BR22997

TECH. MEMO
SPACE 157

TECH. MEMO
SPACE 157

AD 728272

ROYAL AIRCRAFT ESTABLISHMENT

APPLICATIONS AND RESULTS OF ORBIT DETERMINATION

by

R. H. Gooding

November 1970

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

Crown Copyright 1970

UNLIMITED

24

ROYAL AIRCRAFT ESTABLISHMENT

Technical Memorandum Space 157

November 1970

APPLICATIONS AND RESULTS OF ORBIT DETERMINATION

by

R. H. Gooding

SUMMARY

This paper summarizes the reasons for determining the orbits of earth satellites, and describes how knowledge of the earth's gravitational potential and its atmosphere has improved as a result of orbit determination. The paper is based on the second of two lectures given at the ESRO summer school on Spacecraft Operations, held at Gravenbruch, near Frankfurt, West Germany, in August 1970. The first lecture is available as Technical Memorandum Space 156.

1 INTRODUCTION

In the first of my two lectures I gave some technical details of orbit determination by the method of differential correction, with special reference to the computer programs that have been developed at the Royal Aircraft Establishment. In this remaining lecture I propose to list and discuss briefly the main reasons why orbit determination is necessary at all, and then to elaborate on the last two topics in the list.

2 OBJECTIVES OF ORBIT DETERMINATION

Fig.1 is based on Merson's classification[1] of the objectives of orbit determination as operational and non-operational. The distinction between the two classes is that the former are essential to the particular satellite, whereas the latter are not, but this distinction is not always clear-cut.

Orbit achievement may seem too obvious an objective to include in the list. It is too important to omit, however; until a satellite has been tracked, there is no certainty that it is in orbit, and until a rough orbit has been computed from tracking data, there is no certainty that the orbit is satisfactory. Again, the performance of the launching vehicle can only be properly evaluated when the resulting orbit is known.

Guidance and control will be required, in particular, for communication satellites and space probes. The correct commands to change an orbit cannot be made until the current orbit is known.

On-board instrumentation plays a vital role for scientific satellites such as the Ariels. The telemetered data from such instrumentation must usually be correlated with satellite position (for example to plot electron density against height and latitude), and continuous knowledge of satellite position can only come from orbit determination.

Tracking and telemetry scheduling, for the various stations in a network, require knowledge of the orbit, and the stations will need predicted look angles for the scheduled passes. The accuracy desired for this objective is usually much less than for the previous one, but the urgency is of course much greater.

General surveillance, the first non-operational objective listed, is required in both a military and a technical context. The military requirement is to monitor all spacecraft and to detect, as rapidly as possible,

any object newly launched into orbit. The technical requirement is to distinguish one satellite from another, in particular to avoid interference when communicating with a given satellite.

Sensor accuracy, i.e. the accuracy of the instrumentation at a given satellite observing station, can be estimated by consideration of the equipment itself, but such estimation can only be validated by studying residuals of actual satellite observations relative to computed orbits.

Navigation and geodesy could as well appear under 'operational objectives', since a number of satellites have been launched with these objectives specifically in mind. The US Navy has refined a navigation system based on Doppler measurement of range rate, in which the orbit of a satellite is determined from Doppler observations from a standard network of ground stations. The resulting orbital parameters are fed into the satellite itself, which transmits them continuously (until a further set is fed in). Then a ship, using relatively simple equipment, can track the satellite and, from the observed Doppler shift and the orbital parameters, can locate its own position. Geodesy, however, has certainly been mainly associated with non-operational objectives. Here I am not thinking of the 'direct' or 'geometrical' method of approach, which is based on the simultaneous observation of a satellite from two or more stations which have been accurately surveyed relative to the same datum and from a station which it is required to survey to this datum, since this method does not require an orbit to be determined. I am thinking rather of the 'indirect' or 'orbital' method in which one has available a vast number of accurate observations of several satellites, and can therefore carry out a simultaneous differential correction, not merely of the parameters of all the satellites, but also of the sets of station coordinates. The best known contributions to satellite geodesy to date have probably been the 1966 and 1969 Smithsonian Standard Earths[2,9], for which the observational data were obtained mainly from Baker-Nunn cameras.

The last two non-operational objectives, concerned with improved knowledge of the earth's gravitational field and of atmospheric density and rotation, are my two main topics, to be covered in some detail in the following sections. The first of these objectives is related to the geodesy objective, the difference being that an even broader view is taken - we are concerned with the gravitational constants of the solid earth and not merely with the surveying of its surface. Only an 'indirect method' is possible.

The same principle is involved in the orbit-based study of both the earth's gravitational field and its atmosphere, namely that orbits suffer 'perturbations' as I indicated in my first lecture. Such perturbations may be represented by a mathematical theory - the 'orbital model' or 'orbit generator' - and the numerical parameters of this theory may be inferred from observation of the perturbations.

It is important to realise that the method depends not only on the study of the orbits of a number of satellites but also, usually, on a large number of orbit determinations, over a long period, for each satellite. The principle is illustrated in Fig.2. Here an existing orbital model is assumed to be good enough to describe the motion of a satellite during a short period, i.e. good enough for orbit determination over such a period to yield small final residuals. After many orbit determinations for the same satellite, over a long period, the graph of one of the orbital parameters is plotted, its variation being due to perturbations. The existing orbital model gives a poor fit to this long-term variation of the orbital parameter, but the fit may be dramatically improved by use of revised earth constants in the model. Alternatively, it may be that an improved model, incorporating previously neglected perturbations, is required before an adequate fit can be obtained.

It is fortunate that gravity-induced perturbations have a very different character from atmosphere-induced perturbations. This means that the two subjects can essentially be considered independently, as I now proceed to do.

3 GRAVITATIONAL FIELD OF THE EARTH

If the earth were spherically symmetrical its field would be the same as that of a point mass as its centre; i.e. satellite motion would be unperturbed. Thus gravity-induced perturbations yield information about asphericity.

The simplest improvement upon an assumption of sphericity is to take the earth as an oblate spheroid. This is illustrated in Fig.3. The 'flattening' of such a spheroid is defined, in terms of the polar and equatorial diameters, by the relation

$$f = \frac{D_E - D_P}{D_E} \; ; \qquad (1)$$

here flattening must not be confused with eccentricity, a term not usually used in this context but which would be defined by

$$e^2 = \frac{D_E^2 - D_P^2}{D_E^2} .$$

The first estimate of f was by Newton who, by theoretical argument, obtained the value 1/230. After the development of refined geodetic methods, Hayford gave the value 1/297.0 in 1909, and this was adopted as an international standard in 1924. The current best value, based on orbit determination, is 1/298.25, so that Hayford overestimated the difference between D_P and D_E, which is about 43 km, by some 180 metres.

