# ARTIFICIAL SATELLITES AS AN AID TO NAVIGATION

by Dr. H. C. FREIESLEBEN

Artificial satellites can be used in six different ways to fix a position and thus may serve navigational purposes. Measurements of their direction, i.e. of the elevation and azimuth angles, and those of the changes in both angles can be carried out; distances can be determined and, last but not least, the changes of distance can be measured. Out of these six different possibilities — if one disregards some occasional observations [1] — only the last one has hitherto been realised [2]. As with all other methods, this one too assumes satellites of a definite height above the surface of the earth.

The satellite especially recommended for communication purposes with a 24-hour period of revolution round the earth, i.e. always staying over one meridian, would not meet the requirements of a method of measuring the changes of distance. Nor would it serve the purpose of observing the changes of direction but it would be particularly useful for measuring direction. Compared with natural celestial bodies, artificial satellites flying at low altitudes have the advantage of a rapid alteration of their three polar co-ordinates referred to a place of observation. This quality, however, requires an exact knowledge of the satellite's orbit.

All position fixing based on a point of reference depends on a precise knowledge of the position of this point of reference. In the case of terrestrial objects this means that they have to be surveyed geodetically; in the case of celestial bodies the sub-stellar point on the earth has to be known. The ephemerides of the Astronomical Almanac help to find this point. For artificial satellites of short periods of revolution such ephemerides will have to have great dimensions.

Every navigator knows that it is more complicated to compute with the moon than with other planets, as it has a 27-day period of revolution round the earth. The satellites in use up to now need, however, only about 100 minutes to revolve round the earth. Unfortunately they then undergo several perturbations due to anomalies of the earth's shape and the mass distribution as well as the changing density of the exosphere so that a prediction of their orbits cannot be made. In contrast to natural planets, the new orbits have to be determined by a series of new observations. For short periods of time, let us say 10 or 20 revolutions, they may be computed accurately enough as KEPLER's ellipses. It is true that the orientation of the orbit with the earth's axis will change during this period of time, but as this takes place according to rather simple rules no considerable efforts have to be made to consider this.

The above method of measuring the changes of distance is of such accuracy that very precise data on the satellite's position are indispensable. In view of the fact that a prediction, as already stated, is impossible these satellites are equipped with a magnetic memory which is filled from time to time with orbital data. The satellite will emit these every two minutes and will thus enable the observer receiving them to compute what is necessary for the evaluation of his observations. Unfortunately it is not possible to communicate enough orbital positions to the satellite in order to collect and emit them at the right time. The prodigious memory-capacity sufficient for this purpose cannot be provided in the artificial satellite. Thus a considerable amount of computing must be left to the observer. The whole range of the work is already so complicated — to mention only the transformation of the measured changes of velocity into co-ordinates of the place of observation on the earth's surface — that computers of major or minor capacity are necessary. Therefore the idea of carrying out further computations on board comes naturally to mind.

Measurement of the change of distance is based on the DOPPLER principle: a satellite transmitting continuous waves and revolving round the earth at not too great a distance will not receive the frequency transmitted by the satellite, but a higher one when it is approaching and a lower one when it is moving away. When plotting the frequencies received

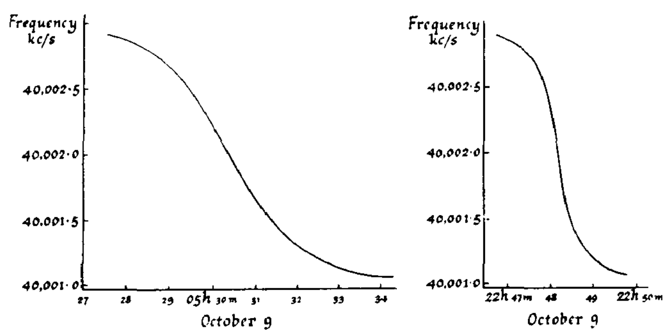


FIG. 1. — Frequency versus time curve for radio transmitter of Sputnik passing at distances of 350 and 135 nautical miles.

