Publications of the Astronomical Department of Eötvös University
No. 6

DYNAMICAL ASTRONOMY

Proceedings of a Workshop
Sponsored by the NSF-U. S. and by
the Hungarian Academy of Sciences

Edited by

Béla A. Balázs

and

Victor Szebehely

Budapest, 1982

Publications of the Astronomical Department of Eötvös University
No. 6

DYNAMICAL ASTRONOMY

Proceedings of a Workshop
Sponsored by the NSF-U. S. and by
the Hungarian Academy of Sciences

Edited by

Béla A. Balázs

and

Victor Szebehely

Budapest, 1982

Proceedings of a Workshop on Dynamical Astronomy
sponsored by the NSF-U.S. and by the Hungarian
Academy of Sciences, August 31-September 4, 1981

Co-Editors and Co-Directors:

Dr. B. Balázs
Eötvös Lorand University
Department of Astronomy
Budapest, Hungary

Dr. V. Szebehely
L. B. Meaders Professor
Department of Aerospace Engineering
University of Texas
Austin, Texas, U.S.A.

Published by the Eötvös University
Budapest, 1982

C O N T E N T S

P R E F A C E

The organizers of this workshop wish to express their deepest appreciation to the participants for their clearly displayed scientific competence, for their cooperation and for their untiring efforts during the long lecture and discussion sessions. The benefits were unquestionably mutual; the personal meetings of scientists interested in the same or similar subjects created natural resonances which will not damp out easily. In fact, we expect that the forcing functions of mutual interests will lead to further exchange of information and to several joint projects. Examples of subjects of common interest to at least two or three participants were basic problems of geo-dynamics, motions of asteroids and quantitative and qualitative behavior of galaxies.

One purpose of publishing the Proceedings of this workshop is to stabilize the cooperations already established and to call the attention to other members of our profession, not so lucky as to attend our workshop, to the proposed cooperative effort. Our invitation is to all those interested in these projects to join our future activities.

We also wish to thank for the support, cooperation and understanding of our sponsors, the Hungarian Academy of Sciences and the United States National Science Foundation.

We offer this volume to our colleagues in both countries with the humbleness of the beginner, hoping that they will realize that any faults are the faults of the organizer and all credits should go to our participants and to our sponsors.

B. Balázs
Eötvös Loránd University
Budapest, Hungary

and

V. Szebehely
University of Texas
Austin, Texas

SELECTED RESEARCH PROJECTS IN CELESTIAL MECHANICS AT THE UNIVERSITY OF TEXAS AT AUSTIN

Dr. Victor Szebehely, L. B. Meaders Professor
Department of Aerospace Engineering, University of Texas
Austin, Texas 78712, U.S.A.

Abstract

This paper reports on recent research on the inverse problem, on stability, on accuracy and on undeterminacy in celestial mechanics.

(1) The inverse problem of dynamics deals with the establishment of the potential function (or in general of the force-field) when a family of orbits is given. The linear partial differential equation for the potential V in the two-dimensional case is

$$A(x,y)\ V + B(x,y)\ \frac{\partial V}{\partial x} + C(x,y)\ \frac{\partial V}{\partial y} = D(x,y) \quad ,$$

where the functions A, B, C and D depend on the given family of orbits. It is known that the inverse problem does not have a unique solution corresponding to the fact that V is the solution of a partial differential equation, which solution is expressed by arbitrary functions of the combinations of the variables describing the system. The linearity of the above equation is lost when rotating (synodic) systems are used as in the restricted problem of three bodies. Other generalizations are the extension to three

dimensions, using generalized coordinates, finding potentials which describe integrable dynamical systems, applications to stellar systems, etc. See References 1-6.

(2) The measure and regions of stability of artificial and natural celestial bodies is established using various stability criteria, such as linearization, Hill's method, Lyapunov's characteristic numbers, etc. The quantitative measure S is represented by the dimensionless deviation of the actual and critical values of a constant of the motion, or by

$$S = \frac{C_{act} - C_{crit}}{C_{crit}} \quad .$$

Stability requires $S > 0$ and the value of $S = 0$ corresponds to bifurcation. This presentation is especially well adapted to Hill's method where C is the Jacobian constant. Stability regions of planetary and satellite systems are established within the solar system and in binary systems using the above criterium reformulated in terms of masses and orbital radii. See References 7-14.

Stability regions in the phase space are established around the Earth-Moon libration points by numerical integration and the effects of initial position and velocity errors are studied. In this work the stability criterium is the requirement for continued libration as opposed to other types of motions, such as chaotic behavior. See References 15-18.

Stability investigations concerning the general problem of three

bodies as applied to triple stellar systems are given in References 18 and 19. This work is aimed at present time to improve the assumptions made in establishing our model.

A study in progress is the stability and long-time behavior of asteroids. Hill's and Lyapunov's methods are used to study the effects of the orbital parameters on the behavior of large numbers of asteroids and comets.

(3) To separate errors of numerical integrations and of analytical approximations from the truly random behavior of dynamical systems, increased accuracy is needed. The approaches to this problem are the method of regularization (introduction of new dependent and independent variables), the introduction of the concept of time element, and forcing the system to remain on some constant surface in the phase space. These techniques allow long-time predictions concerning the solar system and the accurate computations of relative motions and possible collisions. See References 20-29.

(4) The most recent research interest is the problem of the non-deterministic nature of celestial mechanics. Due to the fact that systems of interest in our field are non-linear and non-integrable we must rely on analytical and numerical approximations as mentioned above. To this we must add our incomplete knowledge of the physical laws governing these systems and the uncertainity of the initial conditions used. In this way we might arrive at the conclusion that celestial mechanics is a non-deterministic science and open to approaches used in statistical mechanics. This is a considerable

change from the classical approach to dynamics according to which the knowledge of the initial conditions determine completely the solution. See Reference 30.

(5) Current research in satellite dynamics is discussed in Reference 31.

References

1. Szebehely, V., "The Determination of the Potential" in Proc. Int. Mtg. on the Rotation of the Earth (E. Proverbio, editor) Bologna, 1974, p. 31.
2. Broucke, R., and Lass, H., Cel. Mech., Vol. 16, 1977, p. 215.
3. Morrison, F., Cel. Mech., Vol. 16, 1977, p. 227.
4. Broucke, R., Int. J. of Eng. Sc., Vol. 17, 1979, p. 1151.
5. Szebehely, V., Lundberg, J., and McGahee, W. J., Astrophysical Journal, Vol. 239, 1980, p. 880.
6. Szebehely, V., and Broucke, R., "Determination of the Potential in a Synodic System", Cel. Mech., Vol. 24, 1981, p. 23.
7. Szebehely, V., Proc. Nat. Acad. of Sc., Vol. 75, 1978, p. 5743.
8. Hill, G., Am. J. Math., Vol. 1, part 5, 1878, pp. 129 and 245.
9. Szebehely, V., and McKenzie, R., Astron. Jour., Vol. 82, 1977, p. 303.
10. Szebehely, V., and McKenzie, R., Astron. Jour., Vol. 82, 1977, p. 79.
11. Nacozy, P., and Ferraz-Mello, S., "Natural and Artifical Satellite Motion" (editors), 1979, p. 175.
12. Szebehely, V., ZAMP, Vol. 30, 1979, p. 364.
13. Szebehely, V., Cel. Mech., Vol. 22, 1980, p. 7.
14. Szebehely, V., and McKenzie, R., Cel. Mech., 1981, Vol. 23, p. 3.

15. Szebehely, V., and McKenzie, R., Cel. Mech., 1981, Vol. 23, p. 131.

16. McKenzie, R., and Szebehely, V., Cel. Mech., 1981, Vol. 23, p. 223.

17. Szebehely, V., and Premkumar, L., Cel. Mech., 1982, in print.

18. Zare, K., Cel. Mech., Vol. 16, 1977, p. 35.

19. Szebehely, V., and Zare, K., Astron. and Astrophysics, Vol. 58, 1977, p. 145.

20. Stiefel, E., and Scheifele, G., "Linear and Regular Celestial Mechanics", 1971, Springer Publ., Heidelberg.

21. Sundeman, K., Acta Math., Vol. 36, 1912, p. 105.

22. Nacozy, P., Cel. Mech., Vol. 16, 1977, p. 309.

23. Nacozy, P., Cel. Mech., Vol. 14, 1976, p. 129.

24. Szebehely, V., "Theory of Orbits", 1967, Academic Press, N.Y.

25. Velez, C. E., Cel. Mech., Vol. 10, 1974, p. 405.

26. Nacozy, P., Astron. Journ., Vol. 83, 1978, p. 522.

27. Nacozy, P., Cel. Mech., Vol. 22, 1980, p. 19.

28. Kuiper, G., Cel. Mech., Vol. 9, 1973, p. 321

29. Nacozy, P., and Szebehely, V., Cel. Mech., Vol. 13, 1976, p. 449.

30. Szebehely, V., in "Applications of Modern Dynamics to Celestial Mechanics and Astrodynamics", 1982, Reidel Publ. Co., Dordrecht, Holland, in print.

31. Szebehely, V., and Tapley, B., American Inst. Aeronautics and Astronautics, Reprint 81-0023, 1981.

RECENT PROGRESS IN THE THEORY OF TROJAN ASTEROIDS

B. Garfinkel
Yale University Observatory
New Haven, Connecticut, U.S.A.

ABSTRACT

In previous publications the author has constructed a long-periodic solution of the problem of the motion of the Trojan asteroids, treated as the case of 1:1 resonance in the restricted problem of three bodies. The recent progress reported here is summarized under three headings:

1) The nature on the long-periodic family of orbits is reexamined in the light of the results of the numerical integrations carried out by Deprit and Henrard (1970). In the vicinity of the critical divisor

$$D_k \equiv \omega_1 - k\omega_2,$$

not accessible to our solution, the family is interrupted by bifurcations and short-periodic bridges. Parametrized by the normalized Jacobi constant α^2, our family may, accordingly, be defined as the union of admissible intervals, in the form

$$\mathcal{L} = U_j \left\{ |\alpha - \alpha_j| > \epsilon_j \right\}; \quad j = k, k+1, \ldots \infty .$$

Here, $\{\alpha_j(m)\}$ is the sequence of the critical α_j corresponding to the exact j:1 commensurability between the character-

istic frequencies ω_1 and ω_2 for a given value of the mass parameter m.

Inasmuch as the "critical" intervals $|\alpha - \alpha_j| < \varepsilon_j$ can be shown to be disjoint, it follows that, despite the clustering of the sequence $\{\alpha_j\}$ at $\alpha = 1$, as $j \to \infty$, the family extends into the vicinity of the separatrix $\alpha = 1$, which terminates the "tadpole" branch of the family.

2) Our analysis of the epicyclic terms of the solution, carrying the critical divisor D_k, supports the Deprit and Henrard refutation of the E. W. Brown conjecture (1911) regarding the termination of the tadpole branch at the Lagrangian point L_3.

However, the conjecture may be revived in a refined form, "The separatrix $\alpha = 1$ of the tadpole branch spirals asymptotically toward a limit cycle centered on L_3."

3) The period $T(\alpha, m)$ of the libration in the mean synodic longitude λ in the range

$$\lambda_1 \leq \lambda \leq \lambda_2$$

is given by a hyperelliptic integral. This integral is formally expanded in a power series in m and α^2 or $\beta \equiv \sqrt{1 - \alpha^2}$.

The large amplitude of the libration, peculiar to our solution, is made possible by the mode of the expansion of the disturbing function R. Rather than expanding about

Lagrangian point L_4, with the coordinates $r = 1$, $\theta = \pi/3$, we have expanded R about the circle $r = 1$. This procedure is equivalent to analytic continuation, for it replaces the circle of convergence centered at L_4 by an annulus $|r - 1| < \varepsilon$ with $0 \leq \theta < 2\pi$.

REFERENCES

Brown, E. W. 1911 Mon. Not. RAS 71, 438.

Deprit, A. and Henrard, J. 1970 "Periodic Orbits, Stability, and Resonances", ed. G. E. O. Giacaglia, p. 1 (Reidel, Dordrecht).

Garfinkel, B. 1977 Astron. J. 82, 368.

Garfinkel, B. 1978 Celest. Mech. 18, 259.

Garfinkel, B. 1980 Celest. Mech. 22, 267.

ON THE VARIATION OF THE JACOBI CONSTANT OF TROJAN ASTEROIDS IN THE ELLIPTIC RESTRICTED PROBLEM OF THREE BODIES

B. Érdi
Department of Astronomy, Eötvös University
Budapest, Hungary

Abstract

A relation corresponding to the Jacobian integral of the circular restricted problem of three bodies is derived in cylindrical coordinates for the elliptic restricted three-body problem. The unknown integral appearing in this relation is evaluated for the case of Trojan asteroids using an asymptotic solution for their motion. Analytic expressions for the main variations of a parameter C^* corresponding to the Jacobi constant in the elliptic case are obtained. It is shown that the main variations of C^* depend on two long periods, the period of the libration around the point L_4 and the period of the motion of the perihelion of the asteroids. Upper limits for the amplitudes of the main variations of C^* are also given.

1. Introduction

One of the essential differences between the circular and the elliptic restricted problem of three bodies is that the circular problem possesses an integral, the Jacobian integral, while in the elliptic case it does not exist.

Let us consider for the sake of simplicity the equations of motion of the planar elliptic restricted problem of three bodies. These are (Szebehely, 1967)

$$\begin{aligned} \frac{d^2x}{dv^2} - 2\frac{dy}{dv} &= \frac{1}{1+e\cos v}\frac{\partial \Omega}{\partial x}, \\ \frac{d^2y}{dv^2} + 2\frac{dx}{dv} &= \frac{1}{1+e\cos v}\frac{\partial \Omega}{\partial y} \end{aligned} \qquad (1)$$

where x and y are the rectangular coordinates of the third body with negligible mass, e is the eccentricity of the relative orbit of the primaries, v is the true anomaly of one of the primaries and $\Omega(x,y)$ is the potential function of the problem whose explicit form we do not need now. When $e=0$, Equations (1) turn to the equations of the circular restricted three-body problem.

Multiplying the first of Equations (1) by dx/dv, the second by dy/dv, then adding the results and integrating it we obtain

$$\left(\frac{dx}{dv}\right)^2 + \left(\frac{dy}{dv}\right)^2 = \frac{2\Omega}{1+e\cos v} - 2e\int \frac{\Omega \sin v}{(1+e\cos v)^2}dv - C \qquad (2)$$

where C is a constant. For e=0 Equation (2) reduces to the Jacobian integral

$$\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2=2\Omega-C \qquad (3)$$

where C is the Jacobi constant.

The importance of the Jacobian integral is that it makes possible to establish regions in which motion can or can not occur. However, this can not be done in the elliptic case because of the unknown integral on the right hand side of Equation (2). As the matter of fact, Ω depends on x and y and thus this integral can be evaluated only when the solution for x and y is known.

As the Jacobian integral has an important role in stability problems and in many cases the elliptic restricted three-body problem gives better approximation than the circular one, the analysis of Equation (2), substituting the Jacobian integral in the elliptic case, has been the subject of several papers. For example Ovenden and Roy (1961) discussed its application in long-time predictions. Szebehely and Giacaglia (1964) considered the effects of the unknown integral in Equation (2) for zero velocity curves. Williams and Watts (1978) expanded the unknown integral about a Kepler an solution to the prob em. Invariant relations for the elliptic restricted three-body problem, substituting Equation (2), were derived in forms of series expansions by Vrcelj and de Jonge (1978) and by Delva and Dvorak (1979).

The subject of this paper is also the discussion of Equation (2). The purpose of this paper is to evaluate the unknown integral in Equation (2) in an actual case when the solution for x and y is known. This case is the case of Trojan asteroids for which an asymptotic solution was derived by this author (Érdi, 1978, 1981). The basic idea is to substitute this asymptotic solution into the expression of Ω and then to calculate the integral in Equation (2). However, as the solution for the motion of Trojan asteroids was obtained in cylindrical coordinates, it is suitable to replace Equation (2) by another relation written in cylindrical coordinates also and to carry out the above mentioned calculations in that new equation.

The motivation for this investigation comes from the fact that the stability of Trojan asteroids may be studied by using Equation (2). For example by numerically integrating the equations of motion one can calculate the value C^* of the expression

$$\left(\frac{dx}{dv}\right)^2+\left(\frac{dy}{dv}\right)^2-\frac{2\Omega}{1+e\cos v}=C^*. \qquad (4)$$

From a comparison with Equation (2) it is clear that C^* is not constant - due to the effect of the integral on the right hand side of Equation (2) - but its variation may give an indication about the stability of the motion of the asteroids. The parameter C^* may be regarded as the Jacobian 'constant' in the elliptic restric ed problem of three

bodies. The purpose of this paper is to derive analytic expressions for the variation of C^* (or more precisely for a C^* which will be defined later in a somewhat different way) thus supporting numerical investigations concerning the stability of Trojan asteroids.

2. Equations in Cylindrical Coordinates

In the author's theory of Trojan asteroids the following equations of motion were used (Érdi, 1978)

$$\begin{aligned}
\frac{d^2r}{dv^2} - r\left(\frac{d\alpha}{dv}\right)^2 - 2r\frac{d\alpha}{dv} &= \frac{1}{1+e_J\cos v}\left[r - \frac{1-\mu}{R_1^3}r + \right. \\
&\left. + \mu\left(\frac{\cos\alpha - r}{R_2^3} - \cos\alpha\right)\right], \\
\frac{d}{dv}\left(r^2\frac{d\alpha}{dv} + r^2\right) &= \frac{\mu r\sin\alpha}{1+e_J\cos v}\left[1 - \frac{1}{R_2^3}\right], \qquad (5) \\
\frac{d^2z}{dv^2} + z &= \frac{z}{1+e_J\cos v}\left[1 - \frac{1-\mu}{R_1^3} - \frac{\mu}{R_2^3}\right]
\end{aligned}$$

where r, α, z are the cylindrical coordinates of an asteroid (see Figure 1), v is the true anomaly of Jupiter, e_J is the eccentricity of Jupiter's orbit, μ is the mass of Jupiter divided by the total mass of the Sun-Jupiter system and

$$R_1 = \sqrt{r^2 + z^2}\,, \qquad R_2 = \sqrt{1 + r^2 - 2r\cos\alpha + z^2}\,.$$

The coordinates r and z are dimensionless, the unit distance is the Sun-Jupiter distance.

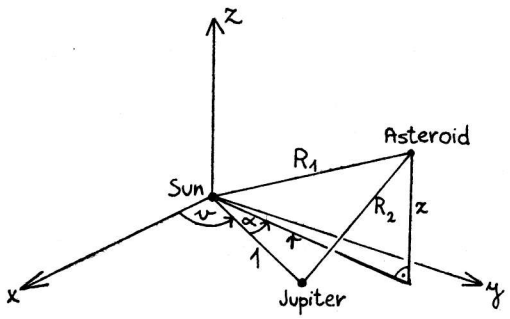


Figure 1

The coordinate system $Sxyz$ in Figure 1 is centered at the Sun, the plane Sxy is the orbital plane of Jupiter and the axis Sx is directed to the perihelion of Jupiter. Equations (5) describe the motion of an asteroid under the assumptions that its orbit around the Sun is perturbed only by Jupiter and Jupiter's orbit is an ellipse.

Multiplying the first of Equations (5) by dr/dv, the second by $d\alpha/dv$, the third by dz/dv, then adding the results and multiplying it by the expression $1+e_j\cos v>0$, we obtain after an integration according to v

$$\frac{1}{2}\left[\left(\frac{dr}{dv}\right)^2+r^2\left(\frac{d\alpha}{dv}\right)^2+\left(\frac{dz}{dv}\right)^2\right]=$$

$$=\frac{1}{2}r^2-\mu r\cos\alpha+\frac{1-\mu}{R_1}+\frac{\mu}{R_2}-C- \tag{6}$$

$$-e_j\int\frac{1}{2}\cos v\left\{\frac{d}{dv}\left[\left(\frac{dr}{dv}\right)^2+r^2\left(\frac{d\alpha}{dv}\right)^2+\left(\frac{dz}{dv}\right)^2+z^2\right]\right\}dv$$

where C is a constant of integration. Putting $e_J=0$ into Equation (6) it reduces to the Jacobian integral of the circular restricted problem of three bodies, written in cylindrical coordinates. Note, that e_J is explicitly present in Equation (6) only as a multiplier of the unknown integral thus this term containes the effect of the eccentricity on the motion in a very compact form. The purpose of this paper is to evaluate the unknown integral in Equation (6) in the case of Trojan asteroids using an asymptotic solution for their coordinates r, α, z.

3. Summary of a Theory of Trojan Asteroids

A solution of Equations (5) was derived for Trojan asteroids in the following three-variable asymptotic expansion form (Érdi, 1981)

$$r = 1 + \sum_{n=1}^{N} \varepsilon^n r_n(v, u, \tau) + O(\varepsilon^{N+1}), \tag{7a}$$

$$\alpha = \alpha_0(u, \tau) + \sum_{n=1}^{N} \varepsilon^n \alpha_n(v, u, \tau) + O(\varepsilon^{N+1}), \tag{7b}$$

$$z = \varepsilon^{1/2}\left[\sum_{n=0}^{N} \varepsilon^n z_n(v, u, \tau) + O(\varepsilon^{N+1})\right] \tag{7c}$$

where

$$\varepsilon = \sqrt{\mu}\ , \tag{8}$$

$$u = \varepsilon(v - v_0), \tag{9}$$

$$\tau = \varepsilon^2(v - v_0) \tag{10}$$

and v_0 is the initial value of v.

The reasons for assuming the solution in the form of Equations (7) are the following.

1. Trojan asteroids are resonant asteroids, they are in a 1:1 resonance with Jupiter as their orbital periods are nearly equals. In case of resonance the largest perturbations are proportional not to the perturbing mass but to the square root of it (Brown and Shook, 1933). Thus the solution (7) is expanded according to the powers of $\varepsilon = \sqrt{\mu}$. Note, that in the Sun-Jupiter system $\mu = 0.000954$ and $\varepsilon = 0.030885$.

2. In Equations (7) the functions λ_n, α_n, z_n depend on the variables v, u, τ representing three different time - scales of the motion of Trojan asteroids. The variable v corresponds to the orbital revolution of the asteroids around the Sun, while u describes the long-periodic librational motion around the Lagrangian points L_4 or L_5. The variable τ is connected with the motion of the perihelion of the asteroids (Érdi, 1978). The periods of the time-scales are approximately 12, 150 and 3600 years, respectively.

3. In Equation (7a) the first term on the right hand side is 1. This is so, because Trojan asteroids are nearly at the same distance from the Sun as the Lagrangian points L_4 and L_5 are and these latters have a dimensionless distance 1 from the Sun.

4. In Equation (7b) the first term in the expansion

of α is $\alpha_0(u,\tau)$. This term is not multiplied by ε so it can take large values. It can be shown that the function $\alpha_0(u)$ describes the main part of the long-periodic, large amplitude librational motion around L_4 or L_5 (Érdi, 1978). On the other hand, from the assumption that α_0 depends on τ too, the dependence of the librational period on the coordinate z can be derived (Érdi, 1981).

5. According to Equation (7c) the coordinate z is proportional to $\varepsilon^{1/2}$. This is so, because for the majority of the known Trojan asteroids $\varepsilon < z < 3\varepsilon^{1/2}$. The assumption $z \approx \varepsilon$ could be applied only in a few cases.

An additional assumption for the solution of Equations (5) was that

$$e_J = \varepsilon e_1 \qquad (11)$$

where the constant e_1 is not very large compared to unity.

Assuming the solution in the form of Equations (7) and substituting them into Equations (5) a system of partial differential equations can be derived for the unknown functions τ_n, α_n, z_n. The solution of these new equations has been determined to $O(\varepsilon^2)$ (Érdi, 1981). At this point it should be mentioned that the three time-scales method, applied in the above investigations, is a generalization of Kevorkian's two-variable method applied by him for the planar motion of Trojan asteroids (Kevorkian, 1970).

In possession of the solution for τ, α and z, the main perturbations of the orbital elements of Trojan asteroids were also derived (Érdi, 1981). The lengthy analytical ex-

pressions will not be repeated here. However, these will be used for the evaluation of the unknown integral in Equation (6).

4. Calculation of the Unknown Integral

Introducing the function F as

$$F = \frac{1}{2}\cos v \frac{d}{dv}\left[\left(\frac{dr}{dv}\right)^2 + r^2\left(\frac{d\alpha}{dv}\right)^2 + \left(\frac{dz}{dv}\right)^2 + z^2\right], \qquad (12)$$

the integral

$$\int F dv = G \qquad (13)$$

is to be determined.

Substituting the solution (7) into Equation (12), F will be a known function of the variables v, u and τ, and thus G will also depend on these variables. However, the calculation of the integral needs care as F depends on v explicitly and according to Equations (9) and (10) through u and τ implicitly. Because of the complicate analytic form of the function F it is suitable to determine G instead of Equation (13) from the equation

$$F = \frac{dG}{dv} . \qquad (14)$$

Expanding F and G according to the powers of ε

$$F = F_0 + \varepsilon F_1 + \varepsilon^2 F_2 + \varepsilon^3 F_3 + \ldots \tag{15}$$

$$G = G_0 + \varepsilon G_1 + \varepsilon^2 G_2 + \varepsilon^3 G_3 + \ldots \tag{16}$$

and substituting these series into Equation (14), a system of partial differential equations can be obtained for the determination of the unknown functions G_n by equating the coefficients of the same powers of ε on both sides of Equation (14). To n=3 these equations are

$$F_0 = \frac{\partial G_0}{\partial v}, \tag{17}$$

$$F_1 = \frac{\partial G_1}{\partial v} + \frac{\partial G_0}{\partial u}, \tag{18}$$

$$F_2 = \frac{\partial G_2}{\partial v} + \frac{\partial G_1}{\partial u} + \frac{\partial G_0}{\partial \tau}, \tag{19}$$

$$F_3 = \frac{\partial G_3}{\partial v} + \frac{\partial G_2}{\partial u} + \frac{\partial G_1}{\partial \tau}. \tag{20}$$

Equations (17) through (20) are enough to determine G_0 and G_1 completely. However, G_2 and G_3 can be calculated only in part.

Substituting the known part of the solution given by Equations (7) (Érdi, 1981) into Equation (12), one obtaines for the functions F_n

$$F_0 = 0, \tag{21}$$

$$F_1 = 0, \tag{22}$$

$$F_2 = F_2^{***}(v, u, \tau) + F_2^{**}(u, \tau) , \tag{23}$$

$$F_3 = F_3^{***}(v, u, \tau) + F_3^{**}(u, \tau) + F_3^{*}(\tau) \tag{24}$$

where as it can be seen F_2 and F_3 consist of different parts according to their dependence on v, u, τ.

From Equations (17) and (21) it follows that G_0 does not depend on v, that is

$$G_0 = G_0^{**}(u, \tau) \tag{25}$$

and G_0^{**} is yet to be determined.

Equations (18), (22) and (25) give

$$\frac{\partial G_1}{\partial v} = -\frac{\partial G_0^{**}}{\partial u} . \tag{26}$$

As in this equation the right hand side does not depend on v, the integration of Equation (26) according to v would give a secular term in G_1. To avoid this let us suppose that

$$\frac{\partial G_0^{**}}{\partial u} = 0 .$$

Then

$$G_1 = G_1(u, \tau) \tag{27}$$

and

$$G_0^{**} = G_0^{*}(\tau) . \tag{28}$$

Equation (19) together with Equations (23), (25), (27) and (28) give

$$F_2^{***} + F_2^{**} = \frac{\partial G_2}{\partial v} + \frac{\partial G_1}{\partial u} + \frac{\partial G_0^*}{\partial \tau} . \tag{29}$$

It is suitable to separate Equation (29) as

$$\frac{\partial G_2}{\partial v} = F_2^{***},$$

$$\frac{\partial G_1}{\partial u} = F_2^{**},$$

$$\frac{\partial G_0^*}{\partial \tau} = 0 .$$

Now it follows that $G_0 = G_0^* =$constant and thus it can be added to the constant C in Equation (6). Furthermore

$$G_2 = G_2^{***}(v,u,\tau) + G_2^{**}(u,\tau) \tag{30}$$

and

$$G_1 = G_1^{**}(u,\tau) + G_1^{*}(\tau) \tag{31}$$

where

$$G_2^{***} = \int F_2^{***} dv \tag{32}$$

and

$$G_1^{**} = \int F_2^{**} du . \tag{33}$$

It can be calculated that F_2^{***} is a periodic function of v and F_2^{**} is a periodic function of u and thus the integrals (32) and (33) can be easily obtained. Thus G_2^{***} and G_1^{**} can be regarded as known functions. However, G_2^{**} and G_1^{*} are yet to be determined.

Finally, Equation (20) is considered. By virtue of Equations (24), (30) and (31) it can be written as

$$F_3^{***} + F_3^{**} + F_3^{*} = \frac{\partial G_3}{\partial v} + \frac{\partial G_2^{***}}{\partial u} + \frac{\partial G_2^{**}}{\partial u} + \frac{\partial G_1^{**}}{\partial \tau} + \frac{\partial G_1^{*}}{\partial \tau} . \tag{34}$$

Equation (34) may be separated into the equations

$$\frac{\partial G_3}{\partial v} = F_3^{***} - \frac{\partial G_2^{***}}{\partial u} , \tag{35}$$

$$\frac{\partial G_2^{**}}{\partial u} = F_3^{**} - \frac{\partial G_1^{**}}{\partial \tau} , \tag{36}$$

$$\frac{\partial G_1^{*}}{\partial \tau} = F_3^{*} . \tag{37}$$

It can be determined from Equation (35) how G_3 depends on v and from Equation (36) the dependence of G_2^{**} on u. Equation (37) gives

$$G_1^{*} = \int F_3^{*} d\tau + \text{constant} \tag{38}$$

and this constant can also be added to C in Equation (6).

By continuing the above method the unknown functions G_n can be determined. Summing up the previous results the main part of G is

$$G = \varepsilon \left[\int F_2^{**} d u + \int F_3^{*} d \tau \right] + O(\varepsilon^2). \qquad (39)$$

Equation (39) shows that the main part of G containes only long-periodic terms depending on u and τ. Short-periodic terms depending on v will occur in higher orders. These however will not be calculated in this paper.

5. Discussion of the First-Order Terms

Calculating the functions F_2^{**} and F_3^{*} and completing the integrals in Equation (39) one obtaines

$$\begin{aligned} G = & -(D_1 + \cos\alpha_0) c \cos(A_0 \tau + \psi_{11}) - \\ & -(D_2 + \sin\alpha_0) c \sin(A_0 \tau + \psi_{11}) - \\ & -e_J \left(\frac{A_1}{A_0} \cos\alpha_0 + \frac{A_2}{A_0} \sin\alpha_0 \right) + O(\varepsilon^2) \end{aligned} \qquad (40)$$

where D_1, D_2, A_0, A_1, A_2, c and ψ_{11} are constants. Moreover

$$D_1 = \frac{1}{2} - \frac{83}{2^4 3} \ell^2 - \frac{1193}{2^8 3^3} \ell^4 + O(\ell^6), \qquad (41a)$$

$$D_2 = \frac{\sqrt{3}}{2} + \frac{91\sqrt{3}}{2^4 3^2} \ell^2 - \frac{11093\sqrt{3}}{2^8 3^4} \ell^4 + O(\ell^6), \qquad (41b)$$

$$A_0 = \frac{27}{2^3} + \frac{129}{2^6}\ell^2 - \frac{87}{2^7}\ell^4 + O(\ell^6), \tag{42a}$$

$$A_1 = -\frac{27}{2^4} - \frac{141}{2^6}\ell^2 - \frac{1095}{2^{11}}\ell^4 + O(\ell^6), \tag{42b}$$

$$A_2 = -\frac{27\sqrt{3}}{2^4} + \frac{45\sqrt{3}}{2^6}\ell^2 + \frac{711\sqrt{3}}{2^{11}}\ell^4 + O(\ell^6) \tag{42c}$$

where ℓ is a constant meaning nearly the amplitude of the librational motion around the Lagrangian point L_4. For the majority of the known Trojan asteroids ℓ is in the range 0.0-0.5.

