GEODETIC APPLICATIONS OF ARTIFICIAL SATELLITES

by Commander J. A. BARAHONA FERNANDES
Portuguese Hydrographic Institute

Article reproduced from Anais do Instituto Hidrográfico, *No. 1, 1964. Ministerio da Marinha, Lisbon.*

1. — GENERAL

The launching of the first Sputnik in October 1957 was a sensational event which caught world-wide attention and ushered in a new era of the exploration of outer space.

Since then the rivalry of the U.S.A. and the U.S.S.R. in the "conquest of space" has given rise to other spectacular excursions, and today there are already plans for sending a man to the Moon or even to the nearest planets.

We have seen yet once more that scientific research in one particular field leads to unlooked for discoveries in other domains, for the exploration of space has proved uncommonly useful to the development of the store of knowledge on the planet on which we live.

It was to be expected that satellites should supply valuable data on the Sun's effulgence and the atmosphere's composition and properties up to an altitude of several hundred kilometres, as well as on the ionosphere, electromagnetic wave propagation in the exosphere, variations in the magnetic and gravity fields, etc. These nevertheless have caused many surprises, among them the relatively high density of air at altitudes where it was thought a near vacuum existed, or again, like the "Van Allen belts" — zones of intense corpuscular radiation extremely dangerous to astronauts.

In addition to this knowledge, and as a by-product of spatial exploration, satellites have contributed in a remarkable way to solving various geophysical and geodetic problems, in particular the determination of the geoid's size and shape and of the relative positions of continents and remote islands, with which the present article deals. The topic is necessarily treated in a general way for the subject is an extremely wide and complex one.

****

Earth's spherical shape has been assumed since at least the time of the great Greek Thinkers, for example, Pythagorus in the VIth Century, B. C., and ingenious estimates about this shape were made : those of

Eratosthenus (276 - 194 B. C.), for instance, who found a very good approximate value for the earth's radius by measuring the distance and the difference in latitude between Syena and Alexandria which are situated on nearly the same meridian.

It was not until the beginning of the XVIIIth Century that the measurement of arcs of meridian at different latitudes was taken up again with more accuracy. The variation of the radius of curvature showed that the Earth, as Newton had already foreseen and computed starting from observed values of gravity intensity, has the form of a flattened ellipsoid of revolution.

Measurements of meridian arcs are affected by the relative plumb line deflection, i.e. by the angle between the normals to the ellipsoid and to the geoid which is the zero equipotential surface of the terrestrial gravity field. Hence we have a lack of accuracy in astro-geodetic determinations arising from the geoid's undulations and irregularities which are due to the unequal distribution of Earth's internal masses and hence also various dimensions for the ellipsoid are obtained.

The plumb line deflection is reckoned by integrating attractions due to the known masses on the Earth's surface, but the method is not very accurate even when applied to Pratt's isostatic compensation layer. The determination of the geoid's shape would be more accurate starting from detailed observation of gravity intensity, either on the continents or over the whole ocean surface, and using the Clairaut and the Stokes formulae. This is one reason that gravimetric surveys are being intensified, for it is necessary to cover the large areas where only few " g " values are known more densely, above all in Africa, Asia and the seas south of the Equator.

The lack of precision in the determination of Earth's shape by arcs of meridian measurement, or by astrogeodetic or gravimetric methods is also reflected in the relative position of remote geodetic nets, although these may have a high internal consistency. This is the case for European and North American geodetic nets which until recently have not been able to be connected in view of the limited visual range in classical triangulation methods.

With Shoran, and its recent improvement Hiran, it became possible to measure distances of up to about 800 km. This system was used for the Bahamas — U.S.A. trilateration, for the Canadian geodetic network, etc.

Until recent years there was no sufficiently accurate measurement system for use beyond these limits.

In addition to their obvious interest for scientific progress, knowledge of Earth's shape and intercontinental geodetic connections are of great practical importance both from the military viewpoint (for guided missile bombing, etc.) and for air and electronic navigation.

Already Radar has shown up large cartographic errors (of about several miles, for example, at some points between Denmark and Norway), and with the advent of Loran many more deviations have been observed. It is understandable for instance that the Loran network, based on stations on the Continent, Madeira and the Azores, should be affected by large errors, for the positioning of these archipelagoes is based on astronomical coordinates that are affected by very high plumb line deflections.

Nowadays new and powerful techniques can be employed in geodesy, and particularly the use of artificial satellites.

Broadly speaking, artificial satellites can be used in geodesy in three different ways :

a) *Dynamic methods,* based on observation of orbit disturbances related to the gravity field variation, thus supplying an evaluation of Earth's shape and size.

b) *Geometric methods,* based on simultaneous observations of the satellite's direction from various points on the Earth's surface, which make it possible to carry out a spatial triangulation between these points.

c) *Radioelectric methods,* based on determination of the distances to the satellite (Secor), or on the law of variation of its distances (Doppler effect), by means of transmitting radioelectric signals from the satellite.

We shall first give a general outline of the procedures for observing satellites and then we shall speak of the principal geodetic methods employed. A brief account of related astronomical methods of occultation and eclipse will serve as introduction.

2. — GEODETIC SATELLITE OBSERVATION

The observation of satellites can be made by various methods that we shall classify as visual, photographic and radioelectric. Visual methods are the simplest, but they are dependent on meteorological conditions; the last have the inconvenience of requiring complex instrumentation. This is why very varying techniques are used, according to the satellite's characteristics, the local conditions and the purpose of the observation.

2.1. Visual Methods

If the satellite gives out a light (either reflected or its own) it can be observed like any planet, but however with a great deal more difficulty in view of the great speed with which it travels. In the " tracking " method, which consists solely of keeping a certain watch on the various orbits, a simple estimation of the satellite's position in relation to neighbouring stars is employed. In spite of the ambiguity entailed the results obtained by thousands of amateurs (there are more than 2 000 in the U.S.A. alone) are of great interest, chiefly during the first few days after launching when the orbits are not yet perfectly defined and can undergo large changes.

Observation by theodolite is difficult, even if photographic records of heights and azimuths are available. It is preferable to use cine-theodolites maintained constantly trained on the satellite by two operators, and recording the successive vertical and azimuthal readings at short intervals.

In this way an accuracy of the order of 30″ in direction and 0.01 seconds in time is obtained, which is however insufficient for geodetic purposes.

