Measurement of the de Sitter Precession of the Moon: A Relativistic Three-Body Effect

I. I. Shapiro, R. D. Reasenberg, J. F. Chandler, and R. W. Babcock
Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138
(Received 27 July 1988)

We analyzed lunar laser-ranging data, accumulated between 1970 and 1986, to estimate the deviation of the precession of the Moon's orbit from the predictions of general relativity. We found no deviation from this predicted de Sitter precession rate of nearly 2 angular sec per century (sec/cy), to within our estimated standard error of 0.04 sec/cy. This standard error, 2% of the predicted effect, incorporates our assessment of the likely contributions of systematic errors, and is about threefold larger than the statistical standard error.

PACS numbers: 95.10.Jk, 96.20.−n

One of the dynamical consequences of Einstein's theory of general relativity, noted first by de Sitter,[1] is a non-Newtonian precession of the Moon as it orbits the Earth in a system freely falling in the gravitational field of the Sun. de Sitter noted, too, that the magnitude of this "geodetic precession"[2] would be approximately 2 angular sec per century (sec/cy). About 50 yr later, in the late 1960's, one of us (I.S.), with Martin Slade and others, undertook the task of reanalyzing the world's accumulation of optical observations of the Moon, about 350000 in all, spanning the period from 1750 to almost 1970. Also included in the analysis were radar observations of the Moon and tracking data of the Surveyor spacecraft. The goal was to detect the predicted geodetic precession reliably. We failed: The result obtained,[3] 1.5 ± 0.6 sec/cy, was of only marginal reliability. We also looked into the possibility of detecting the contribution of the geodetic precession to the total precession of the Earth's spin axis. For this purpose, we considered using a combination of seismic models of the Earth and data from the tracking of artificial satellites to determine accurately the Newtonian contribution, which depends most importantly on the fractional difference of the principal moments of inertia of the Earth. The total precession was already known with sufficient accuracy and would be known even better from the accumulation of very-long-baseline interferometry (VLBI) observations of compact extragalactic radio sources. This approach, too, was unavailing: The Newtonian contribution to the precession could not be determined with the required accuracy. Another possibility considered by one of us (I.S.) was to determine the fractional difference in the Earth's moments of inertia from its effect on the Earth's principal, 18.6-yr, nutation. The amplitude of this nutation term could be measured with the requisite accuracy from the VLBI observations of the extragalactic sources. A covariance analysis,[4] and a consideration of the relevant uncertainties in other aspects of the needed model of the Earth, showed that this approach might be feasible. VLBI data over a sufficient time span and of sufficient accuracy will have been gathered within a few more years for a useful attempt to be made at this determination. On a somewhat longer time scale, the planned National Aeronautics and Space Administration gyroscope experiment holds promise of far higher accuracy.[5]

The main hope for early success lay in laser ranging to the optical corner reflectors[6] placed on the Moon by the Apollo astronauts in the late 1960's and early 1970's. It was clear that an approximately 18-yr accumulation of such ranging data would be required to separate the de Sitter precession reliably from any contribution due to the uncertainty of the amplitude (and phase) of the 18.6-yr period "classical" term in the motion of the lunar node. In the early 1980's, as this accumulation interval drew to a close, the present authors began a special analysis of the lunar laser-ranging data to try to measure this geodetic precession. Meanwhile, as we were completing this analysis, Bertotti, Ciufolini, and Bender[7] (BCB) inferred the presence of the geodetic precession from a lack of any known problems in the comparisons of (1) lunar laser-ranging (LLR) data with the general-relativistic theory which predicts this precession, and (2) the results, deduced from the LLR data, for the time dependence of the orientation of the Earth with independent determinations of this dependence. Their estimated uncertainty for each method was 10%; i.e., they imply that any larger deviation between the predicted geodetic precession and that affecting the observations would have been noticeable from such comparisons.

In the remainder of this paper, we (1) discuss our parametric model of the Moon's motion, with emphasis on the aspect that allows us to distinguish, as directly as feasible, the contribution of any geodetic precession; (2) describe the data sets we analyzed to determine the geodetic precession; (3) present the result we obtained for the geodetic precession, with special emphasis on the contributing source to its overall uncertainty; and (4) comment on the paper by BCB, in particular demonstrating that they significantly underestimated the uncertainties accompanying their approaches. In this last discussion, we show why the uncertainties for their approaches exceed those in ours by at least thirtyfold for the first of their two methods and by about tenfold for the second.