The next improvement in the assumption about the earth's shape, made worthwhile by the advent of satellites, is to take it as axisymmetric, so that the gravitational potential is latitude-dependent but longitude-independent. From the fact that the potential must satisfy Laplace's equation it follows that the most general form, subject to axial symmetry, is given by the Legendre expansion

$$\frac{\mu}{r}\left\{1 - \sum_{\ell=2}^{\infty} J_\ell \left(\frac{R}{r}\right)^\ell P_\ell (\sin \beta)\right\} , \qquad (2)$$

where μ is the product GM of the gravitational constant and the mass of the earth, R is the equatorial radius of the earth, r is distance from the earth's centre, β is geocentric latitude, P_ℓ is the Legendre polynomial of degree ℓ, and the 'zonal harmonic' coefficients J_ℓ are dimensionless constants which represent the shape of the earth. The first coefficient in the series, J_2, is directly related to the flattening, f, but the relation is not as simple as might be expected, because J_2 is associated with the <u>gravitational</u> field, whereas f is associated with the <u>gravity</u> field (i.e. the gravitational field plus the centrifugal field on the surface of the earth due to its rotation).

Fig.4 interprets the first four harmonics (J_2, J_3, J_4 and J_5) as distortions (from circularity) of a meridional cross-section of the earth. The distortions are greatly exaggerated, and the shapes for J_4 and J_5 are for positive values of these constants, whereas they are actually both negative. Although all the zonal harmonics are axisymmetric, only the even harmonics -

i.e. those for which ℓ is even - are symmetric about the earth's equator. Because the perturbations produced by the even harmonics are quite different from those produced by the odd harmonics, it is usual to study them separately.

Fig.5 gives values of the even harmonics, as determined by various authors[3,4,5], and it is seen that there is good agreement as far as J_6. It is striking that, although J_2 is of order 10^{-3}, subsequent even harmonics (and in fact all other harmonics) are of order 10^{-6}; thus it is not surprising that only J_2 could be estimated prior to the satellite era. The fact that J_ℓ $(\ell > 2)$ is of order J_2^2 is important in the various theories of satellite perturbations in the earth's gravitational field.

The main effects of the even zonal harmonics are on the orbital elements Ω (right ascension of the node) and ω (argument of perigee). These have secular perturbations, as I explained in my first lecture, and formulae for their rates of change are

$$\dot{\Omega} = -\frac{3}{2} J_2 \, n \, (R/p)^2 \cos i + 0 \, (J_4 \text{ etc.}) \tag{3}$$

and

$$\dot{\omega} = \frac{3}{4} J_2 \, n \, (R/p)^2 \, (4 - 5 \sin^2 i) + 0 \, (J_4 \text{ etc.}) \quad , \tag{4}$$

where $n = (\mu/a^3)^{\frac{1}{2}}$, i.e. n is the mean motion, $p = a(1 - e^2)$, and the remaining notation has already been given. Fig.6 illustrates $\dot{\Omega}$ as a westward rotation of the orbital plane; this rotation vanishes only for polar orbits. Similarly $\dot{\omega}$, interpreted as the rate of rotation of the orbit within its plane, vanishes for inclinations such that $\sin i = \pm\sqrt{0.8}$. These are the important 'critical inclinations', which are of both theoretical and practical interest and which I referred to in my first lecture. (Many Russian satellites have been launched with near-critical inclination, so that perigee - or apogee for communication satellites - might remain for long periods in the northern hemisphere.) It is from the observed values of $\dot{\Omega}$ and $\dot{\omega}$ (or sometimes just from $\dot{\Omega}$), averaged over long periods for a number of satellites, that the even harmonic coefficients are derived. Thus if the average value of $\dot{\Omega}$ observed for satellite A is $\dot{\Omega}_A$, we have an equation of condition of the form

$$A_2 J_2 + A_4 J_4 + A_6 J_6 + \ldots = \dot{\Omega}_A \quad ,$$

where $A_2 = -1\frac{1}{2}\, n_A\, (R/p_A)^2 \cos i_A$ and $\dot{\Omega}_A$ has been corrected by removal of all the (small) effects of other sources of perturbations. Similar equations of condition for satellites B, C etc. may be set up - it is important that as wide a range of orbital inclinations as possible be covered - and the whole set solved by the method of least squares. The King-Hele/Cook solution[3] was for four J coefficients, using seven satellites.

Fig.7 gives values of the odd harmonics, one set determined by King-Hele, Cook and Scott[6], and another set determined by Kozai[5]. (Exactly zero-values appear in the set of King-Hele et al., because in their first least-squares solution they obtained such small values of these coefficients that they dropped them altogether; having set nine odd J's to zero they determined values of six others, using 22 satellites.) Agreement between the two sets is good up to J_7.

The odd harmonics do not lead to secular perturbations, so they are determined from long-periodic perturbations, usually from the perturbations in e (eccentricity). Fig.8, repeated from my first lecture, shows the secular and long-periodic variation of e for the satellite Ariel 2. If the secular variation is removed, the amplitude of the long-periodic oscillation - of which the period is about 120 days - is easily obtained, and from this an equation of condition for the odd harmonics, in the form

$$A_3 J_3 + A_5 J_5 + A_7 J_7 + \ldots = e_A ,$$

may be derived. As with the even harmonics, a set of such equations may be solved for as many J coefficients as desired.

I have mentioned that the odd zonal harmonics relate to equatorial asymmetry in the shape of the earth, and it is often said that J_3 represents a 'pear-shaped' effect. The actual value of J_3 implies that, at mean sea level, the north pole - which is at the stalk of the pear - is 32 metres further from the equator (which is defined to contain the earth's centre of mass) than the south pole is. Taking into account the other odd harmonics this figure is more like 41 metres. The general picture is given by Fig.9, which shows the height of the geoid (i.e. of mean sea level), relative to a spheroid of flattening 1/298.25, as a function of latitude. Two curves are plotted; both curves relate to a seven-coefficient set of odd harmonics obtained by King-Hele two years before the set listed

in Fig.7, but one curve relates to the King-Hele/Cook set[3] of even harmonics given in Fig.5 and the other to the Smith set[4].

