Teaching the Kepler laws for freshmen.

Maris van Haandel[*]
RSG Pantarijn, Wageningen
marisvanhaandel@wanadoo.nl

Gert Heckman
IMAPP, Radboud Universiteit, Nijmegen
G.Heckman@math.ru.nl

February 11, 2013

Abstract

We present a natural proof of Kepler's law of ellipses in the spirit of Euclidean geometry. Moreover we discuss two existing Euclidean geometric proofs, one by Feynman in hist Lost Lecture from 1964 and the other by Newton in the Principia of 1687.

1   Introduction.

One of the highlights of classical mechanics is the mathematical derivation of the three experimentally observed Kepler laws of planetary motion from Newton's laws of motion and of gravitation. Newton published his theory of gravitation in 1687 in the Principia Mathematica [8]. After two short introductions, one with definitions and the other with axioms (the laws of motion), Newton discussed the Kepler laws in the first three sections of Book 1 (in just 40 pages, without ever mentioning the name of Kepler!)

Kepler's second law (motion is planar and equal areas are swept out in equal times) is an easy consequence of the conservation of angular momentum

---

[*]Supported by NWO.

$\mathbf{L} = \mathbf{r} \times \mathbf{p}$, and holds in greater generality for any central force field. All this is explained well by Newton in Propositions 1 and 2.

Kepler's first law (planetary orbits are ellipses with the center of the force field at a focus) however is peculiar for the attractive $1/r^2$ force field. Using Euclidean geometry Newton derives in Proposition 11 that the Kepler laws can only hold for an attractive $1/r^2$ force field. The reverse statement that an attractive $1/r^2$ force field leads to elliptical orbits Newton concludes in Corollary 1 of Proposition 13. Tacitly he assumes for this argument that the equation of motion $\mathbf{F} = m\mathbf{a}$ has a unique solution for given initial position and initial velocity. Theorems about existence and uniqueness of solutions of such a differential equation have only been formulated and mathematically rigorously been proven in the 19th century. However there can be little doubt that Newton did grasp these properties of his equation $\mathbf{F} = m\mathbf{a}$ [1].

Somewhat later in 1710 Jakob Hermann and Johan Bernoulli gave a direct proof of Kepler's first law, which is still the standard proof for modern text books on classical mechanics [9]. One writes the position vector $\mathbf{r}$ in the plane of motion in polar coordinates r and θ. The trick is to transform the equation of motion $m\mathbf{a} = -k\mathbf{r}/r^3$ with variable the time t into a second order differential equation of the scalar function $u = 1/r$ with variable the angle θ. This differential equation can be exactly solved, and yields the equation of an ellipse in polar coordinates [3].

Another popular proof goes by writing down the so called Laplace-Runge-Lenz vector

$$\mathbf{K} = \mathbf{p} \times \mathbf{L} - km\mathbf{r}/r$$

with $\mathbf{p} = m\mathbf{v}$ momentum and $\mathbf{F}(\mathbf{r}) = -k\mathbf{r}/r^3$ the force field of the Kepler problem. The LRL vector $\mathbf{K}$ turns out to be conserved, i.e. $\dot{\mathbf{K}} = \mathbf{0}$. This result can be derived by a direct computation as indicated in Section 2. An alternative geometric argument is sketched in Section 3. Working out the equation $\mathbf{r} \cdot \mathbf{K} = rK \cos\theta$ gives the equation of an ellipse in polar coordinates [3]. The geometric meaning of the LRL vector becomes clear only a posteriori as a vector pointing in the direction of the long axis of the ellipse. But the start of the proof to write down the LRL vector remains a trick. For a historical account of the LRL vector we refer to Goldstein [4].

The purpose of this note is to present in Section 2 a proof of the Kepler laws for which a priori the reasoning is well motivated in both physical and geometric terms. In Section 3 we review the hodographic proof as given by Feynman in his "Lost Lecture" [5], and in Section 4 we discuss Newton's

2

proof from the Principia[8]. All three proofs are based on Euclidean geometry, although we do use the language of vector calculus in order to make it more readable for people of the 21st century. We feel that our proof is really the simplest of the three, and at the same time it gives more refined information (namely the value of the long axis $2a = -k/H$ of the ellips $\mathcal{E}$). In fact we think that our proof in Section 2 can compete equally well both in transparency and in level of computation with the standard proof of Jakob Hermann and Johann Bernoulli, making it an appropriate alternative to present in a freshman course on classical mechanics.