Our model of the Moon's motion consists of two coupled sets of differential equations, one set describing the orbit and the other the rotation. The coupling is, of course, supplied by the nonspherical figure of the Moon, which affects both motions. The orbital equations are those of a general relativistic three-body system in a parametrized post-Newtonian (PPN) framework.[8] We include the perturbations due to the (1) leading terms in the spherical-harmonic expansion of the gravitational fields of the Earth and Moon, (2) Newtonian effects of the other planets, and (3) drag from tides on Earth. Similarly, the rotational equations include the torques produced by the Earth and the Sun on the leading moments of the Moon's figure and the effect of the anelasticity of the Moon.[3,9] The Earth's torque on the Moon includes the effect of the second zonal harmonic of the Earth's gravitational field.

Our model also includes the equations for the partial derivatives of the motions with respect to the (large) set of relevant adjustable parameters. We integrate the whole ensemble of equations simultaneously by numerical methods[8] to obtain a tabulation as a function of time for the positions and velocities of the relevant bodies, and for the partial derivatives with respect to these parameters. These latter include the mass of the Moon and of the Earth-Moon system; six initial conditions each for the heliocentric motion of the Earth-Moon system, for the Moon's geocentric motion, and for the Moon's rotation; a coefficient for the tidal drag on the Moon; the Moon's moment-of-inertia ratios and its gravitational harmonic coefficients through degree and order three; the PPN metric parameters β and γ; the rate $\dot{G}$ of a possible secular variation in the gravitational coupling constant G; and an *ad hoc* parameter h related to any extra precession of the Moon's orbit about the ecliptic pole that is not included in, or cancels part of, the predicted general-relativistic geodetic precession.

The partial derivatives for other relevant parameters, those that either do not affect at all, or do not affect significantly, the motions of bodies, are easier to compute. These latter parameters include instrumental biases; the coordinates of the laser-ranging sites and of the retroreflector locations; Love numbers for the Earth and Moon[9]; corrections to universal time (UT1) and to the Earth's pole position; corrections to the International Astronomical Union (IAU) values for the constant of general precession and for the coefficients of the series expression for the Earth's nutation[10]; and the scale of the solar system in light travel time units (i.e., the astronomical unit in seconds).

The parameter h is the key to our study since it is a measure of the (dis)agreement between the observed motion of the Moon and the de Sitter precession. We formulated the equations of motion[8] in a coordinate system rotating in the instantaneous ecliptic at a (variable) rate $h\Omega$, where Ω is the de Sitter precession rate. Any departure in the lunar motion from the de Sitter rate is characterized by the numerical factor h, where $h=0$ would be consistent with the prediction from general relativity and $h=1$ would imply a 100% error in this prediction.

The de Sitter precession may be thought of as having contributions from two sources: The first is the effect of mass on the curvature of space, which results in locally measured angles differing from those measured with respect to the fixed stars. The second source, which contributes half as much as the first, is the gravitational analog of the spin-orbit coupling of an electron in an atom. In the PPN formalism, the sum of these two contributions is proportional to $\gamma+\frac{1}{2}$.

Our primary data set was the 1970–1986 collection of LLR data, about 4400 echo-delay measurements for laser signals sent from various sites on the Earth to the various corner reflectors on the Moon. To aid in distinguishing the effects of lunar orbit precession from those of Earth nutation, we included the results of VLBI measurements in the form of (loose) *a priori* constraints on the corrections[10] to some of the coefficients of the shorter-period terms (annual, semiannual, and fortnightly) in the IAU series[11] representing the Earth's nutations. In part of our analysis, the "local" part, we also used the orbits and masses of the other planets and of relevant asteroids as obtained from our separate analysis of radar and spacecraft-tracking data. In the "global" part, we analyzed virtually all of the solar-system data simultaneously. In no case, however, did we include the optical observations of the Moon, since their standard errors are so large as to outweigh the advantage of the long time span they cover.[12]

We started our local analysis with a nominal value of zero for h and appropriate values for the other parameters, and applied our weighted-least-squares filter[13] to obtain estimates for all the relevant parameters, as well as to obtain the corresponding standard deviations and normalized correlation matrix. Of course, in this approach, we set the PPN parameters γ and β to unity and estimated neither. We found no significant departure of the geodetic precession of the Moon from that predicted by general relativity: $h=0.019\pm0.010$ (statistical standard errors are given here and hereafter unless otherwise noted). We also repeated this analysis, with $h\equiv0$, and estimated in its place an *ad hoc* coefficient λ, multiplying the relativistic terms in the equations describing the Moon's orbital motion[8]; we obtained $\lambda=1-0.010\pm0.011$, consistent with general relativity. In both analyses we estimated a total of 335 parameters, including 250 representing corrections to the temporal behavior of the Earth's orientation, and utilized all the LLR observations. We also estimated h and λ simultaneously and found the magnitude of their normalized correlation to be approximately 0.5.