We now remove all restrictive assumptions about the earth's shape. The general expression for the gravitational potential, with no symmetries at all, is[7]

$$\frac{\mu}{r}\left[1 + \sum_{\ell=1}^{\infty}\sum_{m=0}^{\ell}\left(\frac{R}{r}\right)^{\ell} P_{\ell}^{m}(\sin\beta)\,\{C_{\ell,m}\cos m\lambda + S_{\ell,m}\sin m\lambda\}\right], \qquad \ldots (5)$$

where λ is longitude, P_{ℓ}^{m} is the associated Legendre function of degree ℓ and order m, and the harmonic coefficients $C_{\ell,m}$ and $S_{\ell,m}$ are dimensionless constants. When $m = 0$ the harmonics are 'zonal'; $C_{\ell,o}$ is simply $-J_{\ell}$, and $S_{\ell,o}$ does not arise since $\sin 0 = 0$. When $m = \ell$ the harmonics are 'sectorial', and when $0 < m < \ell$ they are 'tesseral', though the term 'tesseral' is often taken to include 'sectorial'. It is sometimes preferred to replace $C_{\ell,m}$ and $S_{\ell,m}$ by coefficients $J_{\ell,m}$ and $\lambda_{\ell,m}$, defined by the relations

$$J_{\ell,m}\cos m\lambda_{\ell,m} = C_{\ell,m}$$

and

$$J_{\ell,m}\sin m\lambda_{\ell,m} = S_{\ell,m} ,$$

where $J_{\ell,m} > 0$. Then $J_{\ell,o} = |J_{\ell}|$.

Certain of the harmonic coefficients can be eliminated without loss of generality. If the 'geocentric' origin of the coordinates β and λ is taken to be the earth's centre of mass, it follows that $C_{1,0} = C_{1,1} = S_{1,1} = 0$. If the 'polar' axis (on which $\beta = \pm\frac{1}{2}\pi$) is taken as a principal axis of the earth, it follows that $C_{2,1} = S_{2,1} = 0$; this is reasonable, since the axis of the earth's rotation <u>is</u> a principal axis. However, it is pointless to try to make $S_{2,2} = 0$ by taking the origin of λ as another (equatorial) principal axis, since only the polar principal axis is known with any precision.

We now rewrite the general potential expression as

$$\frac{\mu}{r} + \sum_{\ell=2}^{\infty}\sum_{m=0}^{\ell} U_{\ell,m} \qquad (6)$$

where

$$U_{\ell,m} = \frac{\mu}{r} J_{\ell,m} \left(\frac{R}{r}\right)^{\ell} P_{\ell}^{m} (\sin \beta) \cos m(\lambda - \lambda_{\ell,m}) \quad . \tag{7}$$

For an arbitrary satellite, the three spherical coordinates r, β and λ may be expressed in terms of the six orbital elements a, e, i, Ω, ω and M. We get[8]

$$U_{\ell,m} = \sum_{p=0}^{\ell} \sum_{q=-\infty}^{\infty} U_{\ell mpq} \quad , \tag{8}$$

where

$$U_{\ell mpq} = \frac{\mu}{a} J_{\ell,m} \left(\frac{R}{a}\right)^{\ell} \Re \left[F_{\ell mp}(i)\, G_{\ell pq}(e) \exp \sqrt{-1}\, \{(\ell - 2p)\, \omega + (\ell - 2p + q)\, M + m\, (\Omega - \nu - \lambda_{\ell,m})\}\right] \quad ; \tag{9}$$

here F and G are standard functions, which fortunately represent inclination and eccentricity effects completely and independently, $\Re$ denotes 'real part', and ν is the sidereal time.

A full study of the perturbing effects of the general $U_{\ell mpq}$ potential term would be an enormous undertaking, so I do not propose to do more than to consider briefly the cases when these effects are most significant. The main criterion for significance is the rate of change of the argument of the exponential function in (9); if we write

$$\Phi = (\ell - 2p)\, \omega + (\ell - 2p + q)\, M + m(\Omega - \nu) \quad , \tag{10}$$

then this argument is $\Phi - m\, \lambda_{\ell,m}$, and the critical quantity is $\dot{\Phi}$. The general situation is that Φ changes at least as rapidly as M, so that the resulting perturbations are of short period and hence (apart from those due to J_2 of course) negligible. It turns out that $G_{\ell pq}(e)$ is of order $e^{|q|}$ and it follows that, unless e is large, significant terms only arise for $q = -1$, 0 or $+1$.

In the more familiar zonal-harmonic situation we have $m = 0$ and there are two important cases: (1) if $\ell - 2p = q = 0$, Φ vanishes and we have secular perturbations, as used to determine the values of the even harmonics: (2) if $\ell - 2p + q = 0$, but $q \neq 0$, then we have long-periodic perturbations, as used to determine the odd harmonics.

If $m \neq 0$, true secular terms cannot arise but there are still two important cases. First, if $\ell - 2p + q = 0$, then $\dot{\Phi} = - m \dot{\nu} +$ smaller terms, and if m is small we have perturbations of period several hours; these will just be significant for close earth satellites. Secondly, if $\dot{M}$ and $\dot{\nu}$ are roughly commensurate, there will be values of ℓ, m, p and q such that $\dot{\Phi}$ is close to zero; this is the condition known as <u>resonance</u>. In an extreme case, if drag is negligible so that $\dot{M}$ is virtually constant, the resonant condition may give rise to terms that are effectively secular.

As an example of important non-resonant perturbations, Fig.10 shows an apparent error, of about 12-hour period, that was detected in residuals associated with RAE orbit determinations for the satellite Ariel 2. An early version of the orbit-determination program was being used, with no representation of tesseral (or sectorial) harmonics, and it was realised that the 'time error' was really an along-track perturbation due to various $U_{\ell mpq}$ with $\ell - 2p + q = 0$ and $m = 2$, but in particular to U_{2210}. When the along-track term due to $J_{2,2}$ was added to the orbital model, the apparent time error was very much reduced. It is from residuals, such as those represented in Fig.10, for a number of satellites, that a general least-squares determination of <u>some</u> set of tesseral harmonics is normally carried out, usually in conjunction with the improvement of station coordinates. Ref.9, for example, gives the set from the 1969 Smithsonian Standard Earth, in which all harmonics up to $(\ell, m) = (16, 16)$ are included (together with 14 pairs of higher degree derived from resonant perturbations).

