Proceedings of the

NATIONAL ACADEMY OF SCIENCES

Volume 51 • Number 5 • May 15, 1964

EVIDENCE FOR A COMET BELT BEYOND NEPTUNE

BY FRED L. WHIPPLE

SMITHSONIAN ASTROPHYSICAL OBSERVATORY AND HARVARD COLLEGE OBSERVATORY, CAMBRIDGE, MASSACHUSETTS

Communicated March 30, 1964

In the condensation of an extended mass of gas and dust to form the Sun and the planetary system, Cameron[1] has suggested that "there must be a tremendous mass of small solid material on the outskirts of the solar system." I concur with this conclusion;[2] the composition of Uranus and Neptune appear close to the expected composition of comets if, following the icy-comet model,[3] the comets were derived from gas comparable in composition to the Sun, but were frozen out of this gas at temperatures perhaps much below 100°K. If the comets are indeed the building blocks of the outermost plants, then the Öpik-Oort[4, 5] great cloud of comets may have been derived from those formed within the planetary region but perturbed gravitationally into extremely elongated orbits. Those "cometesimals" forming beyond this region did not collect into planets because of the low surface density of the preplanetary cloud; they remain in a ring extending indefinitely beyond the "perturbative reach" of Neptune but near the plane of the planets. Pluto may thus be a cometary accumulation or, as suggested by Lyttelton,[6] a lost satellite of Neptune.

If the largest comets in the postulated belt have radii not much exceeding 100 km, or approximately 1/30 Pluto's diameter, their apparent magnitude would be approximately 22 and the chance of discovery remote, even though a number of such comets might exist.

If the total mass of the comet belt is taken as 10 Earth masses at a distance of 50 a.u. from the Sun in a layer only 1 a.u. thick, or approximately a degree perpendicular to the plane of the planetary orbits, the ring cannot contribute much to the Zodiacal Light by reflected sunlight unless the cometesimals average rather small in dimension. For a uniform diameter of 1 km and an albedo of 0.07, the ring would contribute 8.5 magnitudes per square degree, some 4.5 magnitudes fainter than the average night sky, and could not be easily distinguished from the Zodiacal Light unless its plane differed quite appreciably from the plane of the ecliptic. The extremely precise measures that would be necessary to detect such a ring would probably best be made from outside the Earth's atmosphere. Thus we see that, since direct observational evidence for the existence of a cometary belt may not be available for some time to come, we must look for possible gravitational effects.

Perturbative Effects of a Comet Belt.—A comet belt would perturb the motions of a

TABLE 1

VALUES OF B (EQ. 1) AND A (EQS. 2 AND 3)

a_p/a or r/a	B (a_p/a)	A (r/a)
0.00	0.000000	0.000000
0.05	0.001884	0.000063
0.10	0.007643	0.000506
0.15	0.01761	0.001731
0.20	0.03239	0.004188
0.25	0.05291	0.008400
0.30	0.08058	0.01501
0.35	0.1174	0.02482
0.40	0.1666	0.03891
0.45	0.2326	0.05870
0.50	0.3226	0.08622
0.55	0.4480	0.1244
0.60	0.6280	0.1776
0.65	0.8969	0.2529
0.70	1.320	0.3621
0.75	2.036	0.5265
0.80	3.377	0.7895
0.85	6.324	1.253
0.90	14.88	2.227
0.95	61.79	5.274
1.00	∞	∞

planet and should be most effective in producing a retrograde motion of its orbital node or the directions of its pole about a small circle. From Plummer[7] (pp. 147, 173, and 198) a circular comet ring of one dimension, with a total mass m, at a distance a from the Sun (of mass m_s), would impart the following retrograde motion to the node of a planet of mass m_p in a smaller circular orbit of semimajor axis a_p:

$$\frac{d\Omega}{dt} = -\frac{Bm}{a}\left[\frac{G}{a_p(m_s + m_p)}\right]^{1/2}, \tag{1}$$

where G is the constant of gravity, B is a dimensionless function of a_p/a expressible by a Laplace coefficient in the form $B = a_p b^1{}_{3/2}/(4a)$, and the inclination is small. Numerical values of B are given in Table 1.

