# The Kepler Problem 

## Kepler, Newton, Monsieur Bertrand and the Hermann-Bernoulli-Laplace-Hamilton-Gibbs-Runge-Lenz vector: the general solution to the Kepler problem. If the solutions are conic sections, what exactly is the cone? And why does this non-relativistic, classical problem lead us to relativity and quantum mechanics?

Last Update: 24/2/12

Fortuitous, wasn't it? The ancient Greeks had complete knowledge of the geometry of conic sections. Consequently, so did Newton. What jolly good fortune that the motion of the celestial bodies turned out to be just these very conic sections. Good luck or not, it has to be one of the juiciest bits of physics. And the Kepler problem has not run out of surprises yet. Ask yourself, if the planetary motions are conic sections, what exactly is the cone? But first, to Kepler.
By the sweat of his brow, by dint of manual number-crunching over many years, and with no little inspiration, Kepler conjured his famous three laws of planetary motion out of a mountain of observational data. They are,
[1] The planets move in ellipses with the Sun at one focus;
[2] The radius vector from the Sun to a planet sweeps out equal areas in equal times;
[3] The ratio of the square of the period to the cube of the semi-major axis is the same for all planets.

With these laws as the clue, Newton deduced the inverse-square law of gravity. But how many of these laws do you need to deduce the inverse-square law? We can start, as Newton did, from an appreciation that the general law of motion is $\bar{F}=m \ddot{\vec{r}}=\dot{\bar{p}}$, where $\bar{p}$ is the momentum, $\bar{r}$ the time dependent position vector, and $m$ the mass of the orbiting body.

Remarkably only about one-and-a-half of the laws are required to deduce the inverse square law. More precisely we need the second law and part of the first law, plus a little bit more. Specifically what we need is,
(i) The planets move in closed, stable orbits;
(ii) The radius vector from the Sun to a planet sweeps out equal areas in equal times;
(iii) Gravity gets weaker with increasing distance.

The first of these is a reduced version of Kepler's first law. Ellipticity is not mentioned. The last, (iii), is the "little bit more", though most people might regard it as self evident. Hence, Kepler's third law is not necessary, the second law is unchanged, and the first law has been considerably weakened.

Kepler's second law is equivalent to the conservation of angular momentum, which in turn is equivalent to motion in a central force. A "central force" is defined as a force whose magnitude depends only upon the radial distance from some origin, not upon the angular coordinates, and whose direction is radial, thus,

$$
\begin{equation*}
\bar{F}=\dot{\bar{p}}=\hat{r} f(r) \tag{1}
\end{equation*}
$$

This is equivalent to a potential which is a function of the radial distance, $r$ only, $V(r)$, so that,

$$
\begin{equation*}
\bar{F}=-\bar{\nabla} V=-\hat{r} \frac{\partial V}{\partial r}, \text { i.e., } f(r) \equiv-\frac{\partial V}{\partial r} \tag{2}
\end{equation*}
$$

The angular momentum is defined as,

$$
\begin{equation*}
\bar{L}=\bar{r} \times \bar{p} \tag{3}
\end{equation*}
$$

This is a constant of the motion in any central force, as can be seen from,

$$
\begin{equation*}
\frac{d}{d t} \bar{L}=\dot{\bar{r}} \times \bar{p}+\bar{r} \times \dot{\bar{p}}=\dot{\bar{r}} \times(m \dot{\bar{r}})+\bar{r} \times(\hat{r} f)=0 \tag{4}
\end{equation*}
$$

The first term in (4) is identically zero. The second term is zero by virtue of the force being central, i.e., of the form (1).
If the body moves $\delta \bar{r}$ in time $\delta t$, then it sweeps out an area $|\bar{r} \times \delta \bar{r}| / 2$. Consequently the rate of sweeping out area is just the magnitude of $\bar{r} \times \dot{\bar{r}} / 2=\bar{r} \times \bar{p} / 2 m=\bar{L} / 2 \mathrm{~m}$. So the rate of sweeping out area is constant, which is Kepler's second law. What has been shown is that the assumption of a central force is sufficient to derive the second law.