We like to thank Hans Duistermaat and Arnoud van Rooij for useful comments on this article.

2 A Euclidean proof of Kepler's first law.

We shall use inner products $\mathbf{u} \cdot \mathbf{v}$ and outer products $\mathbf{u} \times \mathbf{v}$ of vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with the compatibility conditions

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$$

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$$

and the Leibniz product rules

$$(\mathbf{u} \cdot \mathbf{v})^{\cdot} = \dot{\mathbf{u}} \cdot \mathbf{v} + \mathbf{u} \cdot \dot{\mathbf{v}}$$

$$(\mathbf{u} \times \mathbf{v})^{\cdot} = \dot{\mathbf{u}} \times \mathbf{v} + \mathbf{u} \times \dot{\mathbf{v}}$$
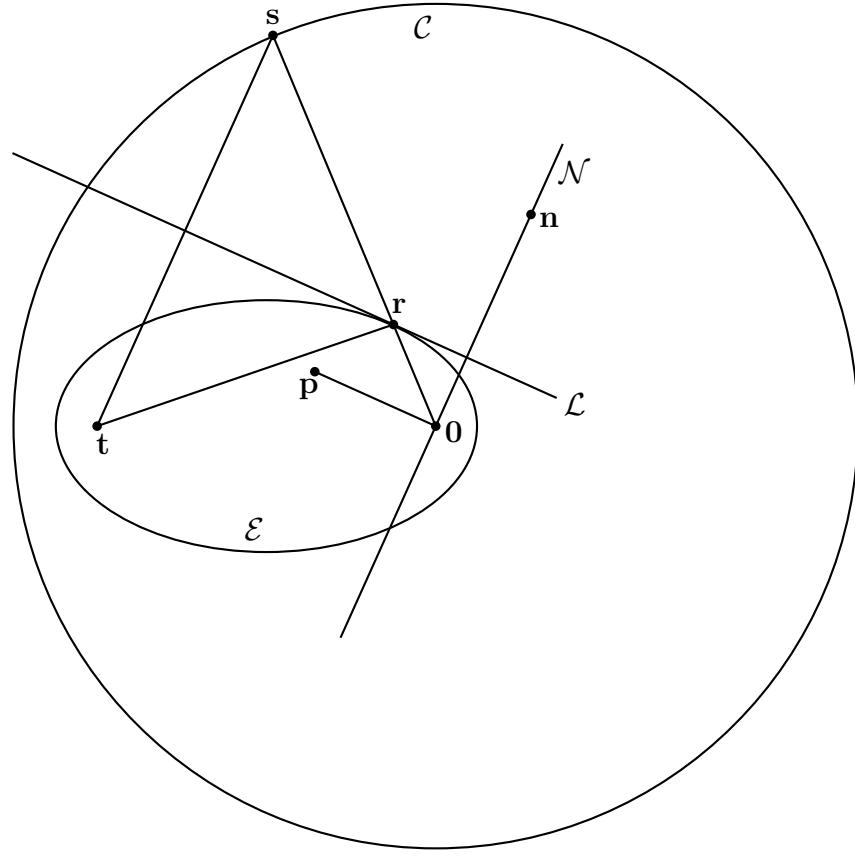
without further explanation.

For a central force field $\mathbf{F}(\mathbf{r}) = f(\mathbf{r})\mathbf{r}/r$ the angular momentum vector $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is conserved by Newton's law of motion $\mathbf{F} = \dot{\mathbf{p}}$, thereby leading to Kepler's second law. For a spherically symmetric central force field $\mathbf{F}(\mathbf{r}) = f(r)\mathbf{r}/r$ the energy

$$H = p^2/2m + V(r) \ , \ V(r) = - \int f(r)\,dr$$

is conserved as well. These are the general initial remarks.

From now on consider the Kepler problem $f(r) = -k/r^2$ en $V(r) = -k/r$ with $k > 0$ a coupling constant. For fixed energy $H < 0$ the motion is bounded inside a sphere with center $\mathbf{0}$ and radius $-k/H$. Consider the following picture of the plane perpendicular to $\mathbf{L}$.