versus the time, a curve (as shown in Figure 1) will result. It has a two-fold curve with a point of inflection. It is only at this moment that the actual transmitted frequency is received. This happens at the moment of the nearest distance. The scientists and technicians of the John Hopkins University in Baltimore therefore called their programme " Transit ". In the case of a satellite passing relatively near to the place of observation the curve will be steep; at a greater distance the curve will assume a flat shape. Thus the distance of the passage can be determined from the inclination. Should the position of the sub-stellar point for the moment of transit, one part of the orbit and the shortest distance be known, the geometric construction of the place of observation will be possible.

The practical realisation of these theoretically represented possibilities is more complicated. It is above all the rotation of the earth which noticeably modifies the curve. It can be taken into account, but it is more difficult to deal with the fact that the change in frequency, due to the curved path through the ionosphere, does not represent the change in the straight distance as would be the case in a vacuum. In this instance it was the fact of observing two frequencies at the same time which helped [3]. A combination of both frequencies permits the influence of the ionosphere to the first order to be taken into account; values of the second order remain. If it were possible to work on higher frequencies than those in use today it would be easier to overcome this difficulty. But this can be neglected should the highest accuracy be unnecessary.

From the theoretical point of view the evaluation of the observations can be carried out by computing a curve (as in Figure 1) for the position on which the observer seems to be, whilst the actually observed curve is being drawn. From the differing values the most probable correction of the estimated position is determined by analysis. When putting this idea into practice it will necessarily be modified to some extent. The very precise method is based on some 50 individual measurements carried out at regular 6-second intervals. This measurement consists of a comparison with an extremely stabilized frequency on board. In this case it is not so important to know the exact mean value as this can be introduced into the analysis as unknown. It is however important that the stability during one transit is guaranteed.

The programme of calculation by this exact method necessitates the computation of data and the evaluation being carried out by an electronic computer, and the obtaining of results without need of any drawing. The all-important exact timing is obtained by time pulses emitted by the satellite at very short intervals. These pulses, too, are controlled and corrected from the earth.

In the case of a reduced demand for accuracy four individual measurements will be sufficient. These will show less precise results because additional velocities caused by the ship's pitch and roll will be introduced into the results of the measurements otherwise smoothed out as the result of a large number of measures. The frequency and the earth's radius for the place of observation must be included in the calculation even in this simplified method. Neither will be explicitly determined.

An impression of the lines of position resulting from this method can be given by a discussion under simplified assumptions. Let us assume a circular orbit of a satellite and a place of observation in which movement can be neglected. Let us, moreover, assume an earth-fixed system of co-ordinates within which the satellite's instantaneous orbit plane forms the XZ-plane, while the satellite has the co-ordinates $X_s = 0$, $Y_s = 0$, $Z_s = r_s$. The place of observation has the co-ordinates X, Y, Z which are connected with its co-ordinates $x$, $y$, $z$, defined in an equatorial system through transformation depending on the momentary orientation of the satellite's orbit. These two triples of co-ordinates correspond to the triples of polar co-ordinates in the equatorial geographical system, i.e. to latitude, longitude and radiusvector on the one hand, and on the other to the same radiusvector but in addition to two angular co-ordinates which may also be called latitude $\Phi$ and longitude $\Lambda$, which are connected with the equatorial geographical ones through well-known equations. The velocity components in a circular satellite's motion are

$$\dot{X}_s \neq 0, \ \dot{Y}_s = 0, \ \dot{Z}_s = 0$$

this last because the satellite, by reference to an instantaneous system of co-ordinates, is at a point of inflection of its orbit.