Conversely if the potential function depended upon the angular position as well as the radial position then the force would have a non-zero component in a direction perpendicular to $\hat{r}$, proportional to $\partial_{\theta} V$, and so $\bar{L}$ would not be a constant of motion, and hence Kepler's second law would not hold. So the second law is also sufficient to deduce that the force is central.

The other thing that the constancy of $\bar{L}$ tells us is that the planetary orbits are planar. It is clear from (3) that $\bar{L}$ is perpendicular to the instantaneous plane of the motion. So the constancy of $\bar{L}$ means that motion is in a fixed plane.
So far, so good, but knowing that gravity is a central force does not assist us with its radial dependence. The function $f(r)$ could be anything at all provided that it depends only upon the radial distance and not upon the angular coordinate. The second law is then respected. So how can we deduce the inverse square law with no extra information than our substitute 'laws' (i) and (iii), above?
That the inverse square force law is the only possibility which obeys these conditions is known as Bertrand's Theorem. It was first proved only in 1873, nearly two centuries after Newton's Principia (1687). An English translation of the French original has been published by Santos et al (2007). The key word in our alternative law No.(i) is "stable". Of course any central force will admit circular orbits if the initial conditions are suitably contrived. However they will always be unstable orbits, becoming non-closed under the smallest of perturbations. Bertrand (1873) proved that there are only two central forces which produce stable, closed orbits. These are the inverse square law and a force which is proportional to the radius (roughly speaking an elastic string). The purpose of our "little bit extra", law (iii), is to exclude the latter. Hence the inverse square law is all that remains.
Newton would not have been aware of this means of deducing the inverse square law. He was, though, well aware of the solution to the equation of motion for both the inverse squared law and the 'elastic string', discussing both in the Principia. Both admit elliptic solutions, though only the inverse square force has the origin (the Sun) at the focus. The elastic string is tied to the centre of the ellipse.

Having arrived at the correct form of the gravitational force law, we can proceed to prove that the motion must be a conic section - specifically ellipses for closed orbits and thence prove Kepler's first and third laws. For amusement, and because it will prove useful later ${ }^{1}$, we shall show that the solutions are the conic sections without integrating the equations of motion. Of course we know that differential equations cannot be necessary, after all Newton (1687) did it using only Euclidean geometry.

The trick is to note that in addition to energy and angular momentum, there is another quantity which is a constant of the motion, namely,

$$
\begin{equation*}
\bar{M}=\frac{1}{m} \bar{p} \times \bar{L}-K \hat{r} \tag{5}
\end{equation*}
$$

where $\hat{r}$ is the unit vector in the radial direction, $\hat{r}=\bar{r} / r$, and $K$ is the constant in the inverse square gravitational force,

$$
\begin{equation*}
\bar{F}=-\frac{K}{r^{2}} \hat{r} \tag{6}
\end{equation*}
$$

The attractive nature of gravity means that $K>0$. The constancy of $\bar{M}$ is easily established as follows,

$$
\begin{align*}
\frac{d}{d t} \bar{M} & =\frac{1}{m} \dot{\bar{p}} \times \bar{L}+\frac{1}{m} \bar{p} \times \dot{\bar{L}}-K \frac{\dot{\vec{r}}}{r}+K \frac{\bar{r}}{r^{2}} \dot{r} \\
& =\frac{1}{m}\left(-\frac{K \hat{r}}{r^{2}}\right) \times(\bar{r} \times m \dot{\bar{r}})-K \frac{\dot{\bar{r}}}{r}+K \frac{\bar{r}}{r^{2}} \dot{r}  \tag{7}\\
& =K\left\{\frac{1}{r^{2}}[r \dot{\bar{r}}-(\hat{r} \cdot \dot{\bar{r}}) \bar{r}]-\frac{\dot{\vec{r}}}{r}+\frac{\dot{r} \bar{r}}{r^{2}}\right\} \equiv 0
\end{align*}
$$