For a planet in the plane of the ring at distance r from the Sun, the ratio, R, of the radial force outward to that inward by the Sun is given by

$$R = A(\gamma)\frac{m}{m_s}, \tag{2}$$

in which it can be shown that

$$A(\gamma) = \frac{\gamma}{\pi}\left[\frac{E(k^2)}{1-\gamma} - \frac{K(k^2)}{1+\gamma}\right], \tag{3}$$

where $\gamma = r/a$, and $E(k^2)$ and $K(k^2)$ are the complete elliptic integrals of $k^2 = 2\gamma^{1/2}/(1 + \gamma)$. Numerical values of $A(\gamma)$ are given in Table 1.

The major effect of a comet ring on planetary longitudes would arise from the reduction in effective solar gravity as given by R in equation (2). An outer planet in a circular orbit would actually move in a smaller orbit a_p than a_p', calculated from the period. To first-order terms we find

$$\frac{a_p' - a_p}{a_p} = \frac{R}{3}. \tag{4}$$

For a planet at a great distance from the Sun, the semimajor axis is, in fact, almost entirely determined by the period because the best observations are made near opposition. The parallactic error, $\Delta\lambda$, made in the longitude observations because of the incorrectly determined value of a_p is then $(O - C)$

$$\Delta\lambda = -\frac{a_p R \rho_s}{3\rho_p^2} \sin H, \tag{5}$$

where H is the heliocentric longitude of the Earth minus that of the planet; ρ_p is the distance from Earth to planet; and ρ_s, the distance to the Sun.

A maximum value of sin H lies near $\rho_p = a_p$. Hence, the maximum value of $\Delta\lambda$ is, for a_p in a.u., approximately

$$\Delta\lambda(\text{max}) = \pm \frac{R}{3a_p}. \tag{5a}$$

As an example, for a comet ring of 10 Earth masses, at a solar distance of 40 a.u. we find for Neptune $R = 1.6 \times 10^{-5}$ and $\Delta\lambda(\text{max}) = \pm 0\rlap{.}''036$; for Uranus $R = 2.6 \times 10^{-6}$ and $\Delta\lambda(\text{max}) = 0\rlap{.}''009$. Thus, the radial effect of our postulated comet belt on any of the planets, except possibly Pluto, is probably unmeasurable at present. We must look to motions of the nodes for evidence of the belt.

Observations of planets when compared with theory will yield residuals in the latitude, β. We wish to calculate the motion in β for a planet that is being perturbed by the comet belt. Let u be the longitude of the planet at time t, measured from its ascending node at the ecliptic, and let Ω be the angle from the ecliptic node to the ascending node with respect to the plane of the comet belt.

The comet belt will exert an acceleration W, normal to the orbital plane, that can be written

$$W = -2n^2 aK \sin (u - \Omega), \tag{6}$$

where n is the mean motion of the planet in a circular orbit, and K is a dimensionless constant to be evaluated from observation.

On integration of the usual equations for perturbations in node and inclination, with respect to the ecliptic, we can calculate the perturbations in β arising from the comet ring,

$$\frac{\Delta\beta}{K} = \frac{r}{\rho} \sin u(P + u \sin \Omega) + \frac{r}{\rho} \cos u(Q + u \cos \Omega), \tag{7}$$

where $P = \sin u_0 \sin (u_0 - \Omega) - u_0 \sin \Omega - \cos \Omega$; $Q = \sin u_0 \cos (u_0 - \Omega) - u_0 \cos \Omega$, ρ is the geocentric distance of the planet, and the perturbation $\Delta\beta$ is integrated from t_0 and u_0.

The constant K can be found from the normal equations for the planetary orbit, solutions being made for various values of Ω to minimize the weighted squares of the remaining residuals in β.

We must then determine the time-averaged value of $d\Omega/dt$ from the usual equation for a circular orbit,

$$\frac{d\Omega}{dt} = \frac{\sin (u - \Omega)}{na \sin i} W, \tag{8}$$

where i is the inclination of the planetary orbit to the plane of the comet belt.