The distance from the satellite to the place of observation is :

$$\rho^2 = (x_s - x)^2 + (y_s - y)^2 + (z_s - z)^2 = X^2 + Y^2 + (r_s - Z)^2$$

The first relation applies to the equatorial geographical system of co-ordinates, the second one to the special system introduced here. Differentiation and translation into the above-mentioned polar co-ordinates $\Phi$ and $\Lambda$ will lead to

$$\dot{\rho} = \frac{\dot{X}_s r \cos\Phi \cos\Lambda}{\rho} \quad \text{where } r^2 = x^2 + y^2 + z^2.$$

An easily understandable presentation of similar reflections without explicit reference to the formula is offered by G. R. MARNER [4].

The equation for the line of position defined by the difference between computed and observed $\dot{\rho}$ will result in :

$$d\dot{\rho} = \dot{X}_s \sin\Lambda \frac{r \cos\Phi}{\rho} d\Lambda + \dot{X}_s \cos\Lambda \frac{r \sin\Phi}{\rho} \left(1 - \frac{r\, r_s}{\sin\Phi} \frac{\cos^2\Phi}{\rho^2}\right) d\Phi$$

The dependence on errors in $r$ and in the frequency may be neglected considering that both of them have already been introduced into the computation for another distance, so that the given equation may be taken as the formula of the line of position in the transit method, where $d\Phi$ and $d\Lambda$ represent the changes related simply to the satellite's motion [5].

To carry out the computation the actual geographical co-ordinates have to be considered. The translation into a formula similar to that mentioned above would involve as much computation as the direct statement of a relation between the measured values of the change in frequency and the corrections of the latitude and longitude of the ship's position. A great deal of computing which arises consists, however, of formulae familiar to the navigator, the exception being the extraction of square roots necessary to determine the values of $\rho$ and $r$ or $r_s$. Methods of approximation would have to be applied here.

It should be stated that the evaluation can be carried out in different ways. Experience has shown that simplifications will arise, when a method is applied over longer periods. In KERSHNER's view [6] the best evaluation nowadays is not to start from a defining equation for $\dot{\rho}$ when four individual measurements have to be evaluated. The equation

$$f_b = -\frac{f}{c} \dot{\rho} + \delta$$

stands for the beat frequency between the received and the transmitted frequencies, where $\delta$ is an unknown but constant difference. When counting the cycles of the beat frequency, integration with respect to the time is being made, and thus may be written :

$$N_b = \int_{t_1}^{t_2} f_b dt = -\frac{f}{c}\int_{t_1}^{t_2} \rho \, dt + (t_2 - t_1)\,\delta = \frac{f}{c}(\rho_1 - \rho_2) + \Delta$$

By means of four observations three intervals of time will be obtained which will then serve for position fixing, i.e. the correction of the latitude and the longitude. The error in $r$ has been considered by the method of computation.

The determination of the rectangular inertial satellite co-ordinates $x_s$, $y_s$, $z_s$ in the geocentric system is somewhat complicated. For low-flying satellites the orbit changes with the changing density of the exosphere. Therefore the calculation by successive steps of new Kepler ellipses is recommended. It is impossible to conceive in theory the curve of the orbit over a long period of time, as unforeseen sunspot eruptions may alter the density. Whether another method for satellites flying at greater distances from the earth, as for example the computation of general or special perturbations [7], is useful and practicable is not unanimously decided. The remoter the satellite, the less the influence of an error in the position of the satellite in space on the position of the sub-satellite point on the earth's surface. For the positioning on the earth only the exact knowledge of the sub-satellite point is of importance. The same applies to the position itself and to the sub-stellar point of low satellites as in the case of those of the Transit programme. This is also the reason why the memory of the satellites transmits only the codified orbit data : the quotation of the co-ordinates in space would necessitate too large a memory capacity. Thus the observer will have the task of carrying out computations hitherto made by astronomers if he is dealing with low-flying satellites.