The physical meaning of resonance is that the ground track of a satellite repeats after some simple fraction of a day. The most familiar example is the once-per-day repetition of synchronous communication satellites (also described as 'geostationary' when the orbit is circular and equatorial). The condition for a synchronous satellite is that $\dot{M} \approx \dot{\nu}$, so it follows from (10) that synchronous resonance occurs for $U_{\ell mpq}$ such that

$$\ell - 2p + q = m \ .$$

If we neglect $O(e)$ perturbations (and also the $O(1)$ eccentricity perturbation that arises for $q = \pm 1$), then we must take $q = 0$, so that $\ell - m = 2p$. It follows that resonance effects are associated with all the $J_{\ell,m}$ for which $\ell - m$ is even. However, due to the distance of a synchronous orbit from the earth, the dominant resonance is that due to $J_{2,2}$. The $J_{2,2}$ resonant perturbation arises through U_{2200}, and I think it is worth repeating that the condition is completely different from the non-resonant condition associated with U_{2210}.

We can interpret $J_{2,2}$ in terms of earth shape, just as we did the J_n zonal harmonics. The interpretation is that the equator is elliptical, with its major axis pointing towards the directions given by $\lambda_{2,2}$ and $\lambda_{2,2} + \pi$. A synchronous satellite which is stationed in either of these directions would be in unstable equilibrium; stationed above either extremity of the equator's minor axis, it would be in stable equilibrium, and above any other point it would start to drift towards the nearer stable point[10]. (One stable point is in the Indian Ocean, and the other is in the Pacific Ocean, west of South America.) The actual value of $J_{2,2}$ is about 1.79×10^{-6} (and of $\lambda_{2,2}$ about -18°) and this is equivalent to a difference between the major and minor equatorial semi-axes of about 69 metres.

Important resonances have occurred for a number of non-synchronous satellites, and determinations of some of the tesseral (and sectorial) harmonics have been based on the study of the resonances. An interesting example was for the inclination of the satellite Ariel 3, plotted (from the RAE orbit determinations by the program PROP) for 840 days in Fig.11. It is immediately clear that, in addition to an oscillatory behaviour due to luni-solar perturbations, there was a marked decrease in i, amounting to about 0.02 degree, early in 1968. In a paper on the Ariel 3 orbit[11], I considered a number of possible explanations of this phenomenon, but unfortunately overlooked the true one entirely, namely that it was due to a $U_{15,15,7,0}$ resonance associated with $J_{15,15}$. The variable ϕ, given by

$$\phi = \omega + M + 15\,(\Omega - \nu) \quad ,$$

became resonant just before 0 hours on MJD 39889 (3 February 1968), and was within 120° of the resonant value during a three-month period centred on this date. By fitting to the values of inclination plotted in Fig.11, an

excellent estimate of the value of $J_{15,15}$ - more precisely of an equation of condition relating $J_{15,15}$, $J_{17,15}$, $J_{19,15}$ etc. - has been obtained. Fig.12 re-plots the Ariel 3 inclinations over 200 days, with a fitted curve representing the $J_{15,15}$ perturbation, and also gives a plot of the resonant variable. It is planned to refine the values of $C_{15,15}$ and $S_{15,15}$, which were obtained in this way, by removing the luni-solar perturbations from the values of i, and then re-fitting.

When a complete set of zonal, tesseral and sectorial harmonics has been derived, it is convenient to summarize it in a contour map, representing the height of the geoid (mean sea level) above a spheroid of some suitable flattening. Fig.13 shows such a contour map based on the 1966 Smithsonian Standard Earth[2]; it differs very little from the latest map for the 1969 Standard Earth[9].

4 THE UPPER ATMOSPHERE

The effects of the upper atmosphere on a satellite orbit are much more difficult to represent, mathematically, than are the effects of the earth's gravitational field. One reason for this is that the gravitational field has a potential function - i.e. the field is conservative - whereas there is no such function for the drag force exerted by the atmosphere. A more important reason, however, is that the earth itself is essentially solid, so that the harmonic coefficients, which specify the potential, are constant (neglecting tides, earthquakes, etc.), whereas the atmosphere is changing all the time. It is still possible to represent the perturbing force mathematically, but the parameters of such a representation - for example the 'density scale height' - can only be treated as constant in a first-order analysis of the orbital perturbations.

In spite of the inherent complexity that we now have to face, the basic expression for the aerodynamic drag force on a satellite is actually very simple, viz.

$$\underset{\sim}{D} = -\tfrac{1}{2}\,\rho\, C_D\, S\, V\, \underset{\sim}{V} \quad , \tag{11}$$

where ρ is the density of the air in the vicinity of the satellite, C_D is the drag coefficient, $\underset{\sim}{V}$ is the velocity of the satellite relative to the surrounding air, and S is the projected cross-sectional area perpendicular to the direction of $\underset{\sim}{V}$. For certain satellites there may also be a lift force, but we ignore this possibility.

If the air is assumed to be stationary, relative to geocentric inertial axes, so that $\underset{\sim}{V}$ is identical with $\dot{\underset{\sim}{r}}$, then $\underset{\sim}{D}$ acts within the orbital plane of the satellite and there is no tendency for the orbital plane to rotate; i.e. with this assumption the motion becomes two-dimensional, within a fixed orbital plane. (An actual orbital plane still rotates, of course, due to gravitational perturbations, but, as I have said, the two classes of perturbations may be considered as essentially independent.) This two-dimensional motion is governed by the air density ρ, as it varies along the orbit, and so information about ρ can be inferred by studying the perturbations in the motion. I propose to give a brief survey of this subject, and then to conclude with an even briefer survey of how atmosphere-induced rotations of the orbital planes of various satellites have been observed, with a surprising corollary about the rotational speed of the atmosphere.

4.1 Air density

Equation (11) showed that the drag force is proportional to the air density, and air density is primarily a function of height. Thus[12] average values of ρ are 10^{-6} kg/m^3 at about 100 km height, but less than 10^{-14} kg/m^3 at 1000 km. The variation with height, h, is approximately exponential, so that we have

$$\rho = \rho_o \exp \frac{h - h_o}{H} , \qquad (12)$$

where H is the 'density scale height' and ρ_o is the density at a reference height h_o. To a first approximation H is taken as constant - and would be about 50 km to reproduce the average variation from h = 100 km to h = 1000 km that has been quoted - but if better accuracy is desired H must itself be taken as a function of height.

Let us consider what happens to a satellite in an orbit of moderate eccentricity, say for which $0.02 < e < 0.2$. (It is the $e > 0.02$ part of the inequality which concerns us here.) It follows that the effect of air drag is effectively concentrated within a small arc of the orbit around perigee. Loss of energy near perigee means that the next apogee is lower than the previous one, and the evolution of the orbit is as shown in Fig.14. The orbit contracts and becomes more circular; i.e. both a and e decrease with time. It is interesting to note the slightly paradoxical fact that the height of perigee, where nearly all the drag is, decreases only very slowly, whereas the height of apogee, where there is virtually no drag, decreases relatively rapidly.

