = \tfrac{1}{2}\|L\|T$$

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

$$\tfrac{1}{2}\|L\|T = \pi a^{3/2} \tfrac{\|L\|}{\sqrt{\mu}}$$

or

$$T = \tfrac{2\pi}{\sqrt{\mu}} a^{3/2}$$

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 tha angle φ as a function of the radius r. In a variant ofthis proofone writes $\sigma = \frac{1}{r}$ as a function of φ. By the change of variables formula in differential calculus and (5)

$$\begin{aligned}\|\dot{q}\|^2 &= \dot{r}^2 + r^2\dot{\varphi}^2 = \dot{\varphi}^2\big((\tfrac{\dot{r}}{\dot{\varphi}})^2 + r^2\big) = \tfrac{\|L\|^2}{r^4}\big((\tfrac{dr}{d\varphi})^2 + r^2\big) = \|L\|^2\big((\tfrac{1}{r^2}\tfrac{dr}{d\varphi})^2 + \tfrac{1}{r^2}\big) \\ &= \|L\|^2\big((\tfrac{d}{d\varphi}\tfrac{1}{r})^2 + \tfrac{1}{r^2}\big)\end{aligned}$$

Then by (3)

$$E = \tfrac{1}{2}\|\dot{q}\|^2 - \tfrac{\mu}{\|q\|} = \tfrac{\|L\|^2}{2}\big((\tfrac{d\sigma}{d\varphi})^2 + \sigma^2\big) - \mu\sigma$$

Differentiating with respect to φ and dividing by $\|L\|^2$ gives

$$0 = \tfrac{d\sigma}{d\varphi}\left(\tfrac{d^2\sigma}{d\varphi^2} + \sigma - \tfrac{\mu}{\|L\|^2}\right)$$

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

$$\tfrac{d^2\sigma}{d\varphi^2} + \sigma - \tfrac{\mu}{\|L\|^2} = 0$$

The general solution of this differential equation is

$$\sigma = \tfrac{\mu}{\|L\|^2}\big(1 + e\,\cos(\varphi - \varphi_0)\big)$$

with integration constants e, φ_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 l 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 detemined 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 Ł 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 φ_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 *Laplace Lenz Runge vector*

$$A = -\frac{q}{|q|} + \frac{1}{\mu}\,\dot{q}\times L = \frac{1}{\mu}\big(\|\dot{q}\|^2 - \frac{\mu}{\|q\|}\big)\,q - \frac{1}{\mu}(q\cdot\dot{q})\,\dot{q} \tag{11}$$

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 $\dot{q}\times L = \dot{q}\times(q\times\dot{q}) = -(q\cdot\dot{q})\dot{q} + \|\dot{q}\|^2 q$. 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^o. So $\frac{1}{\|L\|}\,\dot{q}\times L$ is the vector obtained from $\dot{q}$ by rotating by 90^o.

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

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

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 it is called the *eccentricity vector* in [Cushman]. To prove the statement about the length of A observe that

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

Since $\dot{q}$ and $L = q \times \dot{q}$ are perpendicular, $\|\dot{q} \times L\|^2 = \|\dot{q}\|^2\|L\|^2$. By the standard vector identity $x \cdot (y \times z) = z \cdot (x \times y)$

$$q \cdot (\dot{q} \times L) = L \cdot (q \times \dot{q}) = L \cdot L = \|L\|^2 \tag{12}$$

Therefore

$$\begin{aligned} A \cdot A &= 1 - \frac{2\mu}{\mu^2\|q\|}\|L\|^2 + \frac{1}{\mu^2}\|\dot{q}\|^2\|L\|^2 \\ &= 1 + \frac{\|L\|^2}{\mu^2}\left(\|\dot{q}\|^2 - \frac{2\mu}{\mu^2\|q\|}\right) \\ &= 1 + 2\frac{\|L\|^2}{\mu^2}E = e^2 \end{aligned}$$

by (3).

Before using the Laplace Lenz Runge vector, we describe how one could think of considering this particular vector (11), when 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 the two foci. The origin is one focus of the ellipse, call the other one f. It is known that 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.