These local analyses, as indicated, utilized separately determined values for such other solar-system parameters as the masses and orbital initial conditions for the other planets. We therefore also carried out a global analysis, to be described in detail elsewhere,[8] using all of our data—LLR as well as radar and spacecraft-tracking data for the planets—and solving for all relevant parameters simultaneously: a total of about 20000 observations and about 1300 parameters, including ~ 900 to represent the topography on Mercury and Venus. Although this latter approach is philosophically preferable, in practice it is not distinguishable: We obtained $h = 0.002 \pm 0.006$, not significantly different from the result from the local analysis. However, because of possible systematic errors, discussed below, neither of these statistical standard errors can be accepted as, or simply related to, the true uncertainty.

We attempted to account for the systematic errors in the global analysis by considering first the changes in our estimate of h which would result from large, but not "unreasonable," imposed changes in our estimates of other parameters. The largest effects come from those parameters that are highly correlated with h or that have "realistic" uncertainties significantly larger than the statistical standard errors, because of correspondingly large sensitivities either to systematic errors in the data or to omission from the model. This study showed that the largest of the effects on h was under 0.01 and virtually all were well under this value.

Second, we repeated our weighted-least-squares global analysis, but with different subsets of the data and of the parameters. For example, we split the LLR data into two "halves" according to time of year: approximately January-June (about 1900 observations) and July-December (about 2500 observations), and included one half at a time in otherwise identical analyses. The respective estimates for h were -0.023 ± 0.010 and 0.016 ± 0.008. We also investigated the effect of the use of progressively shorter spans of LLR data in both the global and the local analyses. The results[8] demonstrated that having LLR data over a significant fraction of 18 yr is crucial to detecting the de Sitter precession reliably.

We attempted to quantify any systematic effects that might be present in the LLR postfit residuals from the local and global analyses by (1) forming a histogram of the normalized residuals (each residual divided by the "final" standard error[8] of the observation), and (2) Fourier analyzing the residuals. The results show (1) very nearly Gaussian distributions, with (2) no significant periodicities present. The interpretation of the Fourier transforms, however, is complicated by the data having been taken at irregular intervals, and with fortnightly gaps. In addition, the fortnightly frequency is strongly absorbed by the lunar model, as are the monthly and yearly frequencies.

On the basis of all these studies, we conclude that a reliable estimate of the "true" standard error of our measurement of the de Sitter precession is 2% of the predicted effect, i.e., $h \simeq 0.00 \pm 0.02$.

We turn now to a discussion of the BCB paper. Their argument consisted of two parts; we deal with each in turn. The final argument held that any error in the general-relativistic prediction of geodetic precession would appear as an increase in the LLR postfit residuals, since the model used to analyze the data was consistent with general relativity. In particular, BCB considered the difference between the Moon's predicted mean perigee rate and the Sun's mean motion, and calculated the error in the difference that would increase twofold the rms of the residuals for two selected half-year periods.[7] This approach runs the risk that the full effect of masking (see below) on the estimate of the parameter of interest may be far more severe than anticipated. In this case, considering the full data and parameter set, we find that the cumulative effect of all other parameters leads to the statistical standard error for h being 24 times as large in our global analysis as it would be were only h being estimated. This "masking factor" represents the extent to which a linear combination of changes in the other parameters can be made to reproduce the signature of h and has the following consequence in the analysis of the LLR data: The mean-square residual R^2, which is the sum square of the postfit weighted residuals per degree of freedom, has a quadratic dependence on h: $R^2 \sim 1 + 2h^2$ for the LLR data set. One might expect R to be unity. However, the *a priori* knowledge of the standard errors in the LLR measurements, primarily from instrumental effects, is not precise; only an increase in R above the present level by more than 30% would be a reasonable indication of a problem[14]; in fact, an independent study of the residuals[15] shows that, for different large subsets of the data, R varies, but not monotonically with time, from 0.7 to 1.5. Our own earlier analysis showed variations in R, for somewhat different subsets, from about 0.6 to 2.0, also exceeding $\pm 30\%$. Systematic trends in the residuals would not be discernible easily, in view of the very uneven temporal spacing of the observations. An increase in R of more than 30% corresponds to a value of h greater than 0.6, implying, therefore, an uncertainty of *at least* 0.6 in h if determined *via* this method. BCB indicated that their result had an uncertainty equivalent to 0.1, but, as they stated, this value was not based on any detailed analysis. BCB did not examine any residuals; as noted above, the only statistic they considered was the rms of residuals for certain portions of the data. It would appear that the main problem with their assignment of a standard error was their implicit assumption that the effect of masking would be no more than a factor of 2. Moreover, a method based, as is BCB's, on the rms of residuals could not take significant advantage of a "square root of n"