This average yields, on substituting W from equation (6) and integrating for one period,

$$\frac{d\Omega}{dt_{\text{avg.}}} = -\frac{K}{\sin i}\left[\frac{G(m_s + m_p)}{a^3}\right]^{1/2}. \tag{9}$$

Equating equation (9) to equation (1), we find for the corresponding mass of the comet belt

$$m = (m_s + m_p)\frac{K}{B}\frac{a}{a_p \sin i}, \tag{10}$$

where K is measured in radians.

Note that equation (10) involves three unknowns, a/a_p (directly and in B), m, and $\sin i$. Were the perturbations of the ring measurable for two planets, we could determine Ω, and hence $\sin i$, for each. The corresponding values of K could theoretically provide a solution for m and a/a_p, a "complete solution." In fact, however, an actual comet belt would have a considerable extent in solar radius. More information is thus required for a truly complete solution.

Observed Perturbations on Neptune and Uranus.—From a definitive orbit of Neptune over the interval from 1795 to 1938 (1846–1938 excluding Lalande's prediscovery observations), L. R. Wylie[8] finds that systematic residuals in the heliocentric

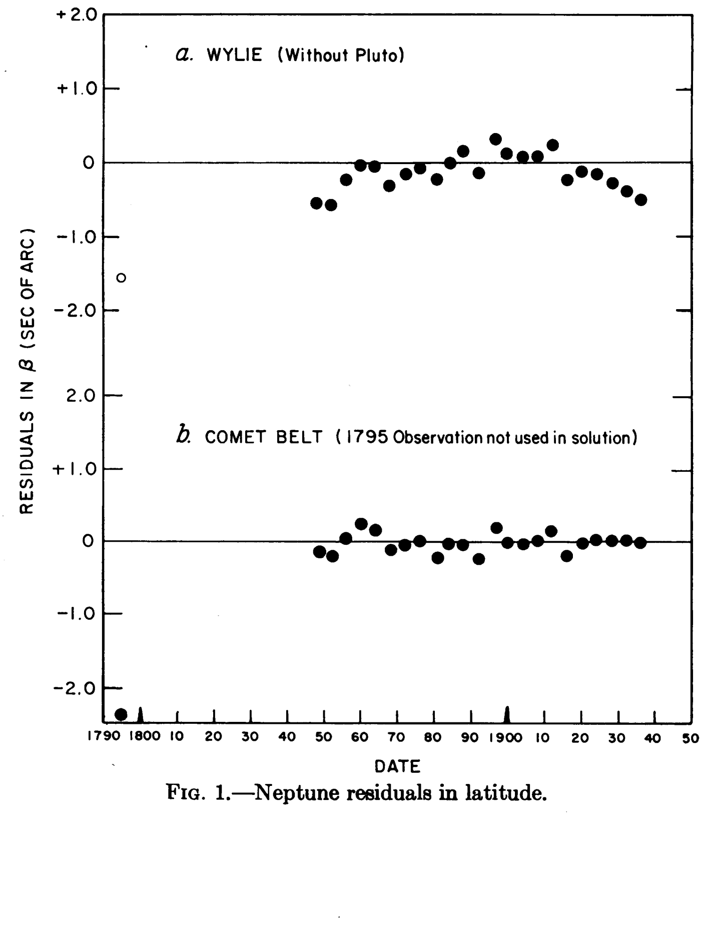

FIG. 1.—Neptune residuals in latitude.

latitude (Fig. 1) can be attributed to perturbations by Pluto. A solution for the mass of Pluto reduces the weighted sum of the squares of the residuals to 25".6, from 83".2. Wylie finds, however, that the latitudes are about as well represented by a secular motion of the node.

From Wylie's[9] normal equations for the positions of Neptune and Uranus over the period 1785–1935 least-squares solutions for the motions of the planes of the orbits have been made; the results are portrayed graphically in Figure 2. Here the ellipses indicate the standard deviations in the solutions for the motions of the orbital poles, the axes of the ellipses representing the standard deviations in Δi and in $\sin i\Delta\Omega$. The fit with observation is slightly better for Neptune than though the effect were attributed to Pluto. For Uranus the solutions areq uite indeterminate; they do not remove the unexplained secular change in latitude.