The satellite may be painted or covered with a highly reflecting metal coating in order to make it visible. It then appears like a planet, but can only be observed from a small part of the earth, i.e. where the sun is below the horizon but still lights up the satellite.

The passive Echo satellites at present in orbit are of this type, and the U.S. plans to launch another similar satellite in 1966 — the Pageos.

The geodetic satellite Anna transmits very strong light flashes and can therefore be observed at night. However its enormous energy consumption and its weight limitations made it necessary to restrict the number of flashes. These are controlled by a memory circuit whose functioning has been erratic.

Finally, the illumination of the satellite by a very narrow " laser "-transmitted beam was visualised. In spite of beam-training difficulties, sights on satellites provided with prismatic reflectors have been successfully obtained.

2.2. **Photographic Methods**

Photography of a satellite against a star background makes it possible to obtain great accuracy in determining its direction, for the stars here act as the reading circles of a gigantic universal theodolite.

This photography requires a highly specialised technique in view of the difference in flashes and the relative satellite-stars motion. Cameras fixed in relation to the earth may be used, thus obtaining photographic recordings of circular arcs representing the stars' diurnal apparent motion and of elongeated arcs for the satellite path. Use may also be made of star-tracking cameras which track in the same way as equatorial telescopes.

The problem of synchronising the photographs is solved automatically for the case of flash-transmitting satellites. For passive satellites a rhythmical interruption of their exposure has to be achieved in order to obtain a reference of their path in relation to time.

As an example of the equatorial type of camera the ballistic Baker-Nunn Schmidt camera may be cited. This weighs more than 3 tons, and its movements are controlled by an electronic computer. It can photograph satellites and stars of magnitudes higher than 12 with an accuracy of $\pm$ 2″ in the sights. Similar to this camera, but easier to handle, is the Wild BC-4 camera in which the T4 theodolite tripod and the optical qualities of lenses used in aerial photography (Astrotar) are combined.

Fixed cameras are simpler, sturdier, steadier and more easy to handle. The cameras used by the French Institut Géographique National are of this type. Further details are given later in this article.

2.3. Radioelectric Methods

Radar can be used for determining the azimuth of, and distance to, a satellite. The method, besides lacking accuracy, is very uncertain in view of the smallness of the target and its great distance.

It is consequently preferable to place transmitters or " transponders " in the satellite which allow its direction, its distance or variation in distance to be determined. As in electronic navigation there exist today manifold techniques, and here we shall but make mention of some of their basic principles.

In the interferometer method (the Minitrac system) the signals put out by the satellite are received on various antennae placed on the ground at regular distances and azimuths and this causes the signals received to be mixed. An interference occurs when two antennae receive signals having a phase difference of an odd multiple of 180°. The direction of these antennae then corresponds to the direction of the signals. This direction is calculated to within $\pm$ 20″.

In the Secor method (similar in principle to the Shoran or the Tellurometer method) the phase of the signal transmitted by a ground-station is compared with the signal retransmitted by the satellite's transponder, and the distance is thus derived.

The Transit system is more accurate and better tested. It is based on Doppler effect measurement which we shall describe later on in more detail, in view of its usefulness for geodetic observations.

There also exist composite systems; such as Microlock which uses the interferometer principle combined with the Doppler effect; the Azuza system — a combination of Secor and Doppler effect — and the Mistram system which derives distances from distance differences and their laws of variation, etc., of which we can only make mention in this article.

3. — GEODETIC CONNECTIONS BY OCCULTATION OR ECLIPSE

Before the existence of submarine cables or of W/T transmissions of time signals, eclipses of the sun or occultations of stars by the moon were used to derive the differences of longitude between distant localities. This method, which has gained much in accuracy, is once more being employed for intercontinental geodetic connections, through the use of either photographic methods or of photoelectric cells for observing occultations.

Basically, a comparison is made between the instant forecast for an astronomic phenomenon and that of its actual observation at a given place, and the difference is a function of the corrections to be made to the coordinates of the locality.

To give an example, let us consider the case of lunar occultations only. At a place having astronomic latitude φ and astronomic longitude λ the instant when a given star of declination δ and right ascension α will appear and disappear behind the Moon's rim can be computed using Bessel's system of rectangular coordinates with origin at the Earth's centre having the z-axis set towards the star, the x-axis being the intersection of the Equator and a plane normal to z, the y-axis being normal to the xz plane.

With the help of the Nautical Almanac, and using complex formulae such as are, for instance, shown in W. CHAUVENET's treatise on astronomy, the projection of the Moon on the xy plane for consecutive hours (GMT) is computed (figure 1). The projection on the same plane of the path described by the observation station, due to Earth's rotation, is computed by other formulae.

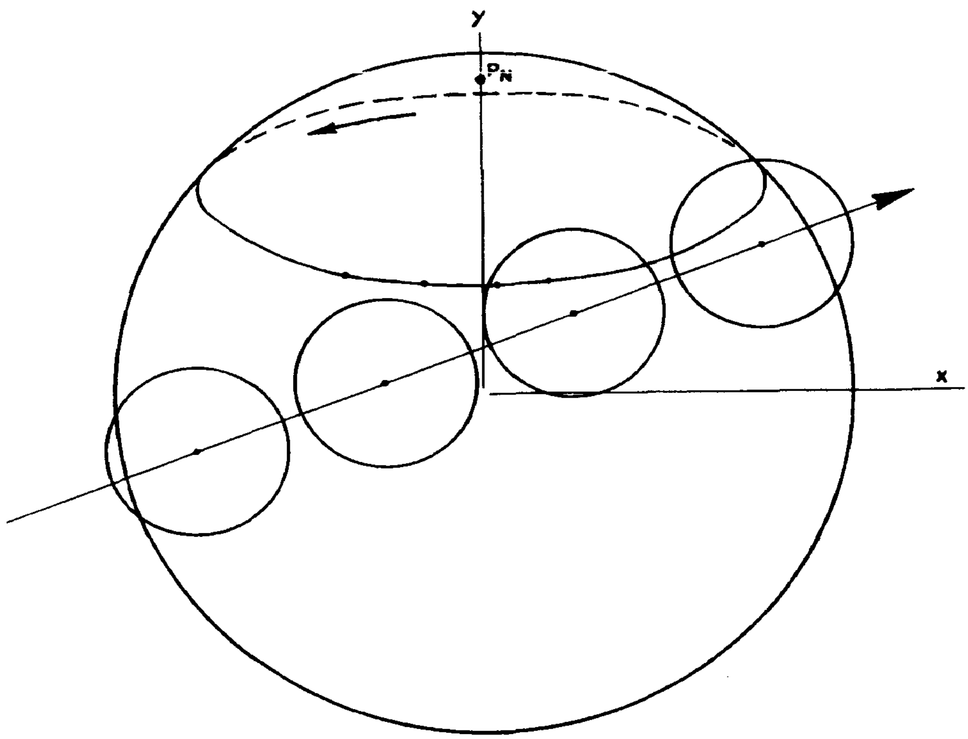

FIG. 1

The instant when the Moon's rim occults the star is determined, and also the instant at which the star reappears from behind the Moon.

