Orbital Dynamics for the COMPLEAT IDIOT

by Alan R. Washburn

1. Introduction

Actually, a *complete* idiot would have trouble with this tutorial. The real intended audience is scientifically oriented, but pressed for time. The objective is to provide a minimal acquaintance with the technology and terminology relevant to Earth orbiting satellites. In spite of this mild ambition, the vocabulary expansion in what follows is nontrivial. Terms that are used in a special sense are italicized when first defined – 24 of them. There is even a formula …

2. Reference system

The coordinate system used in this tutorial has its origin at the center of Earth. In this coordinate system the Sun goes around Earth, contrary to what you may have heard. There is actually no harm in this viewpoint, since the origin of coordinates can always be selected to be whatever is convenient. It is true that Earth would be an awkward origin if we were interested in the motion of other planets, since they have complicated orbits in Earth-centered coordinates. But the planets do not sensibly affect Earth satellites, which are the subject of interest here, so the center of the Earth will do fine.

The problem is how to orient the x, y, and z axes. One is easy – the z axis will point to the North Pole along Earth's axis of rotation. We want to have a direction for the x axis that doesn't rotate with Earth, so that an observer could cling to the axis frame and see all of the stars other than the Sun being stationary – an inertial frame of reference. There are lots of such directions. The convention is to orient the x axis in the direction of the *vernal equinox* ♈, which is by definition the direction to the Sun on the day in Spring when day and night have the same length. Any time you can see the constellation Pleiades, you are looking in approximately the direction ♈. Since ♈ is in the equatorial plane, there is a right angle between the x and z axes. Now orient the y axis to be perpendicular to both of them, and you have our coordinate system. You might wish to imagine a transparent sphere centered on these coordinates, with all of the stars other than the Sun fixed to it. This is the *celestial sphere*. If you sit on top of this sphere and look down at Earth, you will see it rotating counter-clockwise on its z axis down there.

The beauty of this coordinate system is that, neglecting small perturbations that will be described in Section 5, the orbit of any satellite is a fixed ellipse. The satellite moves around the ellipse, but the ellipse itself remains stationary. It takes six numbers to describe such an orbit. The six could be simply position (three numbers) and velocity (three numbers) at some reference time, but a different set is used in practice and is worth getting used to. Two numbers describe the plane and direction of rotation. A satellite's *ascending node* is the direction from the origin to the satellite when it ascends through the equatorial plane going North. The *right ascension* (RA) of the ascending node is an angle measured counter-clockwise (as viewed from the North Pole) from the x axis to the ascending node, an angle between 0 and 360 degrees. The other number is the angle in degrees between the equatorial plane and the plane of rotation, the orbit's *inclination* (IN). For *prograde* (counter-clockwise) orbits, $0 \le IN < 90$. For *retrograde* (clockwise) orbits, $90 < IN \le 180$. All orbits that go over the North Pole are *polar*, and have IN=90. Orbits with IN=0 or IN=180 are all *equatorial*, but the first is prograde and the second retrograde. Figure 1 illustrates RA and IN for a circular, prograde orbit. In this Earth-centered coodinate system, the Sun's orbit is nearly circular and prograde with IN=23.5 and RA=0.

