

The first-order orbital equation

Maurizio M. D'Eliseo^{a)}

Osservatorio S. Elmo, Via A. Caccavello 22, 80129 Napoli, Italy

(Received 27 February 2006; accepted 15 December 2006)

We derive the first-order orbital equation employing a complex variable formalism. We then examine Newton's theorem on precessing orbits and apply it to the perihelion shift of an elliptic orbit in general relativity. It is found that corrections to the inverse-square gravitational force law formally similar to that required by general relativity were suggested by Clairaut in the 18th century. © 2007 American Association of Physics Teachers.

[DOI: 10.1119/1.2432126]

I. INTRODUCTION

Almost all classical mechanics textbooks derive the elliptical orbit of the two-body planetary problem by means of well known methods. In this paper we derive the first-order orbital equation by using the complex variable formalism. The latter is a useful tool for studying this old problem from a new perspective. From the orbital equation we can extract all the properties of elliptic orbits. Newton's theorem of revolving orbits,¹ which establishes the condition for which a closed orbit revolves around the center of force, has a wide range of applicability, and its application to an inverse-square force allows us to apply the first-order orbital equation.

A revolving (precessing) ellipse reminds us of the general relativistic perihelion shift of the planet Mercury. We explain why an approximate general relativistic force found by Levi-Civita² in his lectures on general relativity gives the same perihelion shift of the r^{-4} general relativistic force derived in textbooks. Our results suggest an interesting link with the work of the 18th century scientist Clairaut.^{3,4}

We identify the plane of the motion of the gravitational two-body problem with the complex plane.⁵ An object of mass M is at the origin. The position x, y of the second object of mass m is given by $\mathbf{r}=x+iy$, and the equation of motion can be expressed as

$$\ddot{\mathbf{r}} = -\mu \frac{\mathbf{r}}{r^3} = -\mu \frac{e^{i\theta}}{r^2}, \quad (1)$$

where $\mu=G(m+M)$, $\mathbf{r}=re^{i\theta}=r(\cos \theta+i \sin \theta)$, $\mathbf{r}^*=re^{-i\theta}$ is the complex conjugate of \mathbf{r} , $r=|\mathbf{r}|=\sqrt{\mathbf{r}\mathbf{r}^*}=\sqrt{x^2+y^2}$, and $\theta=\theta(t)$ is the true longitude, that is, the point (x,y) has the polar form $re^{i\theta}$, where r is the modulus and θ is its argument. From $\mathbf{r}=re^{i\theta}$ we have by differentiation with respect to time

$$\dot{\mathbf{r}} = (\dot{r} + ir\dot{\theta})e^{i\theta}, \quad (2)$$

where $\dot{\mathbf{r}}^*=\dot{\mathbf{r}}^*$. The solution of Eq. (1) requires knowledge of the functions $r(\theta)$ and $\theta(t)$, but we are interested here only in the determination of the function $r(\theta)$, which describes the geometry of the orbit.

We denote by $\text{Re}(\mathbf{r})$ and $\text{Im}(\mathbf{r})$ the real and the imaginary parts of \mathbf{r} , respectively. Thus $\text{Re}(\mathbf{r})=(\mathbf{r}+\mathbf{r}^*)/2=x=r \cos \theta$ and $\text{Im}(\mathbf{r})=(\mathbf{r}-\mathbf{r}^*)/2i=y=r \sin \theta$. It is useful to consider complex variables as vectors starting from the origin so that $\text{Re}(\mathbf{A})$ is the component of \mathbf{A} along the x (real) axis, and $\text{Im}(\mathbf{A})$ is the component along the y (imaginary) axis.