The parameter α_0 describes the main part of the librational motion of the asteroids around L_4

$$\begin{aligned}\alpha_0 = {} & \frac{\pi}{3} + \frac{3\sqrt{3}}{2^3}\ell^2 + \frac{13\sqrt{3}}{2^8}\ell^4 + \\ & + \ell\cos\phi - \left(\frac{\sqrt{3}}{2^3}\ell^2 + \frac{\sqrt{3}}{2^8 3^2}\ell^4\right)\cos 2\phi + \\ & + \frac{5}{2^6}\ell^3\cos 3\phi - \frac{25\sqrt{3}}{2^7 3^2}\ell^4\cos 4\phi + O(\ell^5)\end{aligned} \tag{43}$$

where

$$\phi = \sqrt{\frac{27}{2^2}\left(1 - \frac{3}{2^3}\ell^2 - \frac{97}{2^9}\ell^4\right)}\, u + \delta \tag{44}$$

and δ is the function of τ but this dependence can be neglected now.

It follows from Equations (6), (12), (13) and (40) that

$$\frac{1}{2}\left[\left(\frac{dr}{dv}\right)^2 + r^2\left(\frac{d\alpha}{dv}\right)^2 + \left(\frac{dz}{dv}\right)^2\right] = \\ = \frac{1}{2}r^2 - \mu r\cos\alpha + \frac{1-\mu}{R_1} + \frac{\mu}{R_2} - C^* \qquad (45)$$

where

$$C^* = -(D_1 + \cos\alpha_0)e_J c\cos(A_0\tau + \psi_{11}) - \\ -(D_2 + \sin\alpha_0)e_J c\sin(A_0\tau + \psi_{11}) - \qquad (46) \\ -e_J^2\left(\frac{A_1}{A_0}\cos\alpha_0 + \frac{A_2}{A_0}\sin\alpha_0\right) + O(\varepsilon^2 e_J) + C$$

and C is a constant.

Equation (45) may be regarded as an invariant relation of Trojan asteroids in the elliptic restricted problem of three bodies. The parameter C^* corresponds to the Jacobi constant although C^* is not really a constant and its main variations are given by Equation (46).

Using Equations (41), (42) and (43) an estimation for the amplitudes of the terms in Equation (46) may be derived

$$\left|(D_1 + \cos\alpha_0)e_J c\cos(A_0\tau + \psi_{11}) + (D_2 + \sin\alpha_0)e_J c\sin(A_0\tau + \psi_{11})\right| \leq K_1 \qquad (47)$$

where

$$K_1 = e_J c\left(1 + \frac{\sqrt{3}}{2}\ell - \frac{1}{2^3 3}\ell^2 - \frac{875\sqrt{3}}{2^7 3^2}\ell^3 + \frac{685}{2^9}\ell^4 + O(\ell^5)\right) \qquad (48)$$

and

$$\left|e_J^2\left(\frac{A_1}{A_0}\cos\alpha_0 + \frac{A_2}{A_0}\sin\alpha_0\right)\right| \leq K_2 \qquad (49)$$

where

$$K_2 = e_J^2\left(1 - \frac{7}{2^2 3}\ell^2 + \frac{5203}{2^8 3^3}\ell^4 + O(\ell^6)\right). \qquad (50)$$

Table I shows the values of K_1 for 30 known Trojan asteroids obtained from Equation (48) with $e_J = 0.048$ and using the values of c and ℓ given in one of the author's papers (Érdi, 1979).

Table I. Values of K_1 for 30 Trojan asteroids

Around L_4		PL objects around L_4	
Nestor	0.00722	9507	0.00918
Achilles	0.00536	6020	0.00766
Ajax	0.00488	2008	0.00613
Telamon	0.00369	2706	0.00673
Menelaus	0.00295	4523	0.00415
Hector	0.00320	6844	0.00398
Odysseus	0.00272	4596	0.00343
Agamemnon	0.00134	4139	0.00289
Antilochus	0.00117	6541	0.00272
Diomedes	0.00102	6629	0.00250
Around L_5		6540	0.00161
Patroclus	0.00537	4655	0.00154
Anchises	0.00571	6581	0.00115
Priam	0.00493	4572	0.00095
Aeneas	0.00331	6591	0.00050
Troilus	0.00297		

The values of K_2 are in the range 0.00230-0.00208 for $\ell = 0$-0.5.

Note, that the main variations of C^* have two different

periods. One is the period T_L of the libration around L_4 described by α_0 and it can be obtained from Equation (44) as

$$T_L = \frac{T_J}{\varepsilon\sqrt{\frac{27}{2^2}\left(1-\frac{3}{2^3}\ell^2-\frac{97}{2^9}\ell^4\right)}} \qquad (51)$$

where T_J is Jupiter's orbital period. The other is connected with the time-scale of τ and is given by

$$T_{e,\tilde{\omega}} = \frac{T_J}{\varepsilon^2 A_0} . \qquad (52)$$

For ℓ=0-0.5, T_L=147.8-156.3 years and $T_{e,\tilde{\omega}}$=3684-3241 years. Thus in stability investigations of Trojan asteroids numerical integrations should extend at least for about a 3600 years time-interval.

References

Brown, E. W. and Shook, C. A.: 1933, Planetary Theory, Cambridge University Press.

Delva, M. and Dvorak, R.: 1979, Astronomy and Astrophysics, 77, 252.

Érdi, B.: 1978, Celestial Mechanics, 18, 141.

Érdi, B.: 1979, Celestial Mechanics, 20, 59.

Érdi, B.: 1981, Celestial Mechanics, 24, 377.

Ovenden, M. W. and Roy, A. E.: 1961, Monthly Notices of the Royal Astronomical Society, 123, 1.

Szebehely, V. and Giacaglia, E. O.: 1964, Astronomical Journal, 69, 230.

Szebehely, V.: 1967, Theory of Orbits, Academic Press, New York and London.

Vrcelj, Z. and Kiewiet de Jonge, J. H.: 1978, Astronomical Journal, 83, 514.

Williams, C. A. and Watts, J. G.: 1978, in 'Dynamics of Planets and Satellites and Theories of Their Motion' , ed. Szebehely, V., D. Reidel Publishing Company, Dordrecht, Holland, p. 333.

PROBLEMS CONCERNING THE OUTER PLANETS

P. K. Seidelmann
U. S. Naval Observatory
Washington, DC 20390

R. L. Duncombe
University of Texas
Austin, TX 78712

THE DISCOVERY OF URANUS, NEPTUNE AND PLUTO

The Sun, Moon and the planets Mercury through Saturn have been observable objects since ancient times. On 13 March 1781, William Herschel was doing a systematic sky survey for double stars when he discovered an object which was about sixth magnitude and had a planetary appearance (Bennet 1981). Initially, there were questions as to whether or not it was a comet, but subsequent observations indicated that it had to be a planet, Uranus.

On 1 January 1801, the minor planet Ceres was first discovered. This was the beginning of the observational discovery program of minor planets which continues to this day. Currently, over two thousand minor planets have been discovered and their orbits determined.

At the time of the discovery of Neptune in 1846, astronomy was different from the astronomy of today in four significant ways;

1. The ephemerides of the planets were of limited accuracy.

2. There were no Palomar sky survey plates and systematic star charts were of a limited accuracy.

3. Photography was not used.

4. The accuracy of observations was not as good as today.

In the 1840's, the best tables for the motion of Uranus were those of Bouvard (1821). Bessel (1821) pointed out an error in these tables, but it was not large enough to explain the discrepancy between the observations and tables as indicated by Airy (1832). In 1840, Bessel recognized the cause of the problem which he attributed to the presence of an unknown planet. He was ready himself to start work on the problem, an effort prevented by his ill-health and eventual death.

In France, Le Verrier began an investigation of the observation residuals of Uranus and published his first paper on the subject on 10 November 1845. A fortnight earlier, Adams, having accomplished similar work in England, communicated his predictions to Airy. Airy inquired about the radius vector residual and unfortunately receiving no reply, never started the search. On 29 July 1846, Professor Challis of Cambridge University, began a search for this unknown planet based on the predictions of Adams. Since he did not expect the prediction to be accurate, he planned an extended search (Challis, 1846).

Le Verrier sent a request to Dr. Galle at Berlin requesting that he initiate a search for this unknown planet. Le Verrier's letter was received on 23 September 1846 and that night, Dr. Galle discovered the planet Neptune 55 minutes of arc from the geocentric position predicted by Le Verrier. The experience should be a lesson to students and assistants that they can offer useful suggestions and advice even to their professors. Professor Galle was looking for a three second of arc disk, which he could not find. It was the graduate assistant who suggested using the recently published Berlin Academy Star Atlas #21 which had been printed in 1845. Using that atlas an eighth magnitude star was detected that was not present in the atlas. That unlisted star was Neptune.

After the discovery, Challis realized that he had observed the planet Neptune earlier, but had not pursued the comparison of observed positions on different dates far enough to detect his discovery. Adams and Le Verrier had independently predicted the position of the planet Neptune, but the discovery was made in Germany, based on the availability of accurate star charts and confidence in the prediction of the unknown planet. Details concerning the discovery of Neptune are given by Airy (1846), Gould (1850), and Grosser (1962).

Figure 1 shows the difference between the observations and the theory of Uranus that was available at that time. If the theory and the observations agreed, all of the points would be in a horizontal line. The longitudes are indicated on the top of the figure while the dates are given on the bottom. The figure is from Pickering (1909). The lines drawn are tests of Pickering's graphical method which was to be used for the prediction of Pluto. It should be noted that the residuals in this case range from a plus 40 seconds of arc to a minus 100 seconds of arc. Table 1 lists the orbital elements from the predictions of Adams and of Le Verrier and the elements determined by Walker (1847) after the discovery of Neptune. The predicted and discovery elements are in close agreement and the longitude is particularly close to discovery position of Neptune.

Percival Lowell (1915) predicted a planet beyond Neptune based on his calculations. Again, the theory being used becomes a factor. Figure 2 shows the differences in the longitude of Uranus between the 1873 theory of Le Verrier (1877) and the observations of Uranus. The difference in the magnitude of the residuals

should be noted; the maximum scale in this case is only 5 seconds of arc.

Figure 3 shows the residuals in longitude of Uranus based on Gaillots' theory (1910). In this case, the scale maximum is 2.5 seconds of arc. While Lowell made his predictions of the presence of an additional planet based on numerical calculations, Pickering used a graphical method. Figure 4 shows a comparison of the Uranus observations with (a) the Le Verrier theory of Uranus, (b) Pickering's orbit of Uranus with a predicted planet "O", and (c) the theory of Gaillot. (Pickering, 1928(a))

Figure 5 shows the comparison of Uranus's observations with Newcomb's theory of Uranus (1898) as given by Pickering in 1928(b). There is a 1.5 second of arc correction in the origin at 1904 to put Uranus close to the theory of Newcomb. After the discovery of Pluto, Pickering (1931) analyzed a plot (Figure 6) of residuals of Uranus and included a prediction of another planet "P". The observations used by Pickering are from Greenwich and Paris compared to the Le Verrier theory and Washington's observations compared to Newcomb's orbit, with a least squares correction being performed by Morgan and Lyons (1930).

In 1894, Lowell Observatory was established on Mars Hill in Flagstaff, Arizona. The observatory was expected to observe Mars and also to search for a planet beyond Neptune. After Lowell's death and a period of inactivity in the search for another planet, the staff of Lowell Observatory initiated a new search program. They hired a young amateur astronomer Clyde Tombaugh, to carry out the observational program and eventually also to conduct the blink comparison of the plates. Tombaugh's discovery of the planet Pluto was announced on 13 March 1930, exactly on the 75th anniversary of the birth of Percival Lowell and 149 years after Herschel had discovered Uranus.

Figure 7, a and b, (Hoyt, 1980) shows the plates on which Tombaugh discovered Pluto. The plates were taken on 23 and 29 January 1930, and revealed a 15th magnitude planet, which was discovered by blinking the two plates. The area covered by Tombaugh's search is indicated in Figure 8 (Tombaugh 1961). This search was continued for almost 20 years and covered the areas darkly shaded to 16th and 17th magnitude and the lightly shaded areas to the 14th and 15th magnitude. The discovery of Pluto has been reviewed by Hoyt (1980), Tombaugh and Moore (1980) and Whyte (1980). Subsequent research on Pluto is reviewed in a number of papers which were presented in commemoration of the 50th anniversary of its discovery (Icarus, Oct. 1980).

POST PLUTO DISCOVERY:

After the discovery of Pluto, there followed a period of orbit determination and a wide variety of results were obtained (Seidelmann et al. 1980). Eventually, Eckert, Brouwer and Clemence (1951) offered one of the first applications of modern methods of

computation, using a sequential electronic computer, to simultaneously calculate the orbit of Pluto and the other four outer planets.

Figure 9, a and b, shows a comparison between the observations of Pluto and the computation of Eckert, Brouwer, and Clemence, which has been the basis for the ephemeris in the Astonomical Almanac since 1960.

Photometric studies of the planet Pluto have indicated a rotational period of 6.3867 days as given in Figure 10 by Neff, Lane and Fix (1974). In 1950, using micrometer measurements, Kuiper (1950) determined a diameter of 6000 kilometers for Pluto with the possibility that it might be as small as 4000 kilometers. From the near miss of an occultation of a star by Pluto, Halliday et al. (1966) gave an upper limit for the diameter of 6,800 kilometers. With the then accepted mass value of 1/360,000 and Kuiper's diameter for Pluto, a density of 50 grams per cubic centimeter is implied. For comparison, the Earth has a radius of 6378.14 kilometers, a reciprocal mass of 328,900.5 and a resulting density of 5 grams per cubic centimeter. The other planets have densities which range from 0.7 grams per cubic centimeter for Saturn up to the Earth's density. Thus, compared with these values, the density of Pluto was unusually large. Later, Cruickshank, Pilcher, and Morrison (1976) observed evidence of methane frost on Pluto and concluded, based on the implied albedo, that Pluto should have a diameter of only 2,800 to 3,300 kilometers.

In 1978, Jim Christy, of the U. S. Naval Observatory, was measuring Pluto plates which were taken to improve the orbit. He detected an elongation of the image of Pluto, and not of the background stars. Also, the elongation appeared to occur at a regular interval. Figure 11 shows one of these images. It was a new satellite, Charon (Christy and Harrington, 1978). The period of revolution of this satellite appeared to match the light curve period of 6.3867 days, but there is no reasonable way the presence of the satellite can explain the amplitude of the light curve. It must be assumed that the satellite is in synchronous rotation with the planet. Recent investigations of the declination residuals of Pluto by Reis, Duncombe, Van Flandern and Pulkkinen (1981) give a slightly different orbital period for the satellite. Based on this information, Harrington has reexamined the observations of Pluto and its satellite, and determined an orbital period from the position angle and separation measurements of 6.3871 days, $\pm$.0002. The best determination for the value of the separation of the satellite is 190,300 $\pm$ 3000 kilometers or 0.88 $\pm$ .01 arc seconds (Harrington and Christy, 1980). The satellite permits the determination of the mass of Pluto, but first, let us examine the history of the determination of Pluto's mass.

PLUTO'S MASS

Following the investigations of Wylie (1942) and Eckert, Brouwer, and Clemence (1951), a mass of 1/360,000 solar masses was

adopted for Pluto, which is approximately the mass of the Earth. In 1968, faced with the implied density problem for Pluto and the fact that the ephemeris of Neptune was systematically deviating from the observations of Neptune, Duncombe, et al. (1968) used a projection technique as a test for the mass of Pluto. Figure 12 shows the difference between the ephemeris and observations of Neptune at that time. The solution implied a reciprocal mass for Pluto of 1,812,000, which is about 1/6 the Earth's mass. The determination was based on the use of normal points formed from the observations of Neptune. A rediscussion of all the individual observations of Neptune, to eliminate possible systematic star catalog effects, indicated a reciprocal mass of 3 million, approximately 1/10 the mass of the Earth.

The discovery of the satellite of Pluto with its observed period and separation indicates a reciprocal mass of 130 million or 2/1000 the mass of the Earth. The history of the mass determinations for Pluto (Table 2) was presented by Duncombe and Seidelmann (1980) at the Fiftieth anniversary of the discovery of Pluto. These data were put into graphical form by Dessler and Russell (1980) and a curve was fit to the data (Figure 13). Based on this information Pluto will disappear in 1984. Unfortunately, some people took this humorous prediction seriously.

THE PREDICTION OF PLUTO IN HINDSIGHT

The predictions of the location of Pluto by Pickering and Lowell were close in longitude. Table 3 shows a comparison between the predictions and the actual orbit and mass of Pluto. Based on our current knowledge of the mass of Pluto, this body could not have been the cause of the observed residuals in Uranus or Neptune. Thus, the discovery has to be due to an assiduous search rather than to gravitational prediction.

We are left with some questions. Was it serendipity that led to the discovery of the planet close to the predicted place? Have we discovered one of many bodies in the outer solar system? Have we discovered the small body and left a larger body undiscovered? Can the observation residuals be explained in terms of inadequate theories?

Clearly, the predictions and discovery of Neptune were very different than the predictions and discovery of Pluto. The discovery of Neptune should be credited to good celestial mechanics, while the discovery of Pluto should be credited to a thorough observational search.

PRESENT STATUS

What is the current situation with regard to the observations of the outer planets? Pluto has a short history, and the observations are photographs of a 15th magnitude object and stars of a similar magnitude, which are subject to a r.m.s. uncertainty of

about 3 arc seconds. With the 250 year orbital period of Pluto and an observational history of less than 70 years, it is possible to fit the observations but the ephemeris should be expected to have a limited life-time and no conclusions can be drawn from the observation residuals. Figure 14, a and b, shows the comparison of the observations with the most recent ephemeris of Pluto.

For Neptune, the observational history covers 134 years of the approximately 160 year period of the planet. Figure 15, a and b, shows the comparison of observations with the ephemeris in right ascension and declination. Seidelmann, et al. (1971) determined an ephemeris of Neptune which fit the observational data. Ten years later that ephemeris shows a systematic deviation from the observational data. This experience continues a pattern established in the studies of Newcomb (1882), Wylie (1942), and Eckert, Brouwer and Clemence. For some reason, the ephemerides calculated for Neptune are valid for prediction for only a short time period before systematic deviations from the observational data appear. In addition there are two 1795 observations of a star (later identified as Neptune) given in Lalande's diary. These observations, which cannot be reduced rigorously to the adopted catalog reference system, differ from the Neptune ephemeris by approximately 12 arc seconds in right ascension and a little less than 1 arc second in declination. Kowal and Drake (1980) have documented observations by Galileo in 1613 of Neptune with respect to the Jupiter satellite system. These observations are not subject to star catalog errors and are an independent type of observation. They differ from the ephemeris by approximately 30 arc seconds in right ascension and 40 arc seconds in declination. There is some uncertainty as to whether Galileo's observations accurately indicate the position of Neptune or only the direction of Neptune with respect to the satellite system. The uncertainty is based on Galileo's indication of the scale distance for that observation. At this time it is clear that the prediscovery and post discovery observations of Neptune cannot be satisfied by a single ephemeris.

Uranus has an orbital period of approximately 80 years. There are observations dating back to before 1700, so a complete orbital period of observations exist before the 1781 discovery and an orbital period of modern observations exists since 1900. Figure 16, a and b, shows a comparison of the observations since discovery with an ephemeris adjusted to these observations. Every effort has been made to eliminate systematic effects due to instruments and star catalogs. It is evident from Figure 16a that there is a slight run-off between the ephemeris and observations
there is a slight run-off between the ephemeris and observations for the most recent time.

There are observations of the Viking spacecraft in orbit around Mars for a period of six years. While the time period is short, the observations are very accurate and they indicate a different value for the mass of Uranus than is currently adopted by the IAU. It is not possible to satisfy the Mars observations data and all of the observations of Uranus in the same solution. If only the observations of Uranus since 1900 are included, the

Viking observations can also be satisfied.

Figure 17 shows the comparison between an ephemeris fit to the optical observat ons ince 1900 in a solution that included the Viking observational data for Mars. It is evident that a systematic deviation remains for the data before 1900. Thus, the conclusion must be drawn that there are still systematic problems with the observations of Uranus. While the discovery of Neptune and Pluto have significantly reduced the observational residuals of Uranus, there still remains a problem representing the motion of the planet.

Figure 18, a and b, shows the comparison of right ascension and declination observations of Saturn with the recent ephemeris of Saturn. Figure 19, a and b, shows the comparison of right ascension and declination observations of Jupiter with the recent ephemeris of Jupiter.

IMPLICATIONS OF THE CURRENT DATA

It is possible that there are systematic differences between the star catalogs prior to, and after, 1900. These systematic effec s could be affecting the observation data. It is known from observations of the minor planets that systematic effects are present between the observations prior to, and after, 1900. It is also true for the minor planets that visual observations compromise the majority of the observations before 1900, while photographic observations make up most of the data after 1900.

Is it possible that the effect of another planet, undiscovered and existing in the outer part of the solar sytem, is being observed? Such a planet could be sufficiently south of the ecliptic plane that it would have eluded the systematic search of Tomb ugh Lest one adopts this solution too precipitously, however, it should be pointed out that prior to 1900 there were residuals in the observations of Mercury and these data led to the predi tion and naming of a planet, Vulcan, interior to the orbit of Me cury. There were even reports, during solar eclipses, of obser ati ns of Vulcan. The actual answer to the problem of the residuals of Mercury came with the theory of relativity by Einstein. Thus, the possibility that there may be some ulterior correction to our g avitational model canno' be discounted.

There are severa possible approaches to finding another planet. Based on the observation residuals for each of the outer planets, a direction to a perturbing body that would produce these residuals can be determined Then one can s-ek consistency amongst the directions from different planet as a criterion for the validity of such tests. Also, preselected est objects can be included in numerical integrations in the ho e of reducing the residuals in the observations of these planets. Another approach is to note the peculiar nature of the satellite system of Neptune, with one satellite in a very elliptical orbit, one satellite in a retrograde

orbit and, in addition, the planet Pluto in an orbit which intersects the orbit of Neptune. The hypothesis exists that maybe some object came close to this Neptunian system and caused the presently observed state of the satellites of Neptune, while ejecting Pluto from the satellite system.

It is also possible to do a systematic search for an undiscovered planet using current plate files acquired for proper motion studies, to make special observations of limited areas based on predictions from the above approaches, or to make a whole new sky survey. The chances of finding another planet obviously depend on whether such a planet really exists, how bright it is, how fast it might be moving at the time of search, how accurate the predictions are from the celestial mechanics calculations, the quality of the astrometric search observations and the thoroughness of the examination of the search plates.

Finally, the question will arise as to whether this new discovery deserves the notation of being another planet. With the reduced mass of Pluto, the separation between minor and principal planet is not clearly defined. How massive or large does an object have to be to be classified as a principal planet?

CONCLUSION

We have improved the ephemerides of the planets based on the new set of astronomical constants for inclusion in national ephemerides beginning with 1984. However, we still have some remaining discrepancies. Certainly, the ephemerides for Neptune and Pluto will need improvement in about a decade. The ephemeris of Uranus has, of necessity, been based only on observations since 1900, because a consistent ephemeris could not be prepared based on all the available observational data. With the new ephemeris, the observation residuals of Uranus are admittedly smaller. However, two planets have already been discovered based on the motion of Uranus. Are there more?

REFERENCES

Airy, G. B. 1832 Reports of British Association I, 154.

Airy, G. B. 1846 Royal Astr. Soc. Month Not. 7, 9, p 121.

Andersson, L. E. and Fix, J. D. 1973 "Pluto: New Photometry and a Determination of the Axis of Rotation", Icarus 20, 279-283.

Bennett, J. A. 1981 "The Discovery of Uranus", Sky and Telescope, March 1981.

Bessel, F. W. 1840 Populare Vorlesungen p 448.

Bessel, F. W. 1824 Astron. Nachrichten 2 Nr., 48, p 441.

Bouvard, A. 1821 Conn. des Temps p. 341, 342.

Challis, J. 1846 Special Report of Proceedings in the Observatory Relative to the New Planet, Cambridge Observatory.

Christy, J. W. and Harrington, R. S. 1978, "The Satellite of Pluto" Astron J. 83, 1005-1008.

Cruickshank, D. P., Pilcher, C. B. and Morrison, D. 1976. "Pluto: Evidence for Methane Frost", Science 194. 835.

Dessler, S. J. and Russell, C. T. 1980, "From the Ridiculous to the Sublime: The Rending Disappearance of Pluto", EOS 61, p 44, Oct 28, 1980.

Duncombe, R. L. and Seidelmann, P. K. 1980 "A History of the Determination of Pluto's Mass", Icarus 44, 12-18.

Duncombe, R. L., Klepczynski, W. J. and Seidelmann, P. K., 1968 "Orbit of Neptune and Mars of Pluto", Astron. J. 73, 830.

Eckert, W. J., Brouwer, D., and Clemence, G. M. 1951. Astron. Papers of Amer. Ephem. XII.

Gaillot, M. 1910 Annales of the Paris Observatory, 28.

Gould, B. A. 1850 "Report on the History of the Discovery of Neptune," Smithsonian Institution, Washington, D. C.

Grosser, M. 1962. The Discovery of Neptune. Harvard University Press.

Halliday, I., Hardie, R. H., Franz, O. G., and Priser, J. B., 1966. "An Upper Limit for the Diameter of Pluto," Publ. Astr. Soc. Pacific, 78, 113.

Harrington, R. and Christy, J. W. 1980 "The Satellite of Pluto II" Astron. J. 85, 168-170.

Hoyt, W. G. 1980, "Planets X and Pluto," The University of Arizona Press, Tucson, Arizona.

Kowal, C. T., and Drake S. 1980. "Galileo's Observations of Neptune" Nature 287, p 311-3.

Kuiper, G P. 1950 "The Diameter of Pluto", Publ. Astr. Soc. Pacific 62, 133.

Le Verrier, U. -J. 1845 Comptes Rendus XXI, p 1050.

Le Verrier, U. -J. 1849 "Recherches sur le mouvements De La Planete Herschel",. Connaissance des Temps. Paris.

Le Verrier, U. -J. 1877 Annales de L'Observatoire de Paris, XIV.

Lowell, P. 1915. "Memoir on a Trans-Neptunian Planet," Memoires of the Lowell Observatory, 1, 1.

Morgan, H. R. and Lyons, U. S. 1930, "On the Tables of Uranus and Neptune", Astron. J. 40, 97.

Neff, J. S., Lane, W. A. and Fix, J D., 1974. "An Investigation of the Rotational Period of the Planet Pluto," Pub. Astron. Soc. Pacific 86, p 225.

Newcomb, S. 1898 Tables of Neptune, Tables of Uranus, Astronomical Papers of the American Ephemeris, VII.

Pickering. W. H. 1928. (a) "The Next Planet Beyond Neptune", Popular Astronomy XXXVI, 143.

Pickering, W H. 1928. (b) "The Orbit of Uranus," Popular Astronomy XXXVI, 353.

Pickering, W. H. 1931. "Planet P, Its Orbit, Position and Magnitude. Planets S and T ." Popular Astronomy XXXIV p. 385.

Pickering, W. H 1909. "A Search for a Planet Beyond Neptune," Annals of the Astronomical Observatory of Harvard College LXI, pt II, Cambridge, MA.

Reis, J. Duncombe, R. L., Van Flandern, T. C. Pulkkinen, K. F. (1981) "The Period of Charon derived from the Observed Barycentric Motion of Pluto" BAAS 13, 573

Seidelmann P. K., Kaplan, G. H , Pulkkinen, K. F., Santoro, E. J. and Van Flandern, T. C. 1980. "Ephemeris of Pluto," Icarus, 44, 19-28.

Seidelmann, P. K., Klepczynski, W. J., Duncombe, R. L., and Jackson, E. S. 1971. "Determination of Mass of Pluto" Astron J. 76, 488-492.

Tombaugh, C W. and Moore, P. 1980 Out of the Darkness: The Planet Pluto, Stackpole Books, Harrisburg, Pa.

Tombaugh, W W. 1961. "The Trans-Neptunian Planet Search" in Planets and Satellites edited by G. P. Kiuper and B. M. Middlehurst, The University of Chicago Press, Chicago & London

Walker, S C. 1847 "Elements of the planet Neptune" Proceedings of the American Philosophical Society 4 332-335.

Walker, S C. 1847 "Elliptic elements of the planet Neptune," Proceedings of the American Philosophical Society 4, 378.

Whyte, A. J. 1980. The Planet Pluto,Pergamon Press, Oxford

Wylie, L. R. 1942 A Comparison of Newcomb s Tables of Neptune with observations 1795-1938, Publ. U. S. Naval Obs. Ser. 2 15, part 1.

Figure Captions

Fig. 1. The difference between the observations and the theory of Uranus in 1845.

Fig. 2. The difference between the observations and the 1873 theory of Le Verrier for Uranus in 1915.

Fig. 3. The difference between the observations and the 1903 theory of Gaillot for Uranus in 1915.

Fig. 4. Comparison of Uranus observations with (a) Le Verrier's Theory, (b) Pickering's orbit with a predicted planet "O" and (c) Gaillot's theory.

Fig. 5. Comparison of Uranus observations with Newcombs' 1904 theory of Uranus. There is a 1"5 correction in the origin at 1904 to put Uranus close to the theory of Newcomb.

Fig. 6. Comparison of Uranus observations from Paris and Greenwich with Le Verrier's theory and observations from Washington with Newcomb's theory.

Fig. 7. Pluto discovery plates taken on 23 and 29 January 1930.

Fig. 8. The area covered by the trans-Neptunian planet search of Lowell Observatory 1929-1945.

Fig. 9. Comparison of Pluto Observations with the integration of Eckert, Brouwer, and Clemence in (a) right ascension and (b) declination.

Fig. 10. The light curve of Pluto for an assumed period of $6^{d}.3867$; magnitude reduced to mean opposition versus phase.

Fig. 11. The Elongated image of Pluto indicates the existence of a satellite of Pluto.

Fig. 12. Comparison in right ascension of observations with numerical integration for Neptune with a reciprocal mass of Pluto of 1812000.

Fig. 13. Estimated Mass of Pluto versus time. The dots are experimental data, the solid line represents the equation which is the best fit curve on which the theory is developed.

Fig. 14. Comparison between Pluto observations and a recent ephemeris in (a) right ascension and (b) declination.

Fig. 15. Comparison between Neptune observations and the 1984 ephemeris in (a) right ascension and (b) declination.

Fig. 16. Comparison in (a) right ascension and (b) declination between Uranus observations, 1781-1980, and an ephemeris fit to all the data.