reduction factor in the statistical standard error of the result, even were systematic errors not a factor, because of the limit set by the imprecise knowledge of the measurement standard errors.

The second BCB argument held that any (large) departure from the de Sitter precession would have been evident in the comparison between the mean rates of change of UT1, determined separately from LLR and from VLBI data. This idea is basically sound, but the estimated accuracy suffers in the same way, though to a lesser extent, from the lack of data analysis. To test this second method, we performed an analysis that attempted to match the scenario described in BCB. We carried out a series of analyses, corresponding to our local solution, but with h constrained to different values in each. Since we estimate corrections to UT1 values for epochs spread throughout the time span of the LLR observations, we found the weighted-least-squares estimate of the slope of these corrections in each case. Comparison of the results showed that the estimated UT1 rates reflected only about half of the imposed offset ($h \neq 0$) in the geodetic precession rate, thereby yielding a masking factor of about 2. We then estimated the standard deviation of h as that value needed to distinguish, at the 1-standard-deviation level, the difference between the UT1 rates estimated from the LLR data and that estimated from the VLBI data. We obtained $\sigma(h)=0.2$, rather than the equivalent 0.1 given by BCB. Thus, their second approach yields a result still short by about tenfold of the accuracy achievable by actual analysis of the LLR data.

We thank Peter L. Bender for several stimulating discussions, especially on the paper of BCB. We also thank the National Science Foundation, Grant No. PHY-84-09671, and the National Aeronautics and Space Administration (NASA), Grant No. NAGW-967, for partial support.

[1]S. de Sitter, Mon. Not. Roy. Astron. Soc. **77**, 155 (1916).

[2]See, for example, S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972), p. 237. This precession is also sometimes referred to as "geodesic."

[3]M. A. Slade, Ph.D. thesis, Massachusetts Institute of Technology, 1971 (unpublished).

[4]T. Stevenson, B.A. thesis, Massachusetts Institute of Technology, 1979 (unpublished).

[5]C. F. W. Everitt, private communication.

[6]C. O. Alley *et al.*, Science **167**, 458 (1970).

[7]B. Bertotti, I. Ciufolini, and P. L. Bender, Phys. Rev. Lett. **58**, 1062 (1987).

[8]J. F. Chandler, R. W. Babcock, R. D. Reasenberg, and I. I. Shapiro, to be published.

[9]R. J. Cappallo, C. C. Counselman, III, R. W. King, and I. I. Shapiro, J. Geophys. Res. **86**, 7180 (1981); A. J. Ferrari, W. S. Sinclair, W. L. Sjogren, J. G. Williams, and C. L. Yoder, J. Geophys. Res. **85**, 3939 (1980).

[10]T. A. Herring, C. Gwinn, and I. I. Shapiro, J. Geophys. Res. **9**, 11 416 (1986).

[11]P. A. Seidelmann, Celestial. Mech. **27**, 79 (1982).

[12]The LLR data, on the other hand, are so sensitive that we were able to show via a numerical experiment, with the appropriate partial derivatives, that, remarkably, the mass of Jupiter can be estimated from LLR data alone with a statistical standard error barely more than an order of magnitude greater than that obtained from all relevant spacecraft, Jupiter-satellite, and asteroid observations.

[13]See, for example, M. E. Ash, Massachusetts Institute of Technology Lincoln Laboratory Technical Note No. 1972-5, 1972 (unpublished).

[14]R. W. King, private communication.

[15]A. L. Whipple, private communication.