FIGURE

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.

FIGURE

---

* For other arguments, see [Heintz] and [Kaplan]. Remarks on the history of this vector can be found in [Goldstein] and [Goldstein].

A unit tangent vector to the ellipse at the point q is $\frac{\dot q}{\|\dot q\|}$. Therefore $f - q$ has the direction $-q + 2\big(q \cdot \frac{\dot q}{\|\dot q\|}\big)\,\frac{\dot q}{\|\dot q\|}$. This means that there is a scalar function $\alpha(t)$ such that

$$f = q(t) + \alpha(t)\Big(- q(t) + \tfrac{2q(t)\cdot\dot q(t)}{\|\dot q(t)\|^2}\ \dot q(t)\Big) \tag{13}$$

for all t. Differentiating and using Kepler's law (1) gives

$$\begin{aligned}
0 = \dot f &= \dot q + \dot\alpha\big(- q + 2\tfrac{q\cdot\dot q}{\|\dot q\|^2}\ \dot q\big) + \alpha\Big[- \dot q + 2\big(\tfrac{(\dot q\cdot\dot q + q\cdot\ddot q)}{\|\dot q\|^2} - \tfrac{(q\cdot\dot q)\,2(\dot q\cdot\ddot q)}{\|\dot q\|^4}\big)\dot q + \tfrac{2q\cdot\dot q}{\|\dot q\|^2}\ \ddot q\Big] \\
&= -\dot\alpha q + \tfrac{1}{\|\dot q\|^2}\Big\{\|\dot q\|^2 + 2\dot\alpha(q\cdot\dot q) + \alpha\big[- \|\dot q\|^2 + 2\|\dot q\|^2 + 2q\cdot(-\mu\tfrac{q}{\|q\|^3}) - 2\tfrac{q\cdot\dot q}{\|\dot q\|^2}(-\mu\tfrac{\dot q\cdot q}{\|q\|^3})\big]\Big\}\,\dot q \\
&\qquad + 2\alpha\,\tfrac{q\cdot\dot q}{\|\dot q\|^2}\big(- \mu\tfrac{q}{\|q\|^3}\big) \\
&= -\big[\dot\alpha + 2\alpha\mu\,\tfrac{q\cdot\dot q}{\|\dot q\|^2\|q\|^3}\big]\,q + \tfrac{1}{\|\dot q\|^2}\Big\{\|\dot q\|^2 + 2(q\cdot\dot q)\big[\dot\alpha + 2\alpha\mu\,\tfrac{q\cdot\dot q}{\|\dot q\|^2\|q\|^3}\big] + 2\alpha\big[\tfrac12\|\dot q\|^2 - \tfrac{\mu}{\|q\|}\big]\Big\}\,\dot q
\end{aligned}$$

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

$$\dot\alpha + 2\alpha\mu\,\tfrac{q\cdot\dot q}{\|\dot q\|^2\|q\|^3} = 0$$

Inserting this, and the fact that $\frac12\|\dot q\|^2 - \frac{\mu}{\|q\|} = E$ 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

$$\begin{aligned}
f &= q - \tfrac{\|\dot q\|^2}{2E}\Big(- q + \tfrac{2q\cdot\dot q}{\|\dot q\|^2}\ \dot q\Big) \\
&= \tfrac{1}{E}\Big(\big(E + \tfrac{\|\dot q\|^2}{2}\big)\,q - (q\cdot\dot q)\ \dot q\Big) \\
&= \tfrac{1}{E}\Big(\big(\|\dot q\|^2 - \tfrac{\mu}{\|q\|}\big)\,q - (q\cdot\dot q)\ \dot q\Big) \\
&= \tfrac{\mu}{E}\,A
\end{aligned} \tag{14}$$