The formulae are :

$$E - e \sin E = n\,(t - t_o)$$

$$r = a\,(1 - e \cos E)$$

$$\tan \frac{1}{2}\,(v - \Omega - \omega) = \tan \frac{E}{2}\sqrt{\frac{1+e}{1-e}}$$

$$x = r\,[\cos \Omega \cos (v - \Omega) + \sin \Omega \sin (v - \Omega) \cos i]$$

$$y = r\,[\sin \Omega \cos (v - \Omega) + \cos \Omega \sin (v - \Omega) \cos i]$$

$$z = r \sin (v - \Omega) \sin i$$

$E$ = eccentric anomaly
$v$ = longitude in the orbit
$n\,(t - t_o)$ = mean anomaly
$a$ = semi-major axis
$e$ = eccentricity
$\Omega$ = node of the orbit
$i$ = inclination of the orbit
$\omega$ = longitude of perigee
$t_o$ = passage at perigee

Once again it is a question of almost exclusively trigonometric computations which are familiar to the navigator (the enormous amount of computing work not being taken into account). The first equation, however, the so-called Kepler's equation, is a transcendental one to be solved by successive approximation. As in the extraction of square roots the navigator is faced with the successive approximation method, which can also be applied in the latter case using a mathematically different, but similar as to computation, method of calculation.

Kershner [6] states that *given suitable forms of data in the memory* capacity and corresponding transformation of the above formulae for four satellite positions, 43 additions, 57 multiplications and 18 table entries will be necessary. As furthermore we have the evaluation of the measurements to fix the observed latitude and longitude, the question of the necessity for computing machines is undoubted. Computers will be necessary for precise position fixing from as manu as 50 individual measurements. Here an observation instrument of such dimensions and cost is used, so that the additional cost of a computer would hardly be felt. The accuracy achieved amounts to 0.2 n.m. Where such accuracy is required the necessary funds will have to be allotted. In the Merchant Marine this is not to be advocated without reservation. In the less accurate procedure the observation instruments will be less expensive which means that a relatively higher share of the overall cost will be available for the computing machines.

Transit satellites have repeatedly been launched for experimental purposes. Four such satellites in polar orbits will soon be available all over the earth for navigational positioning at intervals of 110 minutes.

Further possibilities of using satellites for navigational purposes have not up to now been realised. Direction measurements are carried out by units too huge to be installed on board (i.e. Minitrack [8] and Microlock [9]). It is now a well known fact that the range of the observation systems is determined by the wave-length of the radiation used. The mirror of an astronomical telescope can be relatively small as the wave-lengths of the light are less than 1/1 000 mm. To get one arc minute of resolving power huge dimensions in radiotelescopes are necessary for the range of 21 cms — one of the most important wave-lengths of space radiation.

The present satellites are for the most part transmitting on longer waves. Even the communication satellites require directional antennae too large to be used on board. There are, of course, radio sextants [10] (Figure 2) for the wave-lengths 1.9 and 0.9 cm, but no satellite has hitherto been transmitting the corresponding frequencies of 16 and 33 Gc. The transmissions would also have to be strong enough to be recognised in the radiation of the atmosphere and space.

The radio sextant was built to observe the radio radiation of the sun and moon under practically all atmospheric conditions. It cannot separate other sources of radio radiation from the atmospheric clutter because its dimensions are restricted by its use on board [11]. Nevertheless, due to the conical scanning of its dipole, or similar provisions, an accuracy of almost 5 minutes is obtained for direction measurements with an antenna of only 1 m in diameter. This value includes inaccuracies resulting from the reference direction.

The radio sextant requires a reference direction for the measurement of the elevation and azimuth angles, and this is provided by a stabilized platform. This additional prerequisite is a disadvantage compared with the Transit programme. It is to be expected that future satellites will transmit on these high frequencies with sufficient amplitude. It will then be possible to carry out direction measurements with the radio sextant tracking the satellite as well as the sun and the moon. This is an advantage of the procedure.