where we have used the identity $\bar{a} \times(\bar{b} \times \bar{c}) \equiv(\bar{a} \cdot \bar{c}) \bar{b}-(\bar{a} \cdot \bar{b}) \bar{c}$ and also the fact that $\hat{r} \cdot \dot{\vec{r}} \equiv \dot{r}$ (which follows from $\dot{\bar{r}}=r \dot{\hat{r}}+\dot{r} \hat{r}$ and the fact that $\hat{r} \cdot \dot{\hat{r}} \equiv 0$ ). The constant vector $\bar{M}$ defined by (5) is generally known as the Runge-Lenz ${ }^{2}$ vector: more of its magic latter. For now we shall use it to solve the Kepler problem without integrating the equations of motion.
We can write $\bar{p} \times \bar{L}=p L \hat{n}$ where $\hat{n}$ is a unit vector perpendicular to the direction of motion (and lying in the plane of motion). Hence (5) can be re-arranged as,

$$
\begin{equation*}
\frac{p L}{m} \hat{n}=\bar{M}+K \hat{r} \tag{8}
\end{equation*}
$$

This shows that $p L / m=|\bar{M}+K \hat{r}|$ and that the trajectory of the heavenly body obeys,

$$
\begin{equation*}
\hat{n}=\frac{\bar{M}+K \hat{r}}{|\bar{M}+K \hat{r}|} \tag{9}
\end{equation*}
$$

[^0]This is effectively an equation for the shape (but not the size) of the trajectory, because $\bar{M}$ and $K$ are constants. It is in a non-standard form which specifies the normal to the trajectory in terms of the angular position. All that remains is to show that (9) is equivalent to an arbitrary conic section. To do so recall that an arbitrary conic section can be written, in polar coordinates,

$$
\begin{equation*}
\frac{1}{r}=A \cos \theta+B \tag{10}
\end{equation*}
$$

Without loss of generality we can assume $A \geq 0$. The different types of conic section occur for,
$B>A>0$ : Ellipse
$B>0 ; A=0$ : Circle

$$
\begin{array}{ll}
B=A: & \text { Parabola } \\
0<B<A: & \text { Hyperbola (+ branch) ; }-A<B<0: \text { Hyperbola (- branch) }
\end{array}
$$

To establish that (10) can be written as (9), take $\bar{M}$ as defining the $x$-axis ( $\theta=0$ ). The gradient of the trajectory is,

$$
\begin{equation*}
\frac{d y}{d x}=-\frac{n_{x}}{n_{y}}=-\frac{M+K \cos \theta}{K \sin \theta} \tag{11}
\end{equation*}
$$

But in polar coordinates for which $x=r \cos \theta, y=r \sin \theta$ the gradient is,

$$
\begin{equation*}
\frac{d y}{d x}=\frac{\sin \theta \cdot d r+r \cos \theta \cdot d \theta}{\cos \theta \cdot d r-r \sin \theta \cdot d \theta} \tag{12}
\end{equation*}
$$

Using (10) we find,

$$
\begin{equation*}
d r=A r^{2} \sin \theta \cdot d \theta \tag{13}
\end{equation*}
$$

So (12) becomes,

$$
\begin{align*}
\frac{d y}{d x} & =\frac{A r^{2} \sin ^{2} \theta+r \cos \theta}{A r^{2} \sin \theta \cos \theta-r \sin \theta} \\
& =\frac{A \sin ^{2} \theta+r^{-1} \cos \theta}{A \sin \theta \cos \theta-r^{-1} \sin \theta}  \tag{14}\\
& =\frac{A \sin ^{2} \theta+(A \cos \theta+B) \cos \theta}{A \sin \theta \cos \theta-(A \cos \theta+B) \sin \theta} \\
& =-\frac{A+B \cos \theta}{B \sin \theta}
\end{align*}
$$

This is exactly as (11) with $A=\lambda M$ and $B=\lambda K$. It remains to find the absolute size of the trajectory by determining the constant $\lambda$. To do so we appeal to the one constant of motion that we have not yet used - the energy. Of course we know that the total energy is just the sum of the potential and kinetic energies and hence is,

$$
\begin{equation*}
E=-\frac{K}{r}+\frac{1}{2} m|\dot{\bar{r}}|^{2} \tag{15}
\end{equation*}
$$