The circle $\mathcal{C}$ with center $\mathbf{0}$ and radius $-k/H$ is the boundary of a disc where motion with energy $H < 0$ takes place. Let $\mathbf{s} = -k\mathbf{r}/rH$ be the projection of $\mathbf{r}$ from the center $\mathbf{0}$ on this circle $\mathcal{C}$. The line $\mathcal{L}$ through $\mathbf{r}$ with direction vector $\mathbf{p}$ is the tangent line of the orbit $\mathcal{E}$ at position $\mathbf{r}$ with velocity $\mathbf{v}$. Let $\mathbf{t}$ be the orthogonal reflection of the point $\mathbf{s}$ in the line $\mathcal{L}$.

Theorem. *The point $\mathbf{t}$ equals $\mathbf{K}/mH$ and therefore is conserved.*

Proof. The line $\mathcal{N}$ spanned by $\mathbf{n} = \mathbf{p} \times \mathbf{L}$ is perpendicular to $\mathcal{L}$. The point $\mathbf{t}$ is obtained from $\mathbf{s}$ by subtracting twice the orthogonal projection of $\mathbf{s} - \mathbf{r}$ on the line $\mathcal{N}$, and therefore

$$\mathbf{t} = \mathbf{s} - 2((\mathbf{s} - \mathbf{r}) \cdot \mathbf{n})\mathbf{n}/n^2.$$

Now

$$\mathbf{s} = -k\mathbf{r}/rH$$

4

$$(\mathbf{s} - \mathbf{r}) \cdot \mathbf{n} = -(H + k/r)\mathbf{r} \cdot (\mathbf{p} \times \mathbf{L})/H = -(H + k/r)L^2/H$$

$$n^2 = p^2 L^2 = 2m(H + k/r)L^2$$

and therefore

$$\mathbf{t} = -k\mathbf{r}/rH + \mathbf{n}/mH = \mathbf{K}/mH$$

with $\mathbf{K} = \mathbf{p} \times \mathbf{L} - km\mathbf{r}/r$ the LRL vector. The final step $\dot{\mathbf{K}} = \mathbf{0}$ is derived by a straightforward computation using the compatibility relations and the Leibniz product rules for inner and outer products of vectors in $\mathbb{R}^3$. □

Corollary. *The orbit $\mathcal{E}$ is an ellipse with foci $\mathbf{0}$ and $\mathbf{t}$, and long axis equal to $2a = -k/H$.*

Proof. Indeed we have

$$|\mathbf{t} - \mathbf{r}| + |\mathbf{r} - \mathbf{0}| = |\mathbf{s} - \mathbf{r}| + |\mathbf{r} - \mathbf{0}| = |\mathbf{s} - \mathbf{0}| = -k/H.$$

Hence $\mathcal{E}$ is an ellipse with foci $\mathbf{0}$ and $\mathbf{t}$, and long axis $2a = -k/H$. □

The above proof has two advantages over the earlier mentioned proofs of Kepler's first law. The conserved vector $\mathbf{t} = \mathbf{K}/mH$ is a priori well motivated in geometric terms. Moreover we use the gardener's definition of an ellipse. The equation of an ellipse in polar coordinates is unknown to most freshmen, and so additional explanation would be needed for that. Another advantage of our proof is that the solution of the equation of motion is achieved by just finding enough constants of motion. The proofs by Feynman and Newton in the next sections on the contrary rely at a crucial point on the existence and uniqueness theorem of differential equations.

We proceed to derive Kepler's third law along standard lines [3]. The ellipse $\mathcal{E}$ has numerical parameters[1] $a, b, c > 0$ given by $2a = -k/H$, $4c^2 = K^2/m^2 H^2 = (2mHL^2 + m^2 k^2)/m^2 H^2$ and $a^2 = b^2 + c^2$. The area of the region bounded by $\mathcal{E}$ equals

$$\pi a b = LT/2m$$

with T the period of the orbit. Indeed L/2m is the area of the sector swept out by the position vector $\mathbf{r}$ per unit time. A straightforward calculation gives

$$T^2/a^3 = 4\pi^2 m/k.$$

---

[1] The long axis equals $2a$, the short axis $2b$, and $a^2 = b^2 + c^2$.

Newton's law of gravitation states that the coupling constant k is proportional to the product of the mass m of the planet and the mass M of the sun, and we find Kepler's third (harmonic) law stating that T^2/a^3 is the same for all planets[2].

3   Feynman's proof of Kepler's first law.

In this section we discuss a different geometric proof of Kepler's first law based on the hodograph $\mathcal{H}$. By definition $\mathcal{H}$ is the curve traced out by the velocity vector $\mathbf{v}$ in the Kepler problem. This proof goes back to Möbius in 1843 and Hamilton in 1845 [2] and has been forgotten and rediscovered several times, among others by Feynman in 1964 in his "Lost Lecture" [5].

