SATELLITE NAVIGATION AND DE-NAVIGATION

F. W. Nagle
NOAA/NESDIS Systems Design and Applications Branch
1225 West Dayton Street
Madison, Wisconsin 53706 USA

ABSTRACT

The term NAVIGATION has several connotations, but among satellite meteorologists it refers to the process whereby one applies a set of known orbital parameters, along with the scanning characteristics of an on-board instrument, to compute the position of a satellite in space at some future instant, and also the point(s) on the ground which the satellite will view at that moment.

One may ask if this process can be inverted. Thus, if we are given a set of points on the Earth's surface viewed by the satellite in the recent past, sometimes called the "footprint", is it then possible to drive the navigation algorithm backward in order to recover the orbital characteristics? And if it is possible, why would anyone wish to derive the orbital characteristics from a set of Earth locations? This inverse process will be termed 'de-navigation' in this paper, and it is our purpose to discuss its feasibility and usefulness.

1. TEMPORAL AND SPATIAL COORDINATES

1.1 Time Units

To forestall confusion involving certain terms, we shall review the temporal and spatial units used below. The term Julian Day Number is a measure of time, and it denotes the numbers of days and fraction thereof which have elapsed since 12 o'clock Universal Time on 1 January 4713 BC. The reader may wonder at this seemingly strange date as the origin of the Julian Day chronology. It is not our purpose here to explain this origin, and the reader is referred to Durant (3) for a full explanation of this choice, as well as the reason for the name 'Julian', which has nothing to do with the Julian calendar or with Julius Caesar. The term Julian Day Number is unfortunately confused with a similar term, Julian Date, which is used both in civilian and military computer systems to mean simply the two-digit year indicator, followed by a three-digit sequential day of the year, e.g. 85365 to mean 31 December 1985. One must keep in mind that a day within the Julian Day Number (JDN) chronology begins at noon, not midnight, the reason being that historically European astronomers preferred that all observations made in the course of a single night be ascribed to the same Julian Day number, which of course would not be so had the date changed at Greenwich midnight rather than noon. The JDN which began at 12 UT on 3 August 1987 was 2447011, from which the reader may deduce the JDN for any other calendar date.

On a computer system such as an IBM 4381, a double-precision quantity has a resolution comparable to about 17 decimal digits, so that by storing time expressed in JDN within double precision words, one can achieve a resolution in time of the order of a billionth of a day, easily adequate for most satellite navigation.

One can readily see the advantages of reducing measurements of time to a single scalar value, allowing us to dispense with months which may have 31, 30, 28 or sometimes 29 days, with leap years and common years, the distinction between the Julian and Gregorian calendars, etc.

1.2 Spatial Coordinates

The point in the sky which the sun occupies at the instant when it crosses the Equator from south to north about 22 March is called the Vernal Equinox, or sometimes the First Point of Aries, an old astrological name. Using the center of the Earth as the origin, the Vernal Equinox to define the x-axis, the North Pole to define the z-axis, and a point on the Equator 90 degrees east of the Vernal Equinox as the y-axis, one can construct a dextral (right-hand) orthonormal coordinate system which we call the celestial coordinate system. Since the Vernal Equinox is almost a fixed point in the sky, much like a seemingly-fixed star, the celestial coordinate system appears to rotate westward with respect to the Earth at the rate of 360 degrees per sidereal day, and in fact it is by the motion of the Vernal Equinox that a sidereal day is defined (1).

The terrestrial coordinate system is analogous to the celestial system, having the same origin and the same z-axis, but with an x-axis fixed to the Earth and directed toward the longitude of Greenwich in the plane of the Equator, and with a y-axis oriented 90 degrees east of Greenwich. Clearly the terrestrial and celestial systems rotate with respect to one another, and a transformation of a position vector in one system to the other at any time T is effected by a matrix equation of the form:

$$V_1 = \begin{vmatrix} \cos(a) & -\sin(a) & 0 \\ \sin(A) & \cos(A) & 0 \\ 0 & 0 & 1 \end{vmatrix} V_2 \qquad (1)$$

where A, the instantaneous longitude of the Vernal Equinox, is given by

$$A = -215.465 \\ -360.9856473(T - 2442348) \qquad (2)$$

The angle A is expressed in degrees, and T in Julian Day Number. The matrix is orthogonal, i.e. its inverse is also its transpose, so that the direction of the transformation depends merely on the sign of the angle A. The constant 360.9856473 is the westward rate of movement of the vernal equinox in degrees per mean solar day.