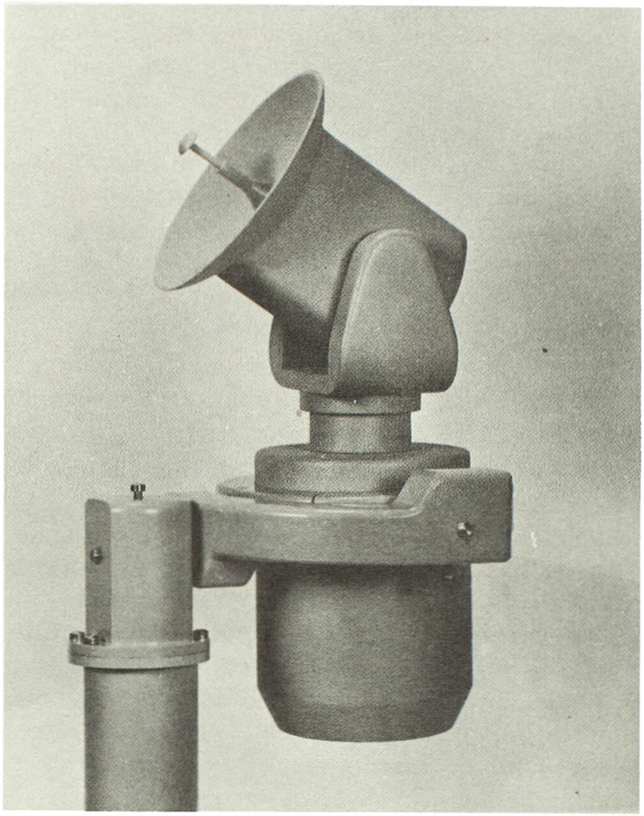

FIG. 2. — Collins Radio Company Radio-Sextant connected to a stabilized platform.

The evaluation of the observations is similar to that for natural celestial bodies. When measuring the angle of elevation position lines will be in the form of small circles centred on the sub-stellar point; isoazimuthal lines will be obtained with the azimuth. Both results are more easily obtained than the evaluation of Doppler measurements with Transit satellites. It depends on the altitude of the satellite above the earth's surface whether it will be difficult or easy to determine the sub-satellite point. For low-flying satellites the same difficulties exist as present themselves in the Transit programme. In all probability the real shape of the earth will also have to be taken into account in the evaluation. The experts expect considerably more favourable conditions for higher flying satellites [12]. There where the change in the position proceeds more slowly, a smaller amount of data is required and the data need not be so precisely related to the accuracy of the observations. So it seems probable that this may be predicted more easily and over a longer period of time.

Rapidly moving satellites have an advantage compared with the slower ones and the natural celestial bodies. Due to their rapid change of position on the sphere it will be possible after a short while, and after a new observation, to obtain a second line of position which will give a complete fix by intersection with the first one. In the case of the sun the same method of advancing a line of position will require an interval of several hours. It can even be employed to measure the changes in the direction of the satellites themselves. The formulae for the lines of position of the angle of elevation and for that of the azimuth can easily be derived [13]. The position line for the change in the elevation angle will then result as a great circle intersecting the lines of position of the elevation angles observable at the same time at right angles. Measuring these changes of angles demands technically a less exact knowledge of the reference directions than the measurement of the angles themselves. For a measurement of change it is only necessary that the reference line be approximately correct and constant over a rather short period of time, as differential and not absolute measurements are to be carried out.

These four ways of measuring directions and changes in direction require directional antennae, reference directions and satellites transmitting on high frequencies. This last demand has hitherto not been met.

As in the case of the measurement of the change in distance, no reference direction is needed for the measurement of distance, the sixth possibility. Perhaps an omnidirectional aerial can be used; a directional antenna is essentially unnecessary. Up to now the measurement of distance has been carried out following the Secor system [14], though not for nautical but for surveying purposes. It suggests itself quite naturally for the application of the principle in measuring the time of propagation of electromagnetic waves. In large radio observatories this has been successfully realised, following the radar echo principle. On board this is impossible because of the required transmitter energy, nor can an antenna capable of receiving the very weak echo be installed.