Let us assume (as in the picture of the previous section) that $i v \mathbf{n}/n = \mathbf{v}$ with i the counterclockwise rotation around $\mathbf{0}$ over $\pi/2$. So the orbit $\mathcal{E}$ is assumed to be traversed counterclockwise around the origin $\mathbf{0}$.

Theorem. *The hodograph $\mathcal{H}$ is a circle with center $\mathbf{c} = i\mathbf{K}/mL$ and radius k/L*

Proof. We shall indicate two proofs of this theorem. The first proof is analytic in nature, and uses conservation of the LRL vector $\mathbf{K}$ by rewriting

$$\mathbf{K} = \mathbf{p} \times \mathbf{L} - km\mathbf{r}/r = mvL\mathbf{n}/n - km\mathbf{r}/r$$

as

$$v\mathbf{n}/n = \mathbf{K}/mL + k\mathbf{r}/rL \ ,$$

or equivalently

$$\mathbf{v} = i\mathbf{K}/mL + ik\mathbf{r}/rL \ .$$

Hence the theorem follows from $\dot{\mathbf{K}} = \mathbf{0}$.

There is a different geometric proof of the theorem, which as a corollary gives the conservation of the LRL vector $\mathbf{K}$. The key point is to reparametrize the velocity vector $\mathbf{v}$ from time t to angle θ of the position vector $\mathbf{r}$. It turns

---

[2]The mass m we have used so far is in fact equal to the reduced mass $mM/(m + M)$, and this almost equals m if $m \ll M$. The coupling constant k is according to Newton equal to GmM with G the universal gravitational constant. We therefore see that Kepler's third law holds only approximately for $m \ll M$.

out that the vector $\mathbf{v}(\theta)$ is traversing the hodograph $\mathcal{H}$ with constant speed k/L. Indeed we have from Newton's laws

$$m\frac{d\mathbf{v}}{d\theta} = m\mathbf{a}\frac{dt}{d\theta} = -\frac{k\mathbf{r}}{r^3}\frac{dt}{d\theta}$$

and Kepler's second law gives

$$r^2 d\theta/2 = Ldt/2m \ .$$

Combining these identities yields

$$\frac{d\mathbf{v}}{d\theta} = -k\mathbf{r}/rL \ ,$$

so indeed $\mathbf{v}(\theta)$ travels along $\mathcal{H}$ with constant speed k/L. Since $\mathbf{r} = re^{i\theta}$ a direct integration yields
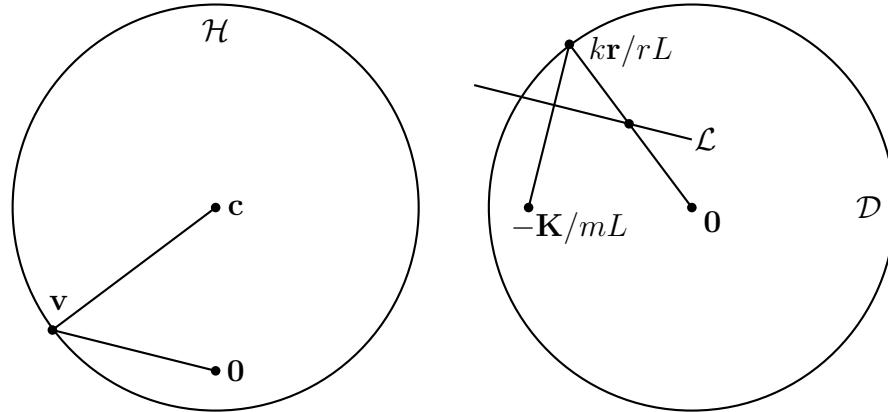
$$\mathbf{v}(\theta) = \mathbf{c} + ik\mathbf{r}/rL \ , \ \dot{\mathbf{c}} = \mathbf{0}$$

and the hodograph becomes a circle with center $\mathbf{c}$ and radius k/L. Comparison with the last formula in the first proof gives

$$\mathbf{c} = i\mathbf{K}/mL$$

and $\dot{\mathbf{K}} = \mathbf{0}$ comes out as a corollary. $\qquad\qquad\square$

All in all, the circular nature of the hodograph $\mathcal{H}$ is more or less equivalent to the conservation of the LRL vector $\mathbf{K}$.