The more common terrestrial coordinates of latitude and longitude are polar variations of the Cartesian terrestrial system, and are readily deducible from the latter by simple trigonometric relationships. The distinction must be kept in mind between geodetic and geocentric latitudes, as described elsewhere (1).

2. DE-NAVIGATION

The path of an orbiting satellite such as the NOAA or DMSP series when expressed in celestial coordinates is very nearly a great circle, at least over a short period of time. The orbit of a sun-synchronous satellite, if it is to remain sun-synchronous, must precess 360 degrees per year in order that an ascending pass always present the same relationship to the sun. This amounts to just less than one degree per day, or about a fourteenth of a degree per orbit. Hence, over any significant fraction of an orbit, the path of the satellite is very close to a true great circle, and if it is necessary to consider the precession of the orbit, the needed adjustment is not difficult.

Let us assume that we are given a set of sub-satellite points on the Earth's surface, presumably expressed in terms of latitude and longitude, together with the times associated with those points, and constituting at least one complete orbit including two Equator crossings, not necessarily consecutive. If the latitudes are geodetic, these must be converted to geocentric, but otherwise the given coordinates are easily converted to terrestrial Cartesian coordinates by straightforward trigonometry. Each of the three-dimensional Cartesian vectors thus obtained is in turn easily changed to celestial coordinates by using (1). Alternatively, the terrestrial longitude of the Vernal Equinox is known as a function of time, and is given by (2) above, so that we could first have converted terrestrial longitude to celestial longitude, and thence to celestial Cartesian coordinates.

From Kepler's third law we know that the semi-major axis of an orbit is proportional to the two-thirds power of the period, where the constant of proportionality estimated by the author from orbital data stored on the NAS 9000 computer at NMC is 330.9982746. The semi-major axis is expressed in kilometers, and the period in minutes. The period itself is found by interpolating a set of known sub-satellite positions to the times of two or more Equator crossings. We now have an accurate estimate of semi-major axis, although as we shall see presently, we shall not use it except perhaps as a reasonableness test of the period. Next, we can compute the cross product of any two non-collinear satellite position vectors which by its orientation in celestial space gives us the orientation of the orbit. This vector cross product, normalized to a length of one, is called hereafter the Vector Orbital Plane (VOP). It is of course perferable if the two non-collinear satellite positions are nearly normal to each other to maximize the accuracy of their cross product. The angle between VOP and its own projection in the plane of the Equator, added to 90 degrees, is the inclination of the orbit. The angle between the same projection and the celestial x-axis, added to 90 degrees, is the right ascension of the ascending node, that is, the angle measured eastward along the Equator from the Vernal Equinox to the point of northward Equator crossing of the satellite.

Of the six classical orbital elements (semi-major axis, eccentricity, orbital inclination, right ascension of the ascending node, argument of perigee, and mean anomaly) we have now determined three. The remaining three in principle could be found, though in practice it is difficult or impossible to do so with much accuracy. Kepler's second law stating that the areal velocity of a satellite is constant can be written as

$$C = r^2 \dot{\theta}$$

and since $\dot{\theta}$ can be estimated from the successive positions of the satellite along the ground, the radius vectors from the center of the Earth to the satellite can be obtained. To these known radii, an ellipse could be fitted by some least-squares means. The point in the orbit with the shortest radius is the perigee, and that with the longest, presumably diametrically opposite to perigee, is the apogee. The sum of these two radii is the major axis. The center of the ellipse evenly divides the major axis, and the distance between the center of the ellipse and the center of the Earth (which is a focus of the ellipse) can be used to determine the eccentricity. If the epoch is taken to be the moment when the satellite is at perigee, then the mean anomaly can be set to zero, and in principle we now know the six classical elements, at least for the case of a pure two-body Keplerian orbit.

In practice, the foregoing procedure is of little value because the known Earth locations are rarely precise enough to allow the needed calculations to be accomplished without introducing large errors in the radius vectors. The eccentricity of a NOAA or DMSP satellite is roughly .0015, and the reader by using this value in the equation of an ellipse may satisfy himself that the difference between the semi-major and semi-minor axes is a mere 8 meters, whereas the mean diameter of the orbit is in the neighborhood of 14000 kilometers. Another technique was therefore preferred.