Observation of this phenomenon gives a value $\Delta\tau$ that is the difference between predicted and observed time. This is a function of the corrections :

$\Delta\varphi$	to be made to the latitude of the place
$\Delta\lambda$	to be made to its longitude
$\Delta\alpha$ and $\Delta\delta$	to the star's right ascension and declination
$\Delta\pi_0$	to the lunar parallax
Δk	to the ratio between the Moon's radius and that of the Earth
ΔN	to the altitude of the place of observation.

This equation has the general form :

$\Delta\tau = a\Delta\varphi + b\Delta\lambda + c\Delta\alpha + d\Delta\delta + f\Delta\pi_0 + l\Delta k + j\Delta N$ where a, b, c, d, f, l and j are coefficients dependent on the coordinates of the place of observation, and those of the star and the Moon.

Equations of the above type are obtained by observing occultations on consecutive nights and in different places on Earth, and these equations easily become more numerous than the unknowns. Using the least squares method, the star and Moon coordinate corrections are derived, as well as corrections to the geodetic coordinates of the place of observation.

Various causes of error exist, as for instance those resulting from the as yet incomplete knowledge of all the Moon's rim irregularities which are due to local humps or troughs on the Moon, to the imperfections of the Nautical Almanac, and to the relative scarcity of stars suitable for observation from the required spot, etc. However the method is capable of supplying coordinates with an accuracy of about several tens of metres, and this has already been attained in the connection of certain Pacific islands to the North American Continent.

Quite recently the idea occurred of using the occultations of stars by artificial satellites for the same purpose.

It is obvious that at the time of the star occultation the satellite's direction is known with great accuracy (about $\pm 1''$), and observations of occultations from various stations make it possible to compute the satellite's path or to carry out distant geodetic connections.

Lenses provided with photo-electric systems for accurate determination of the time of occultation are being studied, but this method's big difficulty lies in the fact that forecasting of occultations is extremly complex and that these occultations are infrequent.

Theoretically the method permits an accuracy of about ± 3 m in the relative position of observational stations, but this development is still in the project stage.

4. — DERIVING THE EARTH'S SHAPE BY OBSERVING ARTIFICAL SATELLITE ORBITS

A satellite's orbital motion is essentially dependent on the law of universal attraction, although it undergoes slight disturbances, due to air resistance, which cannot be neglected even at altitudes of 400 km or so, and to solar radiations. These last influences are small, and they can be taken into account. Therefore we shall here consider only the action of the Earth's gravity field, in view of the great distance to the Moon, the Sun and the planets.

In a first simplified analysis let us consider the Earth as spherical and of uniform mass. Thus the attractive force exerted on the satellite is constantly directed towards the Earth's gravitational centre. This being so, the satellite's orbit remains fixed in space. The irregularities of Earth's gravity field produce disturbances in the satellite's orbit. For instance the equatorial swelling of the Earth, due to its ellipticity, introduces an asymmetry into the attractive forces, resulting in a rotation of the orbit's plane, i.e. a precessional movement of the orbit's nodal line. Moreover, the orbit's position in this plane varies, its perigee undergoing a retrogressive movement.

The precessional speed of the node x (in degree per day) is bound to the ellipsoid's flattening f, to the Earth's equatorial radius a, the satellite orbit's mean radius r_m, the orbit's inclination i and to excentricity e by the approximate expression :

$$x = 6136\,(f - 0.001\,725)\left(\frac{a}{r_m}\right)^{3.5}\frac{\cos i}{(1 - e^2)^2}$$

Thus it is seen that observations of orbit disturbances for various satellites give a series of observational equations in which the values of x, r_m, e and i are known and it is consequently possible to compute the most probable values of f and a.

The most recent values thus computed from observation of Sputnik 2, Vanguard 7 and other satellite orbits supply a value of 1/298.3 for the flattening, f, of the terrestrial ellipsoid, a value which is lower than that adopted in 1924 for Hayford's International Ellipsoid (1/297.0) based on astrogeodetic observations in North America.

This value of 1/298.3 is, furthermore, more in agreement with the measurements based on gravimetric methods (HELMERT, 1901, BOWIE, 1912, HEISKANEN, 1938, UOTILA, 1961) and on the astrogeodetic measurements of HEISKANEN (1929) and KRASSOWSKY (1942).

Earth's equatorial radius a, derived by this method, is 6 387 165 m which is nearly 223 m less than that formerly accepted for Hayford's International Ellipsoid.

In reality mathematical treatment of satellite orbit disturbances is much more complicated, and at present is based on the expansion of potential U of Earth's gravity field into a spherical harmonic series of the type :

$$U = \frac{KM}{a}\left[\frac{a}{r} - \sum_{n=2}^{n=\infty} J_n \left(\frac{a}{r}\right)^{n+1} P_n(\cos\theta)\right]$$

where K is the constant of universal attraction, M the Earth mass, a the equatorial radius, r the distance from a point in the gravity field to Earth's centre of gravity, J_n the constants to be determined, P_n the Legendre polynomials of order n, and θ the colatitude.

The rotation x of the ascending node of the satellite orbit and the precessional movement of the perigee are functions of the Jn constants and of the orbit's characteristics.

Consequently the analysis of disturbances of satellite orbits makes it possible to derive constants J_n which are bound to Earth's shape and size.

The last value derived for J2 is 1082.7×10^{-6}, and the values of J are also known up to the ninth order with good accuracy.

These values have not only made it possible to compute the terrestrial ellipsoid's flattening and its equatorial radius, but have also revealed that Earth's shape is not strictly ellipsoidal, the northern hemisphere being slightly less flattened than the southern hemisphere, and everything points to the terrestrial equator being slightly elliptic, with a difference in radius of about 100 m.

These ellipsoid differences are however unimportant, and have not yet been accurately established.