During geodetic surveys a transponder on the satellite is used; this is a method derived from secondary radar. For navigational purposes the method must be extended to make the reception by many observers without possible interference. The technical equipment could correspond to that used in the Tacan method [15], but it is expensive and cannot be installed aboard the satellites. This is a pity because it is a simple method and one hardly influenced by the ionosphere. The resulting position lines, small circles round the sub-satellite point as centre, can be easily constructed. The computation of the satellite's positions gives rise to a whole series of questions, already fully considered, depending on what height the satellite has its orbit.

It is to be hoped that a method of measuring the time of propagation will be practicable as soon as the progress achieved in the development of atomic clocks has reached such a level that they may be installed aboard a satellite. Once this chapter of technical development is entered it will also be possible to have these clocks aboard the ships. A comparison of the exact time by both clocks with the aid of wireless signals will show distances. The comparison gives an exact difference of time if the two clocks are completely synchronized, or if they have a reliable constant rate. The difference of time is multiplied by the velocity of light to give the distance in length. The ionospheric refraction will give small errors which may have to be taken into account.

Doubts have often been expressed as to whether this simplest of methods of navigation by means of artificial satellites will ever be practised. Attention may, however, be drawn to the fact that the CCIR recommendations IV/18, Washington 1962, considered this possibility when dealing with the distribution of frequencies for satellites. It was expressly stated that the time difference could be measured in terms of the phase difference between low-frequency modulations of continuous wave oscillators to avoid the wide bandwidth and high peak-power requirements of pulse systems. Let us hope that the optimism expressed in these words will be justified in the not too distant future. An effective aid to navigation could then especially be given with high-altitude satellites.

## REFERENCES

[1] FREITAG, R. F. : Project Transit, a Navigation Satellite. *Navigation,* Washington, 7, 106-116, 1960.

[2] SCOTT, W. A. : Determination of Position on the Earth from a Single Visual Observation of an Artificial Satellite. *J. Inst. of Navigation, 14,* 87-93, 1961.

[3] GUIER, W. H., WEIFFENBACH, G. C. : A Satellite Doppler Navigation System. *Proc. I.R.E., 48,* 507-516, 1960.

[4] MARNER, G. R. : Plotting the Future Course of Marine Celestial Navigation. *Navigation,* Washington, 7, 213-225, 1960.

[5] HUGON, P. : Evolution de la technique du Satellite Doppler de Navigation. *Navigation,* Paris, *10,* 133-146, 1962.

[6] KERSHNER, R. B., NEWTON, R. R. : The Transit System. *J. Inst. of Navigation, 15,* 129-144, 1962.

[7] STERNE, T. E. : An Introduction to Celestial Mechanics. New York-London, 1960.

[8] MENGEL, J. T. : Tracking the Earth Satellite, and Data Transmission by Radio. *Proc. I.R.E., 44,* 755-760, 1956.

[9] RICHTER, H. L., SAMPSON, W. F., STEVENS, R. : A Minimum Weight Radio Instrumentation System for a Satellite. Jet Propulsion Laboratory Cal., 1958, *88-111,* Pergamon Press, New York-London-Paris.

[10] McCOY, D. O.: An All-Weather Radio Sextant. *Navigation,* Los Angeles, *4,* 309-319, 1955.

[11] MARNER, G. R. : Automatic Radio Celestial Navigation. *J. Inst. of Navigation, 12,* 249-259, 1959.

[12] MOODY, A. B. : Navigation Using Signals from High-Altitude Satellites. *Proc. I.R.E.*, *48*, 500-506, 1960.

[13] FREIESLEBEN, H. C. : Navigational Aids from other Satellites. *J. Inst. of Navigation*, *15*, 149-155, 1962.

[14] DIERCKS, F. O. : Army Map Service Mapping Activities in the Space Age. *Int. Hydr. Review*, *38*, No. 1, 173-179, 1961.

[15] COLIN, R., DODINGTON, S. H. : Principles and Features of Tacan. Electrical Communication, 1956, March.