NOVEMBER 01 2022

Simple precession calculation for Mercury: A linearization approach

Michael J. W. Hall

Am. J. Phys. 90, 857–860 (2022)
https://doi.org/10.1119/5.0098846

Simple precession calculation for Mercury: A linearization approach

Michael J. W. Hall[a)]
Theoretical Physics, Research School of Physics, Australian National University, Canberra ACT 0200, Australia

(Received 12 May 2022; accepted 28 July 2022)

The additional perihelion precession of Mercury due to general relativity can be calculated by a method that is no more difficult than solving for the Newtonian orbit. This method relies on linearizing the relativistic orbit equation, is simpler than standard textbook methods, and is closely related to Newton's theorem on revolving orbits. The main result is accurate for all values of GM/c^2 for near-circular orbits.
https://doi.org/10.1119/5.0098846

I. INTRODUCTION

An important and early success of general relativity was an explanation of the anomalous precession of the perihelion of Mercury by 43 arcseconds per century.[1] The same analysis also applies to the recent observation of the precession of a star orbiting the supermassive black hole at the centre of our Milky Way galaxy.[2] It is of interest to derive this result in the classroom, but standard derivations typically either apply perturbation theory to the orbit equation,[3–5] or rely on an approximate factorization of the relativistic energy equation[6–8] (formally related to Einstein's original calculation).[1] However, the perturbation method requires finding a particular integral of an inhomogenous second-order differential equation (which is usually supplied to students) and extracting its nonperiodic component, while the energy approach requires evaluating a nontrivial integral via nonobvious changes of variables, which is again usually supplied to students.

Alternative derivations in the literature similarly tend to have challenging features for introductory students. For example, a method based on the Runge–Lenz vector is very elegant but requires sophisticated mathematical machinery.[9] Furthermore, while approaches based on the small eccentricities of near-circular orbits are more elementary in character, they require *ad hoc* assumptions about the typical orbital radius[10,11] and/or angular momentum[12] to obtain the correct result.

The purpose of this note is to point out a particularly simple method for calculating the precession in an introductory course. This is based on linearizing a quadratic term in the relativistic orbit equation. The resulting approximate orbit equation is as easy to solve as the Newtonian orbit equation, is able to estimate the precession for all values of GM/c^2 while avoiding *ad hoc* assumptions, and is related to Newton's method for estimating the precession of near-circular orbits.[13,14]

The outline of this paper is as follows: Necessary elements of Newtonian orbits are briefly reviewed in Sec. II. The linearization approach is presented in Sec. III, and underlying assumptions and comparisons with other methods are discussed in Sec. IV.

II. NEWTONIAN ORBITS

Newton's orbit equation for a particle moving in a given plane about a gravitational centre of mass M is generally solved by writing it in the Binet form[3]

$$\frac{d^2u}{d\phi^2} + u = \frac{GM}{h^2}, \tag{1}$$

where $r = 1/u$ is the orbital distance, ϕ is the orbital angle relative to the perihelion, G is Newton's gravitational constant, and $h = r^2(d\phi/dt)$ is the conserved specific angular momentum. The general solution is found by adding the constant solution $u_0 = GM/h^2$ to any solution of the homogenous equation $d^2u/d\phi^2 + u = 0$ to give

$$\frac{1}{r} = u = u_0(1 + e\cos\phi) \tag{2}$$

for some constant e, with $e \geq 0$ to ensure perihelion at $\phi = 0$. This represents a conic section with $e < 1$ for a bound orbit, corresponding to an ellipse with focus at M and eccentricity e.

The parameters u_0 and e in Eq. (2) are fully determined by the perihelion distance $r_{\min}$ and aphelion distance $r_{\max}$ of the orbit via $1/r_{\min} = u_0(1+e)$ and $1/r_{\max} = u_0(1-e)$, yielding

$$u_0 = \frac{r_{\max} + r_{\min}}{2r_{\max}r_{\min}}, \quad e = \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}}. \tag{3}$$

Thus, observations of the perihelion and aphelion are sufficient to fix the orbit.[15] One also has the useful relation

$$a = \frac{1}{2}(r_{\max} + r_{\min}) = \frac{1}{2u_0}\left(\frac{1}{1-e} + \frac{1}{1+e}\right) = \frac{1}{u_0(1-e^2)}, \tag{4}$$

where a is the semi-major axis of an elliptical orbit.