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 second law is relatively fast. As

$$q\cdot A = q\cdot\big(- \tfrac{q}{|q|} + \tfrac{1}{\mu}\ \dot q\times L\big) = -\|q\| + \tfrac{1}{\mu}\,q\cdot(\dot q\times L) = -\|q\| + \tfrac{1}{\mu}\,L\cdot(q\times\dot q) = -\|q\| + \tfrac{1}{\mu}\,\|L\|^2$$

we have, setting again $e = \|A\|$

$$\|q\| = e\left(\tfrac{1}{e\mu}\,\|L\|^2 - q\cdot\tfrac{A}{\|A\|}\right)$$

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 ???, 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

$$L\times A = \tfrac{1}{|q|}\,L\times(-q) + \tfrac{1}{\mu}\,L\times(\dot q\times L) = \tfrac{1}{|q|}\,q\times L + \tfrac{1}{\mu}\|L\|^2\,\dot q \tag{15}$$

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

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

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

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

* Another derivation of the equation of the hodograph, attributed to Hamilton, can be found in [Hankins], ch.24.

The equation of motion (1) is

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

so that

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

Consequently

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

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

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

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

$$p - u = -\tfrac{\mu}{\|L\|}\,\tfrac{\dot{p}}{\|\dot{p}\|}\times\tfrac{L}{\|L\|} = -\tfrac{\mu}{\|q\|\|L\|^2}\, q\times L \tag{16}$$

and our argument shows that

$$u = p + \tfrac{\mu}{\|q\|\|L\|^2}\, q\times L \tag{17}$$

is a conserved quantity[(*)]. The hodograph has the equation

$$\|p-u\| = \tfrac{\mu}{\|L\|}$$

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

$$\begin{aligned}
\|u\|^2 &= \|p\|^2 + \tfrac{\mu^2}{\|q\|^2\|L\|^4}\,\|q\times L\|^2 + 2\tfrac{\mu}{\|q\|\|L\|^2}\, p\cdot(q\times L)\\
&= \|p\|^2 + \tfrac{\mu^2}{\|L\|^2} + 2\tfrac{\mu}{\|q\|\|L\|^2}\, L\cdot(p\times q)\\
&= \|p\|^2 + \tfrac{\mu^2}{\|L\|^2} + 2\tfrac{\mu}{\|q\|}\\
&= 2E + \tfrac{\mu^2}{\|L\|^2} = \left(\tfrac{\mu e}{\|L\|}\right)^2
\end{aligned}$$

(*) 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.

We now describe the position q in terms of the momentum p. By (16) and the fact that q is perpendicular to 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

$$q = r \, \tfrac{(p-u)\times L}{\|(p-u)\times L\|} = \tfrac{r}{\mu}\,(p-u)\times L \qquad \text{with} \quad r = \pm\|q\|$$

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

$$\tfrac{r}{\mu}\,\big((p-u)\times L\big)\cdot(p\times L) = \|L\|^2$$

Cross product by $\frac{L}{\|L\|}$ corresponds to rotation by 90^o. Therefore this implies

$$\begin{aligned} \mu &= r\,(p-u)\cdot p \\ &= r\,\big((p-u)\cdot(p-u) + (p-u)\cdot u\big) \\ &= r\,\big(\tfrac{\mu^2}{\|L\|^2} + \tfrac{\mu^2 e}{\|L\|^2}\,\tfrac{(p-u)}{\|p-u\|}\cdot\tfrac{u}{\|u\|}\big) \end{aligned}$$

Here we used that $\|p-u\| = \frac{\mu}{\|L\|}$ and that $\|u\| = \frac{\mu e}{\|L\|}$. If we denote the angle between $p-u$ and u by φ, we get

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

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

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$^{\text{st}}$ 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 ???), and conservation of angular momentum (4) implicitly gives the dependence of φ on t by the differential equation $\frac{d\varphi}{dt} = \frac{\|L\|}{r^2}$. To simplify the discussion we put $\varphi_0 = 0$.