II. THE FIRST-ORDER ORBITAL EQUATION

If we multiply Eq. (1) by \mathbf{r}^* , we have

$$\ddot{\mathbf{r}}\mathbf{r}^* = -\mu \frac{\mathbf{r}\mathbf{r}^*}{r^3} = -\mu \frac{r^2}{r^3} = -\frac{\mu}{r}. \quad (3)$$

If we take the imaginary part of both sides of Eq. (3), we find

$$\text{Im}(\ddot{\mathbf{r}}\mathbf{r}^*) = -\text{Im}\left(\frac{\mu}{r}\right) = 0. \quad (4)$$

It is easy to verify that

$$\frac{d}{dt} \text{Im}(\dot{\mathbf{r}}\mathbf{r}^*) = \text{Im}(\ddot{\mathbf{r}}\mathbf{r}^*) + \text{Im}(\dot{\mathbf{r}}\dot{\mathbf{r}}^*). \quad (5)$$

We have $\text{Im}(\dot{\mathbf{r}}\dot{\mathbf{r}}^*)=0$ because the term in brackets is the square of the module $|\dot{\mathbf{r}}|$, a real quantity. Then from Eqs. (5) and (4) we write

$$\frac{d}{dt} \text{Im}(\dot{\mathbf{r}}\mathbf{r}^*) = 0. \quad (6)$$

Equation (6) implies that $\text{Im}(\dot{\mathbf{r}}\mathbf{r}^*)$ is time independent. We denote its real value by ℓ and write

$$\text{Im}(\dot{\mathbf{r}}\mathbf{r}^*) = \frac{\dot{\mathbf{r}}\mathbf{r}^* - \dot{\mathbf{r}}^*\mathbf{r}}{2i} = \ell. \quad (7)$$

Equation (7) is the area integral.

This derivation holds for any central force $f(r)e^{i\theta}=f(r)\times(\mathbf{r}/r)$. We substitute Eq. (2) into Eq. (7) and find the fundamental relation $\ell=r^2\dot{\theta}$, which can be cast in three equivalent forms:

$$\frac{1}{r^2} = \frac{\dot{\theta}}{\ell}, \quad (8a)$$

$$dt = \frac{r^2}{\ell} d\theta, \quad (8b)$$

$$\frac{d}{dt} = \frac{\ell}{r^2} \frac{d}{d\theta}. \quad (8c)$$

We can rewrite Eq. (1) using Eq. (8a) as

$$\ddot{\mathbf{r}} = -\frac{\mu}{\ell} \dot{\theta} e^{i\theta} = \frac{i\mu}{\ell} \frac{d}{dt} e^{i\theta}, \quad (9)$$

or

$$\frac{d}{dt}\left(\dot{\mathbf{r}} - \frac{i\mu}{\ell}e^{i\theta}\right) = 0. \quad (10)$$

The expression in parentheses is complex and constant in time. For convenience we denote it as $(i\mu\mathbf{e})/\ell$. The reason for this choice will soon be apparent. We have thus deduced the Laplace integral⁶

$$\dot{\mathbf{r}} = \frac{i\mu}{\ell}(e^{i\theta} + \mathbf{e}), \quad (11)$$

where $\dot{\mathbf{r}}$ is the orbital velocity, $\mathbf{e} = e \exp(i\omega)$ is a complex constant that we will call the eccentricity vector, and e is the (scalar) eccentricity. The vector \mathbf{e} is directed toward the perihelion, the point on the orbit of nearest approach to the center of force, and ω is the argument of the perihelion.

If we use the area integral to eliminate the explicit presence of t in Eq. (11), we obtain a relation between r and θ . One way to integrate Eq. (1) twice with respect to time is to substitute into Eq. (7) the expression for $\dot{\mathbf{r}}$ given by Eq. (11):

$$\begin{aligned} \ell = \text{Im}(\dot{\mathbf{r}}\mathbf{r}^*) &= \text{Im}\left[\frac{i\mu}{\ell}(r + \mathbf{e}\mathbf{r}^*)\right] \\ &= \frac{\mu}{\ell}[r + \text{Im}(i\mathbf{e}\mathbf{r}^*)] \\ &= \frac{\mu}{\ell}r[1 + \text{Re}(\mathbf{e}e^{-i\theta})], \end{aligned} \quad (12)$$

from which we can solve for r

$$r = \frac{\ell^2/\mu}{1 + e \cos(\theta - \omega)}. \quad (13)$$

Equation (13) is the relation in polar coordinates of the orbit, which is a conic section of eccentricity e with a focus at the origin. If the orbit is an ellipse ($0 < e < 1$), we have the relation $\ell^2/\mu = a(1 - e^2)$, where a is the semi-major axis (which lies on the apse line).⁷ Thus the names given earlier to $|\mathbf{e}| = e$ and ω are justified.