Finally, local undulations of the geoid reduce the advantages of a more accurate knowledge of the ellipsoid but are nevertheless of value for theoretical speculation on the internal composition of the Earth, etc.

The most acute practical problem is at present the intercontinental connection of geodetic networks, and this can be solved not only by astrogeodetic methods but also by spatial triangulation and by Doppler effect measurement using artificial satellites, as will hereafter be shown.

5. — SPATIAL TRIANGULATION BY SATELLITES

5.1. General Principles

Photography of a satellite against a surrounding star background makes it possible to derive its direction by interpolating between the directions of the photographed stars.

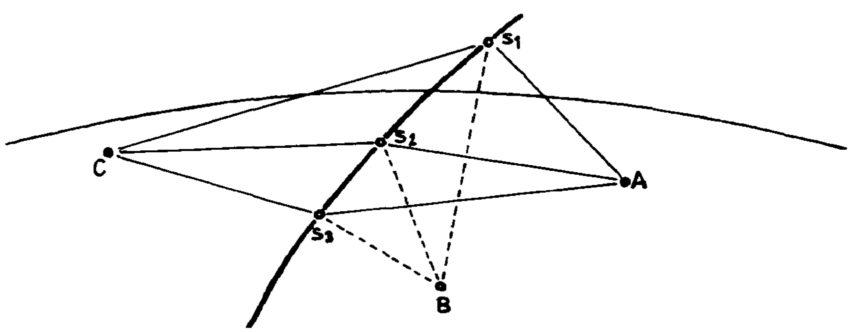

FIG. 2

If the satellite was photographed in three consecutive positions on its orbit, S_1, S_2 and S_3 (figure 2), from two ground stations A and B with known coordinates, the directions AS_1 and BS_1, AS_2 and BS_2, AS_3 and BS_3 allow the positions S_1, S_2 and S_3 to be computed accurately. If the satellite was at the same time photographed from a station C, the intersection of radii S_1C, S_2C and S_3C makes it possible to compute the coordinates of C, and consequently to carry out its geodetic connection to A and B.

The principle is therefore very simple. The operation is only complex because a very high accuracy is desired — of the order of one millisecond in the photograph's synchronization, 1″ in direction, and of several metres in the satellite position and the geodetic positions.

A similar method has already been used. It consisted of observing lighted beacons released by aircraft at great height, or launched by rocket. Satellites however increase the range of such connections.

5.2. **Photographic Technique**

As we have already seen, either mobile cameras or cameras fixed in relation to the earth can be used for satellite photography. Here we shall only deal with the latter, considering, in particular, the photography of Echo type satellites which continuously reflect the Sun's light, thus requiring the camera shutter to be closed periodically so as to obtain a series of pictures strictly defined in time.

For stars, an exposure of several seconds is necessary. The satellite on the contrary must be photographed with very short and accurate exposures. This is why the stars are not photographed at the same time, but rather before and after the satellite's passage. Great stability of the camera must therefore be ensured (of the order of a micron per quarter of an hour). The stars are usually photographed after a two-second exposure followed by an eight-second closing of the shutter.

A chronoscope-controlled electronic timer governs the shutter in order to ensure an accuracy higher than 1 part in 100 seconds. The various exposures of the satellite must be strictly synchronized at the different observational stations which may sometimes be thousands of miles apart. This is why the camera shutters are governed by a rotating disc of small inertia driven by a synchronous quartz-controlled motor as well as by the transmission of the same time signal that continuously puts out a pip per second.

It is obvious that even in this way the various photographs are not strictly synchronous, and that various corrections will have to be made due for instance to the difference in the propagation paths of the time signals between the transmitting station and the various satellite observing stations, or to the difference in light propagation time between the satellite and these stations.

This simple outline gives an idea of the instrumental complexity and of the high accuracy values required.

Several types of camera fulfil these conditions — the one acquired by the French Institut Géographique National which will be used for connecting the Azores is a ballistic camera with a 305 mm focal length, a f/5 aperture, a disc diaphragm completing a revolution per second and controlled by quartz clock and time signals.

The photographs already taken at other places show :

- star paths records consisting of arcs of parallel of the diurnal apparent motion, interrupted at regular intervals to form a time reference;

— a dotted line corresponding to consecutive satellite positions photographed at one-second intervals with an accuracy of a thousandth of a second in timing (figure 4).

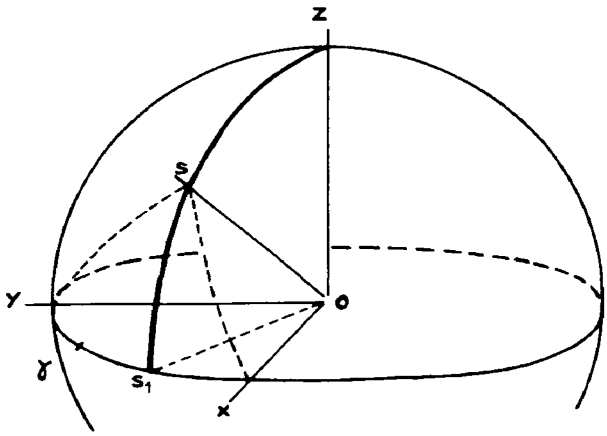

FIG. 3

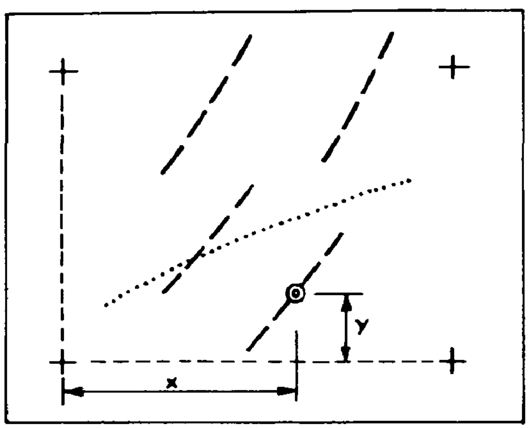

FIG. 4

5.3. Computation of the satellite's direction

Let us take as reference a three-axis system of rectangular coordinates, with the Earth's centre of gravity as origin, the Z-axis being in the direction of the celestial pole, the X-axis at the intersection of the equator with the Greenwich meridian, and the Y-axis at the equator's intersection with the 90° W meridian.

