Kepler's laws derived from Newton's

ANDRZEJ DERDZINSKI

JULY 24, 2014

ABSTRACT. Questions raised during Boming Jia's talk on June 26th concerned planetary orbits that are hyperbolas or parabolas and the precise meaning of the statement that the inverse-square law is equivalent to Kepler's laws. The goal of today's talk is to clarify those issues. No knowledge of differential equations is required.

1. NOTATION

All vector and affine spaces are assumed finite-dimensional. We denote by $\langle\,,\rangle$ the inner product of a given Euclidean vector space Σ, by $|\;|$ the associated norm, and by $(\;)^{\cdot} = d/dt$ the differentiation with respect to the time t.

Given C^1 functions v, w of the variable t, valued in vector spaces Σ, Σ', and a vector-valued bilinear function B on $\Sigma \times \Sigma'$, the Leibniz rule reads

$$[B(v,w)]^{\cdot} = B(\dot v, w) + B(v, \dot w). \tag{1.1}$$

By a *trajectory* in an affine space A we mean any C^3 function $t \mapsto x = x(t) \in \mathrm{A}$ defined on an open interval in $\mathbb{R}$. If a trajectory $t \mapsto x(t)$ in a Euclidean vector space does not pass through 0, (1.1) yields $\langle x, x\rangle^{\cdot} = 2\langle x, \dot x\rangle$, and so

$$\dot r = \langle x, \dot x\rangle / r, \quad \text{where} \quad r = |x|, \tag{1.2}$$

as $r = \langle x, x\rangle^{1/2}$. We say that a trajectory $t \mapsto x(t)$ in a *vector* space is *radially accelerated* if $x(t)$ and $\ddot x(t)$ are linearly dependent for every t. In view of Newton's second law of dynamics,

$$m\ddot x(t) \text{ is the total force acting on the object at time } t, \tag{1.3}$$

for an object of mass m moving along the trajectory $t \mapsto x(t)$. Thus, radially-accelerated trajectories correspond to *central forces*, pushing the object in question towards or away from the fixed location 0.

For a radially-accelerated trajectory $t \mapsto x$ in a Euclidean vector space, let us set

$$\mathrm{E} = \mathrm{T} + \mathrm{U}, \quad \text{with } \mathrm{T} = \langle \dot x, \dot x\rangle/2 \text{ and } \mathrm{U} = \langle x, \ddot x\rangle. \tag{1.4}$$

Assuming that the object moving along the trajectory has unit mass, we call T, U and E its *kinetic, potential* and *total energies*. Since the trajectory is radially accelerated, $\langle \dot x, \ddot x\rangle x = \langle x, \dot x\rangle \ddot x$. With $r = |x|$, this gives, as a trivial consequence of (1.1),

$$\dot z = -\mathrm{E}\dot x, \quad \text{where} \quad z = \langle x, \dot x\rangle \dot x - \langle \dot x, \dot x\rangle x/2, \quad \text{so that} \quad |z| = r\langle \dot x, \dot x\rangle/2. \tag{1.5}$$

Let $\pm\mathrm{E} < 0$, with some sign $\pm$, for the total energy E of a radially-accelerated trajectory $t \mapsto x(t)$ in a Euclidean vector space. Then

(1.6) setting $y = z/\mathrm{E}$, we have $s \pm r = \pm\mathrm{U}r/\mathrm{E}$, where $r = |x|$ and $s = |y|$.

In fact, by (1.4) and (1.5), $\mathrm{U}r = (\mathrm{E}-\mathrm{T})r = \mathrm{E}r - r\langle\dot{x},\dot{x}\rangle/2 = \mathrm{E}r - |z|$, and so $|z| - \mathrm{E}r = -\mathrm{U}r$. Therefore, $s \pm r = |y| \pm |x| = \mp(|z| - \mathrm{E}r)/\mathrm{E} = \pm\mathrm{U}r/\mathrm{E}$.

The time-dependent vector z plays a crucial role in our discussion. For trajectories representing (non-radial) Newtonian planetary orbits, with 0 serving as both a focus and the location the sun, $z \neq 0$ at all times t, and z represents the direction which is either parallel to the line segment joining the planet's position x to the other focus (of an ellipse or hyperbola), or normal to the directrix (of a parabola). Geometrically, $2z$ arises, at times at which $\dot{x} \neq 0$, as the image of $\langle\dot{x},\dot{x}\rangle x$ under the reflecion in the line $\mathbb{R}\dot{x}$. In fact, the square-brackets summand in the decomposition $\langle\dot{x},\dot{x}\rangle x = [\langle\dot{x},\dot{x}\rangle x - \langle x,\dot{x}\rangle\dot{x}] + \langle x,\dot{x}\rangle\dot{x}$ is orthogonal to $\dot{x}$, and so the reflection is the result of replacing that summand with its opposite.