Another way to find the orbit is to transform the time derivative into a θ derivative. We start from Eq. (11), which is a function of θ instead of time. From Eq. (8c) we obtain

$$\dot{\mathbf{r}} = \frac{\ell}{r^2}\mathbf{r}' = \frac{i\mu}{\ell}(e^{i\theta} + \mathbf{e}), \quad (14)$$

where a prime denotes differentiation with respect to θ . Then

$$\mathbf{r}' = (re^{i\theta})' = (r' + ir)e^{i\theta} = \frac{i\mu}{\ell^2}(e^{i\theta} + \mathbf{e})r^2. \quad (15)$$

If we multiply by $e^{-i\theta}$, we obtain the complex Bernoulli equation⁸

$$r' + ir = \frac{i\mu}{\ell^2}(1 + \mathbf{e}e^{-i\theta})r^2. \quad (16)$$

If we take the imaginary and the real parts of both sides of Eq. (16), we obtain the orbit $r(\theta)$ and its derivative $r'(\theta)$, respectively. However, it is better to change the dependent variable, so we divide both sides by r^2 , make the variable change $r(\theta) \rightarrow 1/u(\theta)$, and multiply by i . We find

$$u + iu' = \frac{\mu}{\ell^2}(1 + \mathbf{e}e^{-i\theta}). \quad (17)$$

Despite its heterogeneous nature, it is convenient to write the left-hand side of Eq. (17) in terms of $\mathbf{u} = \mathbf{u}(\theta) \equiv u(\theta) + iu'(\theta)$, so that we have

$$\mathbf{u} = \frac{\mu}{\ell^2}(1 + \mathbf{e}e^{-i\theta}), \quad (18)$$

which we call the first-order orbital equation.

From Eq. (18) we can immediately deduce the orbit and its apsidal points. The orbit is given by the real part of \mathbf{u} ,

$$\text{Re}(\mathbf{u}) = \frac{1}{r} = \frac{\mu}{\ell^2}[1 + \text{Re}(\mathbf{e}e^{-i\theta})] = \frac{\mu}{\ell^2}[1 + e \cos(\theta - \omega)]. \quad (19)$$

The apsidal points are determined from the condition $\text{Im}(\mathbf{u}) = 0$ because $r'(\theta) = 0$ at these points. If $\text{Im}(\mathbf{u}) = 0$ for every value of θ , then $\mathbf{e} = 0$, and we have a circular orbit with $u = \text{Re}(\mathbf{u}) = \mu/\ell^2$. If $0 < |\mathbf{e}| < 1$, then Eq. (18) gives the position of the two apsidal points r_{\min} and r_{\max} . At these points we have

$$\text{Im}(\mathbf{u}) = -\frac{\mu}{\ell^2}e \sin(\theta - \omega) = 0, \quad (20)$$

so that, by considering the derivative $[\text{Im}(\mathbf{u})]' = u''(\theta)$, we find r_{\min} when $\theta + 2n\pi = \omega$ and r_{\max} when $\theta + (2n+1)\pi = \omega$, where $n = 0, 1, 2, \dots$

If we denote by D the differential operator $d/d\theta$, Eq. (18) may be written in operator form as

$$\mathbf{u} = (1 + iD)u = \frac{\mu}{\ell^2}(1 + \mathbf{e}e^{-i\theta}). \quad (21)$$

By multiplying both sides on the left by $(1 - iD)$, we obtain

$$(1 - iD)(1 + iD)u = (D^2 + 1)u = u'' + u = \frac{\mu}{\ell^2}, \quad (22)$$

which is Binet's orbit equation.⁹

III. THE PRECESSING ELLIPSE

The two-body solution we have found together with the appropriate corrections due to the presence of other bodies does not account for the observed residual precession of the planetary perihelia.¹⁰ An explanation in classical terms is that a small additional force acts on all the planets causing precession.