But really we should prove that this is a constant. This is easily done,

$$
\begin{equation*}
\frac{d E}{d t}=\frac{K}{r^{2}} \dot{r}+m \dot{\bar{r}} \cdot \ddot{\vec{r}}=\frac{K}{r^{2}} \dot{r}+\dot{\bar{r}} \cdot\left(-\frac{K}{r^{2}} \hat{r}\right) \equiv 0 \tag{16}
\end{equation*}
$$

by virtue of $\hat{r} \cdot \dot{\bar{r}} \equiv \dot{r}$, as noted already. Now consider the point at which the trajectory is closest to the origin. Call this smallest distance $r_{m}$. At this point we have $\dot{r}=0$ and so the velocity is purely in the $\theta$-direction. The kinetic energy at this point is therefore $\frac{m}{2}|\dot{r}|^{2} \rightarrow \frac{L^{2}}{2 m r_{m}^{2}}$ and so,

$$
\begin{equation*}
E=-\frac{K}{r_{m}}+\frac{L^{2}}{2 m r_{m}^{2}} \tag{17}
\end{equation*}
$$

All terms in (17) are constants, and hence it forms an equation from which to determine $r_{m}$. The solution(s) is/are,

$$
\begin{equation*}
r_{m}=-\frac{K}{2 E} \pm \sqrt{\left(\frac{K}{2 E}\right)^{2}+\frac{L^{2}}{2 m E}} \tag{18}
\end{equation*}
$$

Only solutions with positive radial coordinates are physical. Consequently we can distinguish two cases,
(i) $E>0$ : In this case the first term in (18) is negative and so the positive option must be chosen for the $\pm$ sign, which yields a unique, positive outcome for $r_{m}$;

$$
\begin{equation*}
r_{m}=-\frac{K}{2 E}+\sqrt{\left(\frac{K}{2 E}\right)^{2}+\frac{L^{2}}{2 m E}} \tag{19}
\end{equation*}
$$

(ii) $E<0$ : In this case the first term in (18) is positive and so either option can be chosen for the $\pm$ sign since both result in a positive outcome for $r_{m}$. Writing these two possibilities explicitly,

$$
\begin{equation*}
r_{m}=\frac{K}{2|E|}-\sqrt{\left(\frac{K}{2 E}\right)^{2}-\frac{L^{2}}{2 m|E|}} \quad \text { and } \quad r_{M}=\frac{K}{2|E|}+\sqrt{\left(\frac{K}{2 E}\right)^{2}-\frac{L^{2}}{2 m|E|}} \tag{20}
\end{equation*}
$$

Now we realise that the condition $\dot{r}=0$ which we have used to derive these results is not strictly the condition for a minimum distance, but rather for any turning point of $r$. The two results in (20) are respectively the minimum and maximum distance of the trajectory from the origin. The existence of a finite maximum distance implies that the trajectory is a closed orbit. In contrast, case (i) yields no such maximum and hence corresponds to an open trajectory in which the body escapes to infinity. These conclusions are consistent with the energy requirement for the two cases. In case (ii) escape to infinity is clearly impossible if the total energy is negative ${ }^{3}$. Rather less obviously, case (i) establishes that positive total energy is sufficient to ensure escape. Since we have already deduced that the shape of the trajectories are the conic sections, we can conclude that case (i) gives parabolic or hyperbolic trajectories, and case (ii) elliptic or circular orbits.
To derive the magnitude of $\bar{M}$ in terms of the other two constants of the motion, $L$ and $E$, note that the momentum when $r=r_{m}$ is $p=L / r_{m}$ and so,

$$
\begin{equation*}
\bar{M}=\left(\frac{L^{2}}{m r_{m}}-K\right) \hat{x} \Rightarrow M=\frac{L^{2}}{m r_{m}}-K \tag{21}
\end{equation*}
$$

[^1]and $r_{m}$ is known in terms of $L$ and $E$ from (19) or (20). Note that $\bar{M}$ always points from the origin to the perihelion (the point of closest approach, $r_{m}$ ). [The RHS of (21) can be shown to always be greater than or equal to zero by making use of (19) and (20)].