First we consider the case of **negative energy**, that is, the case of bounded orbits. Write

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

In this case the orbit is described by the equation

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

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

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

by the *eccentric anomaly* s (see appendix ???). To get the dependence of s on the time t, observe that by formula (24) of the appendix and (7)

$$\tfrac{ds}{d\varphi} = \tfrac{\sqrt{1-e^2}}{l}\, r = \tfrac{\sqrt{-2E}}{\|L\|}\, r = \tfrac{\varepsilon}{\|L\|}\, r$$

We just observed that $\frac{d\varphi}{dt} = \frac{\|L\|}{r^2}$. Therefore, by the chain rule

$$\tfrac{ds}{dt} = \tfrac{\varepsilon}{r} \tag{18}$$

By (9), $\varepsilon^2 = 2|E| = \frac{\sqrt{\mu}}{a}$. In formula (25) of the appendix, we show that $r = a(1 - e\cos s)$. Hence (18) gives

$$\tfrac{ds}{dt} = \tfrac{\sqrt{\mu}}{a^{3/2}(1-e\cos s)}$$

or, equivalently

$$\tfrac{dt}{ds} = \tfrac{a^{3/2}(1-e\cos s)}{\sqrt{\mu}}$$

Integrating both sides with respect to s gives *Kepler's equation*

$$t - t_0 = \tfrac{a^{3/2}}{\sqrt{\mu}}(s - e \sin s) \tag{19}$$

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" $s \mapsto s - \sin s$. This inversion cannot be performed by elementary functions. In ???, we discuss Kepler's equation in more detail.

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 (18) in the basic equation of motion (1). Recall that $r = \|q\|$. Then

$$\tfrac{dq}{ds} = \dot{q}\,\tfrac{dt}{ds} = \tfrac{\|q\|}{\varepsilon}\dot{q} \qquad \text{or, equivalently} \qquad \dot{q} = \tfrac{\varepsilon}{\|q\|}\tfrac{dq}{ds}$$

Therefore, using (1)

$$\begin{aligned} \tfrac{d^2q}{ds^2} &= \tfrac{1}{\varepsilon}\tfrac{d\|q\|}{ds}\dot{q} + \tfrac{\|q\|}{\varepsilon}\tfrac{d\dot{q}}{ds} = \tfrac{1}{\|q\|}\tfrac{d\|q\|}{ds}\tfrac{dq}{ds} + \tfrac{\|q\|^2}{\varepsilon^2}\ddot{q} = \tfrac{1}{\|q\|}\tfrac{d\|q\|}{ds}\tfrac{dq}{ds} - \tfrac{\mu}{\varepsilon^2\|q\|}q \\ &= \tfrac{1}{\|q\|^2}\left(q \cdot \tfrac{dq}{ds}\right)\tfrac{dq}{ds} - \tfrac{\mu}{\varepsilon^2\|q\|}q \end{aligned} \tag{20}$$

Once one knows that (18) 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 (3) and (11), the Laplace Lenz Runge vector is

$$\begin{aligned} A &= \tfrac{1}{\mu}\left(2E + \tfrac{\mu}{\|q\|}\right)q - \tfrac{1}{\mu}(q \cdot \dot{q})\,\dot{q} \\ &= -\tfrac{\varepsilon^2}{\mu}\left(\tfrac{1}{\|q\|^2}\left(q \cdot \tfrac{dq}{ds}\right)\tfrac{dq}{ds} - \tfrac{2E}{\varepsilon^2}q - \tfrac{\mu}{\varepsilon^2\|q\|}q\right) \\ &= -\tfrac{\varepsilon^2}{\mu}\left(\tfrac{1}{\|q\|^2}\left(q \cdot \tfrac{dq}{ds}\right)\tfrac{dq}{ds} - \tfrac{\mu}{\varepsilon^2\|q\|}q + q\right) \end{aligned}$$

since $\varepsilon^2 = -2E$. Therefore (20) gives

$$\tfrac{d^2q}{ds^2} + q = -\tfrac{\mu}{\varepsilon^2}A$$

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

$$q(s) = C_1 \cos s + C_2 \sin s - \tfrac{\mu}{\varepsilon^2}A \tag{21}$$

with constant vectors C_1, C_2. This is the parametrization of an ellipse. The vectors C_1, C_2 are easily described in terms of known quantities, see [Cordani].