III. PRECESSION OF RELATIVISTIC ORBITS

In general relativity, the motion of a body orbiting a much larger mass M in a given plane, such as Mercury orbiting the Sun, is well-modeled by the equation for a geodesic orbit around a stationary black hole of mass M,[3–8]

$$\frac{d^2u}{d\phi^2} + u = \frac{GM}{j^2} + \frac{3GM}{c^2}u^2, \tag{5}$$

where $r = 1/u$ and ϕ are the orbital distance and the angle relative to perihelion in Schwarzschild coordinates, respectively, $j = r^2(d\phi/d\tau)$ is the conserved specific angular momentum with respect to proper time, and c is the speed of light. The quadratic term on the right side has no analogue in the Newtonian orbit equation (1) and is responsible for a non-Newtonian precession of bound orbits. The aim of this note is to give a simple method for estimating this precession.

The central idea is to linearize the quadratic term in Eq. (5) about the average value $\bar{u}$ of u, under the assumption that $(u-\bar{u})^2 \ll \bar{u}^2$, which is equivalent to $|r - \bar{u}^{-1}| \ll r$. This is clearly a good approximation for near-circular orbits and is closely related to Newton's theorem on revolving orbits, as discussed below. In particular, linearization gives

$$\begin{aligned} u^2 &= [\bar{u} + (u-\bar{u})]^2 \\ &= \bar{u}^2 + 2\bar{u}(u-\bar{u}) + (u-\bar{u})^2 \\ &\approx \bar{u}^2 + 2\bar{u}(u-\bar{u}) \\ &= 2\bar{u}u - \bar{u}^2. \end{aligned} \tag{6}$$

Substituting this result into Eq. (5) yields the approximate orbit equation

$$\frac{d^2u}{d\phi^2} + Ku = L \tag{7}$$

with the constants

$$K = 1 - \frac{6GM\bar{u}}{c^2}, \quad L = \frac{GM}{j^2} - \frac{3GM}{c^2}\bar{u}^2. \tag{8}$$

Just as for the solution of the Newtonian orbit equation (1), the general solution of Eq. (7) is found by adding the constant solution, $u_0 = L/K$, to any solution of the homogenous equation $d^2u/d\phi^2 + Ku = 0$, to give

$$\frac{1}{r} = u = u_0\left(1 + e\cos\sqrt{K}\phi\right) \tag{9}$$

for a suitable constant e, with $e \in [0,1)$ for bound orbits.[16] Noting that the cosine term varies between 1 and –1, it follows that the parameters u_0 and e in Eq. (9) are related to the observed perihelion and aphelion distances of the orbit as per Eqs. (3) and (4) for Newtonian orbits.[15]

However, whereas the Newtonian orbit in Eq. (2) returns to perihelion at $\phi = 2\pi$, the relativistic orbit in Eq. (9) returns to perihelion at $\sqrt{K}\phi = 2\pi$, i.e., at

$$\phi_p = \frac{2\pi}{\sqrt{K}} = \frac{2\pi}{\sqrt{1 - \frac{6GM\bar{u}}{c^2}}} \approx 2\pi\left(1 + \frac{3GM\bar{u}}{c^2}\right) \tag{10}$$

to first order in $GM\bar{u}/c^2$, using the approximation $(1-x)^{-1/2} \approx 1 + x/2$. Thus, in the relativistic case, there is a precession of the perihelion per orbit by an amount

$$\Delta\phi = \phi_p - 2\pi = 2\pi\left(\frac{1}{\sqrt{K}} - 1\right) \approx \frac{6\pi GM\bar{u}}{c^2}. \tag{11}$$

To determine the average value $\bar{u}$ for use in Eq. (11), note that the orbit equations (5) and (7) are with respect to ϕ, so that an average over ϕ is both appropriate and natural. Then, since $\sqrt{K}\phi$ in Eq. (9) ranges over 0–2π from the initial perihelion to the next, the cosine term is zero on average, giving $\bar{u} = u_0$.[17] Hence, applying Eqs. (3) and (4) gives

$$\bar{u} = u_0 = \frac{r_{\max} + r_{\min}}{2r_{\max}r_{\min}} = \frac{1}{a(1-e^2)}, \tag{12}$$