Applied to $v = x$ and $w = \dot{x}$ in $\Sigma' = \Sigma$, (1.1) gives

(1.7) $B(x,\dot{x})$ is constant for skew-symmetric B and radially-accelerated $t \mapsto x$,

which amounts to *conservation of angular momentum* in the case of central forces.

We refer to a trajectory in a vector space as *radial* if it lies entirely in line containing 0, but does not pass through 0. Every radial trajectory is radially accelerated.

***Remark* 1.1.** For a non-radial, radially-accelerated trajectory $t \mapsto x$ in a vector space Σ, not passing through zero, x and $\dot{x}$ are linearly independent at any t.

In fact, their linear independence for some fixed t implies the same for every t. (To see this, identify Σ with the space $\mathbb{R}^n$ of column vectors, complete $x(t),\dot{x}(t)$ to a basis $x(t),\dot{x}(t),e_3,\ldots,e_n$, and use (1.7) with $B = \det[\,\cdot\,,\,\cdot\,,e_3,\ldots,e_n]$.) If $x(t)$ and $\dot{x}(t)$ were now linearly *dependent* for some (hence every) t, writing $\dot{x} = \dot{\varphi}x$ for a suitable C^1 function $t \mapsto \varphi(t)$, we would get $\dot{w} = 0$, where $w(t) = e^{-\varphi(t)}x(t)$, and all $x(t)$ would lie in the line through 0 spanned by the constant vector w.

***Remark* 1.2.** By Remark 1.1, a non-radial, radially-accelerated trajectory $t \mapsto x(t)$ in a vector space, not passing through zero, has $\dot{x} \neq 0 \neq z$ at all times t, cf. (1.5).

Although the following lemma, as stated, deals with trajectories, it obviously remains true for (time-independent) vectors $x,\dot{x},u$, its geometric content being: if the difference of two unit vectors is nonzero and orthogonal to a vector $\dot{x} \neq 0$, then the unit vectors are each other's images under the reflecion in the line $\mathbb{R}\dot{x}$. See also the discussion of z in the lines preceding (1.7).

***Lemma* 1.3.** *If a non-radial, radially-accelerated trajectory $t \mapsto x(t) \smallsetminus \{0\}$ in a Euclidean vector plane Π has $\langle\dot{x}, u + x/r\rangle = 0$ and $u \neq -x/r$ at all times t for some time-dependent unit vector $u \in \Pi$, where $r = |x|$, then $u = -z/|z|$, with z as in* (1.5).

Proof. As $\langle u+w, u-w\rangle = \langle u,u\rangle - \langle w,w\rangle$, the difference of the unit vectors u and $w = x/r$ is orthogonal to their sum, and so $\dot{x} = \mathrm{G}(r^{-1}x - u)$ for some function G. Taking the norms-squared of both sides of the equivalent relation $\mathrm{G}u = \mathrm{G}r^{-1}x - \dot{x}$, we get, from Remark 1.2, $2\mathrm{G}\langle x,\dot{x}\rangle/r = \langle\dot{x},\dot{x}\rangle \neq 0$ and $2\mathrm{G} = r\langle\dot{x},\dot{x}\rangle/\langle x,\dot{x}\rangle$. By (1.5), the equality $\mathrm{G}u = \mathrm{G}r^{-1}x - \dot{x}$ divided by G yields $u = -z/|z|$. $\square$

***Remark* 1.4.** A real-valued C^1 function γ on an interval I such that $|\dot{\gamma}| \leq |\varphi\gamma|$ for some continuous function φ is either identically 0, or nonzero everywhere in I. In fact, we may assume that I is a closed interval in which $\gamma \neq 0$ somewhere (since a counterexample to our claim—if it existed—would be realized in one). For any fixed a and variable t, both from a maximal open subinterval I' of I on which $|\gamma| > 0$, setting $\lambda(t) = \log|\gamma(t)|$ and denoting by $J[...]$ the integral of ... from a to t, we have $|\lambda(t)| \leq |\lambda(a)| + |\lambda(t) - \lambda(a)|$, while $|\lambda(t) - \lambda(a)| = |J[\dot{\gamma}/\gamma]| \leq |J[|\dot{\gamma}/\gamma|]| \leq \max|\varphi|$, so that $\log|\gamma|$ is bounded on I' and $\gamma \neq 0$ at both endpoints of I'. Due to its maximality, I' must equal the interior of I.