As noted, knowledge of the vector orbital plane (VOP) at a given time provides us with the orbital inclination and right ascension of ascending node. Since the rate of change of the right ascension is roughly known, or better, can be accurately inferred by noting the change in VOP over a period of several days, and since the inclination is almost invariant, we can find the orientation of the orbit at any time in the near future, i.e. for a span of several weeks ahead.

Let us next make the inaccurate assumption that the satellite moves about the Earth with a period found by observing successive

Equator crossings, with a constant angular velocity in a purely circular path, and with an epoch chosen as the moment of one of the northward Equator crossings, i.e. ascending node. The Earth locations resulting from this assumption will contain serious errors, say, of the order of of 50 kilometers. However, the errors are easily decomposed into harmonic components as a function of nodal anomaly, by which we mean the angle in the orbital plane between ascending node and instantaneous satellite position. We show a typical decomposition for the case of the DMSP-7 satellite on 4 August 1987:

Along-track errors:

0	.83	.00	.83	.00
1	1.36	7.41	7.54	79.59
2	.02	3.32	3.32	44.81
3	-.07	-.05	.08	-49.08
4	-.02	.84	.84	22.88
5	.07	.03	.07	5.09
6	-.08	-.42	.43	-16.78
7	-.02	-.08	.08	-14.75
8	.03	.91	.91	11.01
9	.07	.03	.08	2.71

Cross-track errors:

0	.59	.00	.59	.00
1	2.07	-.94	2.27	-24.44
2	.57	.16	.59	7.86
3	-.05	2.85	2.85	30.35
4	-.18	.01	.18	44.52
5	-.09	-1.94	1.94	-18.50
6	.09	.00	.09	.15
7	-.01	1.22	1.22	12.90
8	-.01	-.08	.08	-11.98
9	.01	-.79	.79	-9.89

showing ten harmonics, the zeroth through the ninth, for both along-track and cross-track errors. The fourth column of values given above for along-track amplitudes is of greatest interest. As seen, there is a constant term correction of .83 kilometers, a first harmonic component of 7.54 km, and a second harmonic of 3.32 km. Thereafter the higher harmonics are all very small, suggesting that the assumption of a purely circular motion, corrected by removing the zeroth through the second harmonics of the error, results in errors which are generally of the order of a kilometer. In fact, in this case, the remaining error, compared against the given Earth locations, had an RMS value of only 1.59 kilometers, a discrepancy which is marginally acceptable for a high-resolution Earth-viewing instrument such as AVHRR, and easily acceptable for a sounding instrument such as HIRS, MSU, or SSTI aboard the NOAA or DMSP spacecraft.

The foregoing table shows that cross-track errors tend to be much smaller than the along-track errors, though again they could be easily removed by a harmonic correction. Column 5 shows the phase angles of the along-track and cross-tracks errors.

This case displaying the harmonic error components for the DMSP-7 satellite is of interest, for it represents a practical and necessary instance of de-navigation. The author has experienced great difficulty in obtaining the classical orbital parameters from the Air Force, not because the latter is unwilling to release them as a matter of policy, but simply because their distribution is not the routine responsibility of any individual, and as a result they are usually obtainable only by special request and for a limited time only. On the other hand, the Earth-located micro-wave retrievals from the DMSP-7 can be accessed on a daily basis from the NAS 9000 system at Suitland, Md, through the McIdas system at the University of Wisconsin. Although these data are intended for meteorological use, not for navigational purposes, it has become both feasible and necessary to de-navigate the Earth locations in order to compute future positions of satellite for planning purposes. An immediate disadvantage of the foregoing technique, however, is that it affords no information of satellite altitude.

On the other hand, there is another reason for de-navigating Earth locations even for satellites like those of the NOAA series for which the orbital elements are widely published and distributed. One of the most frequently used orbital prediction models is the Brouwer-Lyddane model (2,4), which is a lengthy and involved program, and which may be unavailable at many computer sites. It is relatively slow in execution owing to its complexity, and is itself not free of error. Its size and complexity are even more disadvantageous on a small computer, such as a personal computer (PC) or lap-top model, than on a larger system such as the NAS-9000 or IBM 4381.