Finally, whilst we have derived the trajectory in the form $\frac{1}{r}=\lambda(M \cos \theta+K)$ we have yet to find the scale factor, $\lambda$. But this is now trivial since $\frac{1}{r_{m}}=\lambda(M+K)=\lambda \cdot \frac{L^{2}}{m r_{m}}$, using (21), and hence $\lambda=\frac{m}{L^{2}}$. So finally we have obtained the general solution for the trajectory as,

$$
\begin{equation*}
\frac{1}{r}=\frac{m K}{L^{2}}+\left(\frac{1}{r_{m}}-\frac{m K}{L^{2}}\right) \cos \theta \tag{22}
\end{equation*}
$$

where $r_{m}$ is given by (19) or the first of (20). Thus the total energy and the angular momentum determine the solution.
Note that time has played no part in this derivation. The geometry of the trajectory can be obtained without considering the equation of motion. But this means, of course, that we have not derived the position or the velocity as a function of time. Nevertheless, the velocity is known as a function of position. Given $\theta$, then the radial coordinate $r$ follows from (22). Then the component $v_{\theta}$ of velocity follows from $L=m r v_{\theta}$ and the radial component of velocity follows from $E=-\frac{K}{r}+\frac{L^{2}}{2 m r^{2}}+\frac{m}{2} v_{r}^{2}$.

Note also that, for closed orbits, (20) and (22) give different but equivalent relations between the perihelion distance $\left(r_{m}\right)$ and the aphelion distance ( $r_{M}$ ),

$$
\begin{equation*}
r_{m}+r_{M}=\frac{K}{|E|} \quad \text { and } \quad \frac{1}{r_{m}}+\frac{1}{r_{M}}=\frac{2 m K}{L^{2}} \tag{23}
\end{equation*}
$$

Only Kepler's third law remains. It may seem that we will have a difficulty with this since it involves the period of the orbit, and we have not solved for the temporal dependencies. But actually it is no problem. We know that the rate of sweeping out area is just $L / 2 m$, so the period is just $\tau=2 m A / L$ where $A$ is the area of the orbit. Because we now know that we are dealing with an ellipse, the area is $A=\pi a b$ where the semi-major axis is $a=\left(r_{m}+r_{M}\right) / 2=K / 2|E|$, from (23), and the semi-minor axis can easily be found to be $b=L / \sqrt{2 m|E|}$. Hence,

$$
\begin{equation*}
\tau^{2}=\left(\frac{2 m A}{L}\right)^{2}=\left(\frac{2 m}{L} \pi \frac{K}{2|E|} \cdot \frac{L}{\sqrt{2 m|E|}}\right)^{2}=\frac{\pi^{2} m K^{2}}{2|E|^{3}}=\frac{\pi^{2} m}{2 K}(2 a)^{3}=4 \pi^{2} a^{3} \cdot \frac{m}{K} \tag{24}
\end{equation*}
$$

This establishes Kepler's third law, namely that $\tau^{2} / a^{3}$ is the same for all planets, only if $m / K$ is independent of the mass of the planet, $m$. This requires that the constant $K$ in the gravitation force equation, (6), is proportional to the planet's mass. But by symmetry the same must be true for the mass of the Sun. So Kepler's third law is seen
to be equivalent to requiring the gravitational force to be proportional to the product of the two masses,

$$
\begin{equation*}
K=G M m \tag{25}
\end{equation*}
$$

where $G$ is a universal gravitational constant. It is salutary to recall that our understanding of the nature of Newton's $G$ is essentially the same now as it was in 1687. Where our knowledge has advanced, thanks to Einstein, is in the appreciation that (25), which expresses the equivalence of gravitational and inertia mass, can be understood to be a consequence of general covariance and the geometrical interpretation of gravity.