I cannot resist remarking on a further paradox which people sometimes find difficult to understand: the immediate effect of drag is that a satellite is retarded and its orbit contracts, but a contracted orbit means (by Kepler's third law) a shorter orbital period, so that the satellite has actually been accelerated. The paradox cannot be resolved by saying that a contracted orbit means the satellite has less far to go; it really does travel - on average - faster. The explanation, of course, is that a (genuine) retardation at perigee involves a loss of energy and that it is at the next apogee that - relative to the preceding one - a speeding up occurs. If, from one apogee to the next, an amount of total energy ΔE has been lost, this corresponds to a loss of potential energy amounting to $2\,\Delta E$, so that there is actually a gain of ΔE in kinetic energy.

It is the decreasing value of T, the orbital period, that is the most easily and accurately measured perturbation of an orbit. The rate of change of T is directly proportional to ρ_o, the reference density in equation (12), so that from observations of $\dot{T}$ it is possible to infer ρ_o. The formula is[13]

$$\rho_o = -0.157\,\frac{\dot{T}}{\delta}\left(\frac{e}{aH}\right)^{\frac{1}{2}}\left\{1 - 2e + 2\tfrac{1}{2}e^2 - \frac{H}{8ae}\left(1 - 10e + \frac{7H}{16ae}\right) + O\left(e^3, \frac{H^3}{a^3e^3}\right)\right\}, \qquad \ldots (13)$$

where $\delta = FSC_D/m$; here F is a factor which can be introduced to allow for atmospheric rotation and m is the mass of the satellite (which no longer cancels out in the equations of motion, as it does for gravitational perturbations). It is assumed, still, that $0.02 < e < 0.2$, and the formula is derived by truncating asymptotic expansions of Bessel functions of argument ae/H. The reference height, h_o, to which this value of ρ_o applies, is not the height of the satellite's perigee, but a height $\frac{1}{2}H$ above perigee. The advantage of a formula which gives density at height $h_p + \frac{1}{2}H$ is that it is less sensitive to errors in H than a formula for density at height h_p; for example, an error of 25% in H gives an error of only 1% in ρ. (Since e is assumed to exceed 0.02, apogee height exceeds perigee height by at least 250 km, so the air at height $h_p + \frac{1}{2}H$ really is being visited!) The overall accuracy for absolute values of density, as given by equation (13), is usually about 10%, due to lack of an exact value for C_D. But variations in density can be measured much more accurately, often to better than 2%.

Discoveries about the density of the upper atmosphere, based entirely on the use of equation (13) for many satellites (in almost every case as a 'non-operational objective'), have been remarkable. The first discovery was that the atmosphere was far less tenuous than had been supposed prior to the satellite era. The second was that the density, at the same height above the same point on the ground, can be incredibly different at one time from another; at 500 km height, for example, the maximum value of ρ can be as much as 200 times greater than the minimum value.

The variation of atmospheric density with time is under the control of the sun, and this control acts in a number of ways, not all of which are fully understood as yet. Some of the control mechanisms may be regarded as acting 'directly', while others act 'indirectly', and an example of each of these may be seen in the plot of T, for the satellite Explorer 1 (h_p about 350 km), in Fig.15. At heights between 200 km and 1000 km density increases during the (local) morning, as a result of solar heating, and decreases during the afternoon and evening, so that there is a 'day-to-night' variation which is 'indirect' in the sense that it is not associated with any change in the sun itself. This effect does not manifest itself diurnally, however, because the earth is rotating underneath the satellite orbit; it is the angle between the sun and the satellite perigee, as seen from the centre of the earth, which is important, and the period of a complete cycle of variation of this angle is usually several months, about nine in the case of Explorer 1. During each cycle the rate of decrease of T - and hence the atmospheric density - is greater during the periods of perigee 'day' than during the periods of perigee 'night'. The 'direct' effect in Fig.15 is visible as the steady reduction in the rate of decrease of T from 1958 to 1962. This can only mean that the satellite perigee, as it very slowly descended, sampled less and less dense air, contrary to what might be expected. The explanation lies in the gradual reduction in solar activity as the sun moved from the maximum of the 10- or 11-year sunspot cycle, in 1958, towards the minimum, in 1964. At heights above 200 km the atmospheric density is directly related to the solar activity; as the extreme ultraviolet radiation increases, and the particles ejected by the sun become more numerous and more energetic, the air heats up and hence becomes denser. The sunspot-cycle variation in density is greater than the day-to-night variation, as can be seen from the four plots of density against height shown in Fig.16.

Besides the correlation with the sunspot cycle, the direct response of the atmosphere to solar activity shows itself in two other ways, both of which may be seen in the plot of $\dot{T}$ for Explorer 9 (h_p about 700 km) in Fig.17. The peaks in this plot correspond to violent storms on the sun, alternative evidence for which is given by the geomagnetic index a_p, which is probably the best indicator of the effect of solar particles impinging on the earth and is also plotted in Fig.17; the correlation between a_p and $\dot{T}$ is very striking. (Similar striking correlations between $\dot{T}$, for various satellites, and the energy of solar radiation at wavelength 107 mm have been observed.) If we ignore the peaks, the other feature of the $\dot{T}$ plot in Fig.17 becomes clear, namely the 27-day cycle. A cycle of this period is characteristic of solar phenomena, since it is the period of the sun's rotation.

Finally, in this summary of solar-induced variations in the density of the upper atmosphere, there is the 'indirect' effect known as 'semi-annual variation'. Though the effect is believed to arise from variations in the lower atmosphere[14], it is most clearly visible high in the upper atmosphere. Fig.18 shows the effect for Echo 2 (h_p about 1100 km). The cause of the semi-annual variation is still uncertain, but its features are well established: the density at a given height, after correction for all other sources of variation, exhibits a residual oscillation which has maxima in early April and late October and minima in mid-January and late July; the maximum in October is usually - though not in Fig.18 - higher than that in April, and the minimum in July is usually lower than that in January.

4.2 Atmospheric rotation

I have already remarked that, if it were not for the rotation of the atmosphere (and for the possibility of lift forces), there would be no tendency for atmospheric forces to rotate the orbital plane of a satellite. Thus if any slight rotation of an orbital plane is observed, after correction for non-atmospheric perturbations, it will provide direct evidence - not merely that the atmosphere rotates, which is obvious - but of the magnitude of the rotation.

Suppose, then, that the atmosphere has an angular velocity which is Λ times greater than that of the earth itself. Here Λ may be assumed to be a function of height, equal to 1.0 when $h = 0$; we might suppose that as h increases Λ would decrease, since the frictional effect of the earth would wear off, but this turns out to be wrong.