A star of declination δ and right ascension α will occupy position S (figure 3) at sidereal time Hs (found from the corresponding G.M.T.) and the direction cosines of OS are deduced from the spherical triangles SS_1X, SS_1Y and from arc SZ :

$$p = \cos \text{SOX} = \cos \delta \, (\text{Hs} - \alpha)$$
$$q = \cos \text{SOY} = \cos \delta \sin (\text{Hs} \quad \alpha)$$
$$r = \cos \text{SOZ} = \sin \delta$$

and thus they may be computed with the required accuracy.

The satellite pictures appear on the photographic plate in the manner shown in figure 4; i.e. as a succession of 59 dots (one per second). The stars appear as circular arcs interrupted at specific known instants. The satellite's direction cosines can be obtained by interpolation using the difference in coordinates measured on the plate and taking one of the plate's reference marks as origin.

To establish the relation between direction cosines $p\ q\ r$, bound to terrestrial system OXYZ (figure 3), and coordinates X and Y measured on the photographic plate it is only necessary to know the coordinates of the observational point and the direction (height and azimuth) of the camera's optical axis.

Since these elements are not accurately known the stars' approximate coordinates Xa and Ya are first of all computed from their direction cosines, making two conversions of axes :

a) computing the stars' heights and azimuths from their direction cosines and the approximate coordinates for the station;

b) computing Xa and Ya from height and azimuth, in terms of the height of the camera's optical axis and its focal length.

Let us now measure, with a comparator and with ± *1 micron accuracy*, the difference between the coordinates (Xa, Ya) and the coordinates (Xo, Yo) on the plate.

These differences arise from :

a) lens distortion — corrected with the camera calibration data;

b) inaccuracy of the camera's presumed orientation — corrected with the help of the following transformation.

These coordinates are indeed related by a homographic transformation of the type :

$$\text{X}a = \frac{a_1 + b_1\text{Xo} + c_1\text{Yo}}{1 + b\text{Xo} + c\text{Yo}}$$

$$\text{Y}a = \frac{a_2 + b_2\text{Xo} + c_2\text{Yo}}{1 + b\text{Xo} + c\text{Yo}}$$

in which coefficients a, b, c, ..., are determined from values (Xa, Ya) and (Xo, Yo) computed and observed for different known stars by applying the least squares method to a set of observational equations obtained from the following homographic transformation equations :

$$a_1 + b_1\text{Xo} + c_1\text{Yo} - b\text{XoX}a - c\text{YoY}a - \text{X}a = v_1$$
$$a_2 + b_2\text{Xo} + c_2\text{Yo} - b\text{XoY}a - c\text{YoY}a - \text{Y}a = v_2$$

. . . .

Thus we wish to find the camera's accurate orientation by measuring the plate coordinates of the satellite pictures and its direction cosines are computed by a process the inverse of the above :

a) by applying homographic transformation equations for which coefficients *a, b, c,* ..., are already known;

b) by converting the coordinates of the photographic plate's system (X*a*, Y*a*) into horizontal coordinates;

c) by computing the direction cosines for the satellite's direction from its horizontal coordinates.

Obviously corrections must also be applied for astronomic refraction, differences in propagation time of the time signals, differences in direction of light paths, for light aberration, etc.

5.4. Computing spatial triangulation

A successive approximations method is here set out, using computed provisional values of the satellite's coordinates inferred by orbital forecast and the approximate coordinates of the stations to be determined. The least squares principle is applied in such a way that the sum of the squares of the corrections between the observed and the adjusted directions is reduced to a minimum.

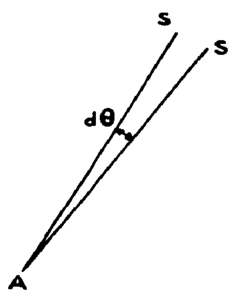

FIG. 5

Let A be the observational station, AS the satellite's observed direction and AS′ the adjusted direction, forming an angle $\delta\theta$ (figure 5). Furthermore let $\hat{p}\ \hat{q}\ \hat{r}$ be the observed directions cosines of S, $\bar{p}$, $\bar{q}$, $\bar{r}$ the adjusted direction cosines of S′, and dp, dq, dr the direction cosine corrections to be determined, i.e.

$$\bar{p} = \hat{p} + dp \qquad \bar{q} = \hat{q} + dq \qquad \bar{r} = \hat{r} + dr$$

The sum of the squares of the direction cosine deviations will be :

$$(\bar{p} - \hat{p})^2 + (\bar{q} - \hat{q})^2 + (\bar{r} - \hat{r})^2 = \bar{p}^2 + \bar{q}^2 + \bar{r}^2 + \hat{p}^2 + \hat{q}^2 + \hat{r}^2 - \\ - 2(\bar{p}\,\hat{p} + \bar{q}\,\hat{q} + \bar{r}\,\hat{r}) = 2(1 - \cos d\theta) = d\theta^2$$

It can be shown that the minimum of the total sum of $d\theta^2$ corresponds to the minimum of the total sum of the squares of corrections affecting the direction cosines.

Then again, let x_1 y_1 z_1 be the approximate coordinates of A; x_0 y_0 z_0 those of S; ρ the distance AS; and dx, dy, dz, $d\rho$ the corrections to be made to these coordinates and to ρ.

The observational equations take the following form :

$$\frac{1}{\rho}(dx_1 - dx_0) - p\frac{d\rho}{\rho} + p - \hat{p} = v_1$$

$$\frac{1}{\rho}(dy_1 - dy_0) - q\frac{d\rho}{\rho} + q - \hat{q} = v_2$$

$$\frac{1}{\rho}(dz_1 - dz_0) - r\frac{d\rho}{\rho} + r - \hat{r} = v_3$$

in which p, q and r are the approximate values for the direction cosines computed from the approximate coordinates.

The application of these equations is straightforward. Taking figure 3 as example : the coordinates of A and B are known, therefore $dx_1 = dy_1 = dz_1 = 0$. It will then be easy to compute dx_0, dy_0 and dz_0 and thus the accurate positions S_1, S_2 and S_3 of the satellite.

For observations CS_1, CS_2 and CS_3 :

$$dx_0 = dy_0 = dz_0 = 0$$

and dx_1, dy_1 and dz_1, which are corrections to the approximate coordinates of C, are computed.

5.5. Results obtained and future plans

In 1961 the French Institut Géographique National inaugurated trials for geodetic connection by satellite Echo. In 1962, following improvements in ballistic cameras, an accuracy of the order of 1/50 000 was obtained for station-to-satellite distances of from 1 500 to 2 000 km, the stations being Brest, Strasbourg, Marseilles and Bordeaux.