By de-navigating the Earth locations, we have obtained period, vector orbital plane, and the zeroth through the second error components, so that we can then apply a very simple and fast algorithm, as described below, to compute future (and past) satellite positions with an accuracy competitive with that of a classical prediction model, and with easy adaptability to a small computer system. Moreover, if our original Earth location data are accompanied with information about satellite altitude, then that too can be harmonically resolved so that we now obtain predictions of height as well as of sub-satellite locations. The computer language in which the simple model was written on the SSEC McIDAS computer, Madison, is High Level Fortran (HLF), which incorporates vectors and matrices as variable types, thus allowing the code to be written concisely in 3-dimensional vector notation (5).

Let us review and quantify the foregoing. Firstly, we are given a set of Earth-locations representing the sub-satellite points for a polar-orbiting satellite, encompassing at least two Equator crossings in the same direction, so that we can estimate the nodal period of the satellite. Normally, none of the sub-satellite points will fall exactly on the Equator, so it may be necessary to do a bit of careful interpolation to find the precise moments of Equator crossing.

Let us introduce the following notation, where the superior arrow will be used to express a vector quantity. We use the unit vectors $\vec{I}$, $\vec{J}$, and $\vec{K}$ to denote a right-hand orthonormal basis in the celestial system, where $\vec{I}$ lies in the equatorial plane pointing from the center of the Earth toward the Vernal Equinox, $\vec{J}$ toward a point 90 degrees eastward of the Vernal Equinox, and $\vec{K}$ toward the North Pole. Let $\vec{N}$ denote the unit vector pointing toward the satellite's ascending node, and $\vec{E}$ a unit vector in the equatorial plane pointed 90 degrees east of the ascending node $\vec{N}$. Clearly, the angle between $\vec{I}$ and $\vec{N}$ is the right ascension. $\vec{P}$ denotes the unit vector orbital plane, and can be computed as the normalized cross product of any two non-collinear satellite positions. Let $\vec{P'}$ be the projection of $\vec{P}$ onto the equatorial plane, and let $\vec{M}$ be a unit vector in the orbital plane pointing toward the northernmost position of the satellite. We denote the epoch T_0 as the time of first equator crossing, and by A the angular rate of change of the right ascension in degrees per day. Finally, let U be a unitizing operator which normalizes any vector by dividing it by its own components of an arbitrary vector. Then at the initial time T

$$\vec{P} = U(\vec{S}_1 \times \vec{S}_2)$$

where S_1 and S_2 are two non-collinear satellite position vectors, perferably chosen so that they are approximately orthogonal, or separated by about 25 minutes in time. Moreover,

$$\vec{P'} = P(I)\vec{I} + P(J)\vec{J} + 0\vec{K}$$

$$\vec{N} = U(\vec{P'} \times \vec{P})$$

$$\vec{E} = \vec{K} \times \vec{N}$$

$$\vec{M} = \vec{P} \times \vec{N}$$

At some later (or earlier) time T, the following relationships obtain:

$$\vec{N}(T) = \text{cosine}(At)\ \vec{N} + \text{sine}(At)\ \vec{E}$$

$$\vec{P'}(T) = \vec{N}(T) \times \vec{K}\ \text{mag}(P')$$

$$\vec{P}(T) = P'(I)\ \vec{I} + P'(J)\ \vec{J} + P(K)\ \vec{K}$$

$$\vec{M}(T) = \vec{P}(T) \times \vec{N}(T)$$

where $t = (T - T_0)$, and the mag( ) operator denotes the magnitude (length) of a vector. The three vectors $\vec{N}(T)$, $\vec{M}(T)$, and $\vec{P}(T)$ constitude an orthonormal basis, with two of them ($\vec{N}(T)$ and $\vec{M}(T)$) lying in the orbital plane. If W is the nodal period of the satellite, inferrable from the times of successive equator crossings, an approximate position of the satellite can be estimated at time T as

$$\vec{S}(T) = \text{cosine}(360\ t/W)\ \vec{N}(T) + \text{sine}(360\ t/W)\ \vec{M}(T)$$

The discrepancy between this crudely-estimated position, and the known Earth location can be obtained, i.e.