All perturbing central forces of the type $F \sim r^{-m}e^{i\theta}$, with $m \geq 3$, produce a secular motion of the apse of an elliptical orbit. Conversely, from the observed planetary apse motion we can deduce by Newton's theorem the presence of a perturbing inverse-cube force $F \sim r^{-3}e^{i\theta}$. This result was obtained by Newton in more general terms using the following reasoning.¹

Consider a closed orbit determined by the centripetal force $-f(r)e^{i\theta}$. If we let $r = r(\sigma\theta)$, where $\sigma \neq 1$ is an arbitrary real constant, we will obtain the same orbit as for $\sigma = 1$, but revolving around the center of force (the two orbits are coincident when $\theta = 0$). From the area integral of the first orbit, $r^2\dot{\theta} = \ell$, we obtain $r^2\sigma\dot{\theta} = \sigma\ell$, which we write as $r^2\dot{\theta} = \tilde{\ell}$. This

integral is for the centripetal force $-\tilde{f}(r)e^{i\theta}$. The radial equations of these two orbits with the same $r(t)$ are

$$\ddot{r} - r\dot{\theta}^2 = -f(r), \quad (23a)$$

$$\ddot{r} - r\tilde{\theta}^2 = -\tilde{f}(r), \quad (23b)$$

from which we obtain

$$\tilde{f}(r) - f(r) = r(\tilde{\theta}^2 - \dot{\theta}^2) = r\left(\frac{\tilde{\ell}^2}{r^4} - \frac{\ell^2}{r^4}\right) = \frac{\ell^2(\sigma^2 - 1)}{r^3}. \quad (24)$$

The extra radial force is outward or inward depending on whether σ is greater or less than unity.

Thus far the force $f(r)$ is arbitrary, but if we specialize to the inverse-square gravitational force, then the first-order orbital equation (18) with the perturbing inverse-cube force

$$\frac{\ell^2(\sigma^2 - 1)}{r^3}e^{i\theta} = \ell^2(\sigma^2 - 1)u^3e^{i\theta} \quad (25)$$

takes the form

$$\mathbf{u} = \frac{\mu}{\ell^2}(1 + \mathbf{e}e^{-i\sigma\theta}). \quad (26)$$

Hence

$$u = \text{Re}(\mathbf{u}) = \frac{\mu}{\ell^2} + \frac{\mu}{\ell^2}e \cos(\sigma\theta - \omega) \quad (27)$$

is an ellipse precessing around the focus with an angular velocity proportional to the radius vector. This description becomes more accurate as σ approaches unity. The apsidal points are given by $\text{Im}(\mathbf{u})=0$. From Eq. (26) and $\mathbf{e} = e \exp(i\omega)$, we have at the apsidal points

$$\sin(\sigma\theta - \omega) = 0. \quad (28)$$

In particular we have r_{\min} when

$$\sigma\theta - \omega = \theta - [\omega + (1 - \sigma)\theta] = 0, \quad (29)$$

so that after one complete revolution the angular perihelion shift is

$$\Delta\omega = 2\pi(1 - \sigma). \quad (30)$$

If $\sigma < 1$, then $\Delta\omega > 0$ and the shift is positive, while for $\sigma > 1$ we have $\Delta\omega < 0$ and the shift is negative.

IV. GENERAL RELATIVITY

The foregoing considerations have a direct application in general relativity. The general relativistic Binet's orbit equation, which is obtained from the geodesic equation in the Schwarzschild space-time, is¹¹

$$u'' + u = \frac{\mu}{\ell^2} + 3\alpha u^2, \quad (31)$$

where $\alpha = GM/c^2 \equiv \mu/c^2$ is the gravitational radius of the central body, and c is the speed of light. The corresponding equation of motion is

$$\ddot{\mathbf{r}} = -\mu \frac{e^{i\theta}}{r^2} - \frac{3\alpha\ell^2}{r^4}e^{i\theta} \quad (32)$$

and we see that general relativity introduces an effective perturbative r^{-4} force.