Fig.19 gives a vector diagram for velocities and drag forces as a typical satellite crosses the equator from south to north. By resolving the drag force $\underset{\sim}{D}$ within and perpendicular to the orbital plane, it is seen that there is a small lateral force which tends to decrease the inclination of the orbit. Half an orbit later, as the satellite crosses the equator from north to south, the force acts in the opposite direction, but this means that its tendency is still to decrease the inclination. In fact $\frac{di}{dt} > 0$ all round the orbit, except at the north and south apexes where $\frac{di}{dt} = 0$ (and except for an equatorial orbit, for which $\frac{di}{dt} = 0$ everywhere). The integrated effect around an orbit can be expressed, like the ordinary atmospheric perturbations, in terms of Bessel functions; the resulting change in the inclination, Δi say, contains ρ_o (the density at height $h_p + \frac{1}{2}H$) as a factor, and ρ_o can then be expressed in terms of $\dot{T}$, by equation (13). The final formula is best expressed in terms of ΔT, the change in T over a long interval of time, and it is given by[13]

$$\Delta i = \frac{\Lambda \, \Delta T \sin i}{3\sqrt{F}} \left\{ (1 - 4e) \cos^2 \omega - \frac{H}{ae} \cos 2\omega + 0 \left(e^2, \frac{H^2}{a^2 e^2} \right) \right\} . \quad (14)$$

Thus if Δi is observed, for a suitable satellite for which an appreciable ΔT has also been observed, a value for Λ can be estimated. (Formulae for $\Delta\Omega$ and $\Delta\omega$ can also be obtained, but are of less practical use.)

Values of Λ for 27 satellites are shown in Fig.20 with the extraordinary indication that Λ increases with height. The standard deviation of each determination of Λ is large, due to the smallness of the values of Δi used, but the effect must be regarded as a genuine one, and it has been confirmed more recently for other satellites[15]. As an example of the care which has to be exercised, however, I can point to the Ariel 3 tesseral-harmonic resonance I described earlier; if this were overlooked, it would be possible to obtain a value for Λ of about 2.3, taking a Δi of about -0.018° over two years from Fig.11, but when the resonant effect is allowed for the value obtained for Λ seems to be less than 1.

A number of attempts have been made to explain the observed increase of Λ with height. Challinor's model[16] of ionospheric winds, for example, leads to $\Lambda = 1.1$ at 200 km, rising to nearly 1.5 at 350 km, and this is in excellent agreement with Fig.20.

Acknowledgement

I am grateful to D. G. King-Hele for allowing me to use a number of illustrations from his publications.

REFERENCES

No.	Author(s)	Title, etc.
1	R. H. Merson	A brief survey of satellite orbit determination. RAE Technical Report 66369 (1966) Phil. Trans. Roy. Soc., A, 262 (1124), p.71 (1967)
2	C. A. Lundquist G. Veis	Geodetic parameters for a 1966 Smithsonian Institution Standard Earth. Smiths. Astr. Obs. Sp. Report 200 (1966)
3	D. G. King-Hele G. E. Cook	The even zonal harmonics of the Earth's gravitational potential. RAE Technical Report 65090 (1965) Geophys. J., 10, p.17 (1965)
4	D. E. Smith	A determination of the even harmonics in the Earth's gravitational potential function. Planet. Space Sci., 13, p.1151 (1965)
5	Y. Kozai	Revised values for coefficients of zonal spherical harmonics in the geopotential. Smiths. Astr. Obs. Sp. Report 295 (1969)
6	D. G. King-Hele G. E. Cook Diana W. Scott	Evaluation of odd zonal harmonics in the geopotential, of degree less than 33, from the analysis of 22 satellite orbits. RAE Technical Report 68202 (1968) Planet. Space Sci., 17, p.629 (1969)
7	Y. Hagihara	Recommendations on notation of the earth potential. Astron. J., 67, (1), p.108 (1962)
8	R. R. Allan	Resonance effects due to the longitude dependence of the gravitational field of a rotating primary. RAE Technical Report 66279 (1966) Planet. Space Sci., 15, p.53 (1967)
9	E. M. Gaposchkin K. Lambeck	1969 Smithsonian Standard Earth (II). Smiths. Astr. Obs. Sp. Report 315 (1970)
10	R. R. Allan	On the motion of nearly synchronous satellites. RAE Technical Report 64078 (1964) Proc. Royal Soc., A, 288, p.60 (1965)

REFERENCES (Contd)

No.	Author(s)	Title, etc.
11	R. H. Gooding	The orbit of Ariel 3 (1967-42A). RAE Technical Report 69275 (1969)
12	D. G. King-Hele	The upper atmosphere and its influence on satellite orbits. RAE Technical Report 69102 (1969) Dynamics of Satellites 1969 (ed. B. Morando), p.249, Springer Verlag, Berlin (1970)
13	D. G. King-Hele	Theory of satellite orbits in an atmosphere. Butterworths, London (1964)
14	G. E. Cook	The semi-annual variation in the upper atmosphere: a review. RAE Technical Report 69074 (1969) Annales de Géophysique, 25, p.451 (1969)
15	D. G. King-Hele Diana W. Scott Doreen M.C. Walker	Upper-atmosphere rotational speed and its variation with height. RAE Technical Report 69263 (1969) Planet. Space Sci., 18, p.1433 (1970)
16	R. A. Challinor	Neutral-air winds in the ionospheric F-region for an asymmetric global pressure pattern. Planet. Space Sci., 17, p.1097 (1969) and 18, p.1485 (1970)

OPERATIONAL OBJECTIVES	NON-OPERATIONAL OBJECTIVES
ORBIT ACHIEVEMENT	GENERAL SURVEILLANCE
GUIDANCE AND CONTROL	SENSOR ACCURACY
ON-BOARD INSTRUMENTATION	NAVIGATION AND GEODESY
TRACKING / TELEMETRY SCHEDULING	EARTH'S GRAVITY FIELD
	AIR DENSITY / ROTATION

