Geodetic Applications of Artificial Satellites

by

David S. Evans

(Based on an address to the Institute of Government Land Surveyors of the Cape Province on 1962 January 26)

Ordinary survey methods rely on the connection by means of triangulation or trilateration of mutually visible points on the surface of the Earth. The connection of separate survey systems across wide water barriers, and, in particular, the linking of surveys of separate continents is difficult or impossible by conventional methods. Each separate system will, of course, be defined in terms of angles at the centre of the Earth by astronomical fixes of reference points in the system, but the study of deviations of the figure of the Earth from sphericity has required the measurement on the surface of long chains of triangulation, and to these, water barriers too wide to be traversed at several points used to prove an obstacle.

Among the motives for improvement, the military has always ranked high. I have been told that, even as late as World War II, the relations between the British and French trigonometric surveys were so uncertain that cross-Channel gunnery was an erratic business. Now, with intercontinental missiles a reality, the aiming of a missile from within one system to a target in another requires an accurate knowledge of the relation between the systems. Less grisly and more peacable needs are those of commercial air navigation by radar systems. When first introduced, the accuracy of the these outran the accuracy of maps. This is no longer true of, for example, the entire continental area of Western Europe, but probably remains true of large ocean areas such as the Pacific, where isolated islands are used for the operating stations.

The idea of using objects outside the surface of the Earth which would be simultaneously visible from mutually invisible stations is not new. To give a local example, Sir John Herschel, during his stay here from 1834 to 1838, discussed the use of rockets, and even considered the use of lunar eclipses where the arrival of the shadow edge at a particular crater could provide a well-defined extraterrestrial point both in space and time.

Proposals for survey linkage across the Gulf of Bothnia using eclipse observations were, if my recollection serves me, made during the

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

'forties. From the same part of the world came a very forward looking proposal by Vaisala of Turku in Finland, when, in 1946 he wrote, "If rocket missiles can be developed to such a degree that it would be possible to realise small moons which would circle the Earth at an altitude of some thousands of kilometres, with a period of only several hours, we should obtain practically eternal light sources for a giant triangulation, and these light sources could also be used for physical measurements of the Earth. A simple calculation reveals that an artificial moon several decimetres in diameter could be followed with medium-sized apparatus. The light flashes necessary for accurate observations would be furnished by the artificial moon if one half of it were white, the other black, and if it were given a suitable rotary motion". I do not know whether, as seems likely, this was the original proposal for the launching of artificial satellites, but it is certainly a fully formed proposal for their use in geodesy, prophetic of much that has now actually happened. In the event, only a very few satellites have been launched with the primary aim of furthering geodesy, but many have served geodetic purposes, even though they have been primarily designed for other physical investigations. The satellite provides an artificial marker in space, and to understand how this may be used we must recapitulate some of the properties of satellite orbits introducing more and more complications as we go along.

The simplest case is that of a perfectly spherical Earth, which may be replaced by a point mass at its centre, round which moves a satellite in a circular orbit. It is easily shown that in c.g.s. units, with the period, $\underline{P}$, in seconds and the orbital radius, $\underline{a}$, in centimetres, we have

$$P = 3.13 \times 10^{-10}\ a^{3/2}$$

We have neglected the resistance of the atmosphere, and the deviations of the Earth from a spherical form. With these simplifications we find that for a satellite just skimming the surface, the period is 83.3 minutes: at a height of 100 kms it becomes 85.9 minutes, while for heights of 500, 1000 and 5000 kms the values are respectively 94.0, 104.1 and 200.0 minutes. The satellite with a period of 24 hours (i.e. one which remains stationary in the sky) moves in a circular orbit of radius 42,390 kms from the centre of the Earth, and our own natural Moon, with an average value of $\underline{a}$ corresponding to 384,400 kms has a sidereal period of 27.3 days, as it should. For the larger values the fact that we have drastically simplified the problem makes little difference, since the height of the atmosphere is limited and at great distances the gravitational approximation improves. Under these simplifying assumptions the orbit is unchanging and eternal. In practice, atmospheric drag and higher terms in the expression for the Earth's gravitational potential both produce changes in the orbital elements of an artificial satellite. Fortunately the effects due to these two causes are largely independent of each other.