and substituting into Eq. (11) gives

$$\Delta\phi \approx \frac{6\pi GM}{c^2}\frac{r_{\max} + r_{\min}}{2r_{\max}r_{\min}} = \frac{6\pi GM}{a(1-e^2)c^2} \tag{13}$$

to first order in $GM/(ac^2)$. This is precisely the same expression for the relativistic precession calculated via standard methods in the literature, and for Mercury's 88-day orbit about the Sun ($M = M_\odot = 1.99 \times 10^{30}$ kg, $r_{\min} = 4.60 \times 10^{10}$ m, $r_{\max} = 6.98 \times 10^{10}$ m[18]) evaluates to the well-known observed value of 43 arcseconds per century.[3–8] Equation (13) further predicts the much larger precession of 12 arcminutes per orbit for the star S2, which orbits the supermassive black hole at the centre of our galaxy every 16 years ($M = 4.26 \times 10^6 M_\odot$, $r_{\min} = 1.8 \times 10^{13}$ m, $r_{\max} = 29.4 \times 10^{13}$ m), which is again consistent with observations.[2,19]

IV. DISCUSSION

The linearization method presented here shows that the relativistic precession of Mercury's orbit may be obtained with no more mathematical sophistication than is required for solving the Newtonian orbit equation. In particular, no higher-order perturbative solutions[3–5] or nontrivial integrals[6–8] need be evaluated. This makes the method very suitable for introductory courses in general relativity. For teachers and advanced students, the additional points discussed below may also be of interest.

First, it is worth noting that while Eq. (13) can be shown to be accurate to first order in $GM/(ac^2)$ for all values of e^2,[20] the general prediction of the linearization method in Eq. (11),

$$\begin{aligned} \Delta\phi &= 2\pi\left(\frac{1}{\sqrt{K}} - 1\right) \\ &= 2\pi\left[\left(1 - \frac{6GM}{a(1-e^2)c^2}\right)^{-1/2} - 1\right], \end{aligned} \tag{14}$$

is, conversely, accurate to first order in e^2 for all values of $GM/(ac^2)$. In particular, substituting Eq. (9) into the central assumption $(u-\bar{u})^2 \ll \bar{u}^2$ used in Eq. (6) and recalling $u_0 = \bar{u}$ gives the equivalent condition

$$e^2 \ll 1 \tag{15}$$

for the validity of Eq. (14) independent of the value of $GM/(ac^2)$, corresponding to near-circular orbits. For Mercury, one has $e^2 \approx 0.04$, so that the method accurately predicts the precession.

The simplicity and generality of condition (15) may be compared to the relatively complicated and restrictive condition required in a related approach for near-circular orbits by Lemmon and Mondragon,[12,21] which linearizes the orbit equation about $u_c = j^2/(GM)$ and makes an *ad hoc* identification of the relativistic angular momentum j with the angular momentum h of a Newtonian orbit. This approach not only requires comparing relativistic and non-relativistic orbits having unequal perihelion distances but moreover, unlike Eq. (14), is accurate only to second order in $GM/(jc)$.

Second, there are two weak constraints, $K>0$ and $L>0$, on the applicability of the linearization method, implicit in the form of the general solution in Eq. (9). These constraints correspond to the existence of solutions with $0<r<\infty$, i.e., to bound orbits, and from Eqs. (8) and (12) they can be written in the form

$$r_S = \frac{2GM}{c^2} < \frac{1}{3}a(1-e^2), \quad v_\parallel = \frac{j}{a} < \frac{1-e^2}{\sqrt{3}}c, \tag{16}$$

where r_S is the Schwarzschild radius of the central mass M and $v_\parallel$ is a measure of orbital speed. These constraints trivially hold for planetary orbits in our solar system, where $r_S \approx 3\,\text{km}$ for the Sun and $v_\parallel \approx 0.0002\,c$ for Mercury.[18]