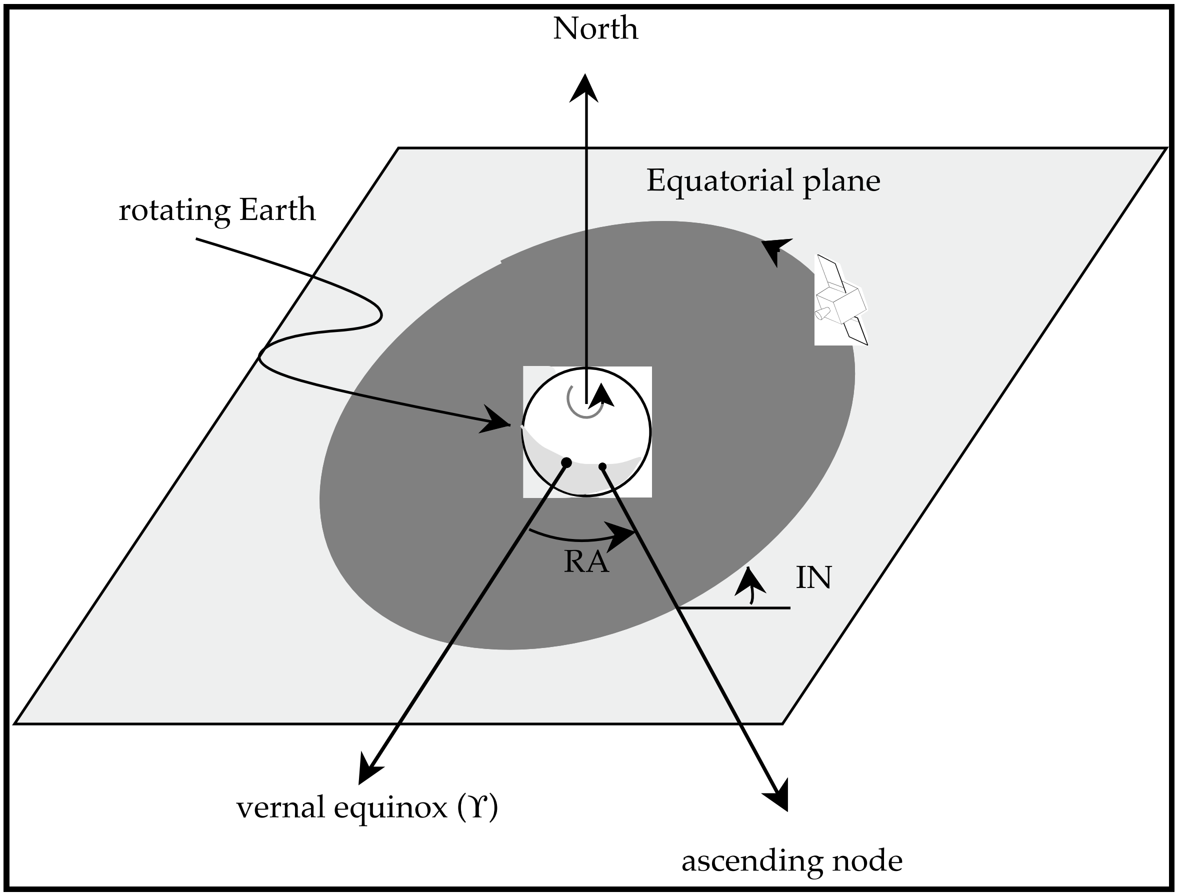

Figure 1. The right ascension (RA) and inclination (IN) of a satellite's orbit

3. Kepler's Laws

As Kepler discovered and as Isaac Newton first showed rigorously in 1685, the solution of the two body problem is that the path of body A with respect to body B is a conic section. When the conic section is a parabola or hyperbola, body A "escapes" and never returns; otherwise, body A (the satellite) repeatedly travels around an ellipse with one focus at body B (Earth, in the present application). This is Kepler's first law. Three additional numbers are required to determine this ellipse within the plane already determined by RA and IN. Systems for doing this vary – the details are not important for the moment. The sixth and last number locates the satellite on the ellipse at some standard time. After the satellite is located on the ellipse, physics takes over and the satellite rotates round and round Earth indefinitely.

Kepler's second law states that the time required to go from one orbital position to another is proportional to the area of the circular wedge formed by the line between Earth and the satellite. When the line is long, the satellite moves slowly. Therefore a satellite moves fastest at *perigee* (closest point to Earth) and slowest at *apogee* (farthest point).

Kepler's third law states that the *period* (T) of the orbit is related to its size in that T^2 is proportional to r^3, where r is the *semimajor axis* of the ellipse (half of its largest diameter). For satellites whose mass is negligible compared to that of Earth, the exact relationship is that

$$T = 2\pi\sqrt{\frac{r^3}{gR^2}},$$

where

g = gravitational attraction at Earth's surface (.00981 km/sec^2)
R = radius of Earth (6378 km).