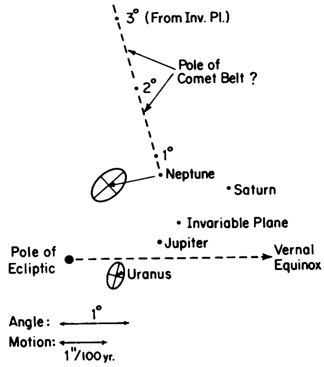

FIG. 2.—Poles of planet orbits (ellipses represent m.e. of motions).

Figure 2 shows that to produce a retrograde motion of the pole of Neptune's orbit, the pole of the comet belt must be oriented approximately along the dashed line at the top of the figure. The position of the pole is only slightly constrained by the requirement that the effect on Uranus must be too small to be detectable. The comet belt must therefore be inclined somewhat more than 1° from the invariable plane of the planets.

Although certain limitations on the position and mass of the comet belt can be obtained from the motion of the orbital planes of Neptune and Uranus, a more precise analysis is justified. Hence, least-squares solutions for the coefficients of K were made from Wylie's normal equations in latitude for both Neptune and Uranus by application of equation (7). The unknowns were Δi_p, $\sin i_p\Delta\Omega_p$ and K. To solve for Ω, solutions were made at 10° intervals in Ω; the value adopted was that for which the weighted sum of the squares of the residuals, Σpv^2, was a minimum.

For Neptune the starting orbit was the same as Wylie's, for which Pluto's effect was neglected ($\Delta B'$ in his Table 5). For Uranus, similarly, Wylie's starting orbit was used (dB' of his Table 15). The coefficients for Δi_p and $\sin i_p\Delta\Omega_p$ were taken from his normal equations, which for Uranus means that the ratio r/ρ was taken as unity, an acceptable approximation.

Table 2 lists the solutions described above. A comparison of the Σpv^2 *if* $K = O$ with the *Optimum* Σpv^2 shows the effect of optimizing Ω. This effect is very significant for Neptune, but not for Uranus. The mean error in K is much underestimated because, in fact, the solution for Ω was not included in the least squares. A better estimate can be made from the solution for the motion of the pole, which for Neptune was 1".04 (± 0".45) per 100 yr. Thus, $K = +0''.28$ (±0".12) is a realistic value. For Uranus $K = +0''.02$ (±0".06) is probably meaningful.

We may conclude that for Neptune the assumption of a comet belt leads to a slightly better fit with the latitude observations than does the assumption of Pluto perturbations; Σpv^2 becomes 23.6 instead of 25.6, even though the 1795 observation was neglected in the solution but not in the residuals. Either assumption reduces Σpv^2 by more than a factor of three, a rather significant amount. For Uranus

TABLE 2

ORBITAL CORRECTIONS

Planet	Neptune	Uranus
Σpv^2 if $K = 0$	83.2	1.14
Optimum Σpv^2	23.6	1.12
Optimum Ω	240°	300°
Δi_p (m.e.)	$-0''.164$ ($\pm 0''.048$)	$-0''.41$ (± 0.11)
$\sin i_p \, \Delta\Omega_p$ (m.e.)	$+0''.498$ ($\pm 0''.037$)	$+0''.68$ ($\pm 0''.07$)
K (m.e.)	$+0''.278$ ($\pm 0''.036$*)	$+0''.020$ ($\pm 0''.025$*)
t_0	1900.08	1851.45

* Too small; see text.

neither assumption improves the fit significantly nor helps remove the secular trend in the residuals in latitude. For neither planet is the fit in longitude improved significantly by either assumption, except for the 1795 observation of Uranus.

These facts provide some support for the existence of a comet belt, particularly as the calculated value of the mass of Pluto falls in the range of 0.5–1.0 Earth masses, which is far greater than physical observation can allow. Note that the observations of Uranus previous to 1781, analyzed by Woolard,[10] have not been included in the solutions here. Brouwer[11] found that "The accidental error of these old observations is disappointingly large."