Fig.1

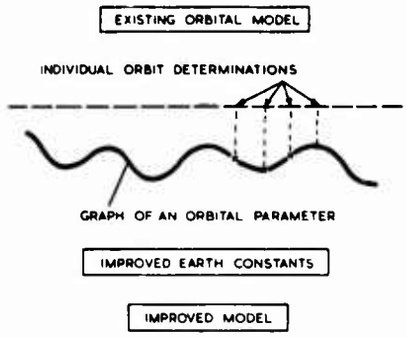
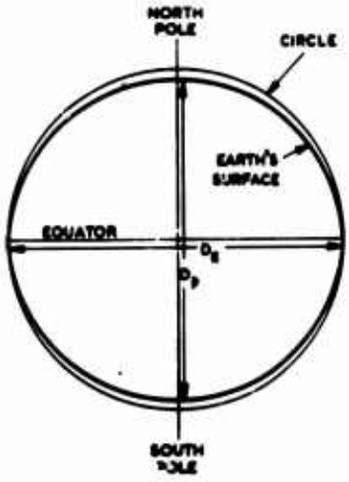
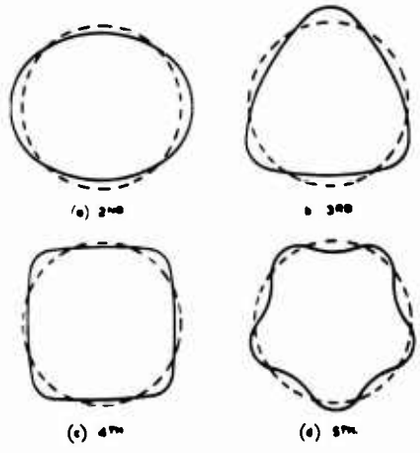


Fig.2



FLATTENING OF THE EARTH (GREATLY EXAGGERATED). THE FLATTENING f MAY BE DEFINED AS $f = \frac{D_E - D_P}{D_E}$

Fig.3



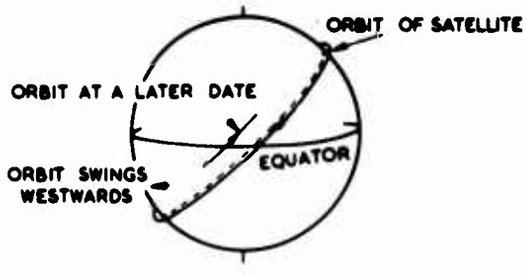
FORM OF THE 2ND-5TH HARMONICS (NOT TO SCALE)

Fig.4

Values of even zonal harmonics

Authors:	King-Hele & Cook (1965)	Smith (1965)	Kozai (1969)
$10^6 J_2$	1082.64	1082.64	1082.63
$10^6 J_4$	-1.52	-1.70	-1.59
$10^6 J_6$	0.57	0.73	0.50
$10^6 J_8$	0.44	-0.46	-0.12
$10^6 J_{10}$		-0.17	-0.35
$10^6 J_{12}$		-0.22	-0.04
$10^6 J_{14}$		0.19	-0.07
$10^6 J_{16}$			0.19
$10^6 J_{18}$			-0.23
$10^6 J_{20}$			-0.01

Fig.5



THE EFFECT OF THE EARTH'S EQUATORIAL BULGE ON A SATELLITE ORBIT

Fig.6

Fig.1-6

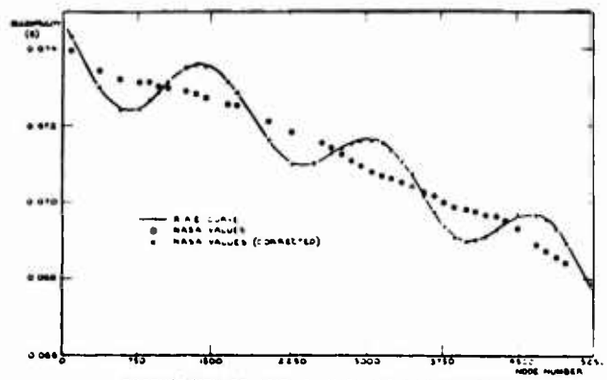
NEG.NO.C7461 T.MEMO SPACE 157

Fig.7-12

Values of odd zonal harmonics (1969)

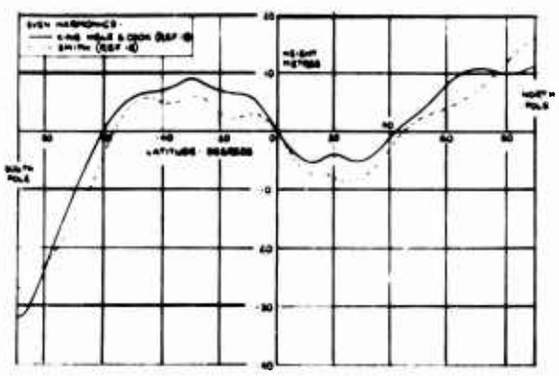
Authors:	King-Hele Cook & Scott	Kozai
$10^6 J_3$	-2.54	-2.54
$10^6 J_5$	-0.21	-0.23
$10^6 J_7$	-0.40	-0.36
$10^6 J_9$	0	-0.10
$10^6 J_{11}$	0	0.20
$10^6 J_{13}$	0	-0.12
$10^6 J_{15}$	-0.20	-0.17
$10^6 J_{17}$	0	0.08
$10^6 J_{19}$	0	-0.22
$10^6 J_{21}$	0.26	0.15
$10^6 J_{27}$	-0.15	
$J_{23}, J_{25}, J_{29}, J_{31}$	0	

Fig.7



ECCENTRICITY OF ARIEL 2 - NASA AND R A E VALUES COMPARED

Fig.8



HEIGHT OF THE GEOID RELATIVE TO A SPHEROID OF FLATTENING 1/298.25, AS GIVEN BY 7-COEFFICIENT SET OF ODD ZONAL HARMONICS

Fig.9

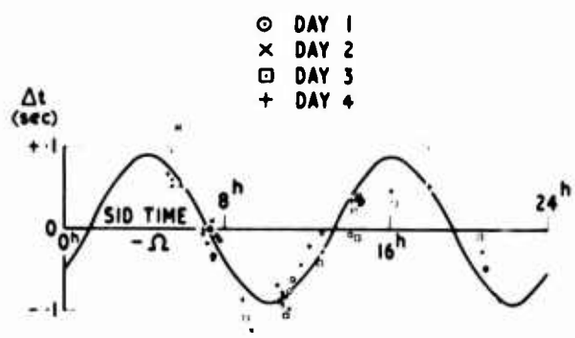


Fig.10

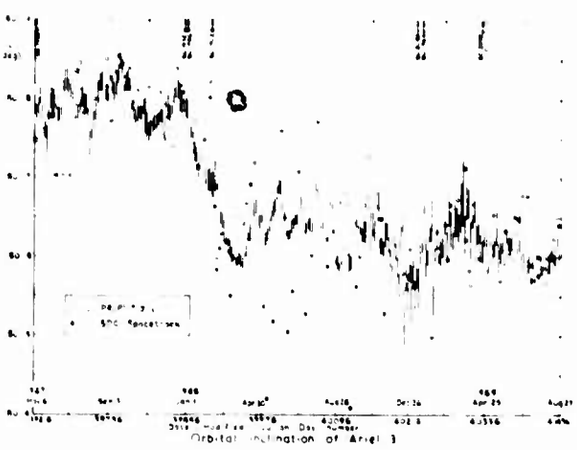
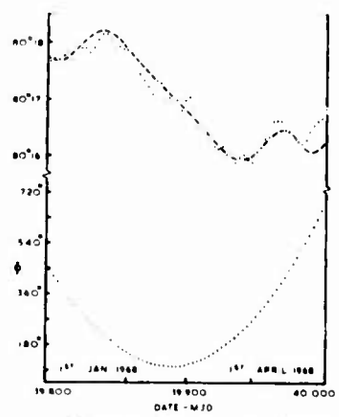