From Eq. (31) or Eqs. (11) and (32) we can deduce¹² the standard formula for the perihelion shift given in general relativity by

$$\Delta\omega = \frac{6\pi\alpha\mu}{\ell^2} = \frac{6\pi\alpha}{a(1 - e^2)}. \quad (33)$$

If we equate Eqs. (30) and (33) and solve for σ , we obtain $\sigma^2 \approx 1 - 6\alpha\mu/\ell^2$, and by using Eq. (25) we obtain the inverse-cube perturbation that gives the same perihelion shift as predicted by general relativity:

$$F = -\mu \frac{6\alpha}{r^3}e^{i\theta}. \quad (34)$$

That is, we can obtain the same precession using either an r^{-3} or r^{-4} perturbative force.

Levi-Civita obtained an effective r^{-3} force in general relativity using a method based on a new form of Hamilton's principle² devised to go smoothly from the classical equation of motion to the Einstein field equation. His approximation is not as general as the usual r^{-4} effective force because it does not produce the bending of light rays, a subject that Levi-Civita treated with another ingenious approximation.

Binet's equation for the r^{-3} perturbative force, obtained from the equation of the motion by the use of Eq. (8c) and the variable change $r(\theta) \rightarrow 1/u(\theta)$, is

$$u'' + \left(1 - 6\alpha \frac{\mu}{\ell^2}\right)u = \frac{\mu}{\ell^2}. \quad (35)$$

The null-geodesic equation of light rays requires that we formally put $\mu/\ell^2=0$ in Eq. (35),¹⁵ so that it becomes

$$u'' + u = 0. \quad (36)$$

The solution of Eq. (36) is $k \sin \theta$ where $k = \text{const}$ and $0 \leq \theta \leq \pi$. In terms of the radius $r = 1/u$, the solution becomes $r \sin \theta = 1/k$. Because $r \sin \theta$ is the Cartesian coordinate y , the solution represents a straight line parallel to the x axis, so that the light ray is not deflected at all by the sun's gravitational field in this approximation.

As we have seen, the weak-field approximation of general relativity adds an effective r^{-3} force or a r^{-4} force (depending on the approximation method used) to Newton's inverse square force to explain the perihelion motion. It is interesting that these results were proposed in the mid-18th century by the mathematician and astronomer Clairaut who proposed the addition of a small r^{-n} force to the r^{-2} gravitational force to explain the swift motion of the lunar perigee. In particular, he examined the influence of both r^{-4} and r^{-3} terms.¹⁴ It was later recognized by Clairaut that this addition was unnecessary, because a purely r^{-2} force law could completely explain the motion of the Moon. The apparently anomalous secular motion of the perigee was due to discarded noncentral force terms in the process of successive approximations.¹⁵ No doubt Clairaut would have again made his suggestion if he had known about the anomalous motion of Mercury's perihelion. Without a new first principles gravitational theory¹⁶

he probably would have employed a phenomenological approach and introduced one of the two forces by empirically adjusting the numerical factors.

^{a)}Electronic mail: s.elmo@mail.com

¹S. Chandrasekhar, *Newton's Principia for the Common Reader* (Clarendon, Oxford, 1995), pp. 184–187. In particular, Proposition XLIV-Theorem XIV: The difference of the forces, by which two bodies may be made to move equally, one in a fixed, the other in the same orbit revolving, varies inversely as the cube of their common altitudes.

²T. Levi-Civita, *Fondamenti di Meccanica Relativistica* (Zanichelli, Bologna, 1929), p. 123.

³A. C. Clairaut, “Du Systeme du Monde, dan les principes de la gravitation universelle,” *Histoires de l'Academie Royale des Sciences*, mem. 1745 and Ref. 4.