Now turn the hodograph $\mathcal{H}$ clockwise around $\mathbf{0}$ over $\pi/2$ and translate over $i\mathbf{c} = -\mathbf{K}/mL$. This gives a circle $\mathcal{D}$ with center $\mathbf{0}$ and radius k/L. Since

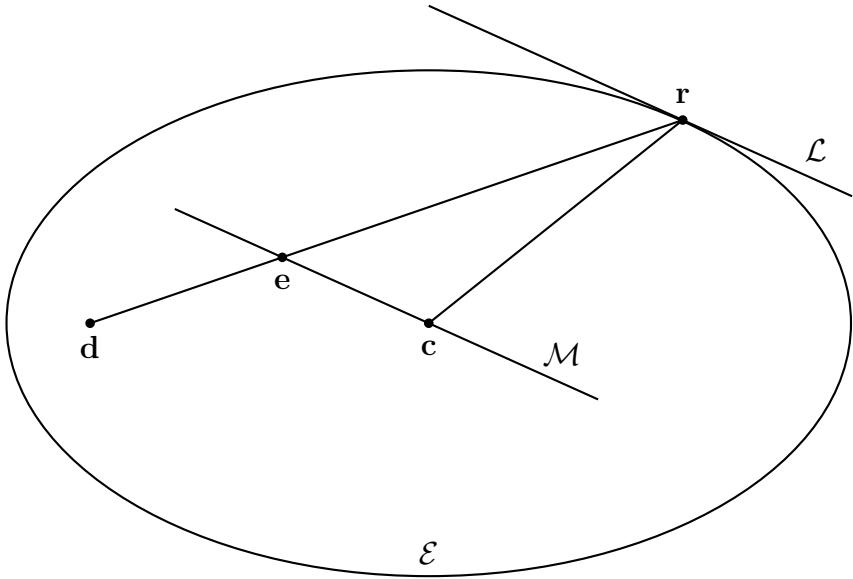$$k\mathbf{r}/rL + \mathbf{K}/mL = v\mathbf{n}/n = -i\mathbf{v}$$

the orbit $\mathcal{E}$ intersects the line through $\mathbf{0}$ and $k\mathbf{r}/rL$ in a point with tangent line $\mathcal{L}$ perpendicular to the line through $k\mathbf{r}/rL$ and $-\mathbf{K}/mL$. For example the ellipse $\mathcal{F}$ with foci $\mathbf{0}$ and $-\mathbf{K}/mL$ and long axis equal to k/L has this property, but any scalar multiple $\lambda\mathcal{F}$ with $\lambda > 0$ equally does so. Since curves with the above property are uniquely charcterized once an initial point on the curve is chosen we conclude that $\mathcal{E} = \lambda\mathcal{F}$ for some $\lambda > 0$. This proves Kepler's first law. A comparison with the picture of the previous section shows that $\mathcal{E} = \lambda\mathcal{F}$ with $\lambda = -L/H$. Indeed $\mathcal{E}$ has foci $\mathbf{0}$ and $-\lambda\mathbf{K}/mL = \mathbf{K}/mH = \mathbf{t}$ and the long axis is equal to $\lambda k/L = -k/H = 2a$.

It is not clear to us if Feynman was aware of his use of the existence and uniqueness theorem for differential equations. On page 164 of [5] he writes: "Therefore, the solution to the problem is an ellipse - or the other way around, really, is what I proved: that the ellipse is a possible solution to the problem. And it is this solution. So the orbits are ellipses."

Apparently Feynman had trouble following Newton's proof of Kepler's first law. On page 111 of [5] the authors write: "In Feynman's lecture, this is the point at which he finds himself unable to follow Newton's line of argument any further, and so sets out to invent one of his own".

4   Newton's proof of Kepler's first law.

In this section we discuss the original proof by Newton of Kepler's first law as given in [8]. The proof starts with a nice general result.