Afterwards the Institute carried out a geodetic connection between the French geodetic network and that of North Africa, with an accuracy of ± 50 m.

In order to obtain an accuracy of 1/100 000 (i.e. 20 m per 2 000 km) the IGN hopes to improve the results by the use of thicker (6 mm) and absolutely flat photographic plates, by augmenting the number of sights, by air-conditioning the camera, and by other improvements dictated by experience.

The IGN plans to follow this up by carrying out the complete connection of the principal vertices of the European network, as well as the connection of the Azores and Madeira which is of great interest to Portugal.

6. — GEODETIC CONNECTIONS BY DOPPLER EFFECT MEASUREMENT

A signal of continuous and steady frequency emitted by a satellite is received at the ground station with a frequency which depends on the Doppler effect caused by the change in relative speed between the satellite and the ground station.

As will later be seen, today it is possible to measure this received frequency with great accuracy, thus enabling the law of variation of satellite-to-station distance to be computed. These measurements carried out at judiciously placed stations make it possible to compute the satellite path as well as the relative positions of the ground stations. This is the theoretical principle of the systems known as TRANSIT and DOPLOC.

6.1. Computing the Doppler effect

The Doppler effect is the frequency change Δf that a signal of frequency f undergoes at a station while the transmitter and the station move with a relative speed so that the distance D varies by $Vr = \frac{dD}{dt}$. This effect is given by the well-known relation :

$$\Delta f = \pm \frac{f}{v} \frac{dD}{dt}$$

v being the propagation speed of electromagnetic waves, the plus sign corresponding to the approach (when the listening frequency increases) and the minus sign to the movement away.

In order to get an idea of this effect a small part of the satellite path $S_1 S_2$ (figure 6), in the neighbourhood of ground station T, may be considered as being circular and having a constant height KR, R being the radius of the Earth, assumed also to be spherical.

Let $S'_1 S'_2$ be the sub-satellite point's path, i.e. the projection of the projection of the satellite path on the Earth's surface, a_1 the minimum distance from the sub-satellite point to T, and a_2 the minimum distance corresponding to the sub-satellite position S'_2.

Let us consider the time origin is at position S_1. Let V be the sub-satellite point's linear speed, and D_1 and D_2 the distances from the satellite to T.

From the triangle OS_2T (O being the centre of the Earth) we have approximately :

$$D_2^2 = R^2 + K^2R^2 - 2\,K\,R^2 \cos \frac{a_2}{R}$$

and from spherical triangle $TS'_1\,S'_2$, where $S'_1\,S'_2 = Vt$, we have :

$$\cos \frac{a_2}{R} = \cos \frac{Vt}{R} \cos \frac{a_1}{R}$$

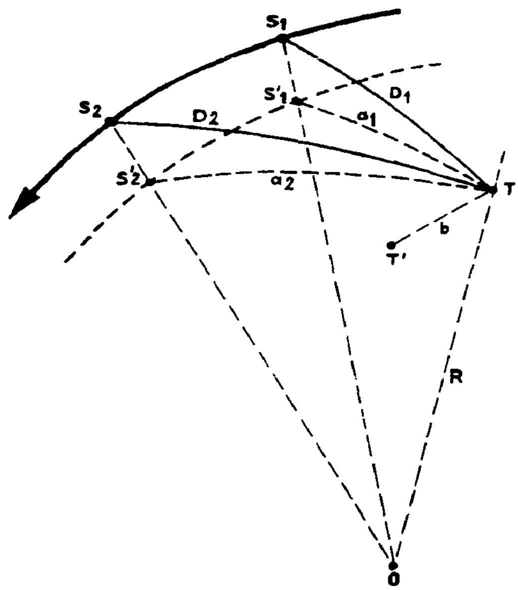

FIG. 6

Whence :

$$D_2 = R\sqrt{1 + K^2 - 2K\cos\frac{a_1}{R}\cos\frac{Vt}{R}}$$

an expression giving approximately the law of variation of distance D as a function of time t.

Its derivative in relation to time makes it possible to compute the Doppler effect :

$$\Delta f = \pm\frac{f}{v}\,\frac{KV\cos\frac{a_1}{R}\sin\frac{Vt}{R}}{\sqrt{1 + K^2 - 2K\cos\frac{a_1}{R}\cos\frac{Vt}{R}}}$$

Analysing this expression we see that the frequency variation of the signal received from a satellite is approximately a sine curve.

The law of frequency variation is given by the variation of Δf as a function of t, i.e. by the derivative of Δf in relation to t, whose value for $t = 0$ is given approximately by :

$$\left.\frac{d\,\Delta f}{dt}\right|_{t=0} = \pm\frac{f}{v}\,\frac{K}{K-1}\,\frac{V^2}{R}\cos\frac{a_1}{R}$$

We therefore see that the nearer the receiving ground station is to the sub-satellite point's path the quicker the frequency varies.

Figure 7 shows the law of variation for the received frequency of a signal of about 40 Mc/s transmitted by Sputnik at distances of 650 km (a)

and 250 km (*b*), and it indicates that it is possible to deduce the distance of this station to the sub-satellite point's path if the Doppler effect is measured accurately.

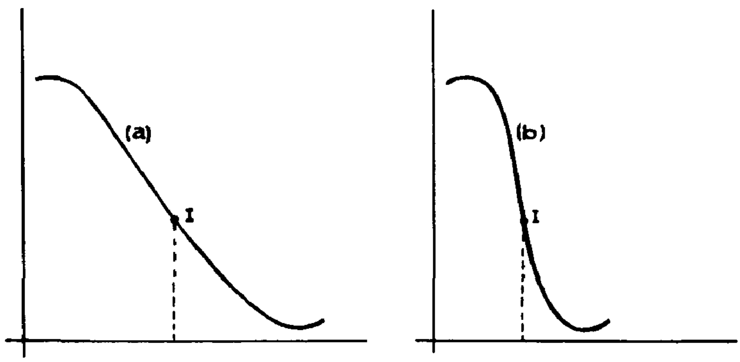

FIG. 7

On the other hand the time corresponding to the point of inflection I of the received frequency variation curve defines the sub-satellite point S_1', and consequently the position of T (figure 6) at distance a on the normal to sub-satellite path.

It is obvious that at point T′, such that TT′ is parallel to $S_1' S_2'$ and at distance b from T, the point of inflection I would be recorded at a time $\frac{b}{v}$ later than at T.