2. The traditional three types of conic sections

In a Euclidean affine plane Π, the *ellipse/hyperbola* associated with two distinct points $p, q \in \Pi$ (the *foci*) and a real number $\mathrm{d} > 0$ such that $\pm(\mathrm{d} - |p-q|) > 0$ is

$$\{x \in \Pi : ||x-p| \pm |x-q|| = \mathrm{d}\}, \tag{2.1}$$

the sign $\pm$ being $+$ for an ellipse and $-$ for a hyperbola. By the *p-branch* of the hyperbola (2.1), with $\pm = -$, we mean the set

$$\{x \in \Pi : |x-p| + \mathrm{d} = |x-q|\}. \tag{2.2}$$

The *parabola* with the *focus* $p \in \Pi$ and the *directrix* Λ is defined to be

$$\{x \in \Pi : \mathrm{dist}(x,\Lambda) = |x-p|\}, \tag{2.3}$$

where Λ is a line in Π such that $p \notin \Lambda$.

In any Euclidean vector space we have the strict triangle inequality

$$|v+w| < |v| + |w| \quad \text{unless } v = \lambda w \text{ or } w = \lambda v \text{ for some } \lambda \geq 0. \tag{2.4}$$

***Lemma* 2.1.** *A non-radial, radially-accelerated trajectory $t \mapsto x(t) \in \Pi \smallsetminus \{0\}$ in a Euclidean vector plane Π lies in an ellipse* (2.1) *or the 0-branch* (2.2) *of a hyperbola* (2.1)*, with $p = 0$ and some $q, \mathrm{d}, \pm$, if and only if there exists a C^3 function $t \mapsto y(t) \in \Pi \smallsetminus \{0\}$ such that, for $r = |x|$ and $s = |y|$,*

$$x + y \text{ is constant, while } s \pm r \text{ is constant and positive.} \tag{2.5}$$

Proof. To establish sufficiency of (2.5), we define the constants q and d by $q = x + y$ and $\mathrm{d} = s \pm r$. The required condition $\pm(\mathrm{d} - |p-q|) > 0$ now holds with $p = 0$:

$$\text{a) } |q| > d \text{ if } \pm \text{ is the minus sign,} \qquad \text{b) } |q| < d \text{ when } \pm \text{ stands for } +. \tag{2.6}$$

In fact, with $\pm$ chosen to be $-$, (2.5) gives $s - r > 0$, and so, obviously,

$$s = |y| = |q-x| \leq |q| + |x| = |q| + r, \quad \text{that is,} \quad |q| \geq \mathrm{d} = s - r > 0. \tag{2.7}$$

All inequalities in (2.7), for any time t, are actually strict: if $\leq$ and $\geq$ were $=$ at some t, the same would be the case for all t, since q and d are constant. As $q \neq 0$ by (2.7), applying (2.4) to $(v,w) = (q,-x)$ we would get $x = \varphi q$ with some function $\varphi \geq 0$, and the trajectory would be radial, contrary to our hypothesis.

Now let $\pm$ be $+$. The triangle inequality $|q| = |x+y| \le r+s = \mathrm{d}$ is strict by (2.4): if we had $y = \varphi x$ with some function $\varphi \ge 0$, then $q = x+y = (\varphi+1)x$ would be nonzero, and so $x = (\varphi+1)^{-1}q$, again making the trajectory radial.

On the other hand, (2.5) is obviously necessary, as we may set $y(t) = q - x(t)$. $\square$

***Lemma* 2.2.** *A non-radial, radially-accelerated trajectory $t \mapsto x = x(t) \in \Pi \smallsetminus \{0\}$ in a Euclidean vector plane Π lies in a parabola* (2.3) *with $p = 0$ and some Λ if and only if*

$$\text{(2.8)} \qquad \langle u, x\rangle + r \text{ is constant for some constant unit vector } u \text{ and } r = |x|.$$

One then has $\Lambda = \{w \in \Pi : \langle u, w\rangle = c\}$, where c is the constant value of $\langle u, x\rangle + r$.