Fig. 17. Comparison in (a) right ascension and (b) declination between Uranus observations 1781-1980 and an ephemeris fit to Uranus observations post 1900 only.

Fig. 18. Comparison between Saturn observations and a recent ephemeris in (a) right ascension and (b) declination.

Fig. 19. Comparison between Jupiter observations and a recent ephemeris in (a) right ascension and (b) declination.

Table 1. Neptune Elements

	Elements	Walker	LeVerrier	Adams
a	(mean distance)	30.25	36.15	37.25
e	(eccentricity)	0.00884	0.10761	0.12062
i	(inclination)	$1^{\circ}54'54''$	*	*
Ω	(longitude of node)	$131^{\circ}17'38''$	*	*
ω	(longitude of Perihelion	$0^{\circ}12'25''$	$284^{\circ}45'$	$299^{\circ}11'$
P	(period, in years)	166.381	217.387	277.3
m	(mass, sun = 1)	0.0000666	0.0001073	0.000150
	(longitude 1847.0)	$328^{\circ}7'57''$	$326^{\circ}32'$	$329^{\circ}57'$

*Not predicted

Table 2. Mass Determinations

Date	Investigator	Observations	Mass in terms of the Earth
1848	J. Babinet	Neptune	12
1899	M. Lau	Uranus	9
1908	W. Pickering	Uranus	2
1909	B. Gaillot	Uranus, Neptune	5
1913	P. Lowell	Uranus, Neptune	6.6
1928	W. Pickering	Uranus, Neptune	0.75
1930	J. Jackson	Neptune	1.0
1931	Nicholson and Mayall	Neptune	0.94
1931	E. Brown	Uranus	0.5
1940	V. Kourganoff	Uranus	1.0
1942	L. Wylie	Neptune	0.91
1951	Eckert, Brouwer, and Clemence	Neptune	1.0
1955	Brouwer	Uranus, Neptune	0.82
1968	Duncombe, Klepczynski, and Seidelmann	Neptune	0.18
1971	Seidelmann, Klepczynski, Duncombe, and Jackson	Neptune	0.11
1971	Ask, Shapiro, and Smith	Uranus, Neptune	0.08
1976	Cruikshank, Pilcher, and Morrison	Albedo	0.004
1978	Christy and Harrington	Satellite	0.002

Table 3. Pluto, Predictions and Actual

	P. Lowell 1915	W. H. Pickering 1919	Actual Pluto
MEAN DISTANCE	43.0	55.1	39.5
PERIOD, YEARS	282	409.1	248
e	0.202	0.31	0.246
LONG PERIHELION	$204^{\circ}.9$	$280^{\circ}.1$	$222^{\circ}.9$
DATE PERIHELION	1991.2	1720.0	1989.9
Ω	———	100°	$109^{\circ}.6$
i	10°	15°	$17^{\circ}.1$
LONG 1930.0	$102^{\circ}.7$	$102^{\circ}.6$	$108^{\circ}.5$
MAGNITUDE	12 to 13	15	15
MASS (EARTH = 1)	6.7	2	0.002

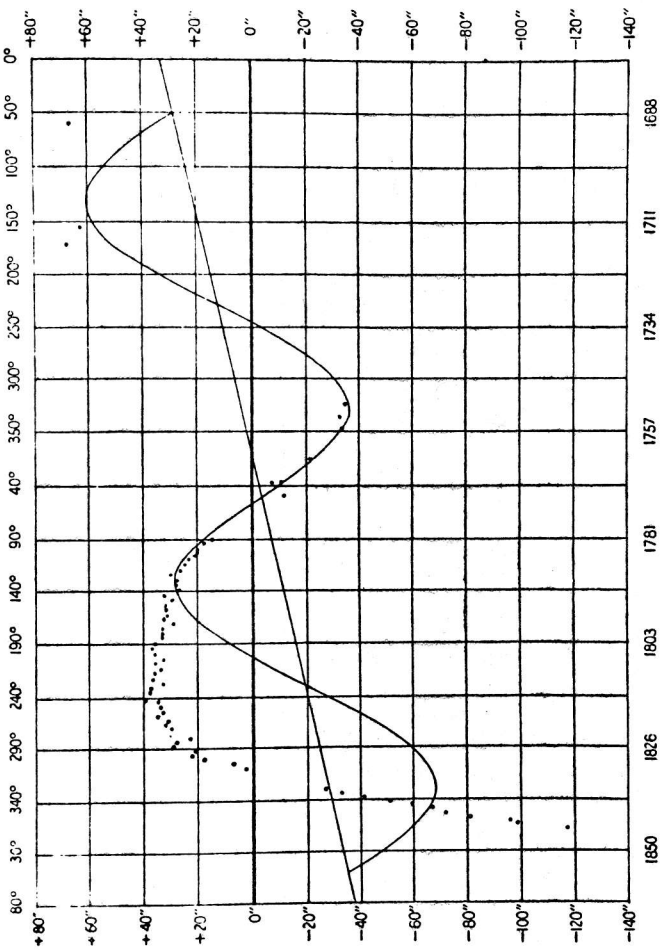

+80"
+60"
+40"
+20"
0"
-20"
-40"
-60"
-80"
-100"
-120"
-140"
0°
50°
100°
150°
200°
250°
300°
350°
40°
90°
140°
190°
240°
290°
340°
30°
1688
1711
1734
1757
1781
1803
1826
1850

Figure 1

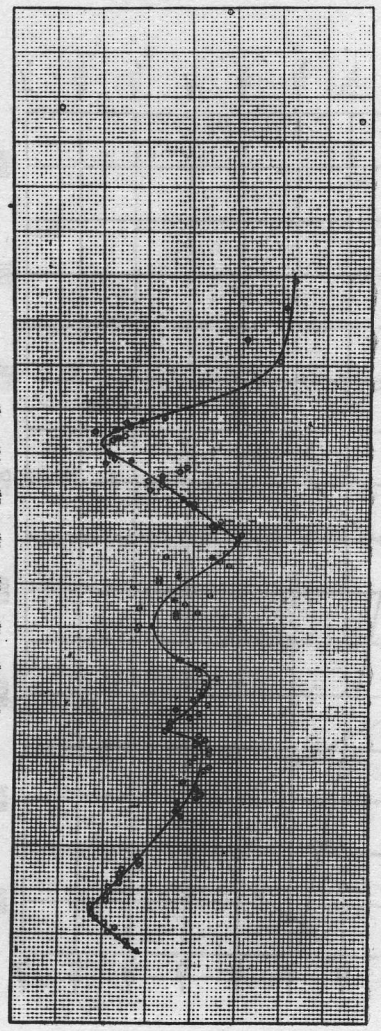


PLATE I

Figure 2

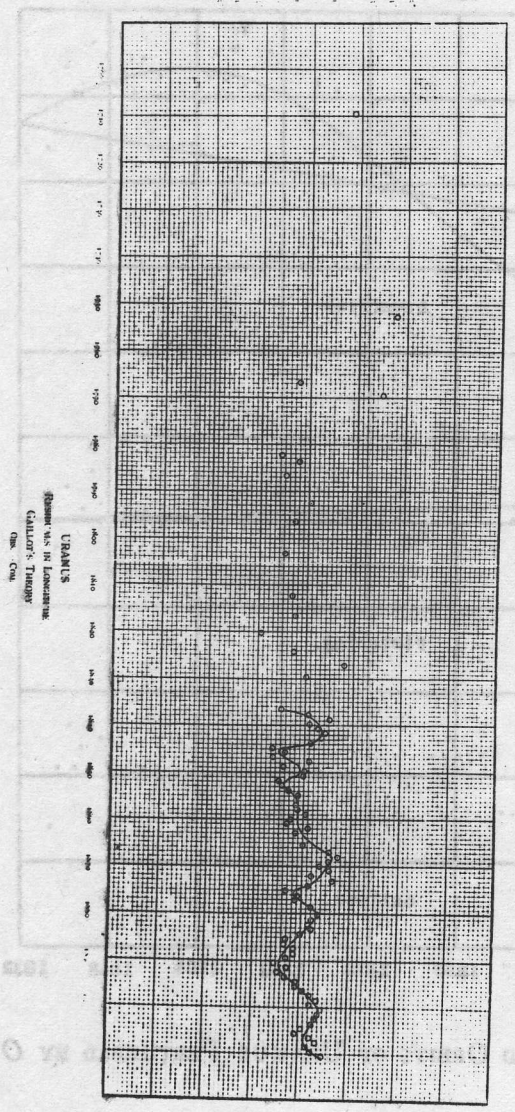
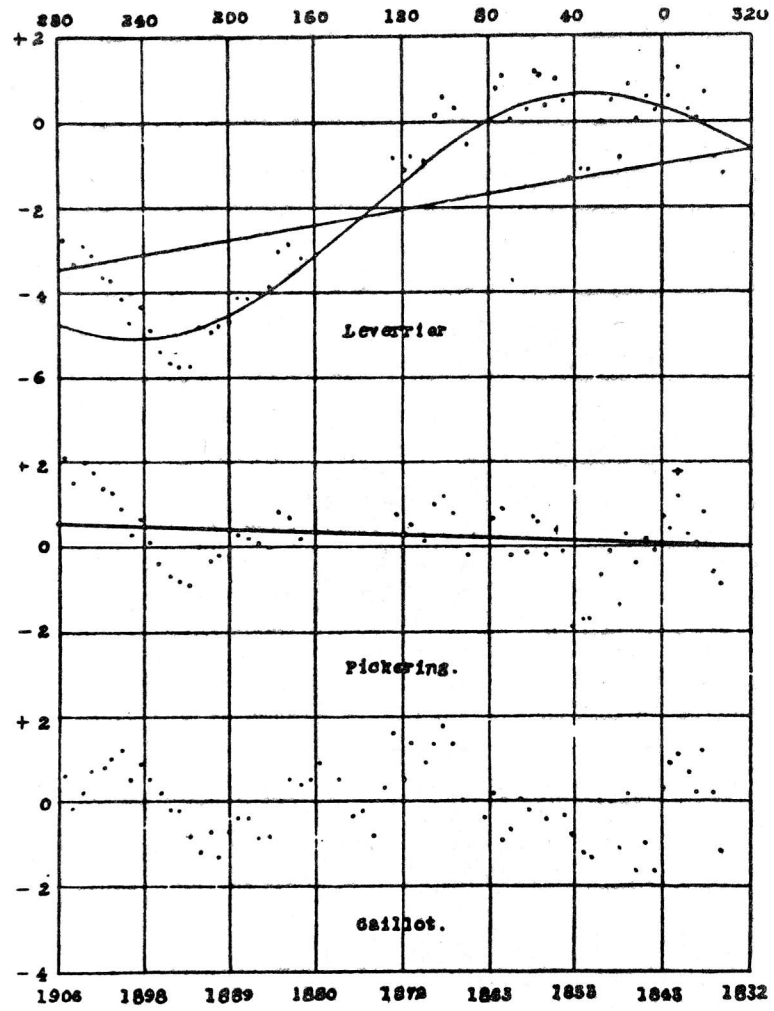
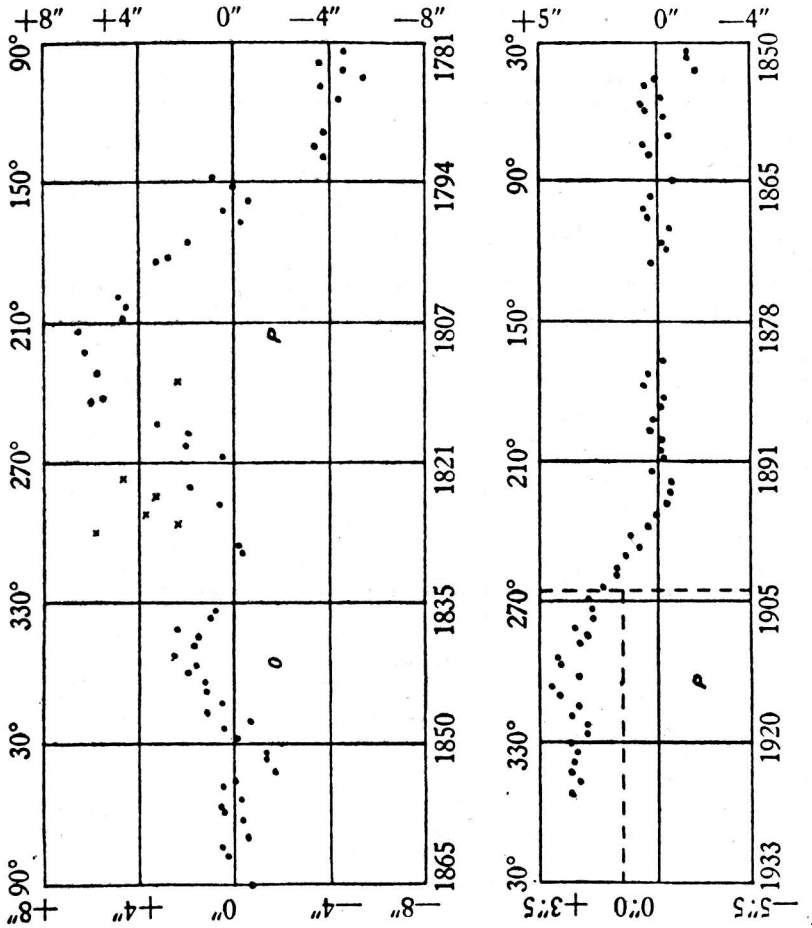


Figure 3



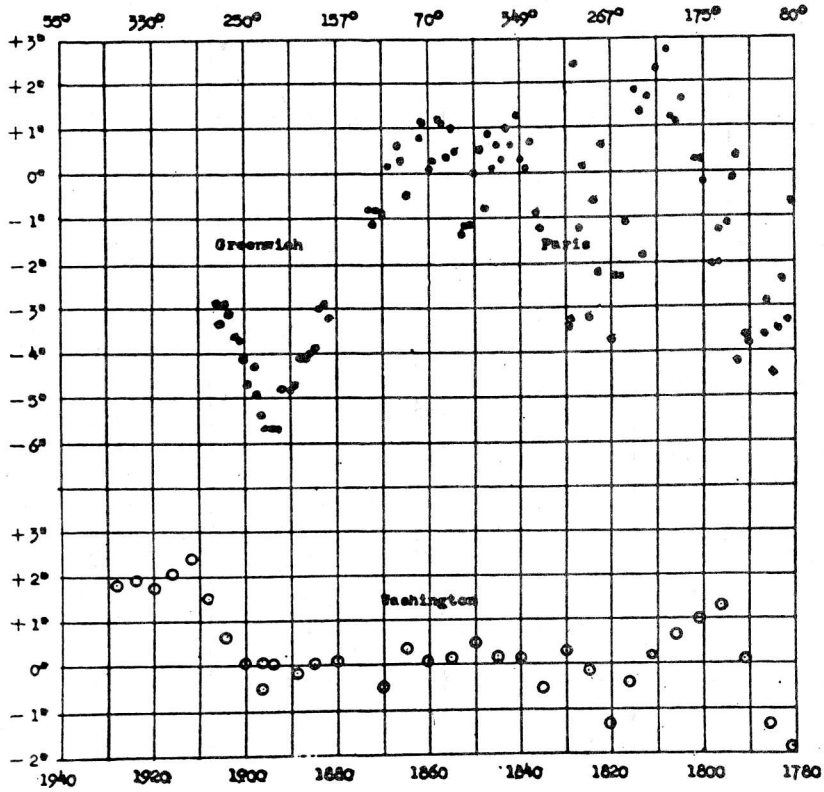
ACCEPTED AND REVISED ORBITS OF URANUS PERTURBED BY *O*.

Figure 4



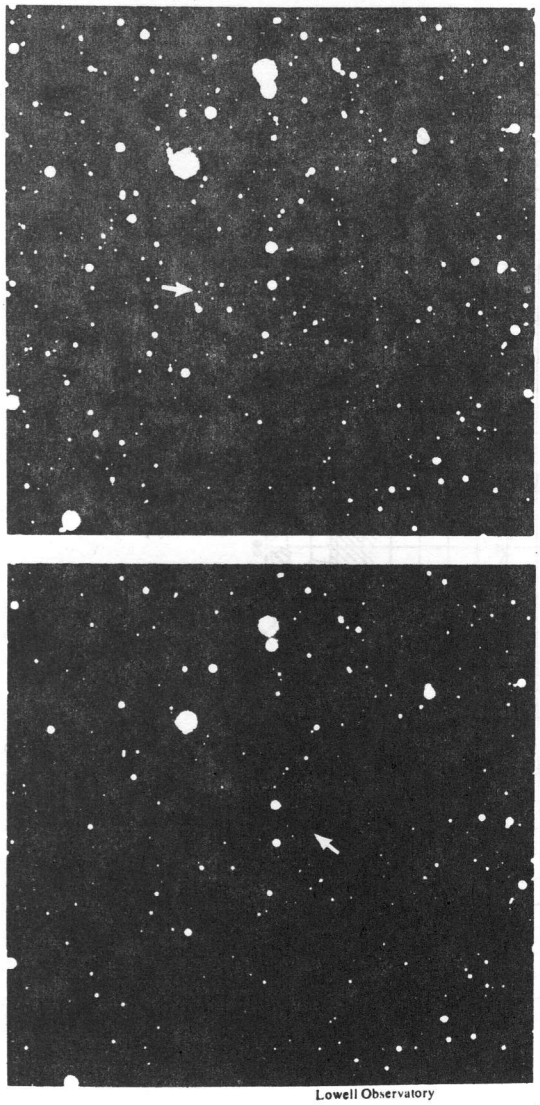
PERTURBATIONS OF URANUS.

Figure 5



PERTURBATIONS OF THE ORBIT OF URANUS.

Figure 6



Lowell Observatory

Figure 7

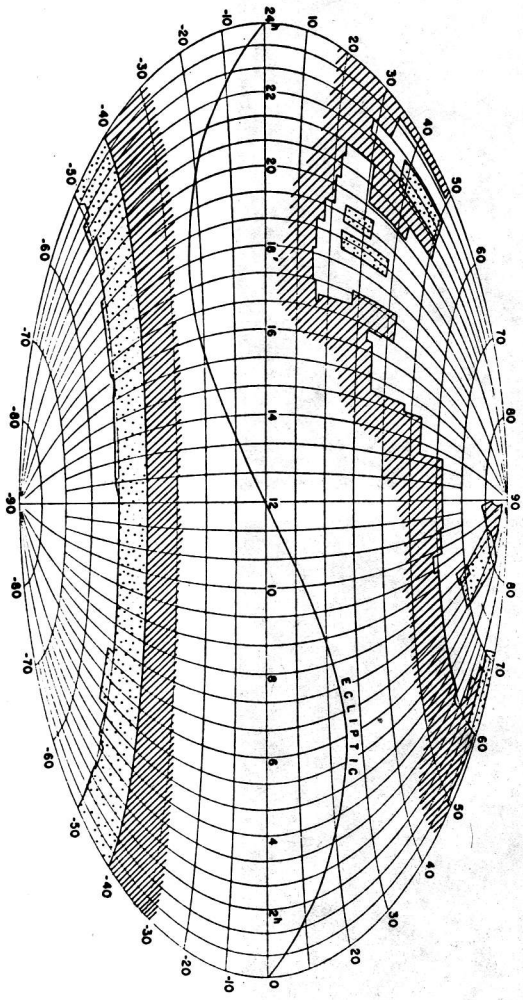


Figure 8

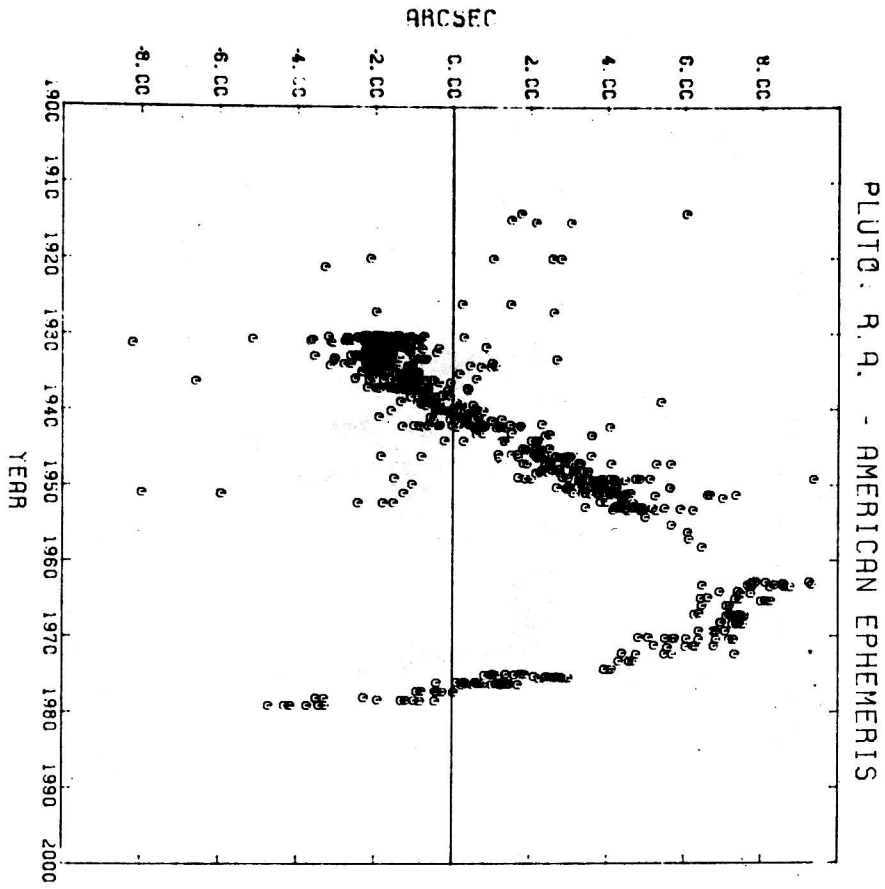

PLUTO. R.A. - AMERICAN EPHEMERIS
ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000

Figure 9a

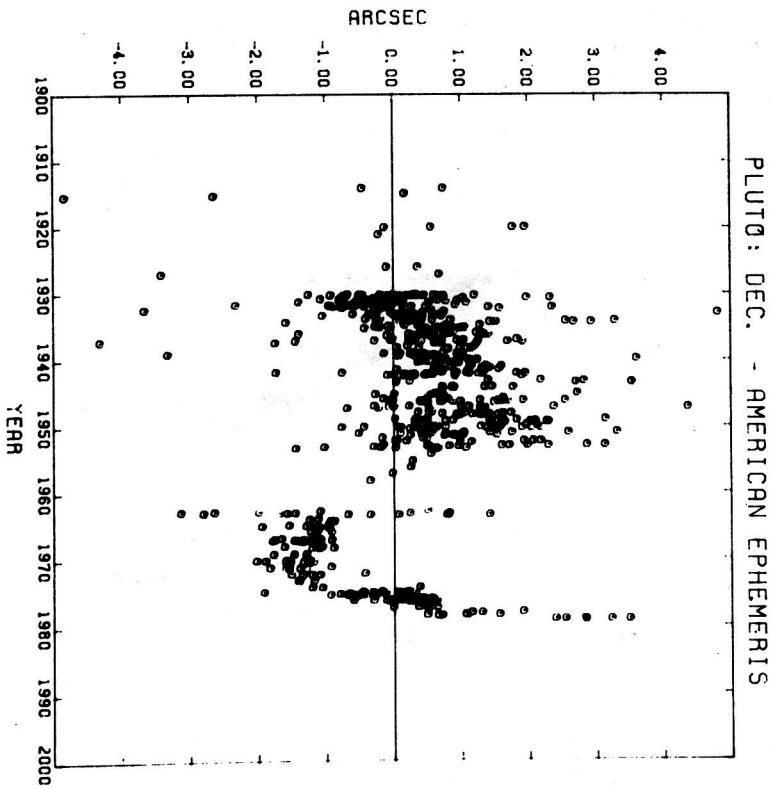

PLUTO: DEC. - AMERICAN EPHEMERIS
ARCSEC
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
YEAR
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000

Figure 9b

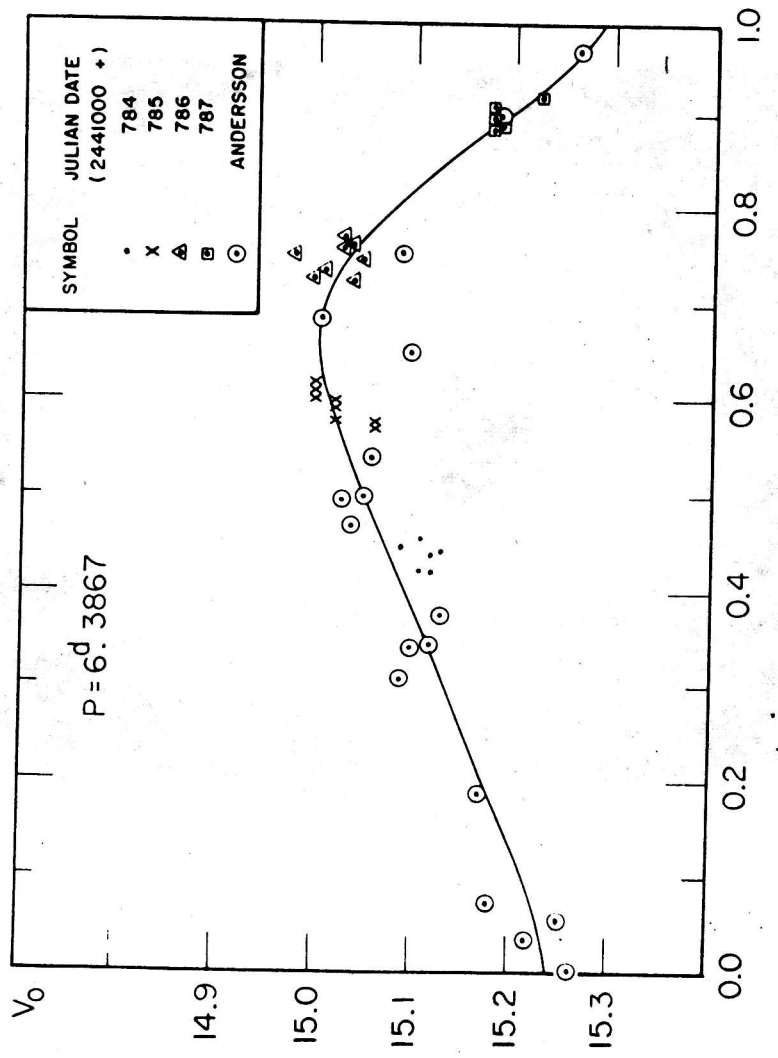

SYMBOL
JULIAN DATE
(2441000 +)
784
785
786
787
ANDERSSON
P=6d.3867
V0
14.9
15.0
15.1
15.2
15.3
0.0
0.2
0.4
0.6
0.8
1.0
φL

Figure 10

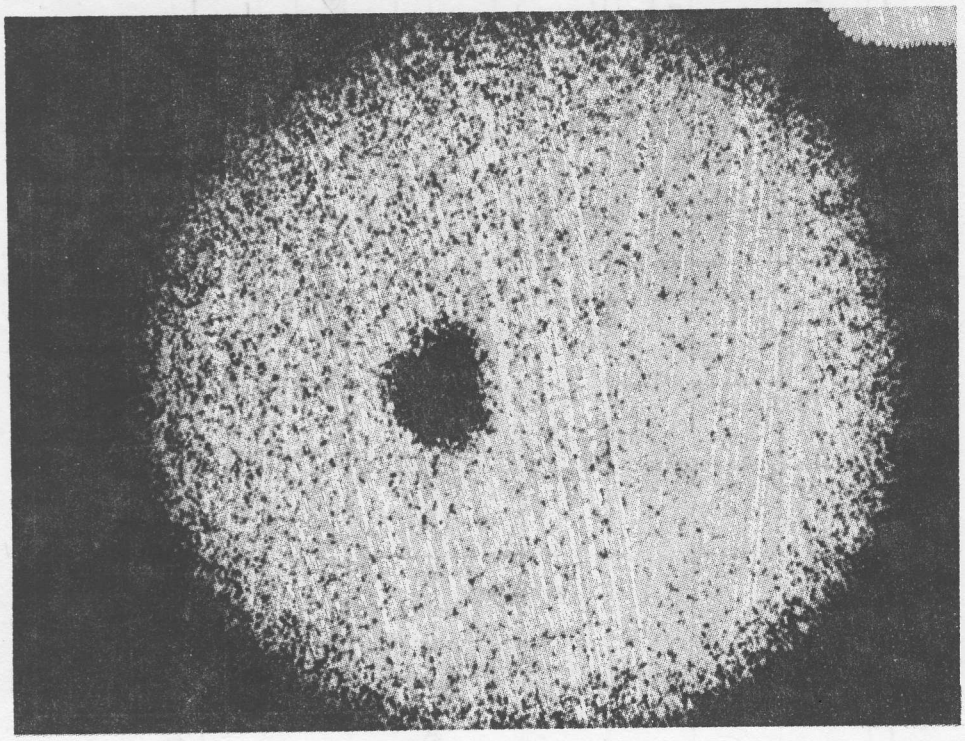


Figure 11

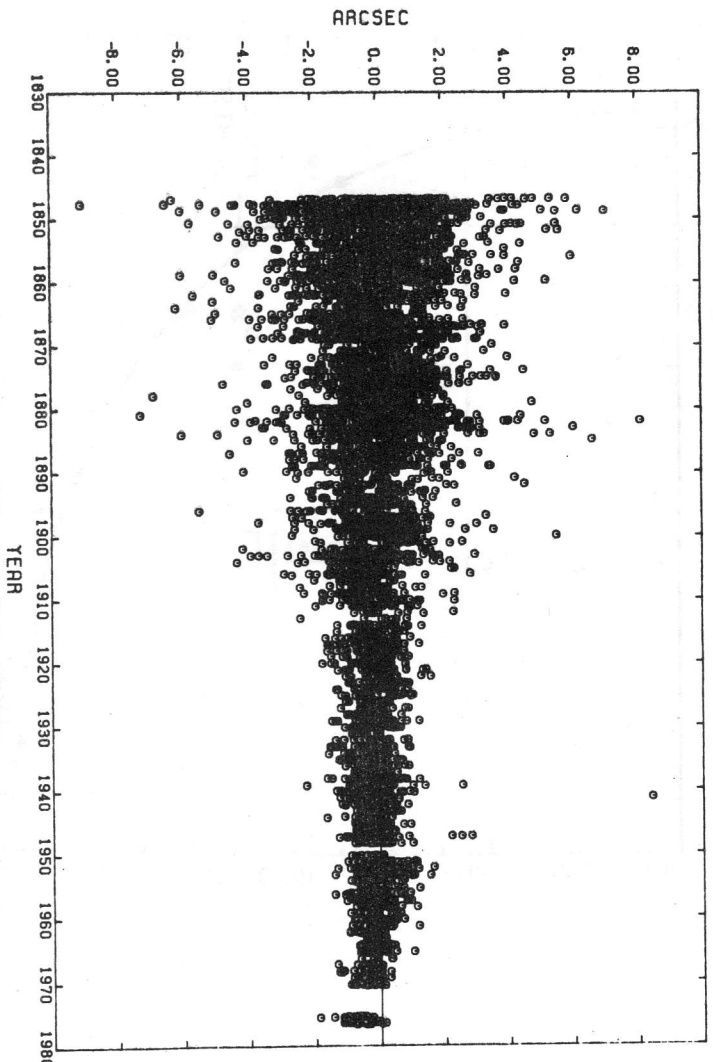

NEPTUNE: R.A. - RESEARCH EPHEMERIS
ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980