It is therefore seen that the satellite's path being known, the station's position may be derived from the curves shown in figure 7.

6.2. Doppler effect integration

The application of the principles just described comes up against great difficulties, for with present day instrumentation it is difficult to measure the law of frequency variation with sufficient accuracy to allow the station (or the satellite) position to be determined with less than a half-mile error.

Therefore the principle has been slightly modified so as to measure the total number of cycles gained or lost by the received signal in a specific time interval, i.e. so as to integrate the Doppler effect in this time interval.

This integral is proportional to the relative satellite-to-receiving station distance variation during the same time interval, for it is easy to see that the number of cycles gained or lost is given by :

$$N_2 - N_1 = \int_{t}^{t_2} \Delta f \, dt = \int_{t_1}^{t_2} \frac{f}{v} \frac{}{dt} \, dt = \frac{f}{v} (Dm_2 - Dm_1)$$

So that the Doppler effect may always remain positive in practice a constant frequency f_0 is added to the received frequency, and this entails the subsequent subtraction of $f_0(T_2 - T_1)$ cycles.

This Doppler effect integration is carried out in a receiver provided with a phasemeter and a local oscillator of strictly stabilized frequency. The heterodynation of the signal generated in the receiver with the signal received from the satellite produces a variable frequency signal which is integrated over short periods by measuring with microsecond accuracy the time interval taken to obtain a definite number of cycles.

During the satellite's passage in the vicinity of the station several measurements may be carried out, each supplying the station-to-satellite distance variation. As the satellite's path is known, each measurement gives a locus of the observed position, which is a hyperboloid of revolution, and three measurements are sufficient to define the station T position completely. This will be the intersection point of the three successively determined hyperboloids.

6.3. **Procedure now in use**

The three measurements just described would be sufficient if the satellite path were accurately known and if no observational errors existed. However, satellite paths undergo changes due to Earth's ellipticity and to gravity field anomalies as well as to the variation in the resistance of the medium in which they move (the exosphere) which is a function of its density, variable for instance with the solar eruption and sunspot variations, and consequently largely unpredictable a long time ahead. The observations therefore remain subject to the influence of the errors that we shall briefly describe later in this article (instrumental errors, ionospheric propagation errors, etc.).

It is therefore necessary to carry out a great number of observations, either at stations with known coordinates in order to correct the satellite path computations, or else at stations where the coordinates are to be determined.

For the various computations a rectangular axes system with its origin at Earth's gravity centre O was adopted, the x-axis being set towards the point of Aries, the z-axis passing through the North Pole, and the y-axis perpendicular to these two directions. Coordinates of a station T_1 can be computed in terms of time, latitude, longitude and altitude. Zt_1 is practically constant, but Xt_1 and Yt_1 vary with Earth's rotation.

The approximate satellite positions (Xs, Ys, Zs) can be computed in the same system, as functions of time and the orbit's elements, of which the principal are:

A : Satellite orbit's semi-major axis;
E : Elliptical orbit's excentricity;
i : Orbit's inclination on the Equator;
π : Right ascension of orbit's ascending node;
W : Longitude of orbit's perigee;
t_p : Transit time at perigee.

These elements are given in the satellite almanac published by NASA (U. S. National Aeronautics and Space Administration).

In this way the distance deviation between the satellite and station T in time interval $t_1 - t_2$ may be computed :

$$\mathrm{D}c_2 - \mathrm{D}c_1 = \sqrt{(\mathrm{X}t_2 - \mathrm{X}s_2)^2 + (\mathrm{Y}t_2 - \mathrm{Y}c_2)^2 + (\mathrm{Z}t_2 - \mathrm{Z}c_2)^2} - \sqrt{(\mathrm{X}t_1 - \mathrm{X}s_1)^2 + (\mathrm{Y}t_1 - \mathrm{Y}c_1)^2 + (\mathrm{Z}t_1 - \mathrm{Z}c_1)^2}$$

which is usually different from the mean distance deviation obtained by Doppler effect integration, mentioned earlier as being $\mathrm{D}m_2 - \mathrm{D}m_1$

At each station from ten to fifteen $\mathrm{D}m_2 - \mathrm{D}m_1$ measurements at the satellite's passage can be carried out, and each measurement will supply an equation of the type:

$$(\mathrm{D}c_2 - \mathrm{D}c_1) - (\mathrm{D}m_2 - \mathrm{D}m_1) = \mathrm{A}_1\Delta\mathrm{A} - \mathrm{B}_1\Delta\mathrm{E} + \mathrm{C}_1\Delta i + \mathrm{D}_1\Delta\pi + \mathrm{E}_1\Delta\mathrm{W} + \mathrm{F}_1\Delta tp$$

where A_1, B_1, C_1, D_1 and F_1 are coefficients computed in function of the orbit, and $\Delta\mathrm{A}$, $\Delta\mathrm{E}$, Δi, $\Delta\pi$, $\Delta\mathrm{W}$ and Δtp are the unknown but generally small corrections to be made to the estimated orbital elements.

By measuring, for example, 10 Doppler effect integrations at three ground stations of known coordinates, we obtain 30 equations each with 6 unknowns which will be solved by the least squares method and these supply the correct orbital elements.

Obviously complexities of computation and the necessity of obtaining results quickly entail the use of electronic computers.

As soon as the satellite path is accurately known the approximate coordinates of another ground station (situated for instance on an island to be connected geodetically to continental stations) may be corrected.

It will suffice to carry out there a series of measurements of the integrated Doppler effect, adopting a time basis strictly synchronized with the other stations by means of time signals.

From this station's approximate coordinates φ, λ, and H (latitude, longitude and altitude) the distances $\mathrm{D}c_2$ and $\mathrm{D}c_1$ to the satellite are computed at instants t_1 and t_2, as before, and they are compared to the distance differences obtained by integration of the $\mathrm{D}m_2 - \mathrm{D}m_1$ Doppler effect.

Each observation supplies an equation of the form :

$$(\mathrm{D}'c_2 - \mathrm{D}'c_1) - (\mathrm{D}'m_2 - \mathrm{D}'m_1) = \mathrm{K}_1\,\Delta\,\varphi + \mathrm{K}_2\,\Delta\,\lambda + \mathrm{K}_3\,\Delta\,\mathrm{H}$$

K_1, K_2 and K_3 being the coefficients computed in function of the approximate station position and the satellite position, and $\Delta\varphi$, $\Delta\lambda$, $\Delta\mathrm{H}$ being the unknown corrections to the station's coordinates.