⁴We have reproduced many papers of historical interest at (gallica.bnf.fr/).

⁵T. Needham, *Visual Complex Analysis* (Oxford U. P., New York, 1999).

⁶The Laplace integral can be found in P. S. Laplace, *Ouvres* (Gauthier-Villars, Paris, 1878), Tome 1, p. 181, formula *P*. To obtain Eq. (11), we need to use $z=0$, $c=x\dot{y}-\dot{x}y=\ell$, $f=\mu \operatorname{Re}(\mathbf{e})$, and $f'=\mu \operatorname{Im}(\mathbf{e})$, and add the first relation to the second one multiplied by $-i$ (see Ref. 4). To keep the customary notation we use the same letter e for the eccentricity and for the complex exponential.

⁷A parallel treatment of the two-body problem with vectorial methods is given by V. R. Bond and M. C. Allman, *Modern Astrodynamics* (Princeton U. P., Princeton, NJ, 1998).

⁸R. E. Williamson, *Introduction to Differential Equations* (McGraw-Hill, New York, 1997), p. 84.

⁹R. d'Inverno, *Introducing Einstein's Relativity* (Oxford U. P., New York, 2001), p. 194.

¹⁰For a complete calculation of all the perturbing effects see M. G. Stewart, “Precession of the perihelion of Mercury's orbit,” *Am. J. Phys.* **73**, 730–734 (2005).

¹¹Reference 9, p. 196.

¹²See, for example, B. Davies, “Elementary theory of perihelion precession,” *Am. J. Phys.* **51**, 909–911 (1983); N. Gauthier, “Periastron precession in general relativity,” *ibid.* **55**, 85–86 (1987); T. Garavaglia, “The Runge-Lenz vector and Einstein perihelion precession,” *ibid.* **55**, 164–165 (1987); C. Farina and M. Machado, “The Rutherford cross section and the perihelion shift of Mercury with the Runge-Lenz vector,” *ibid.* **55**, 921–923 (1987); D. Stump, “Precession of the perihelion of Mercury,” *ibid.* **56**, 1097–1098 (1988); K. T. McDonald, “Right and wrong use of the Lenz vector for non-Newtonian potentials,” *ibid.* **58**, 540–542 (1990); S. Cornbleet, “Elementary derivation of the advance of the perihelion of a planetary orbit,” *ibid.* **61**, 650–651 (1993); B. Dean, “Phase-plane analysis of perihelion precession and Schwarzschild orbital dynamics,” *ibid.* **67**, 78–86 (1999).

¹³R. Adler, M. Bazin, and M. Shiffer, *Introduction to General Relativity*, 2nd ed. (McGraw-Hill, New York, 1975), p. 216, Eq. (6.149).

¹⁴A. C. Clairaut, “Du Systeme du Monde, dan les principes de la gravitation universelle,” *Histoires de l'Academie Royale des Sciences*, mem. 1745, p. 337: Clairaut wrote that “The moon without doubt expresses some other law of attraction than the [inverse] square of the distance, but the principal planets do not require any other law. It is therefore easy to respond to this difficulty, and noting that there are an infinite number of laws which give an attraction which differs very sensibly from the law of the squares for small distances, and which deviates so little for the large, that one cannot perceive it by observations. One might regard, for example, the analytic quantity of the distance composed of two terms, one having the square of the distance as its divisor, and the other having the square square.” On p. 362, Clairaut examined the effect of a perturbing inverse-cube force. This memoir is dated 15 November 1747 and can be found in Ref. 4. Clairaut was also the first to introduce a revolving ellipse as a first approximation to the motion of the moon. This idea is sometimes called Clairaut's device or Clairaut's trick.

¹⁵See F. Tisserand, *Traité de Mécanique Celeste III* (Gauthier-Villars, Paris, 1894), p. 57; reproduced at Ref. 4.

¹⁶Clairaut was also a first-class geometer, specializing in curvature. See his *Recherches sur les courbes a double courbure* at Ref. 4.