Figure 12

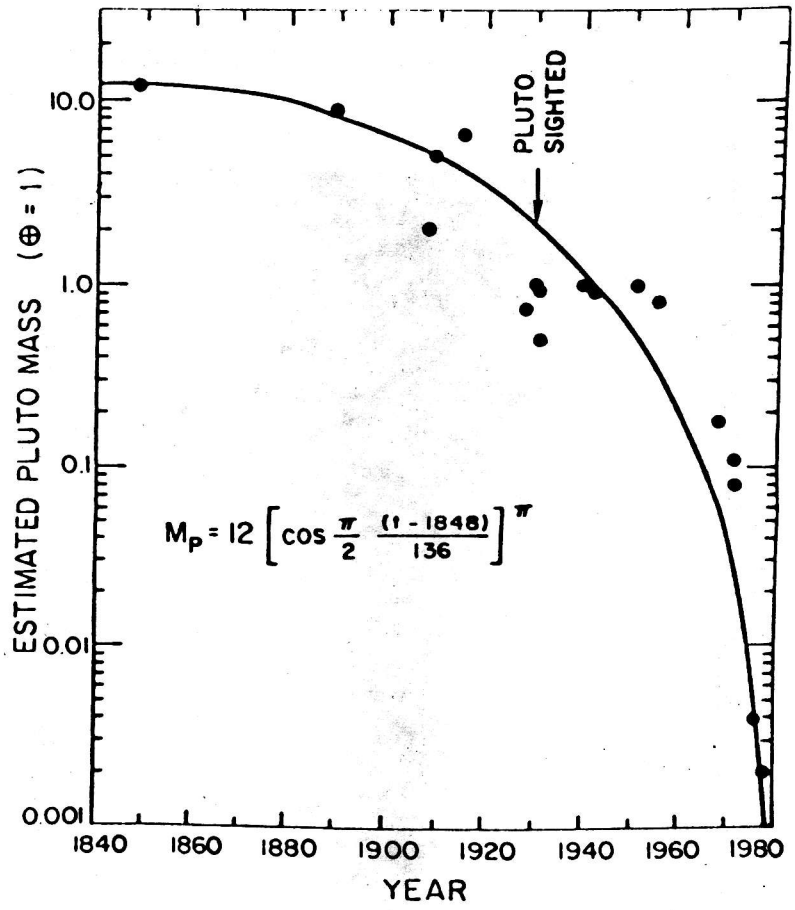

ESTIMATED PLUTO MASS (⊕ = 1)
10.0
1.0
0.1
0.01
0.001
1840
1860
1880
1900
1920
1940
1960
1980
YEAR
PLUTO SIGHTED
$M_P = 12\left[\cos\frac{\pi}{2}\,\frac{(t-1848)}{136}\right]^{\pi}$

Figure 13

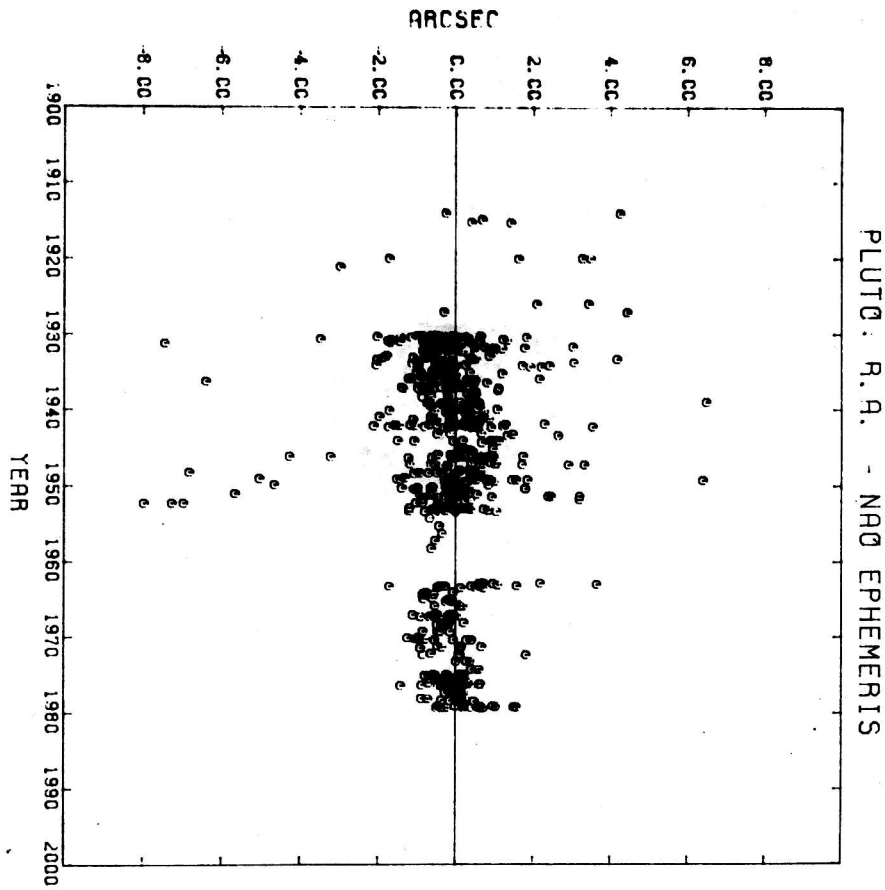

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
YEAR
PLUTO: R.A. - NAO EPHEMERIS

Figure 14a

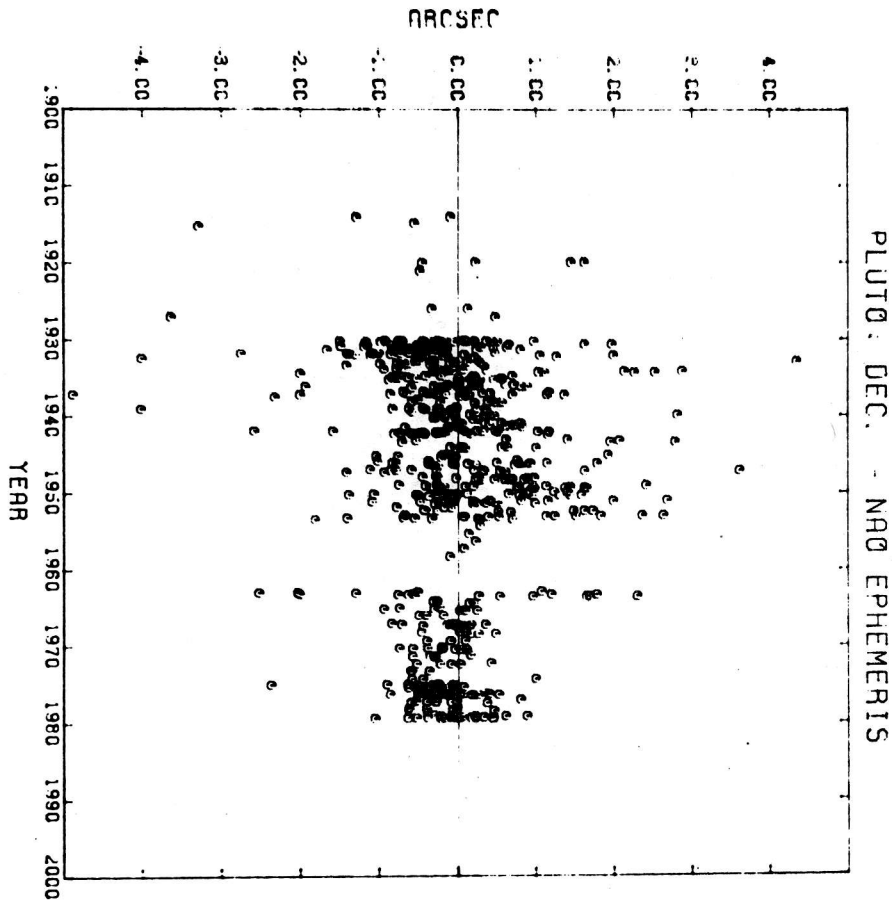

PLUTO: DEC. - NAO EPHEMERIS
ARCSEC
YEAR
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000

Figure 14b

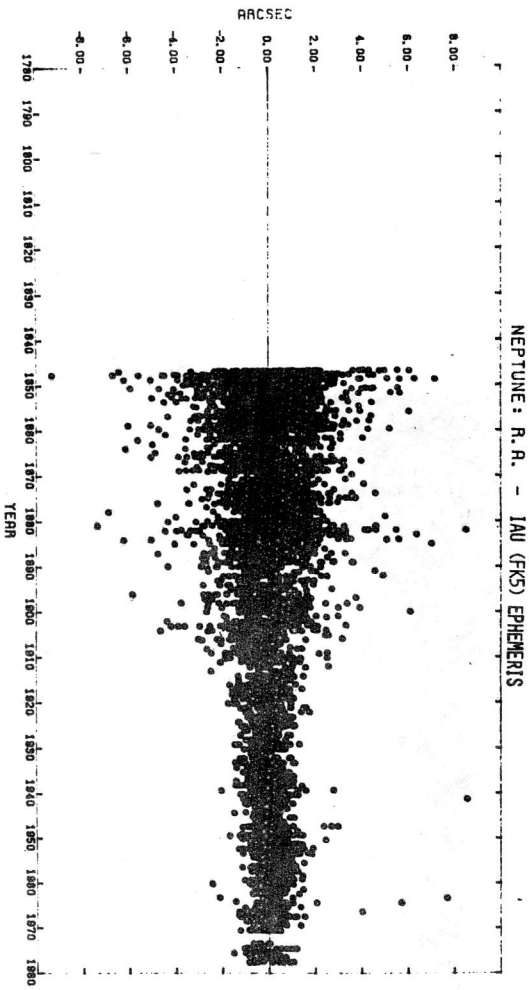

NEPTUNE: R.A. - IAU (FK5) EPHEMERIS
ARCSEC
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
YEAR
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980

Figure 15a

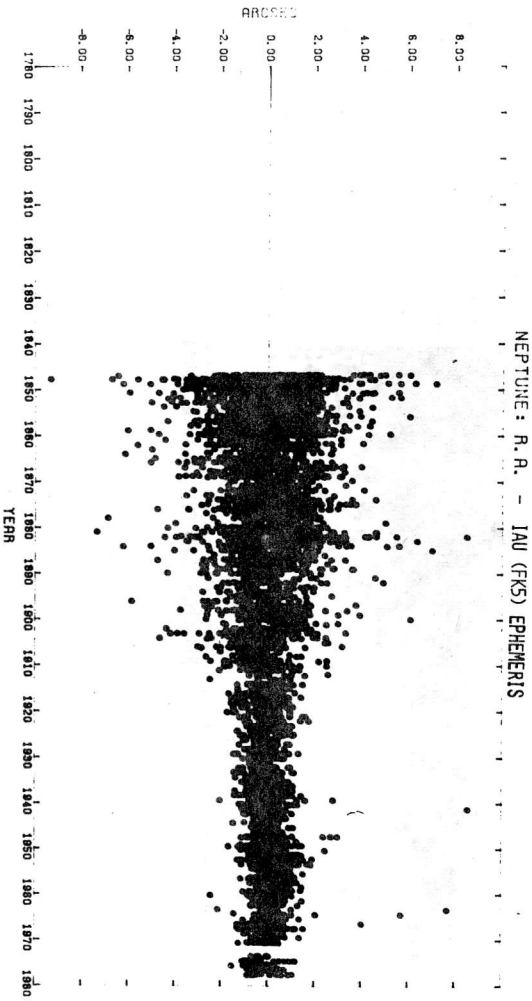

NEPTUNE: R.A. - IAU (FK5) EPHEMERIS
ARCSEC
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
YEAR

Figure 15b

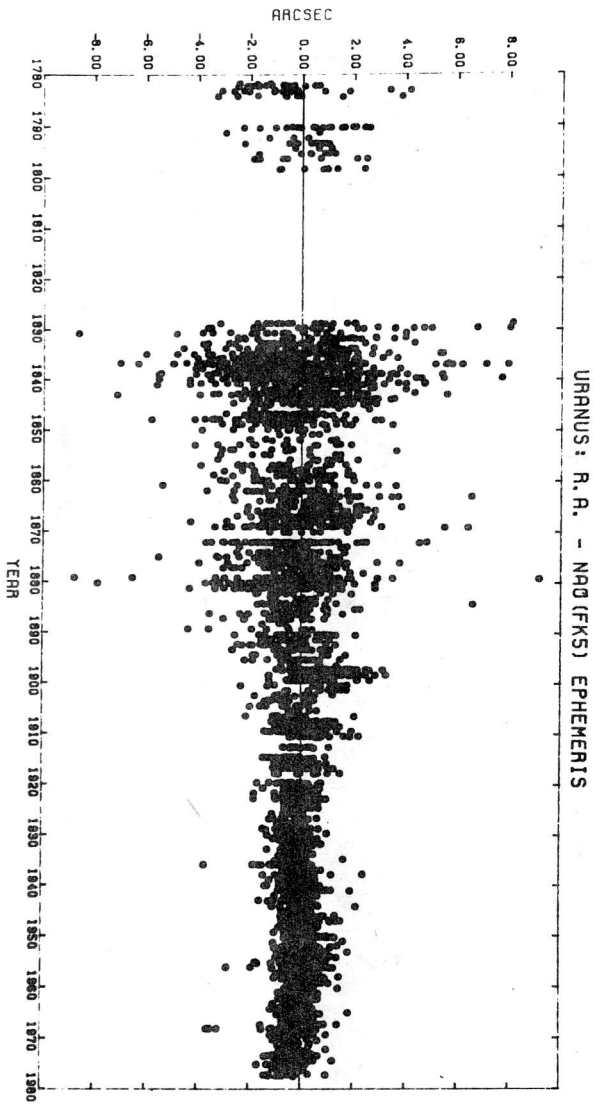

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
YEAR
URANUS: R.A. - NAO(FK5) EPHEMERIS

Figure 16a

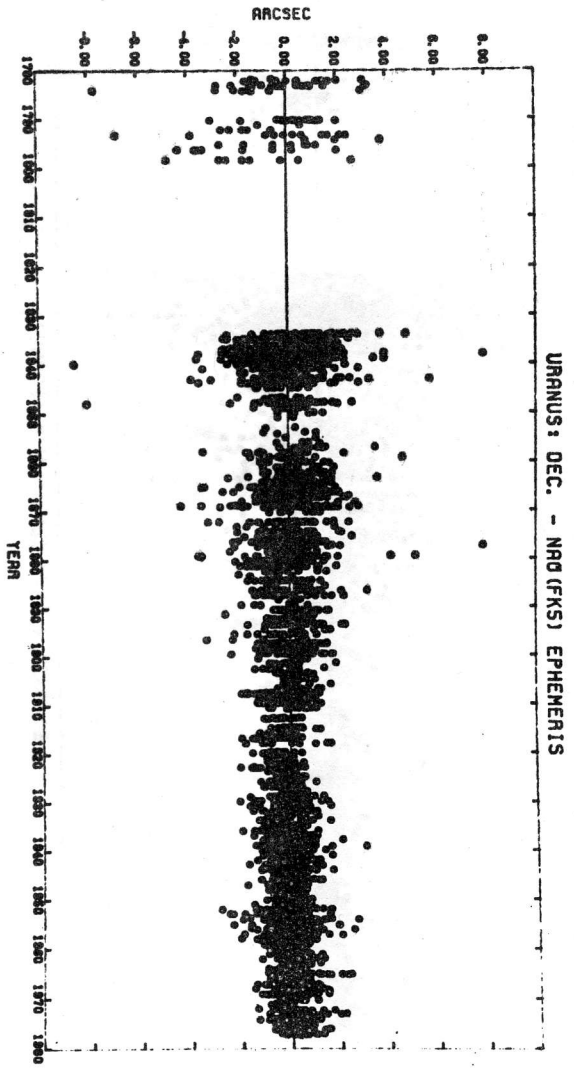

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
URANUS: DEC. - NAO(FK5) EPHEMERIS

Figure 16b

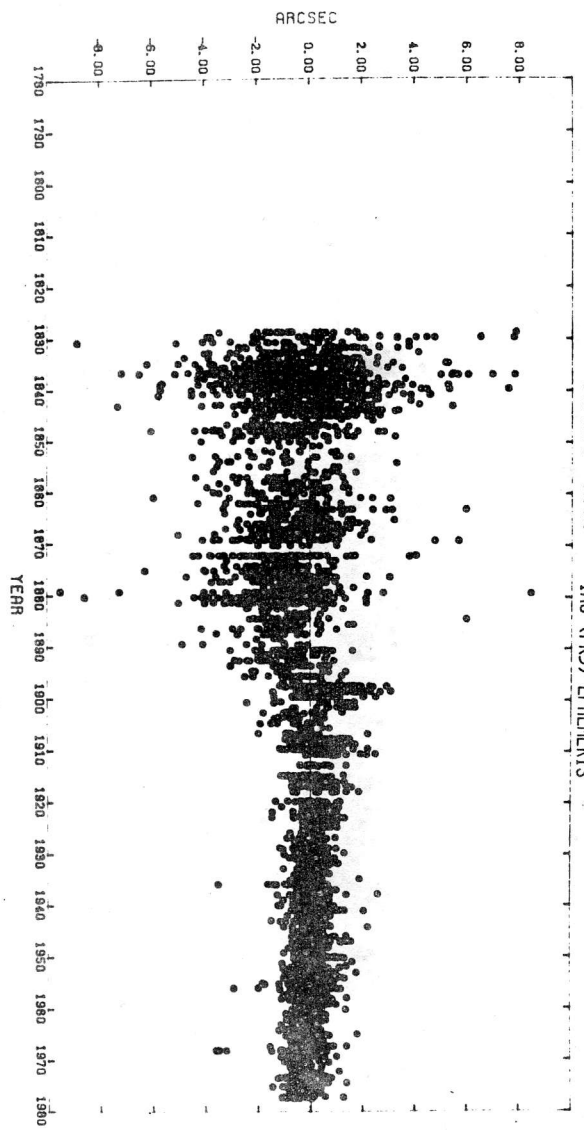

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
URANUS: R.A. - IAU (FK5) EPHEMERIS

Figure 17a

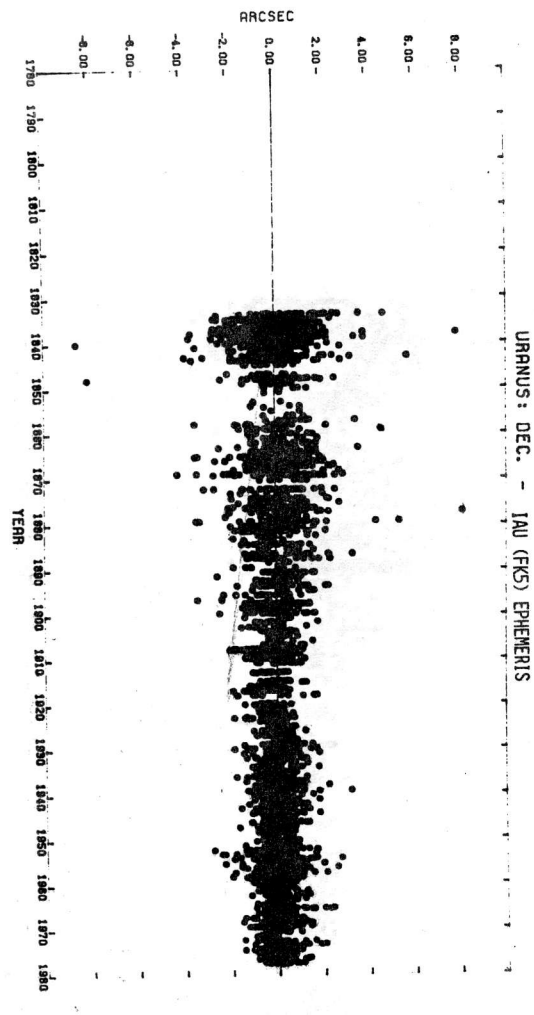

ARCSEC
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
YEAR
URANUS: DEC. - IAU (FK5) EPHEMERIS

Figure 17b

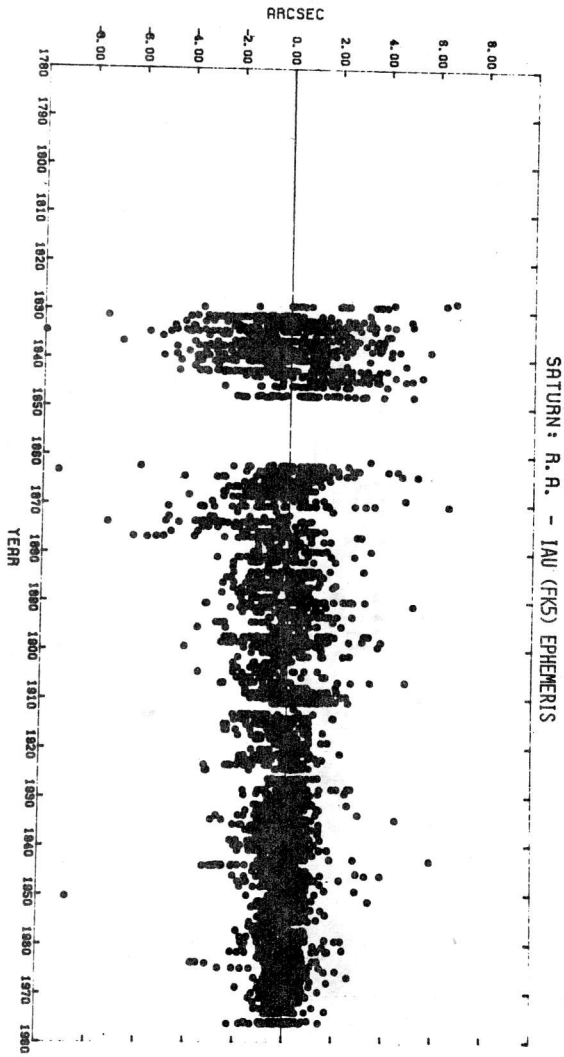

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
SATURN: R.A. - IAU (FK5) EPHEMERIS

Figure 18a

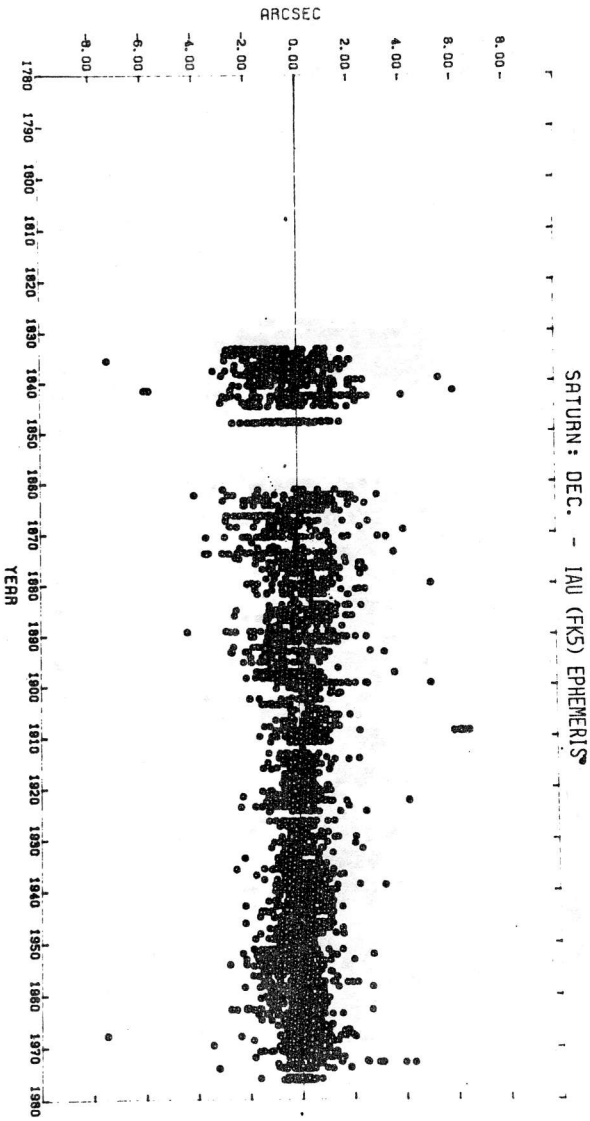

SATURN: DEC. - IAU (FK5) EPHEMERIS
ARCSEC
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
YEAR
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980

Figure 18b

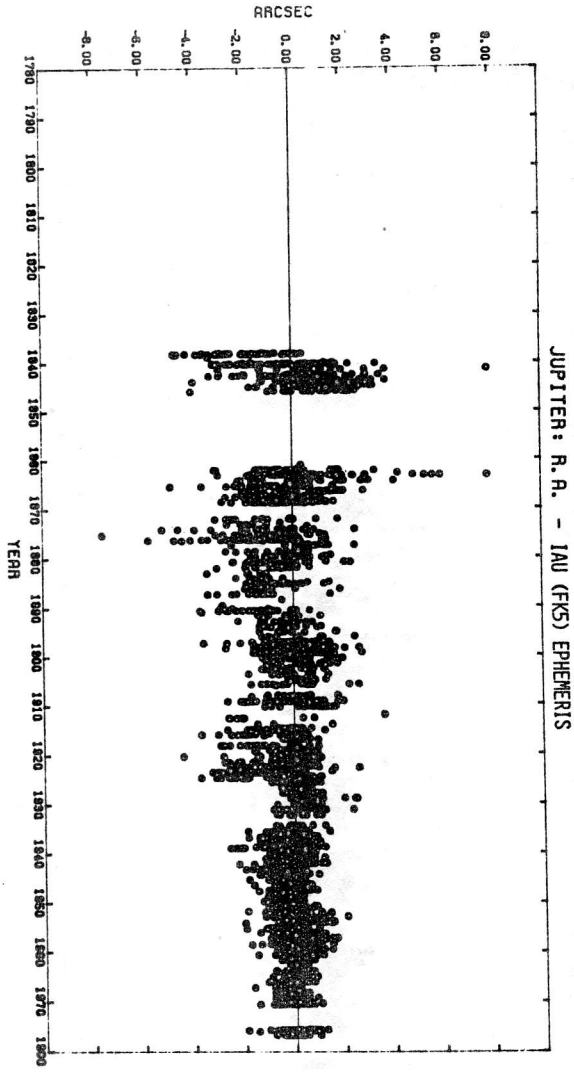

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
YEAR
JUPITER: R.A. - IAU (FK5) EPHEMERIS

Figure 19a

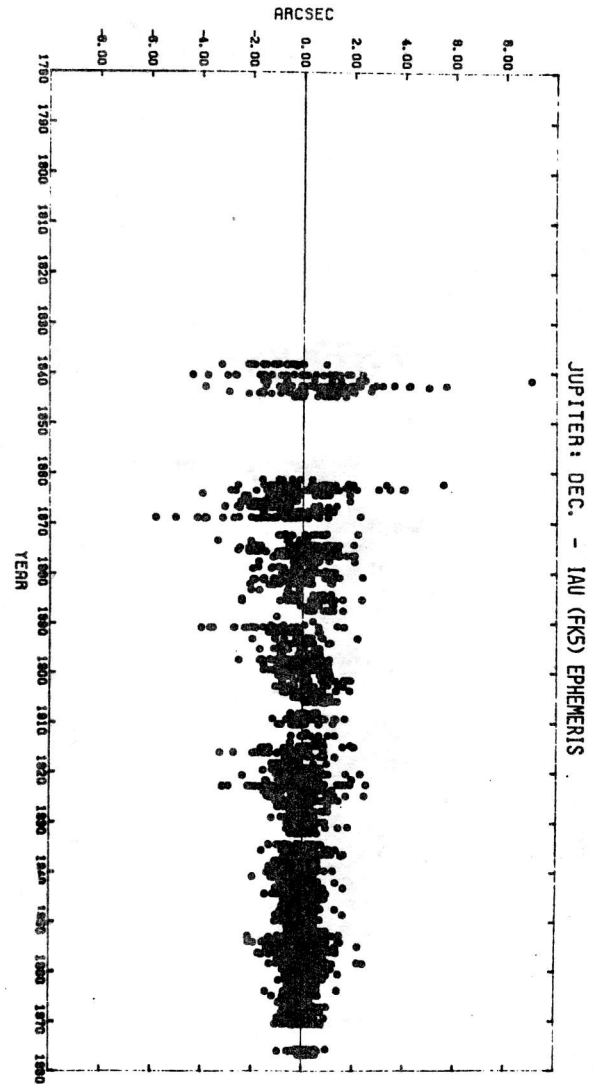

ARCSEC
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
YEAR
1760
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
JUPITER: DEC. - IAU (FK5) EPHEMERIS

Figure 19b

SATURN-ITS RINGS AND SATELLITES, A CELESTIAL MECHANICS LABORATORY

P. K. Seidelmann
U. S. Naval Observatory, Washington, D. C. 20390

Introduction

From February to May 1980, when the rings of Saturn were presented edgewise to the Earth, several observing groups, using special equipment, undertook programs of observation aimed at detecting the faint E-ring (Feibelman, 1967) and inner moons, reported after the 1966 ring-plane crossing (Dollfus 1968, Fountain and Larson 1977, Aksnes and Franklin 1978).

A subset of the Space Telescope Widefield/Planetary Camera Investigation Definition Team sought to make these observations with a Charge Coupled Device (CCD) camera and learn about such a camera at the same time. The people observing in Flagstaff were W. A. Baum, Lowell Observatory, D. Currie, U. of Maryland, D. Pascu and P. K. Seidelmann of U. S. Naval Observatory. J. A. Westphal, the Principal Investigator of the WF/PC Team, and G. E. Danielson of Caltech participated at California Observatories. Other members of the WF/PC team are J. E. Gunn, Princeton, A. D. Code, U. of Wisconsin, T. F. Kelsall, Goddard Space Flight Center, J. Kristian, Mt. Wilson Observatory, C. R. Lynds, Kitt Peak National Observatory and B. A. Smith, U. of Arizona.

Our observations were made on the 61 inch astrometric reflector of the U. S. Naval Observatory Flagstaff Station. D. Pascu and P. K. Seidelmann made further observations in May 1981 with similar instrumentation. This paper describes the observational program and its results, which indicate the interesting celestial mechanics problems presented by the Saturn system. Parts of this paper are included in papers by Baum et al (1981), Seidelmann et al (1981) and Harrington and Seidelmann (1981).

Instrumentation

A coronagraphic system, designed by Baum, (Baum, et al 1981) was used (Figure 1). The Saturn disk, the A and B rings, and the classical satellites were simultaneously masked, leaving only the E ring and its surrounding background to be optically relayed to the detector. This procedure permitted exposures up to 4 minutes in duration (equivalent to several hours photographically) without saturating the CCD and causing an overflow of excess charge. A different mask was made for each night, and it was designed for Saturn and the satellites to be placed behind different parts of the mask as the night progressed. These focal-plane masks, as well as the pupil mask, were made by photo-etching a thin beryllium-copper sheet and coating it afterward with carbon-black.