A single satellite passage will supply 10 — 15 observational equations of this form, and consequently also an accurate value for $\Delta\varphi$, $\Delta\lambda$ and $\Delta\mathrm{H}$ by the least squares method.

6.4. Observational errors and results obtained

The principal errors to be considered are :

a) *Instrumental errors* — resulting from the receiver's lack of stability in frequency comparison, a stability which should be above $1/10^{10}$, and from the uncertainties in time interval measurement for which highly accurate electronic clocks must be used.

b) *Errors in ionospheric and tropospheric propagation* — It is a fact that ionospheric refraction, although not accurately known, is inversely proportional to the square of the frequency transmitted. Consequently, by using two separate frequencies f_1 and f_2 (for instance 500 Mc/s and 80 Mc/s) two different measurements of Doppler effect integration will be obtained, their errors being inversely proportional to these frequencies:

$$\varepsilon_1 = -\frac{f_1}{v}\frac{dD}{dt} + \frac{A}{f_1}$$

$$\varepsilon_2 = -\frac{f_2}{v}\frac{dD}{dt} + \frac{A}{f_2}$$

As f_1 and f_2 are transmitted having a known ratio, the undetermined coefficient A can be eliminated and consequently Δf can be computed independently of ε_1 and ε_2.

The troposphere also gives rise to refraction errors, particularly for frequencies below 1 Mc/s. Air temperature, pressure and humidity measurements, however, enable their influence to be computed with sufficient approximation.

c) *Errors in the satellite transmitted signal* — These arise from the variation in the satellite's transmitted frequency. Today near-perfect stability is obtained for a short period of about 15 minutes — the duration of the observations — i.e. when the satellite is above the observer's horizon.

d) *Computational errors* — arising from the use of digital computers, with the consequent rounding off of the last significant figure, errors which are reduced by augmenting computer capacity.

In spite of these causes of uncertainty the analysis made by ANDERLE and OESTERWINTER in the United States, based on six TRANET stations, and using relatively light and easily transportable equipment, shows that it has been possible to obtain an accuracy well within 20 metres in the determination of relative position, and it is thought that it will be possible to reach an accuracy of about 3 metres before too long, a remarkable and very interesting attainment.

7. — CONCLUSIONS AND OUTLOOK

Spatial research in a very few years has reached a stage of development and a complexity such that artificial satellites are already specialized subjects, i.e.

Interplanetary spatial craft — for transporting m and material;

Satellites for general scientific exploration — for gathering data on exosphere composition, interplanetary radiations, long distance radar communications, magnetic field variations, lunar or planetary research, etc;

Meteorological satellites (TIROS) — for gathering world-wide data on temperature distribution, cloud formations, etc.;

Navigational satellites (TRANSIT) — requiring to be put in orbits suitable for observation, so as to make easy and continuous fixing possible, both at sea and in the air;

Communication satellites (ECHO, COURIER) — for facilitating world-wide radio or television transmissions;

Military satellites (MIDAS, SAMOS) — for detecting missiles, enemy territory reconnaissance, or maybe as launching platforms for atomic weapons;

Geodetic satellites — for improving our knowledge of Earth's shape and the gravity field, and for the inter-continental connection of the various geodetic systems.

Geodetic satellites should for preference have a spherical shape, an orbit of small eccentricity with a distance to the perigee of over 700 km and a large inclination in relation to the Equator, and should have dynamical balance of masses so as to avoid libration. They should preferably have devices for making them easier to observe, such as flashing lights, stabilized transmissions, and an electronic clock giving out time signals, etc.

The main geodetic satellite to date is ANNA, which was launched in 1962 with a quasi-circular orbit at an angle of 50° with the Equator and at an altitude of 1 000 km.

Satellite Anna has a flashing light and is also provided with a transponder for distance determination by the SECOR system; it has a very stable transmitter for Doppler effect observation and emits very accurate time signals.

As Anna had various deficiencies, it is intended to launch other satellites, such as LARGOS, provided with tetrahedric reflectors for facilitating visual or laser-beam observation, and others which will also be equipped with MINITRAC and TRANSIT systems.

The most recent plan is to launch a satellite — GEOS A — with a 60° orbit inclination and a 1 100 km altitude at perigee. This satellite will be stabilized in such a way that it will always have the same face towards the Earth, and it will be provided with various observation instruments; flashing lights, dihedral reflectors for laser beams, two transponders and tranmitters for Doppler effect observation.

Satellite geodesy, of which we have given a necessarily incomplete and superficial account, today offers enormous possibilities. We may expect that before too long a most interesting objective will be attained, i.e. the standardization of world geodetic systems using one single datum throughout the world, and this will greatly simplify the various technical problems of cartography.

The range of geodetic determinations by the dynamic, geometric and radio-electric methods here mentioned is of the order of several thousand kilometres. The accuracy already obtained is of tens of metres, and it is to be hoped that this may be reduced without difficulty to several metres.

There is also a by-product of great value : satellites make it possible to confirm or else to elucidate certain hypotheses about the formation and internal composition of the Earth.

A fascinating prospect for scientific investigation is thus opened up, not only revealing some of the secrets of our planetary system and the constitution of sidereal space but also enabling us to gain a better knowledge of the Earth on which we live, and thus to enrich our lives by means of prospects inconceivable up to a few years ago.

BRIEF BIBLIOGRAPHY

Symposium on *Geodesia Tridimensional,* Venice, July 1959.

Symposium on *Geodesy in the Space Age,* Colombus, Ohio, February, 1961.

Symposium on *Geodesia Tridimensional,* Cortina d'Ampezzo, 1962.

The Use of Artificial Satellites for Geodesy — Proceedings of the First International Symposium on the use of Artificial Satellites for Geodesy, published by G. Veis, Washington, 1962.

Satellites and Scientific Research, Desmond King, Hale, London, 1962.

Geodesy, G. Bomford, 2nd Edition, Oxford, 1962.

XIIIth General Assembly of the International Association of Geodesy, Berkeley, 1963.

Introduction to Satellite Geodesy, Ivan I. Mueller, New York, 1964.

The findings of a Working Group of the Institut Géographique National, Paris, on Space Research.

Various articles published in :
International Hydrographic Review
Navigation
Bulletin Géodésique
The Journal, Coast and Geodetic Survey
Surveying and Mapping.