Possible Mass of the Comet Belt.—We can calculate the mass of the comet belt from the derived value of K and equation (10) if we assume a mean distance from the Sun and inclination to the plane of Neptune's orbit. For the optimum $\Omega = 240°$, the ascending node of Neptune's orbit with respect to the plane of the comet belt lies in longitude $131° + 240° = 11°$. The pole of the comet belt thus lies very nearly along the great circle indicated in Figure 2; its plane is inclined $0^\circ\!.8$ or more to the invariable plane and about $1^\circ\!.4$ more with respect to the orbit of Uranus than to that of Neptune.

For Neptune we may adopt $K = 0''.28$ from Table 2 in equation (10) to derive a mass for a comet ring. Suppose the ring is 40 a.u. from the Sun. Then B (Table 1) $= 2.04$ and

$$\text{Mass (comet ring)} = 8.4 \left(\frac{\sin 2°}{\sin i_N}\right) \text{ Earth Masses,} \tag{11}$$

where i_N is the assumed inclination of Neptune's orbit to that of the comet ring.

For Uranus we may adopt $K = 0''.06$, the estimated error, to represent a maximum value. Roughly, the inclination of Uranus' orbit to that of the comet ring will be $i_N + 1^\circ\!.4$, so that for $a = 40$ a.u., $B = 0.32$, and

$$\text{Mass (comet ring)} \leq 10.1 \frac{\sin 3^\circ\!.4}{\sin (i_N + 1^\circ\!.4)} \text{ Earth Masses.} \tag{12}$$

Placing the comet ring at a distance of 50 a.u. from the Sun, we find that the corresponding coefficients in equations (11) and (12) become, respectively, 27 and 20 Earth masses. As expected, the mass needed in the comet belt increases with solar distance but is of the order of 10–20 Earth masses at distances of 40–50 a.u. for an inclination of 2° to the orbit of Neptune. The calculated mass varies inversely as $\sin i$.

Because of its high inclination and large semimajor axis, the orbit of Pluto should eventually give a definitive answer regarding the existence of a comet belt. At the moment, however, the improved orbit by Sharaf[12] exhibits semiperiodic variations in the declination residuals, with a semiamplitude of about $\pm 2''$ and a period of about 11 years. Thus, any perturbations in declination arising from a comet belt are not conspicuous.

A search for a parallactic longitude effect, as suggested by equation (5), was made for Neptune from the residuals given by Wylie's final orbit. A strong effect appeared among the early observations, 1846–1860, but it was opposite in sign to expectations and an order of magnitude too large ($\sim 0\overset{''}{.}5$). In the remaining residuals, 1861–1906, the effect was greatly reduced, possibly below the level of significance. It probably represents a systematic observational error among evening observations, which are much more numerous than morning observations. The data for Uranus could not be tested simply for a longitude effect.

In conclusion, the evidence for a comet belt lying beyond the planets in a plane a very few degrees from the invariable plane is slightly stronger than the evidence for the mass of Pluto. Brouwer[11] states with respect to the mass of Pluto: "I have now come to the conclusion that the gravitational determination is not entitled to the weight that the formal solution indicates." Definitive orbits of Neptune and possibly of Uranus or Pluto based on currently available observations and theory might enable us to distinguish among the alternatives: (*a*) Pluto is massive, and no appreciable comet belt exists; (*b*) Pluto has a negligible mass, while a comet belt of appreciable mass does exist; (*c*) neither Pluto nor the postulated comet belt has appreciable mass.

It is interesting to note that such a comet belt could be a source of appreciable meteoritic material, particularly finer material, possibly contributing to the Zodiacal Light. Drawn in toward the Sun from large orbits by the Poynting-Robertson Effect, such material would suffer orbital vicissitudes by planetary perturbations much as do comets and cometary debris. Thus, neither its distribution in space nor its orbital characteristics would be distinctive. If we find that comets are indeed inadequate to maintain the zodiacal cloud, perhaps the comet belt will supplement the supply.