In the parametrization (21) all orbits have period 2π. 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".

Appendix on 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 ???. 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

$$FIGURE\ like\ Arnold,\ figure\ 35$$

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

$$\tfrac{x^2}{a^2} + \tfrac{y^2}{b^2} = 1 \tag{22}$$

For a proof see ....The area enclosed by the ellipoid is equal to $\pi\, a\, b$; see for example ????.

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. For a proof see ?????.

FIGURE

This description directly gives the equation of the ellipse in polar coordinates (r, φ) centered at the focus F. Choose the angular variable φ 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, φ) then its distance to F is r. Since the distance from F to the directrix is $\left(\frac{1}{e} - e\right) a$, the distance from P to the directrix is $\left(\frac{1}{e} - e\right) a - r\cos\varphi$. Thus the equation of the ellipse is

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

or, equivalently

$$r = \tfrac{l}{1+e\cos\varphi} \tag{23}$$

where $l = (1 - e^2)a$. The quantity l is called the *parameter* of the ellipse. The angle φ 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 (23) is a parametrization of the ellipse, giving the distance r to the focus as a function of the true anomaly φ. Equation (22) suggests the parametrization

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

of the ellipse. The parameter s is called the *eccentric anomaly.* If we choose the focus $F = (ea, 0)$ as the origin of the polar coordinate system as above$^{(*)}$ then the relation between the eccentric amnomaly s and the true anomaly φ can be seen in the following figure.

FIGURE

$^{(*)}$ This is the direction from the focus F to the closest point on the ellipse.
$^{(*)}$ The other choice would be $F' = (-ea, 0)$

To get this relation in formuli, let (x, y) be a point of the ellipse with true anomaly φ and eccentric anomaly s. Then

$$x = a\ \cos s$$

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

and

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

$$y = r\sin\varphi = \tfrac{(1-e^2)a}{1+e\cos\varphi}\sin\varphi$$

Comparing these two representation, we obtain

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

We differentiate the first equation with respect to φ and insert the second to get

$$\begin{aligned} -\tfrac{ds}{d\varphi}\ \sin s &= -\tfrac{1-e^2}{1+e\cos\varphi}\sin\varphi + \tfrac{1-e^2}{(1+e\cos\varphi)^2}\, e\sin\varphi\ \cos\varphi \\ &= -\sqrt{1-e^2}\ \sin s + \sqrt{1-e^2}\ \tfrac{e\cos\varphi}{1+e\cos\varphi}\ \sin s \\ &= -\sqrt{1-e^2}\ \sin s\ \tfrac{1}{1+e\cos\varphi} \end{aligned}$$

Dividing by $\sin s$ and using (23) gives

$$\tfrac{ds}{d\varphi} = \tfrac{\sqrt{1-e^2}}{l}\, r \tag{24}$$

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

$$\begin{aligned} r^2 &= (x-ea)^2 + y^2 = a^2(\cos s\ - e)^2 + (1-e^2)a^2\sin^2 s \\ &= a^2\big(\cos^2 s - 2e\cos s + e^2 + (1-e^2) - \cos^2 s + e^2\cos^2 s\big) \\ &= a^2(1-e\cos s)^2 \end{aligned}$$

we have

$$r = a\,(1-e\cos s) \tag{25}$$

Hyperboli

FIGURE

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

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

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 $e\, a$ 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

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

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

$$r = \tfrac{(e^2-1)a}{1+e\cos\varphi} \tag{26}$$

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

Paraboli

Finally consider the parabola consisting of all points that have equal distance form 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 ℓ 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

(*) Again, this is the direction from the focus F to the closest point on the ellipse.

If one introduces Cartesian coordinates x, y centered at M with the x–axis being ℓ 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

$$r = \tfrac{l}{1+\cos\varphi} \tag{27}$$

Appendix on 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 $r_1(t), r_2(t)$.