Third, as mentioned in Sec. III, the linearization method is closely related to Newton's theorem on revolving orbits: If $r(\phi)$ describes an orbit corresponding to a central force per unit mass $f(r)$ and angular momentum h, then $r(k\phi)$ describes an orbit corresponding to a central force $f(r)+\alpha/r^3$ and angular momentum $H=h/k$, with $\alpha=(k^2-1)H^2$.[13,14,22] This theorem follows from the Newtonian orbit equation for the latter case[3,14]

$$\frac{d^2u}{d\phi^2}+u = -\frac{1}{H^2u^2}\left[f(u^{-1})+\alpha u^3\right] = -\frac{1}{H^2u^2}f(u^{-1}) - \frac{\alpha}{H^2}u, \tag{17}$$

which simplifies to

$$\frac{d^2u}{d(k\phi)^2}+u = -\frac{1}{h^2u^2}f(u^{-1}), \tag{18}$$

with solution $u=1/r(k\phi)$ following from the definition of $r(\phi)$. The linear term on the right side of Eq. (17), which corresponds to an inverse-cube force, plays the same formal role here as the linear term obtained via linearization in Eq. (6). Thus the simple connection between the forms of the elliptical and precessing orbits in Eqs. (2) and (9) may be viewed as a special case of Newton's theorem on revolving orbits.

Furthermore, Newton used his theorem to estimate the precession of near-circular orbits perturbed by an arbitrary additional force, by approximating the perturbing force by an inverse-cube force.[13,14] However, while this is analogous to linearization of the quadratic term in Eq. (5), previous attempts to apply Newton's method to the relativistic precession of Mercury[10,11] have relied on an *ad hoc* assumption that is avoided in the linearization approach. In particular, the relativistic orbit equation in Eq. (5) corresponds to a perturbing force $-3j^2GM/(c^2r^4)$ [cf. Eq. (17)], which is approximated via the Taylor series expansion $1/r \approx 1/R - (r-R)/R^2$ about some fixed value $r=R$ using $1/r^4 = (1/r)(1/r^3) \approx 2/(Rr^3) - 1/(R^2r^2)$, i.e., by the sum of an inverse-cube term and an inverse-square term. The latter can be absorbed into the Newtonian gravitational force, and applying Newton's theorem for revolving orbits then leads to the corresponding precession[10,11]

$$\Delta\phi \approx \frac{6\pi GM}{Rc^2} \tag{19}$$

to first order in $GM/(Rc^2)$. Comparing with Eq. (13), it follows that one should choose $R=a(1-e^2)$ (to at least first order in e^2) to obtain the correct prediction for near-circular orbits. However, this choice is purely *ad hoc*. Rowlands simply makes this choice *ab initio*, without any physical motivation,[10] while Nguyen instead chooses $R=a/(1-e^2)$, leading to a predicted precession of only 39 arcseconds per century for Mercury.[11] An alternative choice, the average orbital distance

$$R = \bar{r} = \frac{1}{2\pi}\int_0^{2\pi}\frac{d\phi}{u_0(1+e\cos\phi)} = a\sqrt{1-e^2} \tag{20}$$

again leads to an incorrect prediction.[23] In contrast, the linearization method leads naturally and directly to the correct result in Eq. (13) and its generalization in Eq. (14).

It is concluded that the linearization method provides a simple and useful approach for calculating the general relativistic contribution to the precession of Mercury's orbit. The corresponding orbit equation (7) is no more difficult to solve than the Newtonian orbit equation (1), making the method particularly suitable for introductory courses; the precession is accurately estimated for any value of $GM/(ac^2)$ for near-circular orbits ($e^2 \ll 1$); and the natural expansion about the average inverse radius $\bar{u}$ keeps the basic idea of Newton's method for calculating precession intact while removing the need for *ad hoc* assumptions.

ACKNOWLEDGMENTS

The author thanks the referees for several valuable suggestions.

[a]ORCID: 0000-0001-7809-4463.