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

The general orbit of a satellite is an ellipse of eccentricity, e, with the centre of the Earth in one focus. The distance of the satellite from the centre of the Earth ranges from r_{max} at apogee, to r_{min} at perigee, where

$$r_{max} = a (1 + e) \quad \text{and} \quad r_{min} = a (1 - e)$$

and the relation between a and period is that already stated. The orbit is in a plane, and, to a high order of accuracy, the effects of air drag leave this plane unaltered. The density of the atmosphere decreases with height in a complicated way (the details of which at very great heights have been discovered from satellite observations). The nearest simple function describing the variation is an exponential, but the approximation is very crude. However, the reduction with height is rapid, so that at 120 miles above the surface the density is 4×10^{-10} of its sea-level value, and at 260 miles it has fallen to one hundredth of this. Orbital velocity is higher, and air density very sharply higher at perigee than at apogee, so that for an eccentric orbit the effects of air drag may be closely approximated by the model of an impulsive retardation at each perigee. Each successive circuit may be thought of as a new launching from a fixed perigee with a projection velocity which falls with increasing cycle number. This also involves a reduction of period, for if r_{min} is fixed, and e is decreasing, a must also be decreasing. King-Hele who has played an outstanding part in the discussion of satellite orbits has given a simple formula for the further life of a satellite, t_L from the moment when the eccentricity has a value e_o,(this must not be too large or small),and the period a value P_o.

This is

$$t_L = 3 \; e_o \; P_o \; / \; 4 \; x$$

where t_L is in days and x is the rate of decrease of period per day. Once the "rounding-up" process has reduced the eccentricity to near zero, the air drag is operative all the time, and tends to reduce the orbital velocity; this causes the satellite to fall in, picking up more energy from the gravitational field, so that it moves faster and at a smaller height. The satellite thus spirals in, with orbital speed increasing to a maximum, until it is completely braked by air drag and falls to the Earth's surface.

The semi-axis major, a, of the orbit, and the eccentricity, e, define the size and shape of the satellite's path. The further elements needed for complete definition of the orbit are as follows:

The inclination, i: This is the angle between the orbital plane and the equatorial plane of the Earth. It is also the limiting latitude, north or south, reached by the sub-satellite point.

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

The right ascension of the ascending node, Ω. This is the right ascension of the direction from the centre of the Earth to the point on the Earth's equator, (the ascending node), where the sub-satellite point crosses from south to north. It will be noticed that this definition gets over difficulties consequent upon the rotation of the Earth.

The argument of perigee, ω. This is the angle at the centre of the Earth measured in the direction of the satellite's motion, from the ascending node to perigee.

Many of the changes in these elements are due to higher order terms in the expression for the Earth's gravitational potential. Before we discuss them it is well to interpolate two remarks. The first is that the number of physical effects leading to perturbations of the motion of a satellite is very large, so that a satellite is an exceedingly sensitive probe of very many different features of the physical environment, whose effects have to be carefully separated. Merely as instances of the minor effects we may mention radiation pressure from the Sun, perturbations due to the gravitational effects of the Sun and Moon, electro-magnetic forces due to the motion of a metallic body in the magnetic field of the Earth, and sideways drift due to the passage of the satellite through the rotating upper atmosphere. Various methods of analysis can be devised so as to reveal such effects separately, although some are exceedingly complicated.

The second remark is that there are two quite distinct ways of using satellites for geodetic purposes. One is to separate those changes of orbital elements which are due to departures of the Earth from sphericity, and to infer from them data on the figure of the Earth. The other is to make simultaneous observations of the same satellite from widely spaced stations and to use the position of the satellite as a third dimensional marker to link separated geodetic systems. The latter would require no data about the satellite except its instantaneous position, but the method is almost impossible to use in practice, and is of limited utility as well, since the separation of the ground stations from which observations can be made simultaneously, even of a high satellite, is limited. It must be replaced by a much more complex method, namely, to evaluate all the perturbations of the orbit, so that an accurate ephemeris can be produced and used for observations made from stations widely separated on the Earth's surface at different times.

So far the first method has yielded the most important results. The Earth approximates to an oblate spheroid, and the first order effects of the extra mass round the equator are to cause changes both in Ω and ω, which, in terms of the classically accepted numerical values may be expressed as follows:

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

The inclination stays practically constant but the right ascension of the ascending node decreases at the rate of

$$\dot{\Omega} = 9.97 \left(R/a \right)^{3.5} \cos i \, / \, \left(1 - e^2 \right)^2$$

degrees per day, where R is the equatorial radius of the Earth and the value of the numerical constant depends on the assumed value of the polar flattening of the Earth. It will be noted that for a polar orbit the rotation is zero, and that for small inclinations, (Ω is not definable for an exactly equatorial orbit), the rate can approach ten degrees per day. This is the expression for the rate at which the orbital plane rotates. In addition, the orbital ellipse turns in its own plane so that perigee occurs successively at all latitudes accessible. The expression for the rate of change of ω is

$$4.98 \left(R/a \right)^{3.5} \left(5 \cos^2 i - 1 \right)$$

degrees per day, and is thus positive for small inclinations, negative for high, and zero for $i = 63°.4$. The inclinations of the early Russian satellite orbits were all near 65° so that the rate of change of the argument of perigee was very small.