Theorem. *Let $\mathcal{E}$ be a smooth closed curve bounding a convex region containing two points $\mathbf{c}$ and $\mathbf{d}$. Let $\mathbf{r}(t)$ traverse the curve $\mathcal{E}$ counterclockwise in time t, such that the areal speed with respect to the point $\mathbf{c}$ is constant. Likewise let $\mathbf{r}(s)$ traverse the curve $\mathcal{E}$ counterclockwise in time s, such that the areal speed with respect to the point $\mathbf{d}$ is equal to the same constant.*

Let $\mathcal{L}$ be the tangent line to $\mathcal{E}$ at the point $\mathbf{r}$, and let $\mathbf{e}$ be the intersection point of the line $\mathcal{M}$, which is parallel to $\mathcal{L}$ through the point $\mathbf{c}$, and the line through the points $\mathbf{r}$ and $\mathbf{d}$. Then the ratio of the two accelerations is given by

$$|\frac{d^2\mathbf{r}}{ds^2}| : |\frac{d^2\mathbf{r}}{dt^2}| = |\mathbf{r} - \mathbf{e}|^3 : |\mathbf{r} - \mathbf{c}| \cdot |\mathbf{r} - \mathbf{d}|^2.$$

Proof. Using the chain rule we get

$$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt} \cdot \frac{dt}{ds}$$

$$\frac{d^2\mathbf{r}}{ds^2} = \frac{d^2\mathbf{r}}{dt^2} \cdot (\frac{dt}{ds})^2 + \frac{d\mathbf{r}}{dt} \cdot \frac{d^2t}{ds^2} .$$

Because $d^2\mathbf{r}/dt^2$ is proportional to $\mathbf{c} - \mathbf{r}$ and likewise $d^2\mathbf{r}/ds^2$ is proportional

9

to $\mathbf{d} - \mathbf{r}$ we see that

$$|\frac{d^2\mathbf{r}}{ds^2}| : |\frac{d^2\mathbf{r}}{dt^2}| = (\frac{dt}{ds})^2 \cdot |\frac{d^2\mathbf{r}}{dt^2} + \frac{d\mathbf{r}}{dt} \cdot \frac{d^2t}{ds^2}/(\frac{dt}{ds})^2| : |\frac{d^2\mathbf{r}}{dt^2}| = (\frac{dt}{ds})^2 \cdot |\mathbf{r} - \mathbf{e}| : |\mathbf{r} - \mathbf{c}| \ .$$

Since the curve $\mathcal{E}$ is traversed with equal areal speed relative to the two points $\mathbf{c}$ and $\mathbf{d}$ we get

$$|\frac{d\mathbf{r}}{dt}| \cdot |\mathbf{r} - \mathbf{e}| = |\frac{d\mathbf{r}}{ds}| \cdot |\mathbf{r} - \mathbf{d}|$$

and therefore also

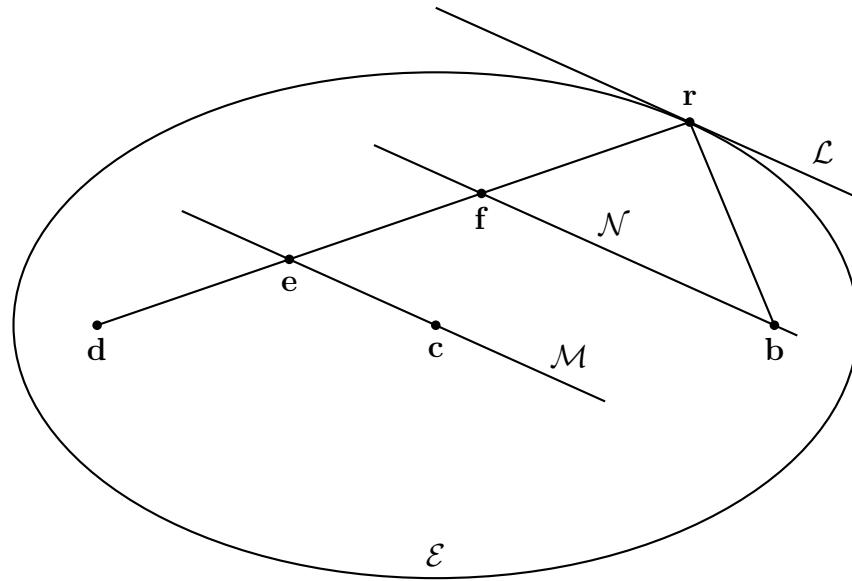$$\frac{dt}{ds} = |\mathbf{r} - \mathbf{e}| : |\mathbf{r} - \mathbf{d}| \ .$$