Proof. First, (2.8) is sufficient: $c = \langle u, y\rangle$ is constant for $y = x + ru$. The trajectory now lies in (2.3) with $\Lambda = \{w \in \Pi : \langle u, w\rangle = c\}$. In fact, $x(t) = y(t) - r(t)u$ for all t, with $y(t) \in \Lambda$ and $r(t)u$ orthogonal to Λ, proving that $\mathrm{dist}(x(t), \Lambda) = |r(t)u| = r(t) = |x(t) - p|$, where $p = 0$. At the same time $p = 0 \notin \Lambda$, or else c would equal 0, that is, $0 = \langle u, y\rangle = \langle u, x\rangle + r$. The resulting equality case $\langle u, x\rangle = -r = -|x||u|$ in the Schwarz inequality would imply that $x(t)$ is, for every t, a negative multiple of u, even though the trajectory was assumed to be non-radial.

Necessity of (2.8): as $\mathrm{dist}(x(t), \Lambda) = r(t)$ for all t, writing $x(t) = y(t) - q(t)$ with $y(t) \in \Lambda$ and $q(t)$ orthogonal to Λ we have $|q(t)| = r(t) > 0$, so that the unit vector $u = q(t)/r(t)$ does not depend on t, and applying $\langle\,\cdot\,, u\rangle$ to the equality $\dot x = \dot y - \dot q = \dot y - \dot r u$ we obtain (2.8), since $\dot r = \langle x, \dot x\rangle / r$ by (1.2). $\square$

3. Kepler's first law

A radially-accelerated trajectory $t \mapsto x$ in a vector space Σ, not passing through zero, must lie in a plane containing 0. To see this, write $\ddot x = (\varphi - 1)x$ for some C^1 function $t \mapsto \varphi(t)$. Hence, for $\alpha = \langle w, x\rangle$ and $\beta = \langle w, \dot x\rangle$ with any fixed vector w, (1.1) implies that $\dot\alpha = \beta$ and $\dot\beta = \varphi\alpha$. Setting $\gamma = \alpha^2 + \beta^2$ we now obtain $\dot\gamma = 2\varphi\alpha\beta$, and so $|\dot\gamma| = 2|\varphi\alpha\beta| \le |\varphi\gamma|$. If $\langle w, x\rangle = \langle w, \dot x\rangle = 0$ at some t, Remark 1.4 thus gives $\langle w, x\rangle = \langle w, \dot x\rangle = 0$ for all t. Applied to linearly independent vectors $w_3, \ldots, w_n$ orthogonal to $x(t)$ and $\dot x(t)$ for some t, where $n = \dim \Sigma$, this shows that the trajectory lies in the plane through 0 orthogonal to $w_3, \ldots, w_n$.

We say that a radially-accelerated trajectory $t \mapsto x = x(t)$ in a Euclidean vector space, not passing through 0, has the *inverse-square property* if $\ddot x$ is a negative constant multiple of $|x|^{-3}x$. With the potential energy U as in (1.4), this amounts to

$$\text{(3.1)} \qquad \ddot x = \mathrm{U}x/r^2, \text{ where } \mathrm{U} = -k/r \text{ for some } k \in \mathbb{R}, \text{ and } r = |x|.$$

***Lemma* 3.1.** *For a radially-accelerated trajectory $t \mapsto x = x(t) \in \Sigma \smallsetminus \{0\}$ in a Euclidean vector space Σ, the inverse-square property is equivalent to conservation of total energy, that is, to constancy of the function* E *in* (1.4).

Proof. With $\ddot x = \mathrm{U}x/r^2$, (1.4), (1.1) and (1.2) give $\dot{\mathrm{T}} = \mathrm{U}\dot r/r$. Hence $\dot{\mathrm{E}} = \dot{\mathrm{T}} + \dot{\mathrm{U}} = (\mathrm{U}r)^{\boldsymbol\cdot}/r$. Consequently, $\mathrm{U}r$ is constant if and only if so is E. $\square$

***Theorem* 3.2.** *For a non-radial, radially-accelerated trajectory $t \mapsto x = x(t) \in \Pi \smallsetminus \{0\}$ in a Euclidean vector plane Π, the following two conditions are equivalent*:

(a) *it lies in an ellipse, the 0-branch of a hyperbola, or a parabola, with a focus at* 0,

(b) *it has the inverse-square property and the constant* k *in* (3.1) *is positive.*

The total energy E *in* (1.4) *is then constant. For an ellipse/hyperbola,* E *is negative/positive and, with* z *as in* (1.5), *the* Π*-valued function* $x + z/\mathrm{E}$ *is also constant, its value being the other focus. For a parabola,* z *is constant, nonzero, and normal to the directrix.*

Proof. Assume (b), that is, (3.1) with $k > 0$. In view of Lemma 3.1, E is constant.