Fig.11

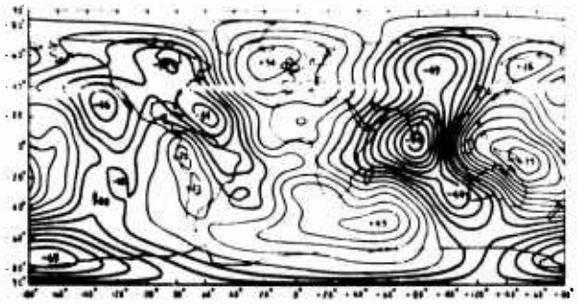


ORBITAL INCLINATION AND RESONANT VARIABLE FOR ARIEL 3 (1967-42A)

Fig.12

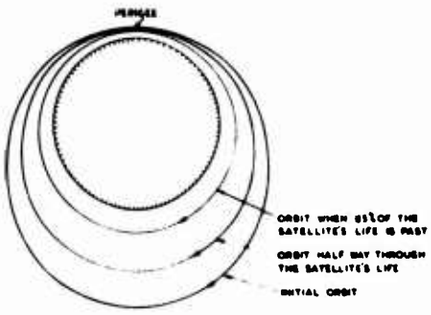
Fig.7-12

NEG.NO.C7462 T.MEMO SPACE 157



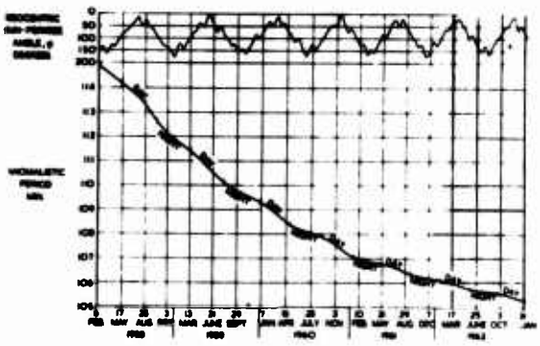
Smithsonian standard Earth 1966 Contours of geoid height at 10-m intervals, relative to spheroid of flattening 1/298 252

Fig.13



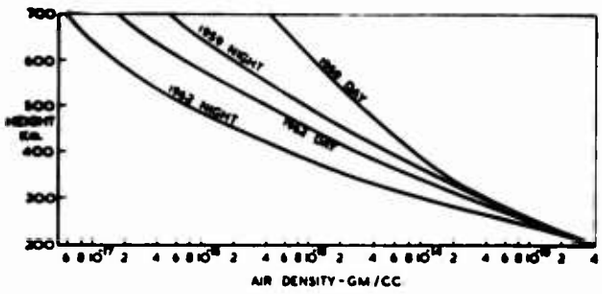
CONTRACTION OF SATELLITE ORBIT UNDER THE ACTION OF AIR DRAG

Fig.14



ORBITAL PERIOD AND GEOCENTRIC SUN-PERIGEE ANGLE FOR EXPLORER I, 1958 α1

Fig.15

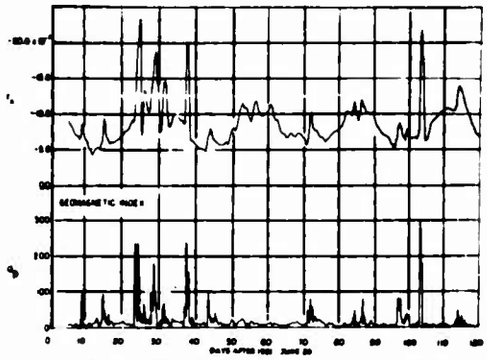


DAY-TO-NIGHT VARIATION FOR 1959 AND 1962.

Fig.16

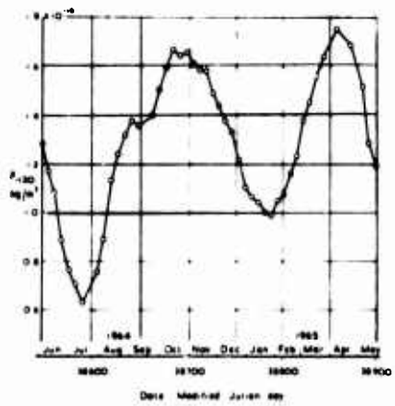
NEG.NO.C7463 T.MEMO SPACE 157

ig.17-20



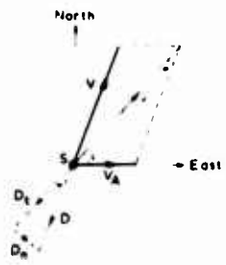
RATE OF CHANGE OF ORBITAL PERIOD $\dot{\tau}_a$ DUE TO AIR DRAG, FOR EXPLORER 9, AND GEOMAGNETIC INDEX a_p

Fig.17



Air density in 1964-5 at a height of 1130 km, ρ_{1130}, from Echo 2, after correction for solar activity

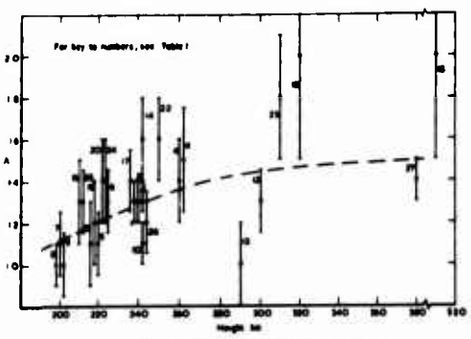
Fig.18



U - Velocity relative to Earth's centre
V = Velocity relative to air

Kinematics of spherical satellite S in a circular orbit as it crosses the equator

Fig.19



Values of A obtained from 27 satellites

Fig.20

Fig.17-20

NEG.NO. C7464 T.MEMO SPACE 157