Since the numerical values in these expressions depend on the assumed flattening of the Earth, the value of this and of higher order terms can be deduced by adopting a general expression and finding the parameters which best fit the observed behaviour of satellites. The analysis is due to King-Hele. He takes the case of an axially symmetric gravitational field and writes the gravitational potential in terms of Legendre polynomials, as follows:

$$U = \left(GM/R \right) \left\{ R/r - \sum_{n=2}^{\infty} J_n \left(R/r \right)^{n+1} P_n \left(\cos \theta \right) \right\}$$

where G is the gravitational constant; M, the mass of the Earth; R, its equatorial radius; P_n, the Legendre polynomial of order n; , the angle between the radius vector r, and the axis of the Earth; and the numbers, J_n, are constants whose values we seek. The Earth's equator is assumed circular. Before the satellites, only J_2 was known with any accuracy, (to about one part in 300), and none of the other constants J_n was known. Using the foregoing formula the rate of rotation of the orbital plane of a satellite can be evaluated as

$$\dot{\Omega} = \left(GM/a^3 \right)^{\frac{1}{2}} \left(R/p \right)^2 \cos i \left\{ \frac{3}{2} J_2 + \frac{9}{4} J_2^2 \left(R/p \right)^2 \left(\frac{19}{12} \sin^2 i - 1 \right) \right\}$$

$$+ \frac{3}{8} J_3 \, e \left(R/p \right) \sin \omega \, \operatorname{cosec} i \left(15 \sin^2 i - 4 \right)$$

$$+ \frac{15}{16} J_4 \left(R/p \right)^2 \left(7 \sin^2 i - 4 \right) + \text{higher terms},$$

where $p = a \left(1 - e^2 \right)$

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

The parameter p changes by only about 10% during the life of a typical satellite and i hardly changes at all, so that the coefficients of J_2, J_2^2 and J_4 are nearly constant for a given satellite. The term in J_3 contains the factor e sin ω, which is small and variable, and averages out to zero over a sufficiently long period. The observed value of λ for a given satellite thus gives a single equation relating J_2 and J_4. To get the values of these constants, satellites with different inclinations are required. Results by King-Hele's group and others based on Sputnik II, Vanguard I and Explorer IV which had orbital inclinations of 65°, 34° and 50° respectively gave values for the constants

$$\begin{aligned} J_2 &= (1083.1 \pm 0.2) \times 10^{-6} \\ J_4 &= (-1.4 \pm 0.2) \times 10^{-6} \\ J_6 &+ (0.1 \pm 1.5) \times 10^{-6} \end{aligned}$$

Somewhat different values have been given by Kozai.

The value of J_3 has to be obtained in a different way. It can be shown that there should be an oscillation in perigee height in the same period as that in which ω changes by 360° and the theoretical expression for this gives a value for J_3 in terms of J_2. Work on this topic has been done by O'Keefe and his associates and the deduced value of J_3 is about $- 2.2 \times 10^{-6}$. The value of polar flattening, f, is expressible in terms of J_2 and the ratio of centrifugal to gravitational force at the Earth's equator. The satellite results lead to a value $1/f = 298.20 \pm 0.03$, corresponding to decidedly less flattening than had previously been supposed. The values previously found from investigations based on gravity surveys, measurement of arcs, the motion of the Moon, and the value of the precessional constant had led to a mean value of 297.1, with considerably less accuracy than provided by the satellites. The value of J_3 was obtained by studies of the perigee heights of Vanguard 1, which, because of its long life and stable orbit, has been the most useful geodetically of all the satellites. The secular decrease in perigee height for this satellite is of order of 5 miles per year but oscillations over a range of 6 miles are detectable during the 82 day period of a complete rotation of perigee and give the value of J_3 quoted above. The existence of this term implies that the Earth deviates from an oblate spheroid, and is more flattened at the south pole than the north. According to King-Hele the best dimensions are

Equatorial radius	3963.18 miles
Centre to N. Pole	3949.90 miles
Centre to S. Pole	3949.88 miles

This means that the North pole lies about 50 feet higher than it would on a spheroidal Earth, and the South pole about 50 feet lower.

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

As O'Keefe, Eckels and Squires remark this implies a considerable crustal load on the Earth and contradicts the "basic hypothesis of geodesy" of Vening-Meinesz and Heiskanen that the Earth's gravitational field is very nearly that of a fluid in equilibrium.