So much for Kepler's laws, but what about the question posed at the start of this Chapter: given that the trajectories are conic sections, what is the cone? Of course we could always contrive a cone which fits any of the orbits. But such cones seem arbitrary and without meaning, and there would be a plethora of different cones for different trajectories. It would seem crazy to suggest that, in fact, there is one single cone whose sections provide every possible orbit, open and closed. This cannot be, you argue, since an ellipse of a given eccentricity could arise in any orientation in space - and all these ellipses obviously cannot be generated as sections of just one fixed cone. Quite true - so long as we confine ourselves to cones in 3D. But let us see what happens when we allow ourselves a fourth dimension. Now we can consider a cone whose right-sections are spheres rather than circles. Calling the extra dimension $t$ the cone is,

$$
\begin{equation*}
t^{2}=x^{2}+y^{2}+z^{2} \quad \text { for } t>0 \tag{26}
\end{equation*}
$$

Amazingly any orbit is a section of this cone - which is, of course, just the forward light cone. Consider the section with respect to the plane $z=0, t=$ constant. Clearly this is just a circle in the $(x, y)$ plane. But circular orbits in any plane can be generated in the same way since (26) involves spheres in the 3 -space $(x, y, z)$ and sections of spheres with respect to any plane are circles.
But how are other conic sections produced? To do so the plane which defines the section must be tilted so that the $t$-direction has a non-zero projection onto it. Suppose the plane is defined by $z=0$ and a normal vector $\hat{n}$ lying in the $(t, x)$ plane. If $\hat{n}$ is time-like then the orbits are ellipses. If $\hat{n}$ is space-like the trajectories are hyperbolae. If $\hat{n}$ is null the trajectory is a parabola. As with the circular orbits, the spatial orientation of the trajectory can be rotated arbitrarily owing to the spherical symmetry of (26). So, we can produce any conic section in the 3 -space $(x, y, z)$ by this means.

But what is going on here? Just what is the light cone, the epitome of relativistic constructs, doing popping up in a non-relativistic problem? You may very well ask! These matters run deep. The interested reader should seek a more competent guide, such as Guowu Meng (2011) who discovered this simple interpretation of the Kepler conic sections, or Guillemin and Sternberg (1990). However, we can give some clue as to what might have brought the light cone into the problem. It is down to the magic of the Runge-Lenz vector.
Recall that the Poisson bracket is defined by,

$$
\begin{equation*}
\{f, g\} \equiv \sum_{i=1}^{3}\left[\frac{\partial f}{\partial x_{i}} \frac{\partial g}{\partial p_{i}}-\frac{\partial f}{\partial p_{i}} \frac{\partial g}{\partial x_{i}}\right] \tag{27}
\end{equation*}
$$

We already know what this will yield for the components of the angular momentum vector,

$$
\begin{equation*}
\left\{L_{j}, L_{k}\right\}=\epsilon_{j k n} L_{n} \tag{28}
\end{equation*}
$$

where $\epsilon_{j k n}$ is the alternating tensor. With a little tedious algebra the definition of the Runge-Lenz vector, (5), leads to,
and,

$$
\begin{equation*}
\left\{M_{j}, L_{k}\right\}=\epsilon_{j k n} M_{n} \tag{29}
\end{equation*}
$$

$$
\begin{equation*}
\left\{M_{j}, M_{k}\right\}=-\frac{2 E}{m} \epsilon_{j k n} L_{n} \tag{30}
\end{equation*}
$$

Instead of deriving (28-30) via the Poisson bracket we could equivalently have converted the momentum to the quantum mechanical operator by the replacement $\bar{p} \rightarrow i \hbar \bar{\nabla}$, so that $\bar{L}$ and $\bar{M}$ also become operators [noting that (5) has to be replaced with the symmetrised form $\bar{M}=(\bar{p} \times \bar{L}-\bar{L} \times \bar{p}) / 2 m-K \hat{r}]$. Equs.(28-30) would then hold for the commutators between these operators, with the minor changes that each RHS would pick up a factor of $i \hbar$ and the term $E$ on the RHS of (30) would be replaced by the Hamiltonian operator, $\hat{H}=\frac{(i \hbar \bar{\nabla})^{2}}{2 m}-\frac{K}{r}$. Even in the quantum case, though, we may confine attention to a sub-space of states with the same energy, $E$, to avoid this latter complication.
In the case of bound states (elliptic orbits) with $E<0$, we can adopt a normalised Runge-Lenz vector (or operator) such that,

$$
\begin{equation*}
\bar{M}^{\prime}=\sqrt{-\frac{m}{2 E}} \cdot \bar{M} \tag{31}
\end{equation*}
$$