For satellites in circular orbit, r is just the radius of the orbit. For example, a satellite in circular orbit 570 km above Earth would have r=6948 km and T=5760 seconds (96 minutes). If atmospheric interactions are ignored, a satellite with r=R would be feasible and would have a period of 84 minutes (it would also have a speed of 7.9 km/sec, providing a real thrill to observers as it skims Earth's surface). In practice most satellites stay at least 200 km above Earth's surface to avoid the atmosphere, so think of an hour and a half as a lower bound on any satellite's period. The satellite's mass does not affect the orbit as long as it is small compared to Earth – the formula works for the Moon, but not for the Sun.

Earth rotates on its axis once every 23 hours and 56 minutes, the length of a *sidereal day*. Remember that the Sun is moving around Earth in a prograde orbit, and that a (solar) day is the length of time required for a given point on Earth to have the Sun overhead again. After Earth has rotated counterclockwise once, it still needs another 365th of a day (about 4 minutes) to catch up to the Sun. A prograde orbit with a period of one sidereal day (circular or not) is *geosynchronous*; that is, the satellite's Earth track will be the same every day. Circular, geosynchronous orbits have a radius of about 42,164 km.

A circular, geosynchronous orbit with inclination IN=0 will appear to hover above a point on the Equator. These orbits are *geostationary*. Specifying the fixed point on the Equator completely defines the location of a geostationary satellite for all time, so there can be only one geostationary orbit for every longitude. Geostationary orbits are popular for communication satellites because the corresponding ground antennas don't have to move much. Fortunately, since the circumference of the geostationary orbit is 267,000 km, there is room for quite a few satellites. From its position at 5.6 Earth radii above the surface, a single geostationary satellite can see as far as 81 degrees north or south latitude – almost a complete hemisphere. Three or four carefully spaced satellites can cover most of the Earth, with the missing parts being polar or ocean regions where very few people live. This is all very convenient. On Jupiter, the corresponding orbit would be only 1.25 radii above the surface, and would consequently not be visible beyond a latitude of 64 degrees. On our sister planet Venus, which hardly rotates, the orbit would be inconveniently far away at a radius of 1.5 million km (244 radii). In yet another way, Earth turns out to be a good place for Man.

4. Effects of Earth's rotation

Of the six satellite parameters, only RA and IN would be of much use if Earth did not rotate. These two alone would determine the track of the satellite on Earth, with the other four parameters affecting only revisit time and the velocity and altitude of the satellite as it passes over. This would have the advantage of conceptual simplicity – most satellite coverage problems could be investigated with a globe and some means of drawing great circles on it. But the operational disadvantages of a nonrotating Earth would be immense, even putting aside the fact that all life would die out. There would not be any geostationary orbits, for example, since one could not count on the Earth to expose itself by rotating under the satellite. In spite of the conceptual complications, Earth's rotation is basically a Good Thing.

The coverage pattern achieved by a satellite on a rotating Earth depends on the relationship between the satellite's period T and Earth's sidereal period E. The ratio $Q \equiv E/T$ is the *repetition factor*. The orbit is synchronous if Q=1. If Q=2, as in the Global Positioning System, each satellite makes two complete orbits in one sidereal day. A satellite with Q=14/3 would repeat itself every 3 sidereal days after 14 revolutions. The ground track of a circular satellite with Q=15 and IN=63 is shown in Figure 2. This satellite appears to weave a kind of "basket" over Earth in the process of making its 15 daily orbits, revisiting Norfolk and every other point on its track once a day. On account of Earth's rotation, a single satellite can come reasonably close to every point on Earth once a day, within the latitudes that its orbit is designed to cover. If the mesh of the basket is too coarse, use a second satellite offset by 360/(15x2)=12 degrees of RA. If the revisit time is too large, put multiple satellites in the same orbit.

If Q is irrational, then the ground track never repeats! Even when the ground track repeats, some remarkably odd shapes are possible. Figure 3 shows the track of a circular synchronous satellite with IN=45. The track is confined to the predictable latitudes, but moves around in a figure 8 pattern. Figure 4 shows the track of an elliptic orbit with Q=2. In Figure 4, apogee occurs twice a day at the two northern cusps; the closeness of the hourly hashmarks means that the satellite is moving slowly there, as is always true in elliptical orbits. There are many Russian Molniya satellites with orbits of this type.