If $\pm\mathrm{E} < 0$ with some sign $\pm$, (1.5) gives $\dot{x} + \dot{y} = 0$ for $y = z/E$. By (1.6) and (3.1), $s \pm r = \mp k/\mathrm{E}$, and (a) follows from Lemma 2.1.

If $\mathrm{E} = 0$, (1.5) and Remark 1.1 imply that z is constant and nonzero. As (2.8) holds for the unit vector $u = -2z/[r\langle \dot{x}, \dot{x}\rangle]$ (cf. (1.5)), Lemma 2.2 implies (a).

In both cases, our discussion also proves the final clause of the theorem.

Conversely, if the trajectory lies in an ellipse or the 0-branch of a hyperbola with a focus at 0, Lemma 2.1 yields (2.5) for some y, where $r = |x|$ and $s = |y|$. Also,

$$r^{-1}x \mp s^{-1}y \neq 0 \quad \text{at all times } t. \tag{3.2}$$

(Otherwise, y would, at some t, be a positive/negative multiple of x, so that a point of an ellipse would lie between the foci or, respectively, one focus would lie in the segment joing the other to a point of a hyperbola.) From (1.2) with $\dot{y} = -\dot{x}$ we now obtain $0 = \dot{r} \pm \dot{s} = \langle \dot{x}, x/r\rangle \pm \langle \dot{y}, y/s\rangle = \langle \dot{x}, r^{-1}x \mp s^{-1}y\rangle$. Applying Lemma 1.3 to $u = \mp s^{-1}y$, which is allowed in view of (3.2), we get $u = -z/|z|$, that is, by (1.5), $y = z/\mathrm{F}$, where $2\mathrm{F} = \mp r\langle \dot{x}, \dot{x}\rangle/s$, and so $\pm\mathrm{F} < 0$. Differentiating $y = z/\mathrm{F}$, we conclude from (1.5) that $-\mathrm{F}\dot{x} = \mathrm{F}\dot{y} = \dot{z} - z\dot{\mathrm{F}}/\mathrm{F} = -\mathrm{E}\dot{x} - \dot{\mathrm{F}}\langle x, \dot{x}/\mathrm{F}\rangle \dot{x} + \dot{\mathrm{F}}\langle \dot{x}, \dot{x}/\mathrm{F}\rangle x/2$. Thus, F is constant (or else, at a time t at which $\dot{\mathrm{F}}$ is nonzero, x and $\dot{x}$ would be linearly dependent, making the trajectory radial in view of Remark 1.1). With $\dot{\mathrm{F}} = 0$, the last equality reads $\mathrm{F} = \mathrm{E}$, cf. Remark 1.2. Therefore, E is constant, and Lemma 3.1 implies (b); positivity of k is immediate since, as we just saw, $y = z/\mathrm{F} = z/\mathrm{E}$ and $\pm\mathrm{E} = \pm\mathrm{F} < 0$, while, from (2.5) and (3.1), $0 < s \pm r = \mp k/\mathrm{E}$.

Finally, let the trajectory lie in a parabola having the focus at 0. Using Lemma 2.2 we choose u with (2.8), and then (1.2) gives $\langle \dot{x}, u + x/r\rangle = 0$. Lemma 1.3 now yields $u = -z/|z|$, that is, $z = -\varphi u$ with $\varphi = |z|$. (Note that $u + x/r$ is nonzero at all times: otherwise we would have $c = \langle u, x\rangle + r = \langle ru, u + x/r\rangle = 0$ in Lemma 2.2, and the focus 0 would lie on the directrix Λ.) Differentiating $-z = \varphi u$ we see that, by (1.5), $\dot{\varphi} u = \mathrm{E}\dot{x}$. Thus, E must be identically zero, or else, on a subinterval of the time interval on which $\mathrm{E} \neq 0$, the nonzero vectors $\dot{x}(t)$ tangent to the parabola at the mutually distinct points $x(t)$ would all be parallel to the same vector u. In view of Lemma 3.1, this proves (b), positivity of k (that is, negativity of U) now being obvious from (1.4) with $\mathrm{E} = 0$, where $\mathrm{T} > 0$ according to Remark 1.2. □

***Remark* 3.3.** Positivity of k in (3.1) means that gravity is a force of attraction. To discuss the case $k < 0$, one only needs to replace d with $-\mathrm{d}$ in (2.2), and 'positive' with 'negative' in (2.5) (for the hyperbola case only). It is now clear that the trajectory then has to lie on the "other" branch of the hyperbola, the ellipse and parabola being excluded in view the line preceding the last paragraph in the proof of Theorem 3.2, along with the last line of that paragraph.