Portions of the Cassegrain image passing the focal-plane mask were relayed to the detector by a Wray transfer lens. The transfer lens was set for a 2:1 image reduction, resulting in a scale of 28 arcsec/mm. The field lens near the Cassegrain focus, together with the front elements of the transfer lens, produced a sharp image of the exit pupil of the telescope (i.e., an image of the primary mirror together with the shadow of the secondary cage and supporting struts) at the iris plane of the transfer lens. When a pupil mask, of the form shown in Figure 1, was put in place of the iris of the transfer lens, the borders of the four quadrants of the pupil were deleted from the beam before it reached the detector. The detector was the ground based CCD camera system developed at the California Institute of Technology for the Space Telescope Wide Field/Planetary Camera Investigation Definition Team. At the time of the Saturn observations of March 1980 this camera contained a developmental 500 x 500 pixel Texas Instruments CCD, housed in a liquid-air dewar and thermostatically maintained at -120°C. Each pixel was approximately 15 x 15 μm, so the total detector area was 7.5 x 7.5 mm. In May 1981 an 800 x 800 pixel CCD was installed and the temperature was maintained at -131°C. Each pixel was 15 x 15 μm. The CCD was oriented such that the pixel array was not aligned with the Saturn ring plane, diffraction axes, nor telescope motion axes. The exposures were made through a filter having a central wavelength (λ_c) of 700 nm and a full-width-at-half-maximum (FWHM) of 260 nm. Exposure times were controlled by an electronic interval timer and a Uniblitz shutter. Images were digitized with a 16-bit A/D converter and were recorded on magnetic tape. tandard procedures were used for bias subtraction and for determining relative pixel sensitivities.

E-Ring

The best E-ring photometry was obtained during the nights immediately following ring-plane passage in 1980. The total field covered by the detector was 215 x 215 arcseconds, and each pixel was 0.43 x 0.43 arcsecond. Saturn was at a distance of 8.448 a.u. from the Earth, so a Saturn radius (R_S) was 9.865 arcseconds, and 1 arcsecond was 6126 km.

Three different versions of part of a sample CCD frame are reproduced in Figure 2. It was a 60-second exposure obtained at 08^h38^m UT on 15 March 1980. The top version represents the original flat-fielded image displayed with six of the full 16 bits of dynamic range, such that the most significant bits in the original scene are clipped, resulting in isophotal contouring of the brightest areas. Except for scattered light, nearly all of the signal in the ring plane belongs to the E ring. The Saturn disk, the A and B rings, and the principal satellites are all occulted. Several star images can be seen, including one at $7.2R_S$ that might be mistaken for a satellite near the western end of the E ring.

In the middle version of the image in the figure, the scattered light due to Saturn has been largely removed within two rectangles containing the E ring. This was done by fitting the radial distribution of scattered light with a smooth function and subtracting it from the top version. The bright peak of the edge-on E-ring profile at 3.8 R_S becomes readily apparent, and a new 18.3-magnitude satellite can easily be seen at $4.8R_S$ on the west side.

The bottom version of the image is similar to the middle version, but the contrast has been increased ("stretched") to show more of the outskirts of the E ring. At the center of the bottom version we have superimposed part of another frame to indicate the location of the edge-on A and B rings (with the disk of Saturn masked) at the same image scale.

Across the bottom of the figure is a scale indicating the distances to the previously known satellites of Saturn.

Figure 3 shows an enlargement of the west side of the E-ring and an image of 1980S25 which is a satellite at the L5 libration point of Tethys.

Each image was sliced into narrow strips (1 x 45 pixels) perpendicular to the ring plane, i.e., 4.6° from a north-south line. They were approximately adjacent to one another outward from Saturn. A traverse of the data within each of these strips yielded a plot of relative E-ring brightness versus distance from the ring plane. Each traverse was then fitted

with a Gaussian curve, expressable in terms of two numbers: an area and a width. The area, after calibration, is a measure of the integrated edge-on brightness of a single-pixel-thick slice across the E ring; while the width (either σ or FWHM), after deconvolution, represents the edge-on spread of the E ring above and below its central plane. One should bear in mind that these edge-on parameters do not represent the spatial distribution of ring material; they are a two-dimensional projection of a three-dimensional spatial distribution.

The traverse areas from several frames, re-expressed in terms of 1-arcsecond-thick slices and translated into stellar magnitudes, are plotted in Figure 4 against radial distance outward from the center of Saturn. Three of these profiles represent the west side of the E ring while the other three represent the east side, but they are all plotted with the radial distance from Saturn increasing to the right so that the west and east sides, as well as profiles from different dates, can readily be compared with one another. The reality of any difference must be judged with reference to the amplitude of the noise indicated by the scatter of the plotted points. As already mentioned, very little of this noise is of instrumental origin; it is primarily due to the fundamental $\sqrt{n}$ statistics of the photoelectrons. As a result, the noise is lowest for the longest exposure (profile W1) and highest for the shortest exposure (profile E1).

Figure 5 is a simple superposition of the six edge-on E-ring profiles of Figure 4. The abscissa is labelled in units of Saturn radii. The few additional points at large radii, plotted as crosses, are means of 7 to 10 traverses each. Semi-major axes of the orbits of Mimas, Enceladus, Tethys, Dione, and Rhea are indicated by their encircled initials. The rectangle labelled A marks the inner and outer edges of the A ring.

The well-defined peak falls at $3.8R_S$, just a bit inside the Enceladus orbit, as would be expected if the true maximum of three dimensional distribution coincides with the Enceladus orbit. Although the profile may change slope in the vicinity of the Tethys and Dione orbits, the relationship of the E ring to Enceladus appears to be of a totally different character. The observational evidence suggests a causal connection between Enceladus and the E-ring. Figure 6 suggests that there may be slight bunching of E-ring particles in the vicinity of the trailing Lagrangian point (L_5), but the evidence is marginal. An alternative interpretation might be to suppose the presence of an "Enceladus B" satellite, embedded in the E ring and librating around the L_5 point.

Satellites

As of IAUC 3574 more than 140 observations of new satellites were reported in the circulars by all observers since November 1979. Identities between many of these "new" satellites were noted in the circulars and three new satellites were recognized. These are 1980S1 (S10) (mag. 14) and 1980S3 (S11) (mag. 15) which move in nearly the same orbit at a distance of 24".5 from Saturn and 1980S6 (S12) (mag. 18), known as "Dione B" because it is located at the L_4 point of Dione-Saturn. Voyager discovered three inner satellites S13, 14, and 15. However, there remained observations that did not fit any of these objects.

a. Orbital computations

Initially, numerical integrations were computed for objects in orbits similar to 1980S1, 1980S3, and 1980S6 and also for 2 day and 6 day orbits, to determine whether perturbations due to the large classical satellites had a significant effect during the short time span of the observations. The dyamical model was planar, included the J_2 and J_4 terms of Saturn's potential function and perturbations due to satellites I through VIII, but excluded perturbations due to the Sun. No significant perturbative effects were detected over a short time period.

Next, approximate ephemerides were computed for these three satellites which were used to make a first pass at selecting those observations associated with each orbit. Improved orbits were derived by differential correction to the selected observations and the selection process was repeated.

b. Co-orbiting 1980S1 and 1980S3

Smith (IAUC 3483) had suggested that 1980S1 and 1980S3 were in the same orbit. The orbit of 1980S1 is about the same as that given for 1966S2 by Fountain and Larson (1977) and by Aksnes and Franklin (1978). The present precision of the motions does not permit confident identification of 1980S1 or 1980S3 with either 1966S1 or 1966S2.

The two satellites known as 1980S1 and 1980S3 have revolution periods of approximately 16 hours 40 minutes and differ by only 28 seconds. Superficially, this indicates that the centers of the satellites would be within 48 kilometers of each other around March 1, 1982, and were in that relative position in mid-January 1978. Either these satellites are very transitory phenomena, or there is some dynamical mechanism, such as libration, that preserves the stability of the system. A series of numerical integrations of the

Saturnian satellite system was undertaken to examine this question.

The dynamical model included Saturn with oblateness (using the values of J2 and J4 given in the 1981 ASTRONOMICAL ALMANAC), the first eight major satellites, plus whatever test satellites were under investigation. All orbits were rotated (rather than projected, to preserve frequencies) into the plane of the rings of Saturn, a significant approximation only for the outer two satellites. Masses and orbital parameters were again taken from the 1981 ASTRONOMICAL ALMANAC. The test satellites were assumed to be isolated, in that the perturbations of the Sun and Jupiter were not considered. The test masses were first deduced from the reported magnitudes, and these were used for most experiments. The masses were later revised downward when Voyager I data on the sizes of the satellites became available.

In the major experiment, masses of 2×10^{-8} and 8×10^{-9} of the mass of Saturn were used, and the system was integrated from February 16, 1980, forward to April 3, 2062, and backwards to November 15, 1867. It was found that there is indeed a libration in the difference in longitudes of the satellites, with a period of 2570 days, such that the satellites never get closer than about 15°, which is about 40,000 kilometers.

Figure 7 shows the motion of 1980S3 with respect to 1980S1 in a coordinate system that is pulsating and rotating such that the later satellite remains fixed on the x-axis. Presented are the two librations prior to February 1980; these are good examples of the extremes of the possibilities. At first glance these resemble the orbit in the restricted three-body problem that is the separatrix between tadpole orbits around L_4 and L_5 and the horseshoe encompassing L_3, L_4, and L_5. (See Rabe[5] 1961 for the first numerical exploration of this problem in the small secondary case, and, for example, Garfinkel 1977 for more recent theoretical work.) However, the period of the separatrix is infinite, and in any case the motion here is not simple periodic.

The stability of this motion is of paramount importance, although it can not be resolved unambiguously for periods comparable to the age of the solar system. However, a few results from this and auxiliary experiments may shed some light on the problem. First, almost 28 complete librations were covered in the main experiment, and there is absolutely no indication of evolution or disintegration of the basic behavior. Second, the shorter period was decreased by up to three standard deviations of its determined value and the

longer one increased by three of its standard deviations and the system was reintegrated. Although the libration period decreased (since the synodic period decreased), the basic librational behavior was maintained. Third, the masses of the test satellites were varied, and again the libration periods changed, but the basic behavior was maintained. Indeed, the finally adopted masses (6×10^{-9} and 1×10^{-9} times that of Saturn), based on the observed sizes and an assumed specific gravity of 1.0, only increased the period of the libration to 3000 days, and decreased the minimum separation to 6°. Fourth, the removal of the major satellites did not affect the apparent stability over the test period. We conclude that this libration is at least quasi-stable and that there is no need to assume that the satellites had to have formed recently or are about to disappear. It should be noted that the analysis of the 1966 observations by Larson et. al. (1981) suggests S11 was following S10 by 137° in December of that year, with the distance decreasing. Assuming S11=1980S3 and S10=1980S1, this is consistent with a libration period of approximately 2500 days, more consistent with the somewhat larger masses used in the initial experiments. Qualitatively, the motions of these two satellites can be considered as undergoing two phases. During the period when they are separated by more than 150° the dominant perturbations are those of the major satellites. Of particular significance in this regard is the 2:1 resonance with Enceladus and the 4:1 resonance with Dione. However, as the satellites approach each other, their mutual perturbations become the significant factor, and the motion is that predicted by standard restricted three body theory, i.e. libration about one of the triangular libration points which prevents a close approach. However, each circulation around a libration point is different, because of the perturbations produced around opposition. Thus, the specific values of relative radius or velocity do not repeat for each libration. In time, therefore, we may see all members of the family of periodic orbits traced out by these two satellites.

The fact that we see these two satellites exhibiting this kind of librational motion suggests that such behavior is neither unstable nor implausible. This in turn gives added credence to the suggestions that a similar behavior between very small particles and large clumps of particles or individual bodies may explain the known ring structure around three planets.

d) 1980S6 ("Dione B")

The discovery of this satellite by Laques and Lecacheux (IAUC3457) came as a surprise considering Kuiper's (1961) completeness limits. Even more surprising was the finding that it was associated with the L_4 point of Dione-Saturn

(Lecacheux et al 1980, Reitsema et al 1980). That conclusion is confirmed and the libration period is approximately 800 days, in agreement with that reported by Reitsema et al (1980).

d) 1980S13, 1980S25 (Tethys librators)

At this point in the identification process in 1980 there remained at least one and possibly three satellites (outside of Mimas' orbit) distinct from Dione B. An attempt was made to fit the observations with preliminary orbits of 2 days and 6 days. Our observational material was re-examined for additional or negative observations. A good image was found in a set of exposures taken on 15 March which had been noted before but not measured because of poor reference images. A new solution favored a two day period.

Subsequently, observat ons of satellites made on 16 March and 19 May were reported (IAUC3545) and the 8 April observations were revised (IAUC3549). A new fit with these additional observations resulted in an orbit of 1.84 days. The observations and the residuals from that solution are listed in Table 1. This orbit suggests an association with the orbit of Tethys which has a period of 1.8878 days. However, an integration of this provisional orbit shows circulation with respect to Tethys, but with destabilizing close encounters.

Assuming that the 19 May observation is not related to 1980S13, we find that a 1.99 day period orbit fits the March and April dates better, as indicated by Table 2, but a re-examination of our observational material disclosed three negative observations; 1980 March 13.33946, 14.21292, and 14.25936.

Significantly, these orbits show the observations to be close to the L_4 point of Tethys-Saturn in April and L_5 point in March and May.[4] (Tables 3 and 4) Thus an alternative and preferred interpretation is that two objects are involved; 1980S13 and the leading triangular libration point (L_4) of Tethys and 1980S25 at the following (L_5).

The objects are either solid bodies or significant accumulations of E-ring material. A provisional analysis shows the April observations to be approximately 65 degrees ahead of Tethys while the March and May observations are about 40 degrees behind. Due to the libration motion, the March and May observations should not necessarily be at the same elongation and thus, a solution for a unique angle is probably incomplete.

The 1981 observations by Veillet at CERGA and Smith at U. of

Arizona have confirmed the L_4 and L_5 point satellites of Tethys. However, Pascu and Seidelmann's observational data, as reduced so far, did not detect the L_5 satellite of Tethys and did detect some images that did not fit the above orbits. These data are still being analyzed. Also unfortunately the reported observations were not measured positions, but interpreted data giving an angular position with respect to Tethys. Voyager II did confirm both 1980S13 and 1980S25 as solid body satellites located at the L4 and L5 points respectively in the Tethys orbit.

e) F-ring satellites?

Table 5 lists those observations which do not fit any of the previous orbits, nor can they be accounted for by a single orbit. Except for 1980S22, the observed magnitudes and separations correspond roughly to what one would expect for the F-ring and its satellites, but the observations do not accurately fit an orbit with the announced periods of the F-ring satellites.

f) Orbital Elements

Table 6 gives the orbital elements for these satellites, listing the synodic period (days), the orbit longitude at 12.0 March 1980=JD2444310.5 uncorrected for light time (greatest eastern elongation corresponds to zero longitude), and the apparent orbital radius at epoch (derived from the period and oblateness). The values of the orbital radii tabulated were consistent (within the errors) with the semi-major axis derived from the differential correction solutions. The quoted errors are the internal standard errors of the linear corrections.

Table 7 gives a comparison between the orbital periods and longitude at epoch given by Reitsema et al (1980), Synnott et al (1981) and this paper. The difference between Synnott et al and the other determinations should be expected for 1980S6. The determinations are based on observations taken at different times and thus different locations in the motion of the satellite around the libration point.

Conclusion

Thus the Saturn system contains rings extending out into the satellite orbits. Voyager has indicated extensive structure in the rings. There are satellites sharing orbits, at librations points, as ring guardians and with resonance relations.

Truly, the Saturn system is a celestial mechanics laboratory.

Figure Captions

Figure 1. Instrument assembled for observing Saturn's rings and satellites. The Cassegrain image of Saturn was focused on the focal-plane mask, and portions of the image passing it were relayed to the CCD by the transfer lens. A coronagraphic mask at the exit pupil suppressed telescopically diffracted light.

Figure 2. Three versions of part of a sample CCD frame (60-second exposure, 0838 UT on 15 March 1980) showing both ansae of the E ring. Top: original "flat-fielded" image in overflow display mode. Middle: scattered light of Saturn subtracted from two rectangular areas containing E-ring data. Bottom: same as middle image, but with more contrast stretch and with superimposed image of the A and B rings. Below the images is a scale marked in units of Saturn radii (R_S). Orbital radii of the classical inner satellites are also indicated.

Figure 3 An enlargement of the west side of the E-ring and an image of 1980S25 which is a satellite at the L5 libration point of Tethys. The image at a distance from the ring is a star.

Figure 4. Six edge-on brightness profiles of the E-ring. Samples of the west ansa: W1=1006 UT March 14, W2=0810 UT March 15, W3=0838 UT March 15. Samples of the east ansa: E1=0909 UT March 13, E2=0810 UT March 15, E3=0838 UT March 15. The gap in W3 is due to a satellite image, while those in E1 and E2 are due to faulty pixel columns in the CCD.

Figure 5. Mean edge-on brightness profile of the E-ring in March 1980. This was produced by superimposing the six profiles of Figure 4 and adding a few points (crosses) at large R_S. The dots were derived from single-pixel-wide traverses across the ring, while each cross is a mean of 7 or more traverses. Orbital radii of the A ring, Mimas, Enceladus, Tethys, Dione, and Rhea are indicated.

Figure 6. Test for variation in the height of the peak of the edge-on brightness profiles. Heights here are based on the points between 37 and 39 arcseconds in the profiles of Figure 4. From left to right, the points here are W2, W3, E1, W1, E2, E3. These data are plotted in a rotating coordinate system tied to the orbital motion of Enceladus. L_4 and L_5 indicate the equilateral libration points that Saturn and Enceladus would have if they constituted an isolated dynamic system. Ordinate units are the same as in Figure 4, and error bars are $\pm 1\ \sigma$ in length.

Figure 7. The motion of 1980S3 in a coordinate system that is pulsating and rotating nonuniformally, such that 1980S1 remains fixed at negative unit distance on the x axis.

REFERENCES

Aksnes, K., and Franklin, F. A. (1978). The evidence for new faint satellites of Saturn reexamined. Icarus 36, 107-118.

Baum, W. A., Kreidl, T., Westphal, J. A., Danielson, G. E., Seidelmann, P. K., Pascu, D., and Currie, D. G. (1981). Saturn's ring. I. CCD Observations of March 1980. Icarus, 47, 84-96.

Dollfus, A. (1968). La decouverte de 10^{e} satellite de Saturne. L'Astrononie 82, 253-262.

Feibelman, W. A. (1967). Concerning the 'D' Ring of Saturn. Nature 214, 793-794.

Fountain, J. W., and Larson, S. M. (1977). A new satellite of Saturn? Science 197, 915-917.

Garfinkel, B. (1977). Theory of the Trojan Asteroids. Part I. Astron. J. 82, 368-379.

Harrington, R. S., and Seidelmann, P. K. (1981) The dynamics of the Saturnian satellites 1980S1 and 1980S3, Icarus, 47, 97-99.

Kuiper, G. P. (1961). Limits of Completeness. In Planets and Satellites (G. P. Kuiper and B. M. Middlehurst, Eds.), pp. 575-591. Chicago: Univ. of Chicago Press.

Larson, S. M., Smith, B. A., Fountain, J. W., and Reitsema, H. J. (1981). The 1966 observations of the coorbiting satellites of Saturn, S10 and S11. Icarus, 46, 175-180.

Lecacheux, J., Laques, P., Vapillon, L., Auge, A., and Despiau, R. (1980). A new satellite of Saturn: "Dione B." Icarus 43, 111-115.

Rabe, E. (1961). Determination and Survey of Periodic Trojan Orbits in the Restricted Problem of Three Bodies. Astron. J. 66, 500-513.

Reitsema, H. J., Smith, B. A., and Larson, S. M. (1980). A new Saturnian satellite near Dione's L_4 point. Icarus 43, 116-119.

Seidelmann, P. K., Harrington, R. S., Pascu, D., Baum, W. A., Currie, D. G., Westphal, J. A. and Danielson, G. E. (1981). Saturn Satellite Observations and Orbi
. Plane Crossing. Icarus, in press.

Synnott, S. P., Peters, C. F., Smith, B. A., and Morabito, L.A. (1981). Orbits of the small satellites of Saturn. Science 212, 191-192.

TABLE 1. OBSERVATIONS OF 1980 S 13

DATE	D	O-C	IAUC	DESIG
MAR 13.36949	-44.10	0.77	3496	80S25
MAR 13.38126	-44.60	0.89	3496	80S25
MAR 15.35059	-46.76	0.62	3605	80S32
MAR 16.01000	36.30	1.39	3549	80S29
APR 8.31764	-46.50	--1.13	3466	80S13
APR 8.35208	-45.20	-2.08	3466	80S13
APR 12.02800	-41.40	1.91	3484	80S24
MAY 19.34000	31.50	--0.07	3549	80S30

TABLE 2. OBSERVATIONS OF 1980 S 13

DATE	D	O-C	IAUC	DESIG
MAR 13.36949	-44.10	0.94	3496	80S25
MAR 13.38126	-44.60	0.32	3496	80S25
MAR 15.35059	-46.76	-1.39	3605	80S32
APR 8.31764	-46.50	0.26	3466	80S13
APR 8.35208	-45.20	-0.20	3466	80S13
APR 12.02800	-41.40	0.08	3484	80S24

TABLE 3. OBSERVATIONS OF 1980 S 13 NEAR TETHYS L4 POINT

DATE	D	O-C	IAUC	DESIG
APR 8.31764	-46.50		3466	80S13
APR 8.35208	-45.20		3466	80S13
APR 12.02800	-41.40		3484	80S24

TABLE 4. OBSERVATIONS OF 1980 S 25 NEAR TETHYS L5 POINT

DATE	D	IAUC	DESIG
MAR 13.36949	-44.10	3496	80S25
MAR 13.38126	-44.60	3496	80S25
MAR 13.35059	-46.76	3605	80S32
MAR 16.01000	36.30	3549	80S29
MAY 19.34000	31.50	3549	80S30

TABLE 5. UNIDENTIFIED OBSERVATIONS

DATE	D	O-C	IAUC	DESIG
MAR 13.29000	23.00		3463	80S07
MAR 15.19000	19.00		3483	80S23
MAR 16.12600	19.75		3574	80S31
MAR 20.88500	-58.05		3474	80S22
MAR 20.89900	18.90		3497	80S20
MAR 20.90800	20.12		3497	80S20
MAR 20.91500	21.89		3497	80S20
MAR 20.92500	21.72		3497	80S20
MAR 20.92800	21.66		3497	80S20
MAR 20.97800	-51.74		3574	80S22
MAR 22.97300	-22.44		3497	80S21
MAR 22.97400	-22.00		3497	80S21
MAR 22.97700	-21.72		3497	80S21
MAR 22.98300	-21.67		3497	80S21
MAR 22.98800	-21.79		3497	80S21
APR 12.01530	-24.80		3470	80S18
APR 12.02710	-25.60		3470	80S18
APR 12.02990	-26.12		3470	80S18
APR 12.03330	-26.35		3470	80S18
APR 12.03890	-26.12		3470	80S18

TABLE 6. SATELLITE ORBITAL ELEMENTS

SATELLITE	SYNODIC PERIOD	ORBITAL LONGITUDE	ORBITAL RADIUS
1980 S1	$0^d.6945 \pm 0.0001$	$306^\circ \pm 1$	24".811
1980 S3	0.6942 ± 0.0001	$134^\circ \pm 2$	24".803
1980 S6	2.7351 ± 0.0005	$2^\circ \pm 1$	61".60
1980 S13	1.8389 ± 0.0003	$246^\circ \pm 2$	47".31
	1.9985 ± 0.0012	$313^\circ \pm 41$	50".00

Epoch 12 March 1980 = J. D. 2444310.5

TABLE 7. COMPARISON OF ORBIT DETERMINATIONS

SATELLITE	SYNODIC PERIOD AND ORBITAL LONGITUDE	THIS PAPER	SYNNOTT ET AL	REITSEMA ET AL
1980S1	P	$0^d.6945$	$0^d.6946$	
	λ_o	306°	301°	
1980S3	P	$0^d.6942$	$0^d.6942$	
	λ_o	134°	132°	
1980S6	P	$2^d.7363$	$2^d.7374$	$2^d.7361$
	λ_o	2°	14°	1°

λ_o for JD2444310.5

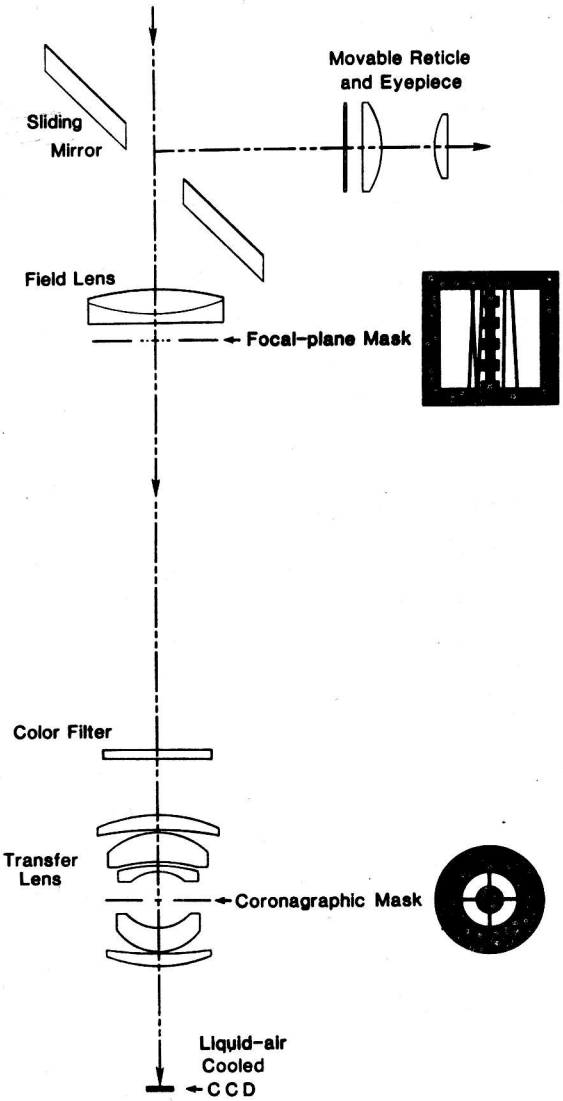

Movable Reticle
and Eyepiece
Sliding
Mirror
Field Lens
Focal-plane Mask
Color Filter
Transfer
Lens
Coronagraphic Mask
Liquid-air
Cooled
CCD

Figure 1

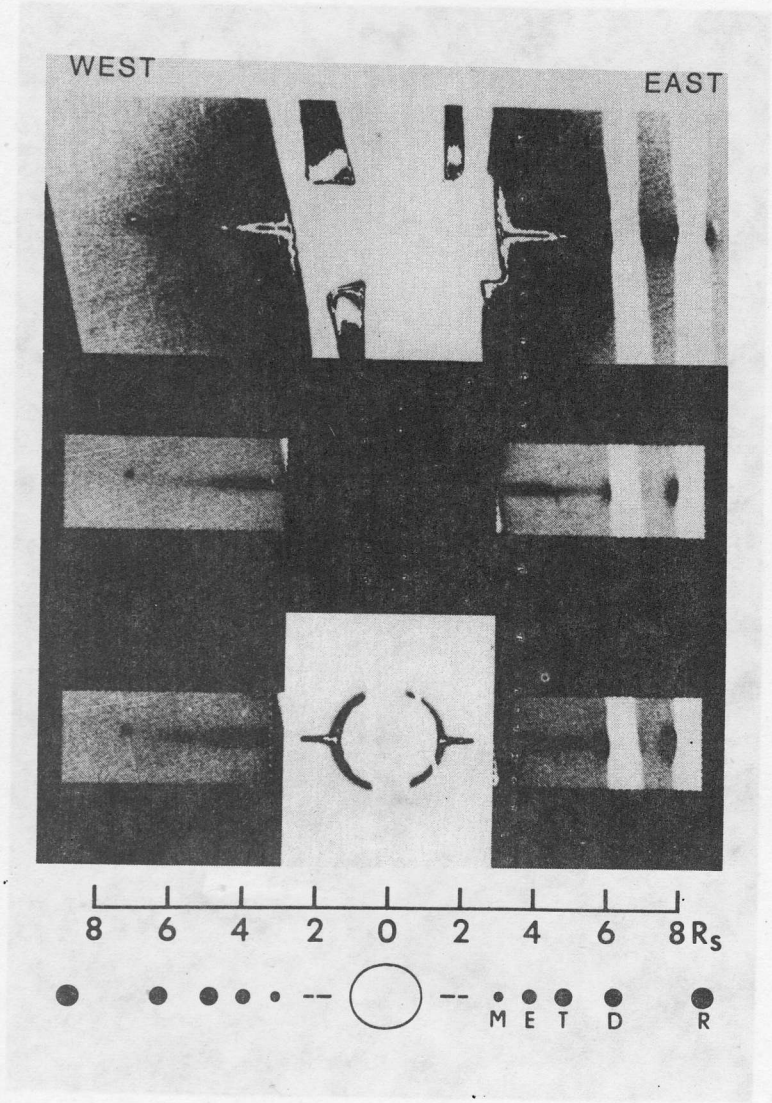


Figure 2

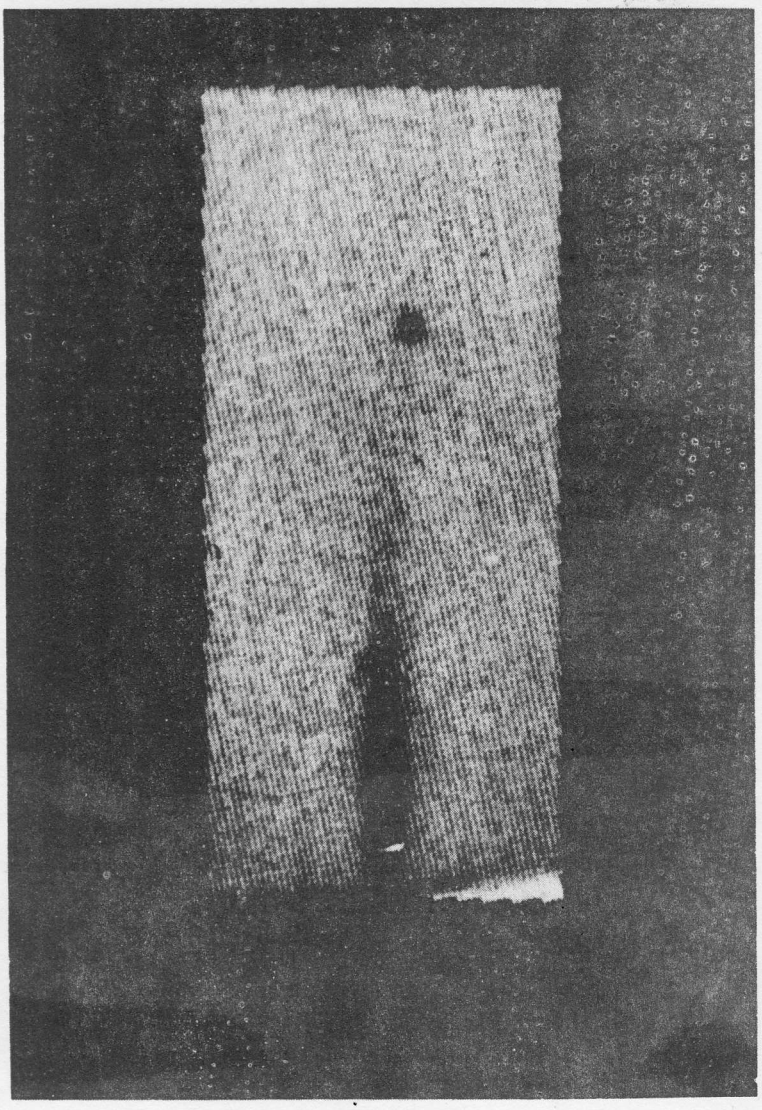


Figure 3

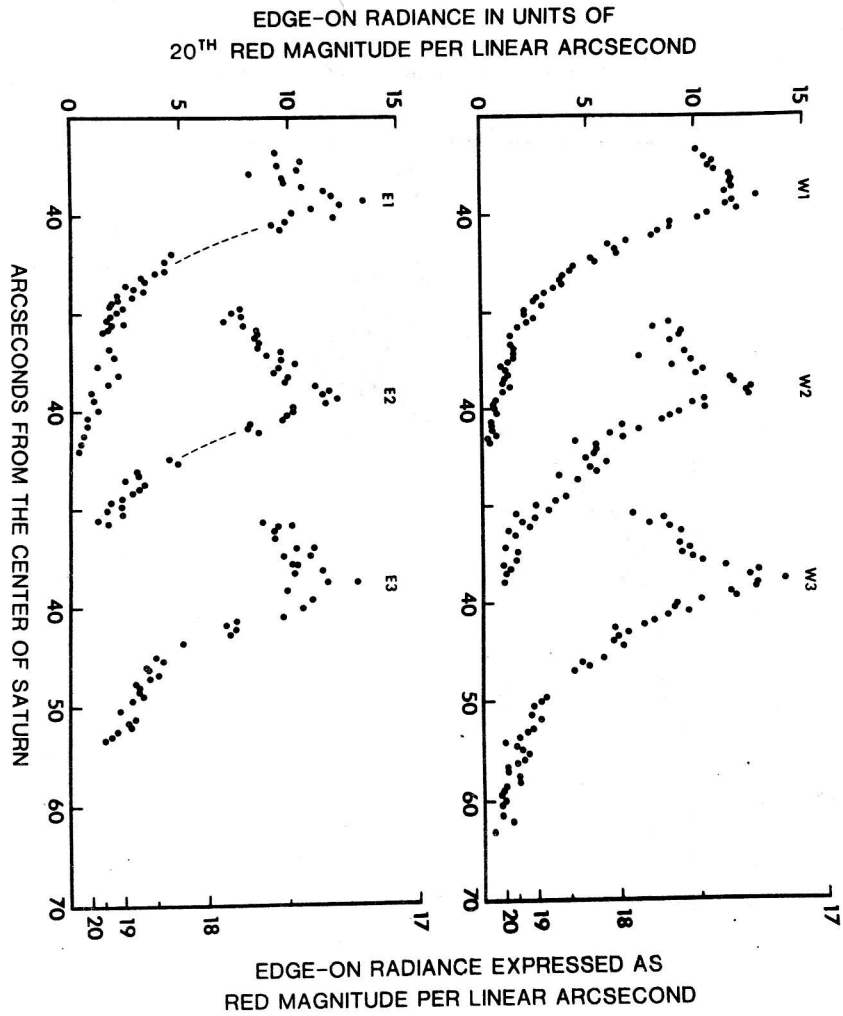

EDGE-ON RADIANCE IN UNITS OF
20TH RED MAGNITUDE PER LINEAR ARCSECOND
0
5
10
15
E1
E2
E3
W1
W2
W3
40
50
60
70
ARCSECONDS FROM THE CENTER OF SATURN
20
19
18
17
EDGE-ON RADIANCE EXPRESSED AS
RED MAGNITUDE PER LINEAR ARCSECOND