I now turn to the problem of the use of satellites as geodetic markers and rely on a comprehensive and complicated memoir by George Veis ("Geodetic Uses of Artificial Satellites", Smithsonian Contributions to Astrophysics, Vol 3, No 9, 1960), and on a more recent undated preprint entitled "Experience in Precision Optical Tracking of Satellites for Geodesy" by George Veis and Fred L. Whipple. Here the main reference is to optical tracking by photographic means, with positions determined to high accuracy. It may be mentioned in passing however that even Minitrack, that is, radio tracking, of Vanguard I, which only gave an accuracy in right ascension measurements of something better than one minute of arc has enabled improvements in map positions of Minitrack stations on Pacific Islands to be made. Map errors of as much as a mile were found from studies carried out on Guam, Wake Island, Samoa, Ponape, Kwajalein and Clark Field in the Philipines. It is thought that map errors can be brought down to 240 feet for these isolated islands. This is not geodetic accuracy, but is important since some of the islands are used as Loran navigational bases. It is worth reminding ourselves that Minitrack observations can be made under much more general conditions than those necessary for photographic observation where the satellite must be in sunlight and the observer in the dark or twilight.

Turning to the problems discussed by Veis, and by Veis and Whipple, I find it impossible to do more than pick out a few of the more important points made. To do justice to the intricate and closely argued discussion of the necessary coordinate transforms reference must be made to Veis' memoir. The first problem is the photography of the satellite and the determination of its apparent place with respect to the star background. A variety of methods is possible. If the instrument follows the stars, a chopper can be used to interrupt the satellite trail for timing purposes, but this may introduce difficulties if the satellite is tumbling and itself flashing in consequence. The most common practice employed with the Baker-Nunn camera, in cases where the apparent track is fairly accurately known, is to train the camera on the satellite and let the stars appear as interrupted trails. This will in particular enhance the apparent brightness of a faint satellite, and has enabled even faint objects such as Vanguard I, with a diameter of only 6 inches, to be photographed at distances of 3000 miles.

Once a photograph has been obtained the determination of the apparent position of a satellite with respect to the stars is very similar to that of plate measurement and reduction for position in ordinary astronomical practice. In some cases where the height of the

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

satellite is small the refraction of the satellite image may be less than that of the stars, since the rays have not traversed all the atmosphere. Veis gives a very full theory of this but the effect is negligible for heights over 100 kms. In addition a correction for the aberration due to the satellite's velocity is required. The timing of the observations may present a problem. Satellites are so fast moving that the utmost accuracy is necessary, and the time system in which timings are made must be that which defines the direction of the radius through longitude zero in terms of the fixed spatial coordinate system. Since time signals are in the system UT2 which is corrected for seasonal variations of the rotation of the Earth they must be corrected back to UT1 which only incorporates corrections for motions of the pole. Alternatively atomic time A1 has been used.

The scale of the Baker-Nunn cameras is 406"/mm and the measuring accuracy attainable is ± 2". Timing accuracy is such that an accuracy of 0s.001 must be considered very good. The result of a reduction is to give the apparent topocentric right ascension and declination of a satellite with reference to the mean equinox and equator of 1950.0. The term, "topocentric", seems to have been coined by Veis and means a system of rectangular coordinates with origin at the place of observation which are parallel to the geocentric coordinates. The relation between the two coordinate systems thus depends on the figure of the Earth. It is stated that the standard deviation of good observations of satellite positions is ± 3".3.

Twelve Baker-Nunn stations in the latitude zone from $-32°$ to $+36°$ are in existence at Organ Pass, Olifantsfontein, Woomera, San Fernando, Tokyo, Naini Tal, Arequipa, Shiraz, Curaçao, Jupiter, Villa Dolores, and Maui, and each of these has been connected to one of the major geodetic systems by ordinary methods of triangulation, traversing and levelling. It has now proved possible by satellite observations to relate all these camera stations to a uniform system based on the International Ellipsoid.

The position of the Australian station has been determined with an uncertainty of about 100 metres, though not with equal accuracy in all directions, and Kozai has obtained the position of the Spanish station with an accuracy of 9 m in both latitude and longitude, with a somewhat poorer determination of the altitude.

It is clear that this work involves an enormous amount of computation with the IBM 7090 computer, and that the work progresses by successive approximations. There seems to be no limit to the accuracy eventually attainable since observations of some long-lived satellites run into many thousands and each, if properly used can add weight to what has gone before. The geodesist can look forward to the establishment of a world wide system with an accuracy undreamed of before the satellites.

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)