Summary.—Certain theories of planetary formation lead to the possibility that a stable belt of comets exists near the plane of the planets beyond the orbit of Neptune. This belt would not be easily visible nor would it contribute appreciably to the observable comets. It would, however, exert a gravitational effect on the outer planets. Calculations indicate that it accounts for observed disturbances of the motion in Neptune's latitude as well as or slightly better than the previously assumed attraction attributed to Pluto. Since Pluto appears far too small to have a mass comparable to the Earth, a belt of comets comprising perhaps 10–20 Earth masses may well exist in the region of 40–50 a.u. from the Sun and beyond. Modern definitive orbits of Neptune, Uranus, and Pluto might permit an answer to this question.

The assistance and advice of Imre G. Izsak have contributed greatly to this paper. I am also grateful to Rudolf Loeser, who programmed and supervised most of the calculations on an IBM-7094 computer.

[1] Cameron, A. G. W., "The formation of the Sun and Planets," *Icarus*, **1**, 13–69 (1962).

[2] Whipple, F. L., Centennial (First Scientific Session), "History of the solar system," these Proceedings, in press.

[3] Whipple, F. L., "A comet model. I. The acceleration of Comet Encke," *Ap. J.*, **111**, 375–394 (1950).

[4] Öpik, E., "Note on stellar perturbations of nearly parabolic orbits," *Proc. Amer. Acad. Arts Sci.*, **67**, 169–182 (1932).

[5] Oort, J. H., "The structure of the cloud of comets surrounding the solar system, and a hypothesis concerning its origin," *Bull. Astron. Inst. Neth.*, **11**, 91–110 (1950).

[6] Lyttleton, R. A., "On the possible results of an encounter of Pluto with the Neptunian system," *Monthly Notices Roy. Astron. Soc.*, **97**, 108–115 (1936).

[7] Plummer, H. C., *An Introductory Treatise on Dynamical Astronomy* (Cambridge Univ. Press, 1918).

[8] Wylie, L. R., "A comparison of Newcomb's tables of Neptune with observation, 1795–1938," *Publ. U.S. Naval Obs.*, Ser. II, vol. 15, Part 1, pp. 1–21 (1942).

[9] Wylie, L. R. "An investigation of Newcomb's theory of Uranus," *Publ. U. S. Naval Obs.*, Ser. II, vol. 15, Part 3, pp. 41–71 (1942).

[10] Woolard, E. W., "Comparison of the observations of Uranus previous to 1781 with theoretical positions obtained by numerical integration," *A.J.*, **57**, 35–38 (1952).

[11] Brouwer, D., "The motions of the outer planets," *Monthly Notices Roy. Astron. Soc.*, **115**, 221–235 (1955).

[12] Sharaf, S. G., "Improvement of the elements of Pluto," chap. 7, *Tr. Inst. Teor. Astron.*, no. 4, 79–94 (1955).

*OUTLINE OF A THEORY OF THE UNIVERSE**

BY T. Y. THOMAS

INDIANA UNIVERSITY, BLOOMINGTON

Communicated March 16, 1964

1. *Introduction.*—The inadequacy of the current cosmological theories for the explanation of the discretized galactic structure found by A. G. Wilson[1] clearly necessitates a modification of our over-all viewpoint of the universe. In the following note we have given an outline of a theory of *epochal cosmology* to provide the theoretical background for these discretization effects and to serve as a basis for later mathematical investigations.

The essential features of this theory can be stated as follows: It is assumed that the universe has existed throughout the infinite past in a state devoid of ponderable matter and that it would have continued to exist in such a state except for the *accidental* creation of material particles. This initial stage of the universe is discussed in more or less detail in section 2. Its termination with the creation of material particles is considered in section 3 and is based on the possibility of such creation within the framework of the Einstein field equations, i.e., the well-known equations

$$B_{AB} - 1/2\, Bh_{AB} = \varkappa T_{AB}, \qquad (A,B = 0,1,2,3), \tag{1}$$

in which h_{AB} and B_{AB} are the components of the metric and Ricci tensors, respectively, the quantity B is the scalar curvature, the T_{AB} are the components of the energy tensor, and finally $\varkappa$ is an absolute or universal constant. The explosion of an accumulation of material created toward the end of the initial stage produced the first cosmological epoch and also gave rise to the series of *epochs of creation* which continue as a creative force in the universe and which are significantly related to the observed discretization in red shifts and the ages of galaxies