Figure 4

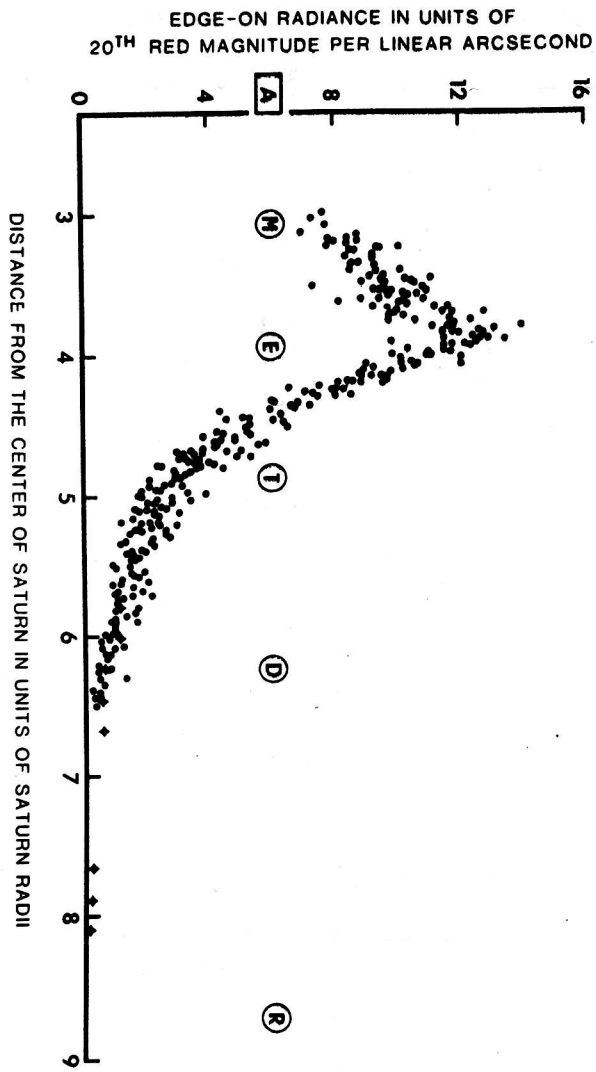

EDGE-ON RADIANCE IN UNITS OF
20TH RED MAGNITUDE PER LINEAR ARCSECOND
0
4
A
8
12
16
3
4
5
6
7
8
9
DISTANCE FROM THE CENTER OF SATURN IN UNITS OF SATURN RADII
M
E
I
D
R

Figure 5

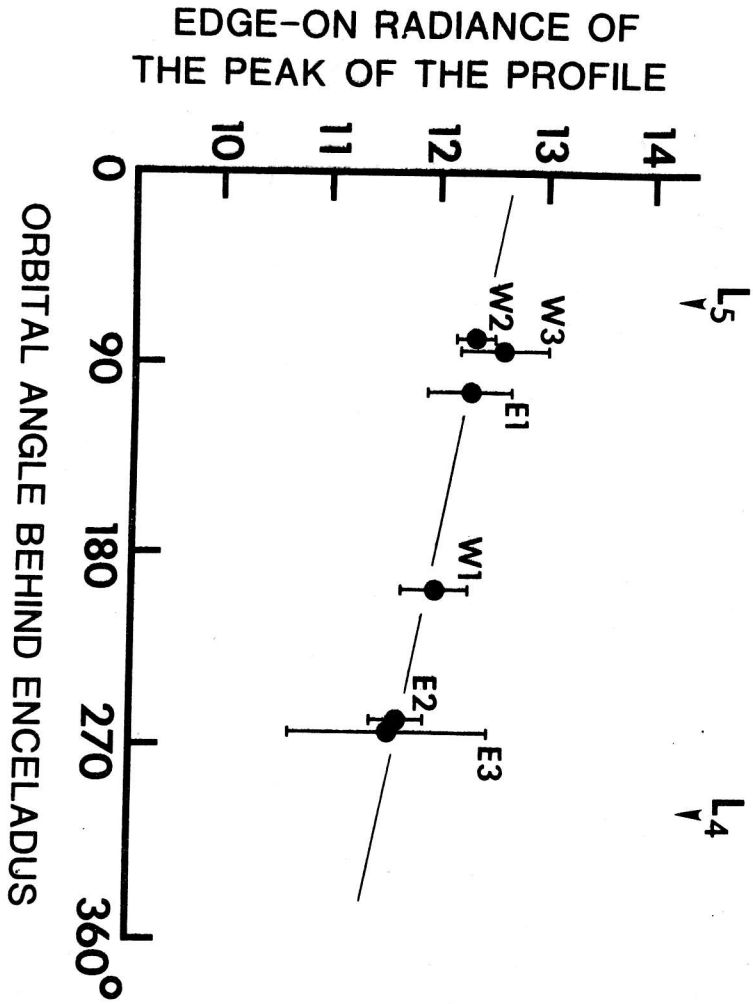

EDGE-ON RADIANCE OF
THE PEAK OF THE PROFILE
10
11
12
13
14
0
90
180
270
360°
ORBITAL ANGLE BEHIND ENCELADUS
W2
W3
E1
W1
E2
E3
L5
L4

Figure 6

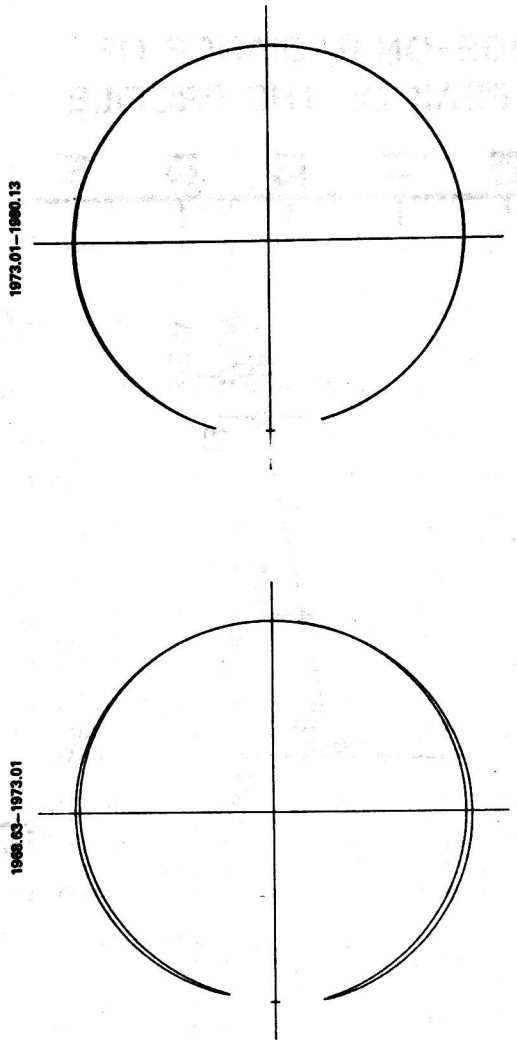


Figure 7

TWICE AVERAGED RESTRICTED THREE-BODY PROBLEM AND ITS APPLICATION TO STUDY THE EVOLUTION OF ORBITS OF ASTEROIDS

F. Veres
Observatory of the Hungarian Academy of Sciences
Baja, Hungary

Abstract

The paper deals with the twice averaged restricted three body problem. First we show a general analitical solution for the planar case and then we give a particular solution for the spatial case. The results of the analitical solution are compared with results obtained by numerical integration.

Introduction

We suppose, that a massive central object and a massive perturbing body revolve around the baricentre on elliptic orbits. Around the central object nearer to it then the perturbing body revolves the particle of negligibly small mass on a slightly perturbed elliptic orbit. Because of the slightness of the perturbations, the disturbances of second order are negligible compared to first order ones. We study the evolution of the orbit of the particle. For this we have to take into account the secular perturbations of first order. To get these we have to average the disturbing function by the mean anomalies of the perturbing body and the particle. For averaging we use the Gauss-scheme, i. e. the two mean anomalies are varied

independently. This method causes the loose of resonant terms. Since in the case of resonance the resonant terms can have a great effect on the evolution of orbit, our method can be applied only for the study of non-resonant motion. First we treat that case, when the particle and the perturbing body revolve in the same plane. In this case after averaging the disturbing function has the following form:

$$R = \frac{f m'}{a'} \sum_{l=0}^{\infty} S_l \cos l\omega$$

$$S_l = \sum_{p=0}^{\infty} \left(\frac{a}{a'}\right)^{l+2p+2} \frac{M^{(l)}_{l+2p+1}(e')}{(1-e'^2)^{l+2p+\frac{3}{2}}} K^{(l)}_{l+2p+4}(e) P_{lp}$$

where f is the constant of gravity; m' is the mass of perturbing body; ω is the longitude of the pericentre of the orbit of particle, measured from the pericentre of the orbit of disturbing body; a', e', a, e are the semi-major axes and eccentricities of the perturbing body and the particle respectively. The $M^{(l)}_n$ and $K^{(l)}_n$ functions are polynomials where the smallest exponent equals to l ; P_{lp} are numerical coefficients.

Since R does not depend on the mean anomaly of the particle and does not depend explicitly on time, the semi-major axis of the orbit a and R itself are prime integrals. Consequently the equations of motion can be integrated in quadratures, i. e. the dependence of e and ω on time can be determined by analitical methods. R is an implicit function of e and ω , so it makes the study of integral curves possible, and by this the classification of the types of motion.

Results

According to our results there exist three qualitativly different types of motion: the libration, the circulation and finite motion. In the first case ω librates around $\omega = 0$, while e varies periodically. The maximum amplitude of the libration is $\pi/2$, the amplitude of variations in e has the order of e'. In the case of circulation ω increases secularly, and e varies periodically. e has its greatest value at $\omega = 0$, and the smallest value at $\omega = \pi$. In the case of finite motion ω increases secularly and e increases up to 1. This means the orbit to become so elliptic, that the particle falls into the central body.

In the course of further computations we suppose the eccentricity of the orbit of perturbing body to be small. Supposing the central body to be the Sun, the perturbing body to be the Jupiter and the particle to be an asteroid this condition is satisfied. Final formulae were determined by an accuracy of e'^2.

In the Sun-Jupiter-asteroid systems the circulation occurs most frequently. The formulae for this case are as follows:

$$\omega = N(t-t_0) + d_1 \sin N(t-t_0) + d_2 \sin 2N(t-t_0)$$

$$e = e_0 + e_1 \cos N(t-t_0) + e_2 \cos 2N(t-t_0)$$

The N, d_1, d_2, e_0, e_1, e_2 constants can be determined from initial conditions. Time t is measured in Jovian years. Applying the obtained formulae for asteroid Vicia (No.1097) we get the following values:

N = 0.16825 degree/Jovian period

$d_1 = -8^{\circ}533$

$d_2 = 0^{\circ}531$

$e_o = 0.28519$

$e_1 = 0.03298$

$e_2 = -0.00157$

Simultaneously the equations of restricted three-body problem were integrated numerically with the input data of Vicia for an interval of 1950 Jovian years. The agreement proved to be good, except on value of N, for which the analitical theory gave a result 6 % less. Further studies proved the difference to be a consequence of neglection of secular terms of second order. Taking into account the latter ones, the difference decreases down to 2 %.

Further on we discussed the spatial case. We supposed the inclination of the orbit to be small and the case of circulation to be realised. We obtained the following results:

$$e = e_o + \Delta e + \delta_1 \cos N(t-t_1) + \delta_2 \cos 2N(t-t_1) - \frac{B}{A} i_c^2 \cos 2M(t-t_2)$$

$$\pi = (N+\beta_1)(t-t_1) + d_1 \sin N(t-t_1) + d_2 \sin 2N(t-t_1) + \frac{\beta_2}{2M} \sin 2M(t-t_2)$$

$$i^2 = i_c^2 \left[1 + \frac{B}{A} \cos 2M(t-t_2)\right]$$

$$\operatorname{tg} \omega = \frac{M}{A+B} \operatorname{tg} M(t-t_2)$$

Applying these formulae for asteroid Apollo (No.1862) the results are as follows:

$$\operatorname{tg}\omega = 0.6594\operatorname{tg}0.12033(t-594.78)$$

$$i^2 = 51.12\left[1+0.39390\cos 0.24076(t-594.78)\right]$$

$$\begin{aligned} e = {} & 0.54362+0.02149\cos 0.03967(t-1414.57)- \\ & -0.00062\cos 0.07934(t-1414.57)- \\ & -0.00393\cos 0.24076(t-594.78) \\ = {} & 0.03967(t-1414.57)-4.17\sin 0.03967(t-1414.57)+ \\ & +0.11\sin 0.07934(t-1414.57)- \\ & -0.42\sin 0.24076(t-594.78) \end{aligned}$$

In the above formulae the angles are measured in degrees and time in Jovian years from the epoch of 26th Aug 1973. The numerical integration was also performed for Apollo for an interval of 4000 Jovian years. According to the obtained results the error of all parameters in the above formulae is less than 2 %.

Acknowledgement

The author wishes to express his gratitude to prof. E. P. Aksonov of Moscow State University for giving significant help and for useful discussions and comments.

References

1. Aksenov, Russian Astronomical Journal, Vol. 56, No. 2, pp. 419-426. 1979.

2. Aksenov, R.A.J., Vol. 56, No. 3, pp. 623-631, 1979.

3. Veres, R.A.J., Vol. 57, No. 1, pp. 182-189, 1980.

4. Veres, R.A.J., Vol. 57, No. 5, pp. 1070-1077, 1980.

5. Veres, R.A.J., Vol. 57, No. 4, pp. 824-832, 1980.

NONSINGULAR OSCILLATOR ELEMENTS

W. T. Kyner
Department of Mathematics and Statistics
The University of New Mexico
Albuquerque, New Mexico 87131

1. Introduction.

The construction of transformations to improve the numerical integration of the perturbed Kepler problem,

$$\ddot{\underline{r}} + \frac{\mu}{r^3}\underline{r} = \underline{P} \tag{1}$$

can be motivated by many reasons such as the need to solve a particular problem, an interest in a mathematical concept such as stability or regularization or, as in this paper, curiosity about the numerical advantages of transformations that were used in studies of the qualitative properties of solutions.

The coordinates of the phase space presented in this paper are called oscillator variables and oscillator elements. They are defined for nonrectilinear orbits and are nonsingular for small eccentricities and inclinations. The true longitude is used as in independent variable. If a time element (see below) is added to the system, then the solutions of the transformed equations of the main problem of artifical satellite theory (J_2 perturbation only) are weakly stable on manifolds of constant energy. Numerical tests indicate that these variables are well suited for high precision computations of bounded orbits subject to general perturbations.

2. The Oscillator Variables.

The oscillator variables are based on the orbital plane geometry in which the motion in the orbital plane is partially decoupled from the rotation of the orbital plane. Since most of the non-Hamiltonian theories of celestial mechanics

employ this geometry, no claim of originality is being made here.

As usual, the orbial plane is defined as the plane through the center of mass of the attracting body whose normal is the angular momentum vector (per unit mass),

$$\underline{c} = \underline{r} \times \dot{\underline{r}} \ .$$

Clearly, the nonvanishing of $\underline{c}$ is essential in this development, i.e., only nonrectilinear orbits are allowed.

A rotating frame is defined by

$$\underline{U} = \underline{r}/r \ , \ W = \underline{c}/c \ , \ V = \underline{W} \times \underline{U} \ .$$

Hence

$$\underline{r} = r\underline{U} \ , \ \dot{\underline{r}} = \dot{r}\underline{U} + (c/r)\underline{V} \ , \ \underline{c} = c\underline{W} \ .$$

The position of the rotating frame relative to an inertial frame is determined by three Eulerian angles

i = the inclination of the orbital plane,

Ω = the longitude of the ascending node,

ψ = the argument of latitude.

Any vector $\underline{G}$ can be represented relative to the two frames by

$$\underline{G} = G_1\underline{i} + G_2\underline{j} + G_3\underline{k} = G_1'\underline{U} + G_2'\underline{V} + G_3'\underline{W} \ .$$

The components relative to the two frames are related by a rotation matrix B ,

$$\begin{pmatrix} G_1 \\ G_2 \\ G_3 \end{pmatrix} = B \begin{pmatrix} G_1' \\ G_2' \\ G_3' \end{pmatrix}$$

where $B = (\underline{U},\underline{V},\underline{W})$. Its components are

$$
\begin{aligned}
B_{11} &= \cos w + (\sin i \sin\psi)(\sin i \sin\Omega)/(1 + \cos i) \\
B_{21} &= \sin w - (\sin i \sin\psi)(\sin i \cos\Omega)/(1 + \cos i) \\
B_{31} &= \sin i \sin \psi \\
B_{12} &= -\sin w + (\sin i \cos\psi)(\sin i \sin\Omega)/(1 + \cos i) \\
B_{22} &= \cos w - (\sin i \cos\psi)(\sin i \cos\Omega)/(1 + \cos i) \\
B_{32} &= \sin i \cos \psi \\
B_{13} &= \sin i \sin \Omega \\
B_{23} &= -\sin i \cos \Omega \\
B_{33} &= \cos i
\end{aligned}
\qquad (2)
$$

with $w = \psi + \Omega$, the true longitude. The matrix B is the inverse of the matrix A given in Goldstein [1] and its components have been rewritten for computational efficiency. This representation is singular for retrograde polar orbits.

By means of standard techniques of rigid body mechanics, the equations of motion can be written

$$
\begin{aligned}
\frac{d^2\underline{r}}{dt^2} &= (c^2/r^3) - (\mu/r^2) + P_1' \\
\frac{dc}{dt} &= rP_3' \ , \ \frac{di}{dt} = (r/c)P_3' \cos \psi \\
\frac{d\Omega}{dt} &= (r \sin\psi/c \sin i)P_3' \\
\frac{d\psi}{dt} &= (c/r^2) - (r/c)P_3' \sin \psi \cot i
\end{aligned}
\qquad (3)
$$

The orbital plane variables $(r,\dot{r},c,i,\Omega,\psi)$ determine the state of the system. If they are known, then $\underline{r}$ and $\dot{\underline{r}}$ can be computed with the aid of the rotation matrix B.

If the perturbing force is zero, then it is easy to derive the linear oscillator equations

$$\frac{d^2}{d^2}(1/r) + (1/r) = (\mu/c^2)$$

$$\frac{d^2}{d\psi^2}(\sin i \sin\psi) + (\sin i \sin\psi) = 0 . \tag{4}$$

These equations motivated the introduction of the oscillator variables that were used by the author [2] to establish the existence of invariant manifolds for the oblate planet problem. Define

$$z_1 = (1/r) - (\mu/c^2) \ , \ z_2 = -(\dot{r}/c)$$

$$z_3 = \sin i \sin \psi \ , \ z_4 = \sin i \cos \psi \ .$$

Then if $\underline{P} = 0$, the equations

$$\frac{d}{d\psi}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} , \frac{d}{d\psi}\begin{pmatrix} z_3 \\ z_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} z_3 \\ z_4 \end{pmatrix} ,$$

are equivalent to (4).

Let

$z_5 = c \cos i$, the polar component of angular momentum,

$z_6 = w = \psi + \Omega$, the true longitude.

Then the oscillator variables $(z_1,z_2,z_3,z_4,z_5,z_6)$ are equivalent to the orbital plane variables $(r,\dot{r},c,i,\Omega,\psi)$ and are defined for all nonrectilinear orbits

with the exception of polar retrograde orbits. Since

$$\frac{dw}{dt} = \frac{d\psi}{dt} + \frac{d\Omega}{dt} = (c/r^2) + (r/c)P_3' \sin\psi \sin i(1 + \cos i)^{-1} \qquad (5)$$

these orbits are excluded. The singlarity $(1 + \cos i)^{-1}$ also occurs in the rotation matrix B. It can be suppressed in several ways, e.g., if Ω is replaced by $-\Omega$, it becomes $(1 - \cos i)^{-1}$.

3. The Oscillator Elements.

In this section five oscillator elements are defined. They are constant if the perturbation is zero and are nonsingular for all but rectilinear orbits. The first two, x_1 and x_2, are defined by the equations

$$\begin{aligned} z_1 &= (1/r) - (\mu/c^2) = (e/p)\cos f = (e/p)\cos(w-\tilde{\omega}) \\ &= (e/p)(\cos\tilde{\omega}\cos w + \sin\tilde{\omega}\sin w) \\ &= x_1 \cos w + x_2 \sin w , \end{aligned}$$

$$\begin{aligned} z_2 &= (-\dot{r}/c) = (-e/p)\sin f = (-e/p)\sin(w-\tilde{\omega}) \\ &= (e/p)(-\cos\tilde{\omega}\sin w + \sin\tilde{\omega}\cos w) \\ &= -x_1 \sin w + x_2 \cos w \end{aligned}$$

where f denotes the true anomaly, $\tilde{\omega}$ the longitude of perigee, e the eccentricity, p the parameter of the conic, $w = f + \tilde{\omega}$, and

$$x_1 = (e/p)\cos\tilde{\omega} , \quad x_2 = (e/p)\sin\tilde{\omega} .$$

The variables x_3 and x_4 could be defined by similar formulas in terms of z_3 and z_4 but it is more convenient to set

$$x_3 = c \sin i \sin \Omega$$

$$x_4 = -c \sin i \cos \Omega$$

$$x_5 = c \cos i \ ,$$

the components of $\underline{c}$, the angular momentum vector. Then

$$z_3 = \sin i \sin \psi = \sin i \sin(w-\Omega)$$

$$= (-1/c)(x_3 \cos w + x_4 \sin w)$$

$$z_4 = \sin i \cos \psi = \sin i \cos(w-\Omega)$$

$$= (-1/c)(-x_3 \sin w + x_4 \cos w) \ .$$

Let

$$\Phi(w) = \begin{pmatrix} \cos w & -\sin w \\ \sin w & \cos w \end{pmatrix} ,$$

then

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \Phi(w) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} , \quad \begin{pmatrix} z_3 \\ z_4 \end{pmatrix} = (-1/c)\Phi(w) \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} .$$

and

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Phi(-w) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} , \quad \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = -c\Phi(-w) \begin{pmatrix} z_3 \\ z_4 \end{pmatrix} .$$

The oscillator elements (x_1,x_2,x_3,x_4,x_5) are well defined for all nonrectilinear orbits with a modification required for polar retrograde orbits.

4. The Time Element.

The set (x_1,x_2,x_3,x_4,x_5,w) has five slow variables and one fast variable. The instability of solutions to the Kepler problem is well displayed in this coordinate system since a small change in the initial value of w can result in a large change in position and velocity. However, the solutions on manifolds of constant energy are stable in coordinate systems that employ a time element (a slow variable) as a sixth variable. We shall use a time element that has no first order (in J_2) secular terms for the elliptic orbits of the main problem of artifical satellite theory.

Let us first consider

$$Q = M - \bar{n}t$$

where M is the mean anomaly and

$$\bar{n} = |2(\text{total energy}|^{3/2}/\mu$$

is the Cunningham mean motion [4]. For the Kepler problem, $Q = M_0$, the mean anomaly at epoch. For the main problem, Q is constant (to first order in J_2) at successive perigees, hence, it can be considered to be the oscillatory part of M (for more details see [3]). Therefore, on manifolds of constant energy, the solutions to the main problem are weakly stable.

The variable Q is not well defined for nearly circular orbits so we shall modify it slightly. Let

$$x_6 = M + \tilde{\omega} - \bar{n}t = w - \bar{n}t + F \qquad (5)$$

where

$$F(x_1,x_2,x_3,x_4,x_5,w) = M - f = -2 \arctan \frac{\beta(e) \sin f}{1 + \beta(e) \cos f} - e(1-e^2)^{1/2} \frac{e \sin f}{1 + e \cos f},$$

$$\beta(e) = \frac{(1+e)^{1/2} - (1-e)^{1/2}}{(1+e)^{1/2} + (1-e)^{1/2}} .$$

The variable x_6 is well defined for nonrectilinear orbits with $0 \le e < 1$.

If t is the independent variable and $x_1, x_2, \ldots, x_6$ are known then w can be found by iteration (if the perturbation is zero, (5) is equivalent to Kepler's equation). To avoid this w is taken as the independent variable and

$$t = (x_6 - w - F)/\bar{n}$$

is computed as needed. The position and velocity at prescribed times can be found by interpolation. The equations of motion in terms of these variables is given in an appendix.

5. Conclusion.

The oscillator elements

$$x_1 = (e/p) \cos \tilde{\omega}$$
$$x_2 = (e/p) \sin \tilde{\omega}$$
$$x_3 = c \sin i \sin \Omega$$
$$x_4 = -c \sin i \cos \Omega$$
$$x_5 = c \cos i$$
$$x_6 = w - \bar{n}t + F$$

are well defined for nonrectilinear orbits with $0 \le e < 1$. Since the rotation matrix B is an algebraic function of the x's, $\cos w$ and $\sin w$, it is easy to compute. Hence, perturbations given relative to the Earth fixed frame are efficiently represented relative to the moving frame. The time variable effectively supresses instabilities and permits highly accurate computations over many revolutions. The element set was tested numerically with

perturbations that included the effects of atmospheric drag, Sun-Moon perturbations and with realistic Earth potentials. They were found suitable for those near-Earth orbits that must be computed with high precision.

Appendix: The Equations of Motion.

From

$$\frac{dw}{dt} = (c/r^2) + D$$

$$\frac{dc}{dt} = rP_3'$$

$$\frac{d^2r}{dt^2} = (c^2/r^3) - (\mu/r^2) + P_1'$$

where $D = (r/c)P_3' \sin\psi \sin i \,(1 + \cos i)^{-1}$, one obtains

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = J\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}$$

where

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad S_1 = (2\mu r/c^2)P_3'$$
$$S_2 = -rz_2P_3' - P_1'/c \,.$$

set

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \Phi(w)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Phi(w) = \begin{bmatrix} \cos w & \sin w \\ -\sin w & \cos w \end{bmatrix}.$$

Then

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -D\Phi(-w)J\Phi(w)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \Phi(-w)\begin{bmatrix} S_1 \\ S_2 \end{bmatrix} .$$

From

$$\underline{c} = x_3\underline{i} + x_4\underline{j} + x_5\underline{k}$$

and

$$\frac{d\underline{c}}{dt} = \underline{r} \times \underline{P} = r\underline{U} \times (P'_1\underline{U} + P'_2\underline{V} + P'_3\underline{W}) = r(P'_2\underline{W} - P'_3\underline{V})$$

one obtains

$$\frac{dx_3}{dt} = r(P'_2B_{13} - P'_3B_{12})$$

$$\frac{dx_4}{dt} = r(P'_2B_{23} - P'_3B_{22})$$

$$\frac{dx_5}{dt} = r(P'_2B_{33} - P'_3B_{32}) .$$

There seems to be no easy way to derive

$$\frac{dx_6}{dt} = (n-\bar{n}) + P'_1\left[\frac{(1-e^2)^{1/2} - 1}{e}\right]\frac{(1-e^2)^{1/2}}{na}\cos f$$

$$- \frac{2r}{na^2}P'_1 - P'_2\left[\frac{(1-e^2)^{1/2} - 1}{e}\right]\frac{(1-e^2)^{1/2}}{na}\left(1 + \frac{r}{a(1-e^2)}\right)\sin f$$

$$+ \frac{r \sin i \sin \psi}{a^2 n}\left(\frac{1}{1 + \cos i}\right)P'_3 - t\frac{d\bar{n}}{dt}$$

where instantaneous elements are employed. The factor

$$\frac{(1-e^2)^{1/2} - 1}{e} = \left[\frac{(1-e^2)^{1/2}-1}{e}\right]\left[\frac{(1-e^2)^{1/2}+1}{(1-e^2)^{1/2}+1}\right] = \frac{-e}{(1-e^2)^{1/2} + 1}$$

is well defined for nearly circular orbits. There are no small inclination singularities.

References.