$$\vec{X}(T) = \vec{S}(T) - \vec{S}$$

Let us use the lower case letters $\vec{i}$, $\vec{j}$, and $\vec{k}$ to denote an orthonormal coordinate system attached to the moving satellite, where $\vec{k}$ is a unit vector pointing downward toward the center of the Earth, i points in the direction of motion of the satellite, and j points to the right of the motion. We have:

$$\vec{k}(T) = -U(\vec{S}(T))$$

$$\vec{i}(T) = \vec{k}(T) \times \vec{P}(T)$$

$$\vec{j}(T) = \vec{k}(T) \times \vec{i}(T)$$

Then the discrepancy vector can be resolved into vertical, along-track, and cross-track components by:

$$x(T) = \vec{X}(T) * \vec{i}(T)$$

$$y(T) = \vec{X}(T) * \vec{j}(T)$$

$$z(T) = \vec{X}(T) * \vec{k}(T)$$

where * is the dot product operator. These three error components can be resolved by a Fourier analysis as a function of the nodal anomaly (360 t/T), and the resulting coefficients then used to correct the approximate position S(T) with a more precise one:

$$\vec{S}(T) = \vec{S}(T) + \vec{i}\ (c0 + c1\ \text{cosine}(A-e_1)$$

$$+ c2\ \text{cosine}((2(A-e_2)) \qquad (3)$$

The values c0, c1, and c2 are cosine coefficients obtained from the Fourier analysis, and e and e are corresponding phase angles. We have now corrected the satellite position vector S(T) for the along-track error associated with the first approximation. Similar corrections may be applied to remove cross-track and vertical errors as well, though these are usually far smaller than the along-track error. In fact, for the NOAA-10 satellite on 10 September, the Fourier error components were:

harmonic	x	phase	y	phase	z	phase
0	5.33	180.00	0.02	0.00	4.30	0.00
1	14.65	107.04	0.74	61.88	8.24	16.50
2	0.72	-46.29	0.01	39.36	1.47	-0.21
3	0.04	-9.75	0.02	40.44	0.01	29.90
4	0.03	13.94	0.01	37.48	0.01	24.16
5	0.01	29.63	0.01	-21.90	0.01	-25.34
6	0.02	-26.19	0.01	-5.25	0.01	-7.44
7	0.02	-8.70	0.01	4.12	0.01	2.75
8	0.02	-6.02	0.01	-12.45	0.01	-7.60
9	0.02	3.61	0.00	19.44	0.01	11.21

showing that the amplitude of the along-track error was 14.65 kilometers, whereas the cross-track amplitude was .74 km. The maximum vertical amplitude, if we care to remove it, is 8.24 km. The biases and RMS errors of the satellite positions, predicted five days into the future, i.e. for 15 Sepetmebr 1987, expressed in kilometers, were:

x,y,z	BIAS:	4.09	0.02	-0.04
	RMS:	5.55	0.56	2.06

where a positive x-bias means that the satellite position predicted by equation (3) was ahead (further along the orbit) than the true position. This result compares favorably with the position predicted by the Brouwer-Lyddane model, which had 10.38 km as the along-track RMS error, versus 5.55 km for the simple model.

3. SUMMARY

De-navigation of Earth-located satellite data affords not only a means to obtain orbital characteristics which may otherwise be difficult to obtain, but also allows the use of a much simplified navigation algorithm, suitable for small computers, faster than the classical Keplerian-Newtonian models, and competitive with the latter in accuracy.

REFERENCES

1. Bowditch, N., 1977, 'American Practical Navigator', Pub. No. 9, Vol I, Defense Mapping Agency Hydrographic Center, p. 371

2. Brouwer, D., November 1959, 'Astronomical Journal', Vol 64, No. 1274, p.378

3. Durant, W., 1935, 'The Story of Civilization', Simon and Schuster, New York

4. Lyddane, R.H., October 1963, 'Astronomical Journal', Vol 68, p. 555

5. Nagle, F., 1987, 'High-Level Fortran', unpublished document available on request, Madison.

The Technical Proceedings of

The Fourth International TOVS Study Conference

Igls, Austria

March 16-22, 1988

Edited by

W. P. Menzel

Cooperative Institute for Meteorological Satellite Studies
Space Science and Engineering Center
University of Wisconsin
1225 West Dayton Street
Madison, Wisconsin 53706
(608) 262-0544

October 1988