FIGURE

In this situation, Newton's laws give

$$\begin{aligned} m_1\, \ddot{r}_1(t) &= G\, m_1 m_2\, \tfrac{r_2(t)-r_1(t)}{|r_2(t)-r_1(t)|^3} \\ m_2\, \ddot{r}_2(t) &= G\, m_1 m_2\, \tfrac{r_1(t)-r_2(t)}{|r_1(t)-r_2(t)|^3} \end{aligned} \tag{28}$$

Denote by $R(t) = \frac{1}{m_1+m_2}\big(m_1 r_1(t) + m_2 r_2(t)\big)$ the center of gravity. Adding the two equations of (28) gives $\ddot{R}(t) = 0$. That is, the center of gravity moves with constant speed. Also, set $q(t) = r_2(t) - r_1(t)$. (28) gives

$$\ddot{r}_1(t) = G m_2\, \tfrac{q(t)}{|q(t)|^3}\ , \qquad \ddot{r}_2(t) = -G m_1\, \tfrac{q(t)}{|q(t)|^3}$$

so that

$$\ddot{q}(t) = -\mu\, \tfrac{q(t)}{|q(t)|^3}$$

This is the basic equation (1).

Appendix on the Lagrangian formalism

The Euler Lagrange equations associated to a function $\mathcal{L}(t,q,p)$ is

$$\tfrac{d}{dt}\big(\tfrac{\partial\mathcal{L}}{\partial p_i}\big|_{p=\dot q}\big) - \tfrac{\partial\mathcal{L}}{\partial q_i}\big|_{p=\dot q} = 0 \tag{29}$$

Conventionally one views $\mathcal{L}$ as a function of the variables $t, q, \dot q$ and uses the shorthand formulation

$$\tfrac{d}{dt}\big(\tfrac{\partial\mathcal{L}}{\partial \dot q_i}\big) - \tfrac{\partial\mathcal{L}}{\partial q_i} = 0$$

The Lagrange function for the Kepler problem is the difference between the cinetic and the potential energy

$$\mathcal{L}_K(q,p) = \tfrac{1}{2}\|p\|^2 + \tfrac{\mu}{\|q\|}$$

It is independent of t. Then

$$\tfrac{\partial\mathcal{L}_K}{\partial p_i} = p_i \ , \qquad \tfrac{d}{dt}\big(\tfrac{\partial\mathcal{L}_K}{\partial p_i}\big|_{p=\dot q}\big) = \ddot q_i \ , \qquad \tfrac{\partial\mathcal{L}_K}{\partial q_i} = -\mu\tfrac{q_i}{\|q\|^3}$$

so that the associated Euler Lagrange equation are

$$\ddot q_i + \mu\tfrac{q_i}{\|q\|^3} = 0$$

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

$$\begin{aligned} \gamma :\ [t',t''] &\to \mathbb{R}^3 \\ t &\mapsto q(t) \qquad\qquad \text{with}\ \ q(t') = p',\ q(t'') = q'' \end{aligned}$$

define the action

$$\Phi(\gamma) = \int_{t'}^{t''} \mathcal{L}\big(t, q(t), \dot q(t)\big)\ dt$$

If γ 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 "(cinetic 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 $G(q) = \big(g_{ab}(q)\big)_{a,b=1,\cdots,n}$. For simplicity we assume that the coefficients $g_{ab}(q)$ are C^∞ 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

$$\text{length}(\gamma) = \int_{t'}^{t''} \Big(\dot q(t)^\top G\big(q(t)\big)\, \dot q(t)\Big)^{\frac{1}{2}}\ dt$$

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 $\varepsilon > 0$ such that for all s with $t < s\ < t + \varepsilon$, the curve is a shortest connection between the points $q(t)$ and $q(s)$. It follows that geodesics fulfil the Euler Lagrange equations (29) with

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

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 $\dot q(t)^\top G\big(q(t)\big)\, \dot q(t) = \text{const}$ for all t. If one has a minimizer for the variational problem that is parametrized by arclength, then it also fulfils the equation

$$\tfrac{d}{dt}\big(\tfrac{\partial \mathcal{L}^2}{\partial p_i}\big|_{p=\dot q}\big) - \tfrac{\partial \mathcal{L}^2}{\partial q_i}\big|_{p=\dot q} = 0 \tag{30}$$

Indeed, multiplying the Euler Lagrange equation (29) by $\text{const} = \mathcal{L}(q, \dot q)$, we get

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

which implies (30).