1. H. Goldstein, Classical Mechanics, Addison-Wesley, Reading, Mass., 1980.

2. W. T. Kyner, Qualitative Properties of Orbits about an Oblate Planet, Communications of Pure and Applied Mathematics, Vol. 17, pp. 227-236 (1964).

3. W. T. Kyner, A New Application of the Cunningham Integral to the Oblate Earth Problem, In Differential Equations and Dynamical Systems, pp. 35-42, Academic Press, New York, 1967.

4. T. E. Sterne, An Introduction to Celestial Mechanics, Interscience Publishers, Inc. New York, 1960.

A COMPARATIVE STUDY OF THE EARTH'S MAGNETIC AND GRAVITY FIELDS

G. Barta

Eötvös University, Geophysical Department, Budapest
Hungary

Abstract

The article contains a comprehensive summary of the results of investigations into the connections between magnetic and gravity fields of the Earth. It was shown more than 30 years ago by means of studies concerning long period variations of the permanent magnetic field of the Earth that to the secular variation a wave of about 50 years period has been superimposed. The origin of the phenomenon should be found in the eccentricity of the inner core of the Earth being similar to the eccentricity showing itself in the magnetic dipole.

The more accurate expansions into series of the geoidal figure - carried out since the appearence of satellite observations - made it possible to study the problem from the gravity side too. Some ten years ago it was shown that the 6 big anomalies determining the geoid can be written as sums of two influences and the position of the two sources can be deduced from the characteristical points of the magnetic eccentricity. These similarities of the magnetic and gravity field of the Earth raise - of course - a lof of new problems /e.g. due to the shifting of eccentric dipole the gravity acceleration observable on the Earth must show changes too/. Thus the study of connections between the two fields of force can not be regarded as being closed down in spite of the many similarities detected as yet.

...

The original aim of my investigations taken up more than 30 years ago on the magnetic secular variation was to make up a uniform magnetic observation series which, in the Carpathian basin, had begun in 1871, but had been interrupted several times for technical and political reasons. During the investigations it turned out that there is a periodic pulsation of about half a century in the magnetic annual mean value series measured in the observatories of the Earth's northern temperature zone, considering either the declination or the vertical and horizontal intensities. This recognition pointed forward beyond the original goal. Several characteristics of global validity have been found in the secular variation recorded in the magnetic observatories. The phenomena observed in the magnetic components can be more closely related to the physical reality if we use a vector diagram representation. It may be proved with the aid of the vector diagrams that the half-a-century period in the components comes from a helicoid-like spatial motion of the end point of the magnetic vector. The sense of rotation of the helicoid is clockwise - seen from the direction of the main secular variation - and the rotation period is about fifty years. The component of the secular variation which is perpendicular to the direction of the main variation is called "transversal effect".

The transversal effect takes place in the normal plane of the variation. The actually measured magnetic point now goes ahead, then falls behind the averaged point. This is called the "longitudinal effect" of the magnetic secular variation. The reason for this is that the measured variation is occasionally greater at other times less than the average one. Its period is also half a century. It is interesting that either the speeding up or the slowing down of the secular variation takes place at the same time all over the world.

It is obvious then that there is a secondary effect of a period of about

fifty years superposed on the magnetic secular variation. This effect is characterized by features of global validity. Similarly, a half-a-century period may be found in the amplitude of the Chandler period and in the angular velocity of the Earth's rotation, that is in some phenomena closely related to the rotation of our planet. From the similarity of the periods of secondary phenomena we may conclude that the magnetic secular variation is connected with significant mass motions. The internal mass distribution of the Earth influences the Earth's figure, in other words the sea level, which represents the surface. Examining the mean sea level data of several coastal stations certain relations suggesting a global character have been found a period of about 50 years is recognizable in these data, too, but in general this phenomenon depends so much on local and perhaps on regional effects that only very few stations are suitable to obtain data from which conclusions of general validity can be drawn.

The magnetic secular variation has a pulsation of a period of about 50 years - as a component. If in a component part of a variation one recognizes the common general features outlined above, then the main variation must well be of general nature. However, the so-called isoporic charts, which show the change of the magnetic field, give a rather incoherent picture of the variation. The reason for this is that when an isoporic chart is compiled the changes of each of the components are represented separately on different sheets of map, while the phenomenon itself takes place not in separate components but in such a way that the end-point of the magnetic vector recorded in an observatory traces a spatial curve. If the properties of this path are being studied in separate components, the particular local coordinate system of the observatory will be an inherent part of the investigations. Obviously in such a treatment the vertical of European or Asiatic or American observatories will mean different di-

rections. In order to properly interprete the phenomenon we must choose a standard coordinate system.

About 25 years ago we investigated the vectors of secular variation of the geomagnetic field in a coordinate system fixed to the Earth's centre. First the variation vectors were orthogonally projected to the equatorial plane and to two other planes perpendicular to the equator, with the aim of searching for some general feature. After several attempts it was found that if an axis of the coordinate system is chosen to point from the Earth's centre towards Pakistan as a projection centre, then in this projection the vectors of secular variation converge near the projection centre. Thus the symmetry center of the secular change is somewhere near to Pakistan. It is to be noted that the magnetic dipole has long been known to be located eccentrically towards the Marshall islands. This direction is at right angles to the coordinate axis connecting the Earth's centre and Pakistan, and again, both of these directions are perpendicular to the magnetic dipole axis. These three directions seem to have an important role in the description of the magnetic field and its secular change. Indeed, the vectors of secular variation - seen from the directions of magnetic eccentricity and of the dipole axis - point out certain characteristic directions.

It has also been known long since that the Earth's magnetic dipole - now being eccentric towards the Marshall islands - has been drifting westwards at a velocity of $0,2^{\circ}$/year. If - according to what is said above - beside the magnetic secular variation the eccentricity of the magnetic dipole is also assumed to originate from some kind of mass asymmetry in the "background", then this must be apparent in the Earth's figure, too. When these investigations began - about 20 years ago - we had only certain forms of hypotheses about the triaxiality of the Earth and a general belief in geodesy was that the equatorial major axis of the Earth

pointed towards Australia. Although this direction coincided with the direction of eccentricity of the magnetic dipole, this conception was not more than a guess owing to the inaccuracy of primary measurement data.

The first reliable geoid heights were computed from the perturbations of satellite orbits 15 years ago. At the time we made an attempt at calculating the equatorial ellipticity from the equatorial section of the 1966 geoid approximating it with a zonal spherical harmonic the axis of which was directed towards Australia. The calculations gave the surprising result that a spherical harmonic of this type could not fit well the equatorial section. Varying the axis of approximation we determined the direction of best-fitting ellipse and, subtracting it from the measured data, obtained a characteristic antisymmetry in the residual map. Thus we came to the conclusion that the equatorial section could be approximated not by one, but two zonal harmonics with axes nearly at right angles to each other. This approximation is really very good. Assuming that the two reflection-symmetric figures obtained in the approximation were rotation-symmetric, we drew a map of the combined body. Surprisingly this combination reproduced the six well-known anomalies of geoid, that is the origin of the six large geoid anomalies can not be six separate, independent mass inhomogeneities, but the geoid is the sum of two great effects.

The pairs of anomalies in the northern and southern temperate zones are the antipodal superpositions of the two main effects. The axes of the best-fitting component figures are not far away from those mentioned above /Australia, Pakistan/.

It is very demonstrative that this geoid map has been computed solely from the equatorial geoid heights, in other words the anomaly pattern of temperate zones is implied in the equatorial data. If a separate density inhomogeneity is attributed to every geoid anomaly, these inhomogeneities

should satisfy this very peculiar condition the probability of which is extremely small.

It is also worth mentioning that according to a schematic calculation, if the inner core is eccentric and of high density, the level surface becomes egg-shaped, peaked towards the direction of eccentricity. In our case the zonal harmonic approximation gives an egg-shaped surface with its axis and peak pointing towards Australia. It is interesting that the rotation-symmetric component the axis of which is directed towards India contains no even spherical harmonics, so it has no ellipticity. The rotation symmetry is represented by harmonics of order 3, 5 and 7. In this stage of investigation, since we use only the equatorial data, the lines connecting the oceanic anomaly pairs of temperate zones are perpendicular to the plane of equator. In reality these lines are not perpendicular to the equator, instead they are inclined to each other northward. This suggests that the mass inhomogeneity lying behind the geoidal figure is located north of the equatorial plane. And really, the magnetic dipole is shifted north of this plane.

In the next stage of research - to explain this distortion of anomalies - we had to abandon the equator as a plane of approximation and search for a better fit. This generalization of computations made the formulae extremely complicated. The best approximation was difficult to determine because the surficial "inhomogeneities" of topography have also an influence on the geoidal figure, so not only the two rotation-symmetric effect but the irregular surficial mass inhomogeneities are also represented in the geoid. Setting out from the consideration that the influence of the surficial source bodies diminishes with the altitude more rapidly than that of the global sources, moving away from the Earth's surface we can separate the different types of anomalies. To achieve this separation we have calculated the geoid for 1000, 2000. 3000, 6000, 10000, 20000 and 50000 km altitudes with the

well-known method. The pictures at 20000 and 50000 km heights are rather schematic and more than the equatorial ellipticity can not be observed. Between 5000 and 10000 km altitudes the effect of the Indian source is clearly discernible although the maps are still schematic. Their schematic pattern means that the effect of surficial disturbances is negligibly small at this height. Therefore the data of the map corresponding to 6000 km altitude were chosen for the purposes of further research. During the calculations a mathematical method was found which determined the axial directions of the two approximating rotation-symmetric figures quite uniquely. The calculations can be carried out in two different ways: either from the data along the main circle given by the two axes or from the complete map of data measured all over the world. If the source body is a composition of rotation-symmetric figures, the two definition will give the same set of spherical harmonic coefficients when and only when the two axes of the approximation are properly chosen. In this way we determined the best fit directions excluding the effect of surficial inhomogeneities. Then the geoid so obtained was extrapolated back to the Earth's surface. Finally the result - which was supposed to reflect the effect of deep seated sources - was subtracted from the measured geoid surface, thus giving the effect of the surficial sources.

The residual map reveals a correlation between the anomaly pattern and the relief of topography. The rows of positive anomalies of the residual picture usually coincide with the big mountain ranges of the Earth and similarly the system of 0 lines with the mid-oceanic ridges. Furthermore the zonal asymmetry of geoid /its pear-shaped figure/ can be attributed only to the deep-seated sources. The surficial sources have no zonal harmonic components. The polar ellipticity of the Earth may also be imagined as the sum of two parts corresponding to our conception. The major term is due to the Earth's rotation, and about 0.2 percent is attributed to the internal mass

inhomogeneities.

The connection of magnetic and geoidal anomalies naturally calls forth the conclusion that just like the magnetic field the gravity field also undergoes a secular change. The degree of this change is very small, at most several times 10 μ gal a year.

THE NEXT DEVELOPMENT IN SATELLITE DETERMINATION OF THE EARTH'S GRAVITY FIELD

Professor W. M. Kaula
Department of Earth and Space Sciences
University of California, Los Angeles, California 90024

Abstract

Satellite-to-satellite tracking is expected to achieve $\pm 10^{-6}$ m/sec accuracy between co-orbiting satellites a few 100 km apart. The minimum sustainable altitude is about 160 km. It is estimated that features of the gravity field $\sim$110 km in extent should be resolvable. To achieve this resolution economically, the analysis process should be iterative, treating the residuals with respect to the best previous model as a time series for each orbit of limited duration (e.g., one revolution), and determining the mean correction to the potential for each duration by adjustments at trajectory crossings.

Introduction

The GEOS-3 altimeter satellite has yielded an order-of-magnitude improvement in knowledge of the Earth's gravitational field over the oceans: after adjustment of orbits for agreement at track crossings, the accuracy is better than ± 1 meter. A more accurate altimeter, SEASAT, was placed in orbit, but failed after about one month. Global solutions based on satellite altimetry have been published in the form of maps of the ocean geoid (Brace, 1977; Rapp, 1979) and as spherical harmonic coefficients of potential (Gaposchkin, 1980; Lerch et al., 1981).

However, to obtain the gravity field over the continents, as well as to obtain the geoid distinct from the mean sea level (as desirable for ocean dynamics studies), it is necessary to obtain more accurate measurements of satellite orbit accelerations. The method most likely to achieve such accuracy is satellite-to-satellite range-rate, for which μ/sec = 10^{-6}m/sec accuracy at ranges of a few 100 km appears attainable (Pisacane et al., 1981). The system proposed by the Applied Physics Laboratory, Johns Hopkins University, will transmit at two frequencies, 94.0 and 40.5 GHz, to overcome ionospheric refraction. It is expected that the stable frequency standard will have better than ± 0.2 μ/sec accuracy for four-second averages, and that the main error will arise from noise in the circuits translating the stable reference to the transmission frequencies. The minimum altitude at which the spacecraft are sustainable long enough for a global coverage is 160 km. At this altitude, variations in acceleration due to atmosphere drag can be ± 10 μ/sec^2. Hence it is essential that

the satellite be made drag-free. It is planned to use the Disturbance Compensation System (DISCOS) of Stanford University. In this system, each spacecraft has a proof mass within an evacuated cavity. Capacitors detect motion of the spacecraft with respect to the shielded proof mass, and thrust is applied to maintain a purely gravitational orbit.

The gravity satellite (GRAVSAT) spacecraft is now expected to be the same as for the next magnetic field survey (MAGSAT). The combination will not only be more economic, but will improve the accuracy of the magnetic data and produce coordinated data sets. It is hoped that a launch can be made before the 1988 increase in solar activity.

Resolution

The effective "resolution" of the satellite-to-satellite range rate--the size of the features in the gravity field it can measure--depends not only on the overall accuracy of $\pm$ 1 μ/sec, but also on the spectral distribution of this total error, and on the magnitudes of the spectral components of the gravity field compared to these components of the error. In the 4-second averaging time of the system, the spacecraft traverses 31.2 km, equivalent to a maximum harmonic $\ell_x = 657$. If the error spectrum is "white"--assumed to be equally distributed among all frequencies--the error per coefficient is $\pm 10^{-6}/(2\times657)^{1/2} = \pm 3\times10^{-8}$m/sec in velocity, or (using the 4 second interval) $\pm 7\times10^{-9}$m/sec^2 in acceleration at spacecraft altitude. At the surface, the error will be $[(6371+160)/6371]^{\ell}$ times as great: $\pm 1.025^{\ell}\times7\times10^{-9}$m/sec^2 in acceleration or $\pm(1.025^{\ell}/\ell)\times7\times10^{-10}$ in dimensionless potential coefficient. The magnitude of an ℓth degree coefficient is about $\pm 10^{-5}/\ell^2$. Equating these two expressions gives $\ell \approx 180$, equivalent to a resolution of 110 km.

Various other estimates have been made (Breakwell, 1979; Douglas et al., 1980; Pisacane & Yionoulis, 1980; Jekeli & Rapp, 1980; Rummel, 1980), most of them for the accuracy of determination of mean values for squares, and there has been some debate about how the different estimates should be compared. However, the dominance of the 1.025^{ℓ} damping factor leads to all the estimates falling within the range 100-200 km.

It therefore is the consensus that satellite-to-satellite range-rate will undoubtedly yield a significant improvement in gravity field determination, and that a greater task is to develop data analysis procedures which will economically and reliably extract the information from the data. Improved estimates of accuracy, or resolution, should be by-products of these procedures.

Data Analysis

The following appear to informed consenses:

1. the reference surface for which the global potential is to be calculated should be approximately at satellite orbit altitude: to avoid the expense of integrating over a wide surface area to get accelerations at the spacecraft, and to separate clearly the problem of potential determination from that of downward continuation;

2. the determination from satellite-to-satellite range-rate of variations in potential should be done orbit-by-orbit, primarily to have better insight into what is happening and secondarily for economy; and

3. the data set for communication among investigators should be earth-fixed locations (r_1, ϕ_1, λ_1; r_2, ϕ_2) and all three forms of differences: velocity ($\delta v_t = \Delta \dot{r}$), anomalous potential ($\Delta T$), and acceleration ($\Delta \ddot{r}_d$), so that investigators do not have to differentiate numerically or duplicate Fourier analyses.

Consensus 2, improvement of the gravity field using orbit-by-orbit analyses, is a considerable departure from classical methods in satellite geodesy, where normal matries on the order of 1000x1000 may be solved. However, not only would such global solutions from satellite-to-satellite range-rate be inordinately expensive, but they are unnecessary and inappropriate. The satellite-to-satellite range-rate is most sensitive to wavelengths in the gravity field comparable to the satellite spacing, while it is insensitive to very long wavelengths and to tracking station locations. Hence the range-rates constitute a data set quite complementary to the ground station tracking data, and analyses of the range-rates are most effectively treated as analyses of time series which are residuals with respect to models of the gravity field determined by the classical techniques. Important to this mode of treatment, of course, is that the data constitute as continuous a time series as possible.

A logical sequence of steps which conforms to these principles follows.

1. Select a reference surface which is a close approximation to the complete set of orbits used for a global solution. The most evidence surface is a sphere. However, the perturbations arising from the zonal harmonics will cause any orbit to depart some kilometers from circularity. Hence an axisymmetric figure would be more appropriate, approximating:

$$r_o(\phi) = a + \Delta a(J_2, \phi)$$

$$= a + \frac{3}{2} J_2 \frac{a_e^2}{a} \cos 2\phi, \qquad (1)$$

where a is semi-major axis, J_2 is the oblateness factor (1.082×10^{-3}), a_e is the equatorial radius, and ϕ is latitude (Kaula, 1966: p. 40, with inclination $i = \pi/2$). The resulting oscillations are thus ± 10.5 km. Bender (1981) suggests further an offset of the center of the figure ~ 4 km southward, to accommodate odd zonal effects:

$$\Delta r(J_3) \simeq -a\, e(J_3) \cos M = -\frac{15}{8} a \left(\frac{a_e}{a}\right)^3 J_3 \frac{n}{\dot{\omega}} \sin\omega \cos M$$

$$= \frac{a_e}{4} \cdot \frac{J_3}{J_2} \{\sin\phi + \sin(2\omega - \phi)\} \qquad (2)$$

where M is mean anomaly, n is mean motion, ω is argument of perigee, and $J_3 \simeq -2.5 \times 10^{-6}$.

2. Select a reference potential field V_o which is the best currently available. This field is partly to calculate departures δr of the satellites from the reference surface. Hence it should include enough

terms to assure that

$$\frac{2g_o}{r_o}\,\varepsilon\{\delta r(V_o)\}\frac{\psi}{2} < \varepsilon\{\Delta\dot{r}\}/\delta t_D \qquad (3)$$

where ψ is the angular separation of the satellites, $\varepsilon(\Delta\dot{r})$ is the error in range rate between the satellites, and δt_D is the time interval between data points. (See step 6 below). Thus a separation of 300 km, or $\psi \simeq 0.05$, $\varepsilon(\Delta\dot{r}) \simeq 1\mu$/sec, and $\delta t_D \simeq 4$ secs require location error $\varepsilon\{\delta r\} < 3$ meters.

3. Determine orbits using the reference potential, other appropriate model parameters, and tracking data, for the purpose of locating the satellites at the data points selected for analysis, step 6 below. The primary criterion for the duration τ of the orbits is minimizing $\varepsilon\{\delta r\}$.

4. Fourier analyse the satellite-to-satellite range-rates $\Delta\dot{r}$ for each duration τ. The data intervals δt_I for the analysis should be determined mainly by instrumental considerations. The purpose of this analysis is to obtain reliable accelerations $\Delta\ddot{r}_d$, step 5 below. Hence it may be economic to break up the orbit duration τ into several overlapping blocks. With Fast Fourier Transform (FFT) techniques this step should not be a major element of expense.

5. Obtain satellite-satellite accelerations $\Delta\ddot{r}_d$ by differentiating the Fourier analysis of $\Delta\dot{r}$ with respect to time and transforming back to the temporal domain. The time intervals δt_d for calculation of accelerations should be determined mainly by what is thought to be resolvable from the satellite-to-satellite range rate at satellite altitudes: i.e., smaller than the 14 seconds it takes to traverse 110 km.

6. Set up (O-C)'s for accelerations. At each time t for which there is an "observed" acceleration $\Delta\ddot{r}_d(t)$ we also have:

a. the earth-fixed satellite locations $\underset{\sim}{r}_1(t)$, $\underset{\sim}{r}_2(t)$ from the orbits calculated at step 3;

b. the reference potentials $V_o(\underset{\sim}{r}_1)$, $V_o(\underset{\sim}{r}_2)$ and their gradients, prescribed at step 2.

Define the complete potential at $\underset{\sim}{r}_i$, i = 1 or 2, as:

$$\begin{aligned} V(\underset{\sim}{r}_i) &= V_o(\underset{\sim}{r}_i) + T(\underset{\sim}{r}_i) \\ &\simeq V_o(\underset{\sim}{r}_{oi}) + \left(\frac{\partial V_o}{\partial r}\right)_i \delta r_i + T(\underset{\sim}{r}_{oi}) \\ &\simeq V_o(\underset{\sim}{r}_{oi}) + \left(\frac{\partial V_o}{\partial r}\right)_i \{\delta r_i(V_o) + \delta r_i(T)\} + T(\underset{\sim}{r}_{oi}) \\ &\simeq V_{oo}(\underset{\sim}{r}_i) - g(r_o)\delta r_i(T) + T(\underset{\sim}{r}_{oi}) \end{aligned} \qquad (4)$$

Here, $V_{oo}(\underset{\sim}{r}_i)$ is the reference potential V_o at location $\underset{\sim}{r}_i$ calculated for time t from the orbit determined at step 3 plus the reference potential V_o; $T(\underset{\sim}{r}_{oi})$ is the "disturbed" potential on the reference sphere r_{oi}, negligibly different from that at r_i; and $\delta r_i(T)$ is the difference of the actual location of the satellite from that obtained using the reference potential V_o.

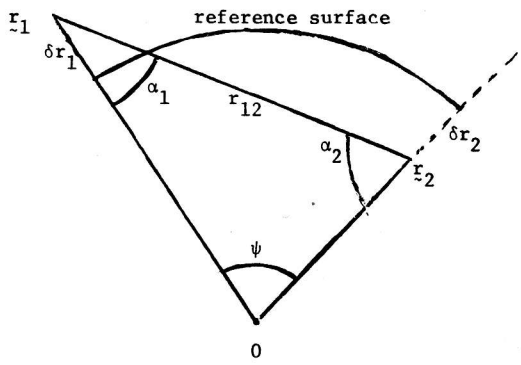
The "calculated" acceleration $\Delta\ddot{\underset{\sim}{r}}_c(t)$ is:

$$\Delta\ddot{\underset{\sim}{r}}_c(t) = \frac{(\ddot{\underset{\sim}{r}}_1 - \ddot{\underset{\sim}{r}}_2)\cdot(\underset{\sim}{r}_1 - \underset{\sim}{r}_2)}{|\underset{\sim}{r}_1 - \underset{\sim}{r}_2|} = (\nabla V_1 - \nabla V_2)\cdot(\underset{\sim}{r}_1 - \underset{\sim}{r}_2)/r_{12}, \qquad (5)$$

using astronomic/geodetic sign convention,

$$\ddot{\underset{\sim}{r}}_i = \nabla V_i. \qquad (6)$$

The figure is a sketch of the geometry, in which δr_i and ψ have been greatly exaggerated compared to what would exist in the actual system.



We thus can write in the plane defined by the Earth's center of mass 0 and the two satellites:

$$\Delta\ddot{r}_c(t) = \left[\frac{1}{r_1}\frac{\partial V_1}{\partial\psi}\sin\alpha_1 - \frac{1}{r_2}\frac{\partial V_2}{\partial\psi}\sin\alpha_2\right]$$

$$+ \frac{\partial V_1}{\partial r}\cos\alpha_1 + \frac{\partial V_2}{\partial r}\cos\alpha_2$$

$$= \Delta\ddot{r}_{co}(t) + \frac{\sin\alpha_{10}}{r_{10}}\frac{\partial T(r_o,\psi_1)}{\partial\psi} - \frac{\sin\alpha_{20}}{r_{20}}\cdot\frac{\partial T(r_o,\psi_2)}{\partial\psi}$$

$$+ \left[\frac{\psi}{2} - O(\psi^3)\right]\left[\frac{\partial T_1}{\partial r} + \frac{\partial T_2}{\partial r}\right] + D_2\{g\delta r_2(T)\} - D_1\{g\delta r_1(T)\} \qquad (7)$$

where

$$\Delta\ddot{r}_{co} = D_1(V_{01}) - D_2(V_{02}), \tag{8}$$

the operator

$$D_i = \frac{\sin\alpha_i}{r_i}\cdot\frac{\partial}{\partial\psi} + (-1)^i \cos\alpha_i \frac{\partial}{\partial r_i}, \tag{9}$$

and the subscripts on α_{io}, r_{io} denote dependence on the orbit calculation using the reference potential V_o at step 3. (Given $\underline{r}_1$, $\underline{r}_2$, determination of the orientation of the 0-1-2 plane necessary to get $\partial V_i/\partial\psi$ from $\partial V_i/\partial\phi$ & $\partial V/\partial\lambda$ is a problem of spherical trigonometry, while α_1 α_2 are obtained from r_1, r_2, r_{12}, etc.).

For an orbital span τ one can write the disturbing potential T(t) as a Fourier series in time. Then for the horizontal derivatives

$$\frac{\partial T}{\partial\psi} = \frac{1}{n_o}\frac{\partial T}{\partial t} \tag{10}$$

where

$$n_o = [GM/r_o^3]^{1/2} \tag{11}$$

However, the radial derivatives $\partial T/\partial r$ and the orbital perturbations $\delta r(T)$ cannot be evaluated from T(t) because this representation is not harmonic: i e., not a solution of $\nabla^2 T = 0$. Hence the $\partial T/\partial r$ and $\delta r(T)$ terms in (7) must be set zero, and a determination of the Fourier coefficients for T made by least squares from the residuals $\Delta\ddot{r}_d - \Delta\ddot{r}_{co}$. In complex notation

$$T = \sum_\omega T_\omega e^{i\omega t} \tag{12}$$

$$\Delta\ddot{r}_d - \Delta\ddot{r}_{co} = \frac{\sin\alpha_1}{r_1}\left[\frac{\partial T}{\partial\psi}\right]_1 - \frac{\sin\alpha_2}{r_2}\left[\frac{\partial T}{\partial\psi}\right]_2$$

$$\approx \frac{i}{n_o}\sum_\omega \omega\{e^{i\omega t_1} - e^{i\omega t_2}\}T_\omega \tag{13}$$

Not determined by this procedure is an additive constant T_{oj}, where the subscript j pertains to the orbit. If the duration τ corresponds to a half revolution π/n_o pole-to-pole, then in a single such duration nothing would be learned to discriminate east-west variations of the fie d from north-south. Also for $\omega << 1/\delta t_d = n_o/\delta\psi$, the determination of T_ω will be weak. Con ideration should be given to a priori variances, since these low frequency terms will be the best known in the reference fi ld V_o.

7. Determine orbital additive constants for potential corrections. Using the locations $\underline{r}_{oi}$, from a set of orbits j a globa set of corrections to poten ial $T_j(r_o,\phi\ \lambda)$ is obtained. Next an adjustment must be made from condition equations at orbit crossings to determine the additive constants T_{oj} for eac orbit. The adjusted set $T(r_o\ \phi,\lambda)$ w ll still have a global unknown additive constant, which should be set so as to make the global mean zero.

8. Harmonically analyse the global field $T(r_o,\phi,\lambda)$ to determine whether the corrections to the reference field V_o implied thereby are plausible. If they are, a new reference field should be generated by adding T to V_o and steps 2-7 iterated. Hence at each iteration the $\partial T/\partial r$ and $\delta r(T)$ from the previous iteration are used, in effect. For the calculation of $\delta r(T)$ a spherical harmonic analysis of T is probably most efficient. However, the $\partial T/\partial r$ may require a considerably higher degree, so consideration should be given to inverse Stokes' integrations of the residual potential over limited caps.

At UCLA, we are undertaking a simple simulation to improve insight and procedures. A total potential $V = V_o + T$ is calculated from a limited number of spherical harmonic coefficients. The orbits are assumed to be polar and circular plus perturbations δr as calculated from V. The $\Delta\ddot{r}_{co}$ are calculated from V_o using (8), but the $\Delta\ddot{r}_d$ are calculated from the complete potential V. After procedures are developed to recover the correction T from simple cases, the work will move to more complicated situations.

Acknowledgements. This work is supported by NASA grant NSG-5263 and NOAA grant NA81-AA-D-00071.

References

Bender, P.L. (1981). Informal communication.

Bruce, K.L. (1977). Preliminary ocean-area geoid from GEOS-III satellite radar altimetry. DMA Aerospace Ctr., St Louis, Rep., 26 pp.

Breakwell, J.V. (1979). Satellite determinations of short wavelength gravity variations. J. Astronaut. Sci. 27(4).

Douglas, B.C., C.C. Goad, & F.F. Morrison (1980). Determination of the geopotential from satellite-to-satellite tracking data. J. Geophys. Res. 85, 5471-5480.

Gaposchkin, E.M. (1980). Global gravity field to degree and order 30. J. Geophys. Res. 85, 7221.

Jekeli, C. & R.H. Rapp (1980). Accuracy of the determination of mean anomalies and mean geoid undulations from a satellite gravity field mapping mission. Ohio State Univ. Dept. Geod. Sci., Columbus, Rep. 307, 22 pp.

Kaula, W.M. (1966). Theory of Satellite Geodesy. Blaisdell, Waltham, 124 pp.

Lerch, F.J., B.H. Putney, C.A Wagner & S.M. Klosko (1981). Goddard earth models for oceanographic application. Marine Geodesy 5, 145-187.

Pisacane, V.L. & S.M. Yionoulis (1980). Recovery of gravity variations from satellite-to-satellite tracking. Johns Hopkins Univ., Applied

Phys. Lab, Baltimore, Rep. SDO-5583.

Pisacane, V.L., J.L. MacArthur, J.C. Ray, & S.E. Bergson-Willis (1981). Description of the dedicated gravitational satellite mission (GRAVSAT). Int. Geosci. & Remote Sensing Symp., Washington.

Rapp, R.H. (1979). Global anomaly and undulation recovery using GEOS-3 altimeter data. Ohio State Univ. Dept. Geod. Sci., Columbus, Rep. 285, 49 pp.

Rummel, R. (1980). Geoid heights, geoid height differences, and mean gravity anomalies from "low-low" satellite-to-satellite tracking--an error analysis. Ohio State Univ. Dept. Geod. Sci., Columbus, Rep. 306, 44 pp.

ORBIT DETERMINATION METHODS USED IN THE SATELLITE GEODETIC OBSERVATORY

I. Almár - Sz. Mihály - T. Borza - J. Ádám
Satellite Geodetic Observatory
Penc, Hungary

Abstract

The Satellite Geodetic Observatory of the Hungarian Institute of Geodesy and Cartography is adopting and developing different programs of satellite orbit determination connected with its tracking activity /photographic, laser and Doppler observations/.

The PREDICE prediction program is used in laser tracking of geodetic satellites. It serves also for the improvement of certain orbital elements.

Different short-arc programs have been developed for station positioning. The SAMULPO program, which is based on a first approximation orbit by PREDICE, makes use of synchronous laser and photographic observations to improve the original 3 vector-components in several points of a short arc. The result is utilized to derive coordinates of another laser station observing simultaneously /FOTOLASER method currently in use in the Intercosmos countries/.