In turn this implies that

$$|\frac{d^2\mathbf{r}}{ds^2}| : |\frac{d^2\mathbf{r}}{dt^2}| = (\frac{dt}{ds})^2 \cdot |\mathbf{r} - \mathbf{e}| : |\mathbf{r} - \mathbf{c}| = |\mathbf{r} - \mathbf{e}|^3 : |\mathbf{r} - \mathbf{c}| \cdot |\mathbf{r} - \mathbf{d}|^2$$

which proves the theorem. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

We shall apply this theorem in case $\mathcal{E}$ is an ellipse with center $\mathbf{c}$ and focus $\mathbf{d}$. Assume that $\mathbf{r}(\mathbf{t})$ traverses the ellipse $\mathcal{E}$ in harmonic motion, say

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{c} - \mathbf{r} \ .$$

Let $\mathbf{b}$ be the other focus of $\mathcal{E}$, and let $\mathbf{f}$ be the intersection point of the line $\mathcal{N}$, passing through $\mathbf{b}$ and parallel to $\mathcal{L}$, with the line through the points $\mathbf{d}$ and $\mathbf{r}$. Then we find

$$|\mathbf{d} - \mathbf{e}| = |\mathbf{e} - \mathbf{f}| \ , \ |\mathbf{f} - \mathbf{r}| = |\mathbf{b} - \mathbf{r}|$$

which in turn implies that $|\mathbf{e} - \mathbf{r}|$ is equal to the half long axis a of the ellipse $\mathcal{E}$. We conclude from the formula in the above theorem that the motion in time s along an ellipse with constant areal speed with respect to a focus is only possible in an attractive inverse square force field. The converse statement that an inverse square force field (for negative energy H) indeed yields ellipses as orbits follows from existence and uniqueness theorems for solutions of Newton's equation $\mathbf{F} = m\mathbf{a}$ and the above reasoning. This is Newton's line of argument for proving Kepler's first law.

5 Conclusion.

There exist other proofs of Kepler's law of ellipses from a higher view point. One such proof by Arnold uses complex analysis, and is somewhat reminiscent to Newton's proof in Section 4 by comparing harmonic motion with motion under an $1/r^2$ force field [1]. Another proof by Moser is also very elegant, and uses the language of symplectic geometry and canonical transformations [6],[7]. However our goal here has been to give a proof which is as basic as possible, and at the same time is well motivated in terms of Euclidean geometry.

It is hard to exaggerate the importance of the role of the Principia Mathematica in the history of science. The year 1687 marks the birth of both modern mathematical analysis and modern theoretical physics. As such the derivation of the Kepler laws from Newton's law of motion and law of universal gravitation is a rewarding subject to teach to freshmen students. In fact we were motivated for our work, because we plan to teach this material to high school students in their final grade. Of course, the high school students first need to get acquainted with the basics of vector geometry and vector calculus. But once this is achieved there is nothing in the way of understanding our proof of Kepler's law of ellipses.

For physics or mathematics freshmen students in the university who are already familiar with vector calculus our proof in Section 2 is fairly short and

geometrically well motivated. In our opinion of all proofs it qualifies best to be discussed in a freshman course on mechanics.

References

[1] V.I. Arnold, Huygens and Barrow, Newton and Hooke, Birkhauser, Boston, 1990.

[2] D. Derbes, Reinventing the wheel: Hodographic solutions to the Kepler problems, Am. J. Phys. 69 (4), April 2001, pp 481-489.

[3] H. Goldstein, Classical Mechanics, Addison-Wesley, 1980 (2nd edition).

[4] H. Goldstein, Prehistory of the "Runge-Lenz" vector, Am. J. Phys. 43 (8), August 1975, pp 737-738.

   H. Goldstein, More on the prehistory of the Laplace or Runge-Lenz vector, Am. J. Phys. 44 (11), November 1976, pp 1123-1124.

[5] D.L. Goodstein and J.R. Goodstein, Feynman's Lost Lecture: The motion of planets around the sun, Norton, 1996.

[6] V. Guillemin and S. Sternberg, Variations on a theme of Kepler, Colloquium Publications AMS, volume 42, 1990.

[7] J. Moser, Regularization of the Kepler problem and the averaging method on a manifold, Comm. Pure Appl. Math. 23, 1970, pp 609-636.

[8] I. Newton, Principia Mathematica, New translation by I.B. Cohen and A. Whitman, University of California Press, Berkeley, 1999.

[9] D. Speiser, The Kepler problem from Newton to Johann Bernoulli, Archive for History of Exact Sciences 50 (2), August 1996, pp 103-116.