Equation (30) is simpler than the Euler Lagrange equations associated to the original Lagrange function. We now derive the equation of motion associated to (30). To see that we really get geodesics, we then have to verify that the resulting curves are indeed parametrized with constant speed. Equation (30) gives

$$2\tfrac{d}{dt}\Big(\sum_{b=1}^{n} g_{ab}(q)\, \dot q_b\Big) - \sum_{b,c=1}^{n} \tfrac{\partial g_{bc}}{\partial q_a}\, \dot q_b \dot q_c = 0 \qquad \text{for } a = 1, \cdots, n$$

Since $\frac{d}{dt}\, g_{ab}(q) = \sum_{c=1}^{n} \frac{\partial g_{ab}}{\partial q_c}\, \dot q_c$, (30) is equivalent to

$$\begin{aligned}\sum_{b=1}^{n} g_{ab}\, \ddot q_b &= \tfrac{1}{2} \sum_{b,c=1}^{n} \tfrac{\partial g_{bc}}{\partial q_a}\, \dot q_b \dot q_c - \sum_{b,c=1}^{n} \tfrac{\partial g_{ab}}{\partial q_c}\, \dot q_b \dot q_c \\ &= \tfrac{1}{2}\Big(\sum_{b,c=1}^{n} \big(\tfrac{\partial g_{bc}}{\partial q_a} - \tfrac{\partial g_{ab}}{\partial q_c} - \tfrac{\partial g_{ca}}{\partial q_b} \big)\Big)\, \dot q_b \dot q_c\end{aligned}$$

for $a = 1, \cdots, n$. 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 $G\,\ddot q$ is equal to the vector with entries $\frac{1}{2}\Big(\sum_{b,c=1}^{n} \big(\frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \big)\Big)\, \dot q_b \dot q_c$. We denote by g^{ab} the entries of the inverse matrix G^{-1}. Then the equations are equivalent to

$$\ddot q_a + \sum_{b,c=1}^{n} \Gamma^a_{bc}\, \dot q_b \dot q_c = 0 \qquad (31)$$

where Γ^a_{bc} are the *Christoffel symbols*

$$\Gamma^a_{bc} = \tfrac{1}{2} \sum_{d=1}^{n}\, g^{ad} \sum_{b,c=1}^{n} \big(\tfrac{\partial g_{db}}{\partial q_c} + \tfrac{\partial g_{cd}}{\partial q_b} - \tfrac{\partial g_{bc}}{\partial q_d} \big)$$

To verify that (31) really describes geodesics, we still have to verify that solutions of (31) are curves that are parametrized by constant speed. So let $t \mapsto q(t)$ be a solution of (31). Reversing the calculation above, we see that shows $2\,G\,\ddot q$ is the vector with entries $\sum_{b,c=1}^{n} \big(\frac{\partial g_{bc}}{\partial q_a} - \frac{\partial g_{ab}}{\partial q_c} - \frac{\partial g_{ca}}{\partial q_b} \big)\, \dot q_b \dot q_c$. Therefore

$$\begin{aligned}\frac{d}{dt}\, \dot q\, G(q)\, \dot q &= 2\, \dot q\, G(q)\, \ddot q + \dot q\, \dot G\, \dot q \\ &= \sum_{a,b,c=1}^{n} \big(\tfrac{\partial g_{bc}}{\partial q_a} - \tfrac{\partial g_{ab}}{\partial q_c} - \tfrac{\partial g_{ca}}{\partial q_b} \big)\, \dot q_a \dot q_b \dot q_c + \sum_{a,b,c=1}^{n} \tfrac{\partial g_{bc}}{\partial q_a}\, \dot q_a \dot q_b \dot q_c = 0\end{aligned}$$

and $\dot q\, G(q)\, \dot q$ is indeed constant.