So that (28-30) become, adopting the quantum mechanical operator interpretation now,
and,

$$
\begin{align*}
& \left\lfloor L_{j}, L_{k}\right\rfloor=i \hbar \epsilon_{j k n} L_{n}  \tag{32}\\
& \left\lfloor M_{j}^{\prime}, L_{k}\right\rfloor=i \hbar \epsilon_{j k n} M_{n}^{\prime}  \tag{33}\\
& \left\lfloor M_{j}^{\prime}, M_{k}^{\prime}\right\rfloor=i \hbar \in_{j k n} L_{n} \tag{34}
\end{align*}
$$

In (32-34) the reader may recognise the commutators of the six generators of the group of rotations in 4-dimensional Euclidean space, the group $S O(4)$. If not, the generators of this group can be represented by $\hat{L}_{\alpha \beta}=x_{\alpha} \hat{p}_{\beta}-x_{\beta} \hat{p}_{\alpha}$ with the usual commutator $\left\lfloor x_{\alpha}, p_{\beta}\right\rfloor=i \hbar \delta_{\alpha \beta}$ and where the Greek subscripts here span the range $[1,4]$. The reader can check for himself that these generators reproduce the commutation structure of (32-34), specifically with $L_{i}=\epsilon_{i j k} \hat{L}_{j k} / 2$ and $M_{i}^{\prime}=\hat{L}_{i 4}$ (Latin subscripts taking values $1,2,3$ only).
In the case of parabolic or hyperbolic trajectories, with $E>0$, the normalisation of the Runge-Lenz vector (or operator) is,

$$
\begin{equation*}
\bar{M}^{\prime}=\sqrt{\frac{m}{2 E}} \cdot \bar{M} \tag{35}
\end{equation*}
$$

This leads to a minus sign on the RHS of (34) and the resulting commutator structure indicates that the group has changed from $S O(4)$, for the bound states, to $S O(3,1)$ for the free trajectories. The significance of this is that $\operatorname{SO}(3,1)$ is just the connected part of the homogeneous Lorentz group. This is the result that we have been working towards. The suggestion is that it is the algebraic structure of the $\bar{L}$ and $\bar{M}$ vectors (or operators) which creates the geometrical link with relativity - despite the Kepler problem itself being non-relativistic.
The rather stunning conclusion that the Kepler trajectories are conic sections because they are sections of the forward light cone is matched by the equally elegant observation that this fact was always lurking in the magic of the Runge-Lenz vector.

## References

J.Bertrand $\mathbf{J}$ (1873) "Théorème relatif au mouvement d'un point attiré vers un centre fixe", C.R. Acad. Sci. Paris 77849

F C Santos, V Soares and A C Tort, "An English translation of Bertrand's theorem", Lat.Am. J. Phys. Educ. 5 (2011) 684.
I.Newton (1687) "Philosophiae Naturalis Principia Mathematica", (Royal Society of London).

Guowu Meng (2011): "Generalized Kepler Problems I: Without Magnetic Charges", arXiv:1104.2585
Victor Guillemin and Shlomo Sternberg (1990): "Variations on a Theme by Kepler", Providence, R.I., American Mathematical Society.

This document was created with Win2PDF available at http://www.win2pdf.com. The unregistered version of Win2PDF is for evaluation or non-commercial use only. This page will not be added after purchasing Win2PDF.


[^0]:    ${ }^{1}$ That is, in the derivation of the energy levels of the hydrogen atom. Just as in this Chapter the trajectories are found without directly solving the equations of motion, so in Chapter ? we shall show how the energy levels of the hydrogen atom can be found without solving the Schrodinger equation, in both cases the magic being provided by the Runge-Lenz vector, $\bar{M}$.
    ${ }^{2}$ The Runge-Lenz vector appears to bear this name only because it was discussed by these authors in the early $20^{\text {th }}$ century and thence picked up by Pauli who attached their name to it. However it was first discovered in primitive from by Jakob Hermann in 1710 and found its way via Bernoulli to Laplace later the same century. It was rediscovered several times in the next two centuries.

[^1]:    ${ }^{3}$ This is because, at infinity, the potential energy is zero so the kinetic energy equals the total energy. But kinetic energy cannot be negative.