[1]A. Einstein, "Erklärung der perihelbewegung des merkur aus der allgemeinen relativitätstheorie," Sitzungsber. K. Preuss. Akad. Wiss. **47**, 831–839 (1915), https://nms.kcl.ac.uk/gerard.watts/Einstein-1915.pdf.
[2]GRAVITY Collaboration, "Detection of the Schwarzschild precession in the orbit of the star S2 near the Galactic centre massive black hole," Astron. Astrophys. **636**, L5–L19 (2020).
[3]S. T. Thornton and J. B. Marion, *Classical Dynamics of Particles and Systems*, 5th ed. (Brooks/Cole, Belmont, 2004), Secs. 8.4 and 8.9, pp. 291–295 and 312–316.
[4]W. Rindler, *Relativity: Special, General, and Cosmological*, 2nd ed. (Oxford U. P., Oxford, 2006), Sec. 11.9, pp. 241–245.
[5]P. J. O'Donnell, *Essential Dynamics and Relativity* (CRC Press, Boca Raton, FL, 2014), Sec. 5.8, pp. 127–129.
[6]S. Weinberg, *Gravitational and Cosmology* (Wiley, New York, 1972), Sec. 8.6, pp. 194–201.
[7]J. B. Hartle, *Gravity: An Introduction to Einstein's General Relativity* (Addison-Wesley, San Francisco, CA, 2003), pp. 216–217.
[8]M. J. W. Hall, *General Relativity: An Introduction to Black Holes, Gravitational Waves, and Cosmology* (Morgan & Claypool/IOP Publishing, San Rafael, CA, 2018), pp. 5-13–5-14.
[9]M. G. Stewart, "Precession of the perihelion of Mercury's orbit," Am. J. Phys. **73**, 730–734 (2005).
[10]P. Rowlands, "A simple approach to the experimental consequences of general relativity," Phys. Educ. **32**, 49–55 (1997).
[11]H. D. Nguyen, "Rearing its ugly head: The cosmological constant and Newton's greatest blunder," Am. Math. Mon. **115**, 415–430 (2008).
[12]T. J. Lemmon and A. R. Mondragon, "Alternative derivation of the relativistic contribution to perihelic precession," Am. J. Phys. **77**, 890–893 (2009).
[13]I. Newton, *Mathematical Principles of Natural Philosophy*, translated by A. Motte and F. Cajori (University of California Press, Berkeley, CA, 1962), Book I, Sec. IX, pp. 135–147.

[14]S. Chandrasekhar, *Newton's Principia for the Common Reader* (Clarendon, Oxford, 1995), pp. 184–199.

[15]Note also in both Eqs. (2) and (9) that $1/u_0$ is equal to the orbital distance at the angle halfway from perihelion to aphelion (corresponding to the semilatus rectum for elliptical orbits).

[16]One can alternatively obtain this result by dividing Eq. (7) by K to give $\frac{d^2u}{d(\sqrt{K}\phi)^2} + u = L/K$, which is equivalent to the Newtonian orbit equation (1) with ϕ replaced by $\sqrt{K}\phi$.

[17]The same result is obtained if one averages over many orbits, corresponding to $\bar{u} = \lim_{\Phi\to\infty} \frac{1}{\Phi}\int_0^{\Phi} d\phi\, u(\phi)$.

[18]Planetary data are available at, e.g., the NASA website <https://nssdc.gsfc.nasa.gov/planetary/factsheet/> (accessed on July 2022).

[19]The given values of $r_{\min}$ and $r_{\max} = r_{\min}(1+e)/(1-e)$ for the orbit of S2 follow directly from the parameters $r_{\min} = 120$ AU and $e = 0.885$ given in: GRAVITY Collaboration, "Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole," Astron. Astrophys. **615**, L15 (2018).

[20]K.-H. Lo, K. Young, and B. Y. P. Lee, "Advance of perihelion," Am J. Phys. **81**, 695–702 (2013).

[21]The self-consistency condition in Eq. (31) of Ref. 12 is $e + 3(\frac{GM}{jc})^2(e+2) \ll 1$.

[22]M. M. D'Eliseo, "The first-order orbital equation," Am. J. Phys. **75**, 352–355 (2007).

[23]Similarly, choosing R to be the time average of r between successive perihelia P and P' and using $r^2\frac{d\phi}{d\tau} = j$ gives $R = \int_P^{P'} r\, d\tau / \int_P^{P'} d\tau = \int_0^{2\pi} \frac{d\phi}{[u_0(1+e\cos\phi)]^3} / \int_0^{2\pi} \frac{d\phi}{[u_0(1+e\cos\phi)]^2} = a\left(1+\frac{1}{2}e^2\right)$, thus also leading to an incorrect prediction.

Color Wheel

In the latter years of the 17th century Isaac Newton showed that a thin beam of white light could be broken up into the colors of the spectrum by passing it through a triangular prism. The inverse process relies on the fact that the human eye cannot resolve pieces of information that appear oftener than about thirty times per second. The circular disk has a series of segments painted with the various colors in the spectrum. When it is spun rapidly the eye-brain combination runs all of the colors together and you see white again. The apparatus is in the Amherst College Collection. (Amherst College picture and text by Thomas B. Greenslade, Jr., Kenyon College)