Another short-arc program, called SADOSA, serves the purpose of Doppler positioning by multilocation. It enables the

user to provide rigorous geodetic adjustment of coordinates of up-to 15 stations.

Among the unknowns the positional and velocity bias parameters of every orbit are also determined. Either range differences or pseudo-range equations can be used. The coordinate system can be defined optionally. The SADOSA program which forms a part of a research agreement between JMR Instruments Inc, California and the Satellite Geodetic Observatory provided practical results in the first West-East Doppler Observation Campaign in 1980 as well as in national and international network computations.

Several programs were developed to investigate the connection and transformation between different geodetic coordinate systems. These programs are used mostly in the interpretation of station coordinates determined by different technics as well as in orbit determination and geodynamical investigations.

Finally it is emphasized that the Satellite Geodetic Observatory is ready to serve as a monitoring station of geodetic satellites providing continuously precise observations in order to improve orbit calculation methods and procedures.

EFFECT OF UPPER ATMOSPHERIC VARIATIONS ON SATELLITE LIFETIMES

E. M. Both
Observatory of Hungarian Academy of Sciences
Baja, Hungary

Without any perturbations an Earth satellite would remain on a Keplerian orbit endless. Of the perturbing forces of different origin, atmospheric drag causes a secular decrease of the semi-major axis and orbital eccentricity, so satellites have finite lifetimes.

The theoretical lifetime of a satellite (L) can be determined by King-Hele's formulae, as a function of orbital decay rate ($\dot{T}$). The latter can be derived either from observations or from any atmospheric model. Determining L for decayed satellites and comparing this to the observed lifetime, we may not get any difference. However, there are such systematic differences because of the exactly unpredictable variations in the Earth's atmosphere. The systematic errors enabled us to demonstrate the effect of 11-year, semiannual and 27-day variations on satellite lifetimes.

First we examined the dependence of King-Hele's formulae on orbital elements and atmospheric parameters. We can state, that L depends mainly on orbital eccentricity and perigee height. The exact determination of the mass/area ratio of the

satellite is also of great importance. Comparing these results to the observed lifetimes of nearly 1000 satellites we experienced a good agreement, especially in the case of some groups of satellites of the same mass, shape and size.

Demonstrating the 11-year cycle we used 48 satellites of the same type, and their rockets. The satellites had nearly the same orbit. The L values showed a strong correlation with $\overline{F}$, the weighted mean of F solar decimeter flux. The error of computation, i. e. the difference between computed and observed lifetimes showed no correlation with $\overline{F}$. This proved, that the 11-year variation can be considered as a variation of the exospheric temperature, and this effect is taken into account satisfactorily by King-Hele's method.

The semiannual effect could be demonstrated in the case of 13 satellites, for which a long series of orbital elements was available. The relative error of lifetime determination showed strong correlation with density variations. Correcting the observed $\dot{T}$ values by the semiannual density variation ratio of CIRA-72 atmospheric model, the correlation ceased. This proves that King-Hele's formulae do not take into account the semiannual effect.

The same situation is at 27-day cycle. The relative errors show strong correlation with F index, characterizing the active area component of solar radiation. Applying a 27 day or longer running mean, the correlation coefficient decreased, while the correlation with $\overline{F}$ increased.

Summarizing the results we can establish, that King-Hele's

method for satellite lifetime determination can be improved by using a semiannual correction of any upper atmospheric model and by the use of a running mean for at least 27 days.

References

1. Both E.: The study of the Earth's upper atmosphere by the use of satellite lifetimes (Ph.D. dissertation,1981)
2. CIRA-72 atmospheric model (Akademie-Verlag Berlin 1972)
3. Jacchia, L. G.: Thermospheric temperature, density and composition: new models (SAO Spec. Rep. 375. 1977)
4. King-Hele, D. G.: Theory of satellite orbits in an atmosphere (Butterworths, London 1964)

DYNAMIC PROCESSES IN BINARY SYSTEMS

L. Patkós
Konkoly Observatory
Budapest, Hungary

Abstract

According to the theory of the evolution of close binary systems the initially more massive component fills up the Roche lobe in the course of its evolution, and mass transfer is started. It usually proceeds in two steps. In the rapid phase the mass ratio is more than reve sed. The subsequent evolution proceeds on a slow, nuclear time scale while the originally more massive, but now less massive star continues to give further masses to the other component.

The eclipsing binary star SV Cam has been observed with the 50 cm Cassegrain telescope of Konkoly Observatory's Mountain Station since 1973. Analysis of the obtained UBV light curves shows the system to have three interesting peculiarities. First of all there is a migrating distortion wave which is characteristic of RS CVn type stars. This wave is caused by dark spots on the surface and it migrates towards increasing orbital phase on the light curve because of the differential rotation of the stars. Flare activity in connection with the spotted region - as the second peculiarity - was also observed.

Moreover, a large-scale dynamic process also seems to exist in the system. Judging by my observations the starspot-activity of SV Cam was at its minimum at the end of the year 1974. There were no spots present, an undistorted "normal" light curve could be observed. This was confirmed by another observation a month later. Yet a further month later a brightness increase of about $0^m.05$ between phases 0.25 and 0.65 appeared. The existence and the shape of this brightness increase were confirmed by another observation four days later. On the light curve ob-

tained two months later the light-up had almost disappeared, but remains of it were still recognizable between phase 0.45 and 0.55.

An analysis of the times of minima obtained between 1964 and 1981 /among them more than 50 personal observations/ shows that there is a break in the O-C curve at the beginning of 1975. The direction of the break indicates a period increase. For a conservative system it means that the smaller secondary component gave mass to the primary.

The observed phenomena could be interpreted as follows: The secondary star of the system was in its quiet star-activity phase at the end of 1974. This tranquil phase was suddenly interrupted by the appearance of a stream of gas which landed on the surface of the primary causing a hot spot there. That the light-up originated from an area hotter than the stars themselves was confirmed by the B - V and U - B curves. The simultaneous appearance of dark spots on the surface of the secondary marks the beginning of a new starspot-activity cycle.

The appearance of the new distortion wave was later confirmed by further observations. Even the migrating direction was the same as for previous waves.

ROLE OF RANDOM FORCES IN STELLAR DYNAMICS

L.G. Balázs

Konkoly Observatory - Budapest - Hungary

Abstract

Our Galaxy contains some 10^{11} stars and diffuse material. A mechanical description of the stellar component would require the solution of a system of some 10^{11} second order differential equations. There are several approximations that can be utilized to solve this problem: numerical integration, numerical simulation, and the statistical approach. The force experienced by a star has some stochastic nature and so has the motion itself. The random effects of star - star encounters probably do not play an important role in the life of the Galaxy but cooperative phenomena and encounters with huge clouds of diffuse material and stars may be very significant. The time derivative of stochastic processes in the equation of motion is interpreted by mean square differentiation. In the case of harmonic motion the approximate solution, which is consistent with observations is a random walk. In time scales longer than some 10^{8} years one probably cannot avoid taking into account the effect of random forces.

Our Galaxy is a highly complex system in which widely different processes are going on at the same time. Such processes are: the shaping of the spatial distribution of the matter of our stellar system, mechanical motions of many kinds, star formation, the building up of chemical elements and, at least in one case, biological evolution. If we restrict ourselves to describing the spatial distribution and

motion of matter of our Galaxy then we describe it as a mechanical system in terms of the laws of mechanics.

The Galaxy as a mechanical system

The main interaction in our stellar system, as in general in the world of cosmic bodies, is the gravitation keeping together the different components of our Galaxy. The main component is represented in the form of some 10^{11} stars mutually interacting gravitationally and besides these we find a few per cent of diffuse material, gas and dust and, probably, matter of an unspecified nature ("black dwarfs", massive neutrinos, etc.).

A mechanical description of the stellar component requires our solving a system of second order differential equations in the form of

$$\ddot{x}_i = \frac{\partial}{\partial x_i} \sum_{i \neq j}^{n} \Phi_{ij} + F_{x_i} \qquad i=1,2,.....,n \qquad (1)$$

and similar equations for the Y and Z coordinates. In these equations Φ_{ij}, F_{x_i}, n mean the gravitational interaction between stars, between stars and the remaining part of the system, and the number of stars, respectively. We may add further equations to this system describing the mechanical behaviour of the remaining parts which are coupled to the equation of motion of stars by perturbing forces of F_i. It is obvious in the case of stars that the integration of 10^{11} unperturbed equations of motion is a hopeless task. Until now one has integrated such systems numerically up to n being approximately a few thousand (Aarseth et al. 1979). Numerical simulations (Miller 1978) proceed in n much further: one divides the space into cells by a spatial grid and avoids the problem of close pair encounters by assuming that two stars in the same cell do not interact with each other. In this way the influences of long range forces are treated. These numerical methods have resulted in significant suc-

cesses in studying the mechanical behaviour of stellar systems. At the same time however, they set up a difficult problem: To what degree do the solutions represent the mechanical behaviour of the real Galaxy? To answer this question we need to study other independent descriptions of our stellar system and to compare them with the results of numerical methods.

Suppose we have a system of particles interacting gravitationally. Now we may ask: Is it meaningful to look for the probability of finding a star in a given volume of the 6 dimensional phase space? The answer of statistical mechanics is yes and we can get dN, the expected number of objects in a unit volume, from the expression:

$$dN = Nf(p_i,q_i,t)dV \qquad i=1,2,3 \qquad (2)$$

where N,p_i,q_i,t,dV are the number of objects, the impulse coordinates, the spatial coordinates, the time and unit volume in phase space, respectively and $f(p_i,q_i,t)$ is the probability density function. In general the objects have different masses. Therefore $f(p_i,q_i,t)$ can be decomposed in the form of

$$f(p_i,q_i,t)=\int f(p_i,q_i,t|m)\ g(m)\ dm \qquad (3)$$

where $f(p_i,q_i,t|m)$ is the probability density function supposing that mass (m) has a given value and g(m) is the function characterizing the distribution according to m. We can now interpret the spatial mass density by the expression:

$$\rho(q_i,t)= N \int\int mf(p_i,q_i,t|m)\ g(m)\ dmdp \qquad (4)$$

Using this density we can define a gravitational potential by the Poisson equation:

$$\Delta\bar{\phi}= -4\P G\ (\rho + \rho_d) \qquad (5)$$

where G and ρ_d are the gravitational constant and the density of the diffuse material, respectively. The bar over ϕ discriminates this potential from the true potential of the Galaxy which is "grainy" because of the presence of stars. The statistical description of our stellar system means the finding of the probability density function in phase space or its moments. The statistical approach has the advantage that is works well even in the case of a very large number of objects.

Motions in a smoothed field

The mass density introduced in the preceding paragraph may be treated as the smoothed density of stars belonging to the same unit volume. The potential therefore defined by this smoothed density is a smoothed version of the true potential. The task of stellar dynamics is to find a solution to the equations of motion in this smoothed graviational field or in other words to solve the following system of second order differential equations

$$\ddot{x} = \frac{\partial \bar{\Phi}}{\partial x} \quad ; \qquad \ddot{y} = \frac{\partial \bar{\Phi}}{\partial y} \quad ; \qquad \ddot{z} = \frac{\partial \bar{\Phi}}{\partial z} \qquad (6)$$

The solution, in general, is very complicated and needs numerical methods. In some limiting cases, however, we can get rather simple analytic expressions.

If the star moves on a nearly circular orbit around the centre of our Galaxy and the inclination of the orbital plane is small then the motion perpendicular to the galactic plane is harmonic and epicyclic in the galactic plane. The epicyclic motion means harmonic motions in radial and tangential directions relative to a point (the so called epicentre) revolving around the centre of the Galaxy at the star's mean distance to the centre and with its orbital period.

By a statistical description we have tried to construct the probability density function in phase space. The probability density function, however, relates by equation (2) to the expected number of stars in a given u-

nit volume. The true number of stars generally deviates from this expected value and this deviation also has a probabilistic nature. The true gravitational potential relates to this true density by an equation similar to (5). Additionally, matter does not fill the unit volume uniformly because it is mostly condensed into stars and the distribution of diffuse material is uneven too. All these have the consequence that the deviation of the true force field from the smoothed field has some probabilistic nature. It means that we can not predict exactly the force acting on a star at a given moment. The force experienced by a star in the course of time, therefore, is a stochastic process and so is the motion itself. Before entering into the details of mathematical description of stochastic forces it is worth estimating their significance in stellar dynamical problems.

The nature of random forces

The most striking difference between the smoothed and true distribution of matter in the Galaxy is the "grainness" caused by the presence of stars. The dynamic influence of the "grainy" structure of the potential field is generally studied by means of star - star encounters which is a common topic of textbooks dealing with stellar dynamics. Even a very crude model is sufficent for estimating the order of magnitude of the dynamic effects of star - star encounters (see e.g. Woltjer 1967).

Let a star with a mass m encounter another star having a mass M. Let the initial velocity and impact parameter be equal to v and D, respectively. As a consequence of encounter the star will have a velocity component perpendicular to the initial velocity. The duration of encounter (T), the deflecting force acting on the star (F), and the velocity increment ($\Delta V_\perp$) can be estimated by the following formulae:

$$T \sim 2D/V$$
$$F \sim GmM/D^2$$
$$\Delta V_\perp \sim 2GM/DV \qquad (7)$$

(The meaning of these quantities is displayed in Fig.1)

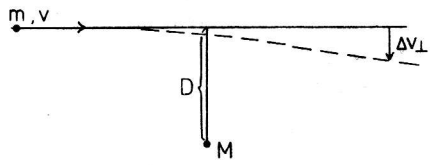


Figure 1

Due to the encounters the average kinetic energy associated with V⊥ steadily increases. By definition the time required to equal the kinetic energy of motion in the direction of deflection with the initial kinetic energy is called the time of relaxation. Using the simple formulae given above for T, F and ΔV⊥ the time of relaxation (T_r) can be written as

$$T_r = \frac{V^3}{8\pi G^2 M \rho \ln\Lambda} \qquad (8)$$

where Λ = Dmax/Dmin and ρ = mass density of perturbing objects. Assuming that the perturbing objects are common pop I. disc stars $T_r = 10^{12} - 10^{13}$ years, in the disc of the Galaxy. This time is 2 - 3 orders of magnitude greater than the 10^{10} years age of our stellar system. The effect of star-star encounters, therefore, can be neglected. The situation changes drastically, however, if we insert massive clouds instead of stars in place of the perturbing bodies (Spitzer, Schwarschild 1953). Assuming $M \sim 10^6 M_\odot$ and $\rho \sim 0.01 M_\odot/pc^3$ we can get $T_r \sim 10^8$ years.

Besides the effect of encounters the so-called cooperative phenomena such as streams, waves, instabilities contribute to the relaxation processes as well (Lynden - Bell 1967). In the next paragraph we shall give a more rigorous mathematical treatment of the effect of stochastic forces on the motion.

Mathematical description of the effect of stochastic forces

As it was mentioned in the preceding paragraph, if we were to insert stochastic forces in the equation of motion the motion itself would also become stochastic. Originally the equation of motion contained normal functions of time (explicit or implicit) and their time derivatives. If the motion has a stochastic character then we must interpret the time derivative of a stochastic process. One of the possible interpretations is the so-called mean square derivation (see e.g. Soong 1973).
By definition

$$\lim_{\tau \to 0} \| x(t+\tau) - x(t) \| / \tau = \dot{x}(t) \qquad (9)$$

where $\| x(t) \| = E\{x(t)^2\}$ and $E\{\ \}$ means the expected value. This definition has the very suitable property

$$E\left\{\frac{d^n x(t)}{dt^n}\right\} = \frac{d^n}{dt^n} E\{x(t)\} \qquad (10)$$

It means that the derivation and the operation of expected value are changeable. The time derivative of mean motion, therefore, equals the expected value of the stochastic time derivative of the real motion.

It was mentioned already that in the case of orbits with small eccentricity and inclination the motion is decomposed into harmonic motions. It is worth while, consequently, to investigate how stochastic perturbation affects the motion in a harmonic force field In the one dimensional case the equation of motion is

$$\ddot{x} + \omega^2 x = Y \qquad (11)$$

where ω equals the frequency of unperturbed motion and Y is a stochastic force having the properties:

$$E\{Y(t)\} = 0$$

and (12)

$$E\{Y(t)\ Y(s)\} = \Gamma_{yy}(t-s)$$

The first property means that Y describes only the deviation from mean force, thought to be harmonic, and the second property expresses the stationarity of Y. The solution is given by the very plausible formula of

$$x(t) = \int_0^t h(t-s)\ Y(s)\ ds \qquad (13)$$

where h (t) is called the impulse response which is the solution of the deterministic equation with Dirac's delta perturbation

$$\ddot{h}(t) + \omega^2 h(t) = \delta(t) \qquad (14)$$

assuming $t \geq 0$, $\dot{h}(o) = 0$, $h(o) = 0$.

The first property in (12) leads to $E\{x(t)\} = 0$ and the standard deviation can be expressed in the form of

$$\sigma_x^2 = E\{x(t)^2\} = \int_0^t\int_0^t h(t-u)h(t-v)\Gamma_{yy}(u-v)dudv \qquad (15)$$

An approximate evaluation of the integral in (15) is possible if $S_{yy}(\omega)$, viz. the power spectral density of Y(t), is relatively smooth with no sharp peaks. Under these conditions (15) reduces to

$$\sigma_x^2(t) \cong \frac{\pi}{4}\frac{S_{yy}(\omega)}{\omega^3}(2\omega t - \sin 2\omega t) \qquad (16)$$

or in the case of large values of t to

$$\sigma_x^2(t) \cong \frac{\pi}{2}\frac{S_{yy}(\omega)}{\omega^2}t \qquad (17/a)$$

and similarly

$$\sigma^2_{\dot{x}}(t) \cong \frac{\pi}{2} S_{yy}(\omega)\, t \qquad (17/b)$$

Expressions (17) show that the standard deviation of coordinates and velocities grows monotonically in the course of time as in the case of random walk. In the next paragraph we shall discuss how these predicted effects will be recognized by observations.

Observational evidence of random forces

Formulae (17) can be tested in terms of observations if we have objects in our Galaxy for which age, spatial position, or velocities are available. Such objects are open clusters and nearby stars. In Figure 2 the dispersion of distances form the galactic plane is displayed as a function of time in the case of open clusters. Figure 3 shows the increase of velocity dispersions of nearby stars in the course of time. As one can infer the points displayed in Figure 2 and 3 can be fitted very well by straight lines corresponding to (17).

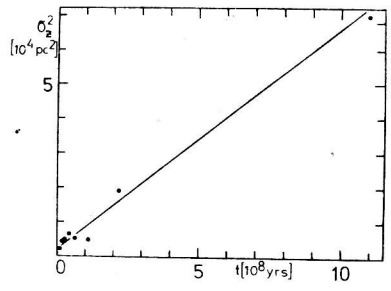


Fig.2. Dispersions of distances from the galactic plane as a function of time in the case of open clusters. Each point represents 20 clusters. (Data from Lyngå 1980)

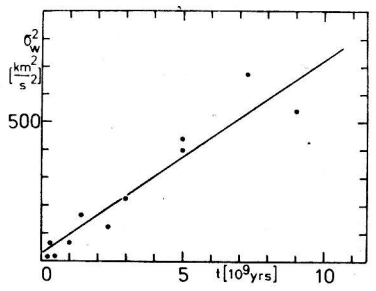


Fig.3. Increase of dispersions of velocity components perpendicular to the galactic plane as a function of time in the case of nearby stars. (Data from Wielen 1977)

As is well known, star formation takes place near to the galactic plane and the newly born stars have nearly circular orbits and small velocity dispersions. It means that even at t = 0 the dispersions do not vanish. The measured variances, therefore, consist of two parts: the initial dispersion and an increment described by formulae (17).

At this point we need to reflect on an important problem: Is the observed increase of dispersions really caused by random forces or are there other physically completely different phenomena leading to the same observable result? Following Wielen (1977) we shall consider three possibilities for the increase of dispersions:

1. variation of typical velocity at birth,
2. acceleration by global gravitational field,
3. acceleration by stochastic fluctuation of forces.

As to the first possibility it seems probable that the conditions in which stars are born today are different from those at the early phase of our stellar system. This means that the longterm variation of dispersions could be explained but it is difficult to believe that this effect could account for the rapid increase over some 10^8 years. The second effect, on the contrary, is capable of explaining the short term incrase in the dispersions but, as computations have shown, it is insufficient for longer time scales. The best explanation appears to be the third possibility. In this way we may conclude that when describing the dynamic behaviour of our Galaxy in time scales longer than some 10^8 years one must take into account the effect of random forces.

References:

Aarseth, S.J., Gott, J.R. and Turner, E.L. 1979, Ap.J., 228,664
Lynden-Bell, D. 1967, Cooperative Phenomena in Stellar Dynamics, in 'Relativity Theory and Astrophysics 2.', ed. J. Ehlers, Am. Math. Soc., Providence, Rhode Island

Lyngå, G. 1980, privat comm.
Miller, R.H. 1978, Ap.J., 223, 22
Spitzer, L., Schwarzschild, M. 1953, Ap.J., 118,106
Soong, T.T. 1973, Random Differential Equations in Science and Engineering Academic Press New York and London pp92-99 and pp161-163
Wielen, R. 1977, Astr. Ap., 60,263
Woltjer, L. 1967, Structure and Dynamics of Galaxies, in 'Relativity Theory and Astrophysics 2.', ed. J. Ehlers, Am. Math. Soc., Providence, Rhode Island

DYNAMICS OF GALAXIES

R. H. Miller
University of Chicago, U S A

Abstract

A large computational project to study the dynamics of galaxies has been in progress at NASA-Ames Research Center in cooperation with Dr. Bruce F. Smith for the past 5 years. Galaxies are represented as self-consistent self-gravitating batches of particles whose responses are computed in a fully three-dimensional n-body treatment using 100 000 particles. Results are frequently unexpected and usually differ from prior guesses. The principal discoveries from this work in the past few years include: (1) Demonstration that "cold" axisymmetric disks, like our Galaxy, are dynamically unstable; (2) Discovery that a prolate bar, rotating end-over-end in space about a short axis, is the dynamically preferred form for rapidly rotating stellar systems; (3) Discovery of sharp contractions of both members as two galaxies pass near each other in a galaxy collision. This sharp contraction precedes the "explosion" of the galaxy and is responsible for most of the dynamical effects seen following a collision; (4) Demonstration that the internal dynamics of a galaxy in a cluster of galaxies is affected on

the timescale of the cluster crossing time; (5) Discovery that tidal braking of galaxies rotating in the force field of a cluster of galaxies can account for the observed slow rotation of elliptical galaxies; and (6) Demonstration that perturbations present as the Universe becomes matter-growth rates to 1-2% accuracy, and that fluctuations 10^{-4} to 10^{-3} at decoupling are sufficient to produce present-day galaxy clusters and superclusters. Results from these numerical experiments are usually so complex that motion pictures are the only practical way to understand the dynamics. Motion pictures were shown for two experiments. Collisions of Disk galaxies in massive halos: Each halo is represented by 50 000 particles in a self-consistent steady-state galaxy model. Disks, like the halos, are represented by particles. Disk particles orbit along circular paths centered in the halo potential. The initial disks represent only 1% of the mass, but contribute all the light. Two disk-halo combinations are thrown at each other along a specified orbit. Several experiments with different initial orbital parameters (parabolic, hyperbolic, different initial orbital angular momentum) are shown, with different disk orientations. Disk and halo particles are distinguished by different colors in the film. Disks rapidly distort into barlike forms in collisions with the disks initially in the orbital plane. Damage to the disks is about equally severe whether the disks rotate prograde or retrograde to the

collision. With initial disk normals along the line between the two initial halos, rings form in the disks even at surprisingly large impact parameters.

Gravitational Clustering of Galaxies in an Expanding Universe: Smaller perturbations to the initial state and integrations that represent greater total expansions can be obtained with the galaxy dynamics programs than have been possible previously. Forms that look like present-day galaxy clusters and superclusters develop naturally from a wide range of initial conditions. Large empty regions between the superclusters, giving a cellular appearance, are characteristic of the well-developed final states. This lends support to the idea that present-day clustering developed through gravitational processes, but it implies that present-day clustering does not carry much information about details of the Universe shortly after decoupling.

ON THE OBSERVATIONAL VULNERABILITY OF THE MODELS FOR THE GALAXY

B. A. Balázs
Department of Astronomy, Eötvös University
Budapest, Hungary

Abstract

Any scientific model must be as simple as possible and vulnerable by measurements or observations. Herein some possible ways of checking the models for the Galaxy are discussed.

In contrast to Kapteyn's distorted, small and "heliocentric" galactic system, the models of Oort and Schmidt are the first successful quantitative attempts to model the Galaxy (see References 1-2). The fundamental new data and ideas entering into these models are of dynamical nature and concern the gravitational force perpendicular to the galactic plane and the differential galactic rotation. The basic idea behind this set up is that, collectively, the motions of the stars make up the dynamical structure of the galactic system, which in fact determines the system's geometrical structure and its evolution.

In the last 20 years or so authors were making more and more frequently use of a sizeable number of analogies between our Galaxy and similar extragalactic systems. It stands to reason that the basic parameters and properties of

any model for the Galaxy must comply with the rotation curves, surface brightness distributions, scale lengths, mass/luminosity ratios and other caracteristics of similar stellar systems. For detailed discussions the excellent model of Ref. 3 is selected because practically all the significant parameters and predictions of this model can be checked by observations feasible with the aid of contemporary astronomical technics.

As far as the application of star counts to the determination of galactic structure parameters is concerned, using an analytic smooth curve approximation of the empirical luminosity function and an exponential disk plus a de Vaucouleurs spheroid for the global distribution of matter the variations of the calculated star densities with apparent magnitude, latitude and longitude agree well with the star count data available for the observationally thoroughly covered range of $4 \leq m_V \leq 22$. The luminosity function may increase strongly outside the available empirical limits. If the logarithmic slope of the increase is around the not unreasonable value of 0.2 dimwards of 16 absolute magnitude, the increment in counts to 21 mag is negligible, but we can count with an approx. 20% increase in the total counts up to 28th mag. The further extension of the empirical luminosity function is therefore very important for checking galax models. This is one of the reasons why the multicolour observation of the luminosity distribution up to the faintest stars in quite a number of properly selected galactic clusters is planned with the aid of the 6 m telescope (in the framework of an eastern european

multilateral cooperation called "Stellar Physics and Evolution").

Making use of dinamical considerations similar to those introduced by Oort and the empirical rotation curve of our stellar system (see Reference 4), it is possible to show that the observational data require the existence of a third major mass component in the Galaxy. On the one hand the computed rotation curves of all the two-component models of the Galaxy fall monotonically beyond 12 kpc, while the empirical one (see References 3-4) is flat to at least 25 kpc, on the other there is an irreconcilable difference in the circular velocity of the sun, which can be derived empirically from the solar motion with respect to the local group of galaxies (see References 4-5), halo stars and also from the escape velocity suggested by peculiar velocities in the solar neighbourhood. The observational restrictions do not allow such low a value as 170 kms^{-1}, which is the maximum rotation velocity of the standard two-component model. These discrepancies can be removed by invoking a third mass component of the Galaxy: a massive halo the stellar content of which is essentially detectable by observations up to $m_V = 28$. (The halo can make only a small contribution to the star counts in the currently available data region $m_V \leq 21$.) It is therefore expected, that the existence of a massive halo (with a mass of approx. $5.10^{11} M_\odot$) will be confirmed or disproved in the near future by direct (primarily Space Telescope) observations. (See References 6-7.)

In the concluding part some problems concerning the models for the spiral structure of our Galaxy are discussed having regard for the circumstance that at the present time only the density wave theory of C. C. Lin has been developed toward a coherent model to provide a quantitative viewpoint from which it is possible to explain the spiral pattern (see Reference 8). From the model parameters the value of the galactocentric distance of the sun ($R_\odot$) and the angular pattern speed (Ω_p) is discussed. The so called standard IAU value of $R_\odot$ is 10 kpc but in the last years there is a strong tendency to correct it downwards (see References 5, 9-13). Herein $R_\odot$= 7 kpc is preferred. As far as the angular pattern speed is concerned, Ω_p is usually chosen so that the corotation radius defined by $\Omega(R_c) = \Omega_p$ is equal to the distance of the outhermost visible H II regions to the galactic centre (see Ref. 14). It is possible to check the result of this procedure on an independent way using young open clusters ($\tau < 3.2\ 10^7$y) as spiral tracers and moderately old ones ($\tau < 10^8$y) as indicators of the pattern speed (see References 15-16), which turns out to lie somewhere between 33 and 36 km/s/kpc i. e. considerably larger than the value originally accepted (see also References 17-22). Therefore not the outhermost H II regions, but the stars in the solar neighborhood lie in the zone of corotation and the spiral waves propagate from the center to the periphery in the Galaxy, quite in accordance with the results of Reference 23.

References

1. Oort, J.H.: 1932, B.A.N. 6, 238, 249.
2. Schmidt, M.: 1965, "Rotation Parameters and Distribution of Mass in the Galaxy" in "Stars and Stellar Systems" V, (Univ. of Chicago Press), p.513.
3. Bahcall, J.N., Soneira, R.M.: 1980, Astrophys. J. Suppl. 44, 73.
4. Gunn, J.E. et al.: 1979, Astron. J. 84, 1181.
5. Oort, J.H., Plaut, L.: 1975, Astron. Astrophys. 41, 71.
6. Bahcall, J.N., O'Dell, C.R.: 1979, in IAU Colloq. No. 54, "Scientific Research With the Space Telescope", (US Gov. Printing Office), p. 5.
7. Longair, M.S.: 1979, Quart. J.R.A.S. 20, 5.
8. Lin, C.C., Shu, F.: 1964, Astrophys. J. 140, 646.
9. Balázs, B.A.: 1977, Mitt. Astron. Ges. 43, 254.
10. Clube, S.V.M., Dawe, J.A.: 1978, IAU Symp. No. 80, 53.
11. Clube, S.V.M,, Dawe, J.A.: 1980, M.N.R.A.S. 190, 591.
12. de Vaucouleurs, G., Buta, R.: 1978, Astron. J. 83, 1383.
13. Frenk, C.S., White, S.D.M.: 1982, M.N.R.A.S. 198, 173.
14. Feitzinger, J.V., Schmidt-Kaler, Th.: 1980, Astron. Astrophys. 88, 41.
15. Balázs, B.A.: 1981, Mitt. Astron. Ges. 52, 21.
16. Balázs, B.A.: 1982, in print (in Soobshch. Spets. Astrofiz. Obs.).
17. Burton, W.: 1976, Ann. Rev. Astron. Astrophys. 14, 275.
18. Stecker, F.: 1976, Nature 260, 412.
19. Nelson, A., Matsuda, T.: 1977, M.N.R.A.S. 179, 663.
20. Pavlovskaya, E.D., Suchkov, A.A.: 1978, Pis'ma Astron. Zh. 4, 450.

21. Palouš, J.: 1980 Astron. Astrophys. 87, 361.

22. Talbot, R.J.: 1980, Astrophys. J. 235, 821.

23. Marochnik, L.S. et al.: 1972, Astrophys. Space. Sci. 19, 285.

Készült az ELTE Sokszorosítóüzemében
400 példányban
Felelős kiadó: Dr. Kubovics Imre
Felelős vezető: Arató Tamás
ELTE 82167