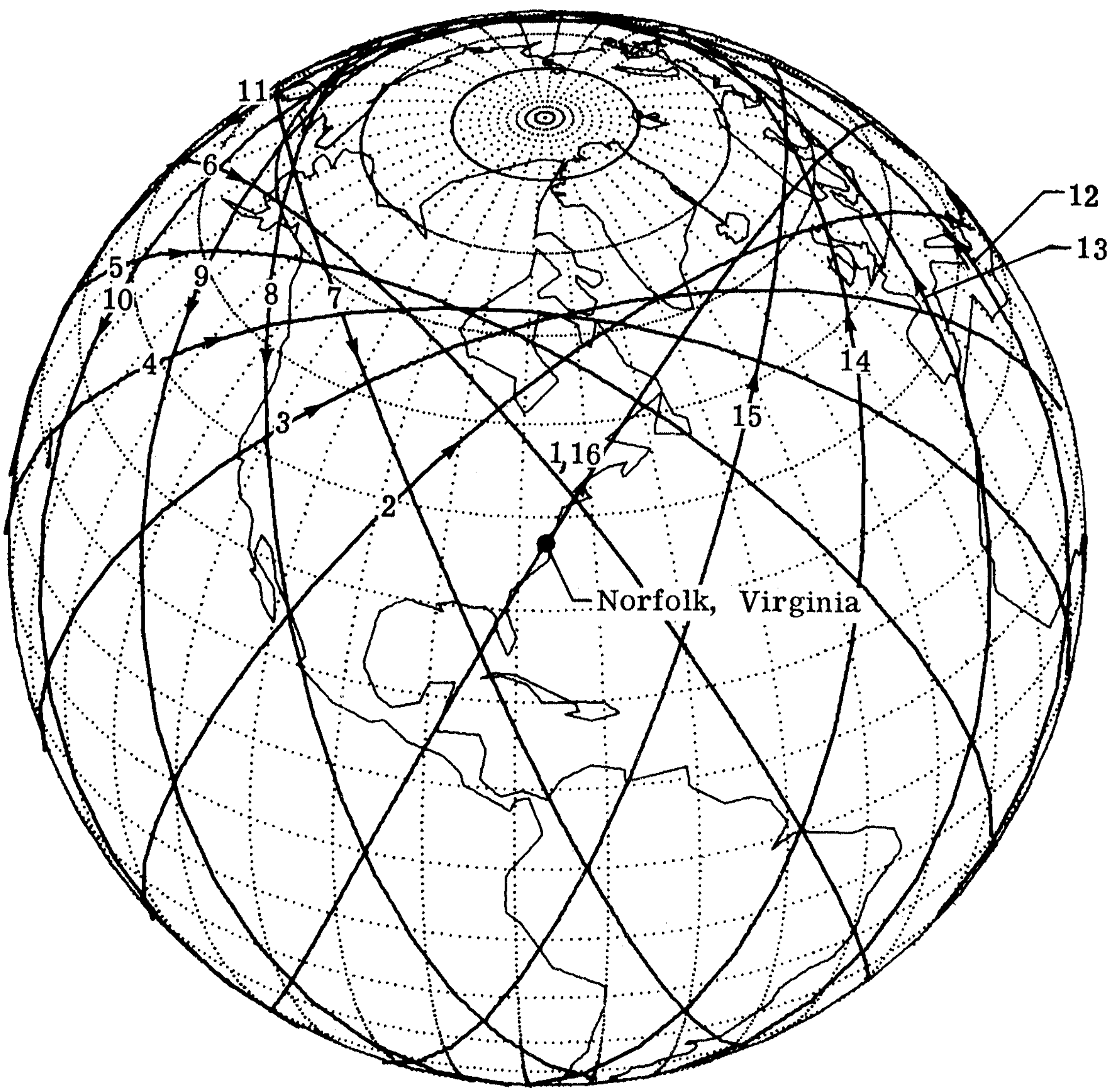

Figure 2. Ground track of a satellite with repetition factor Q=15 and inclination IN=63. Norfolk is viewed once a day.

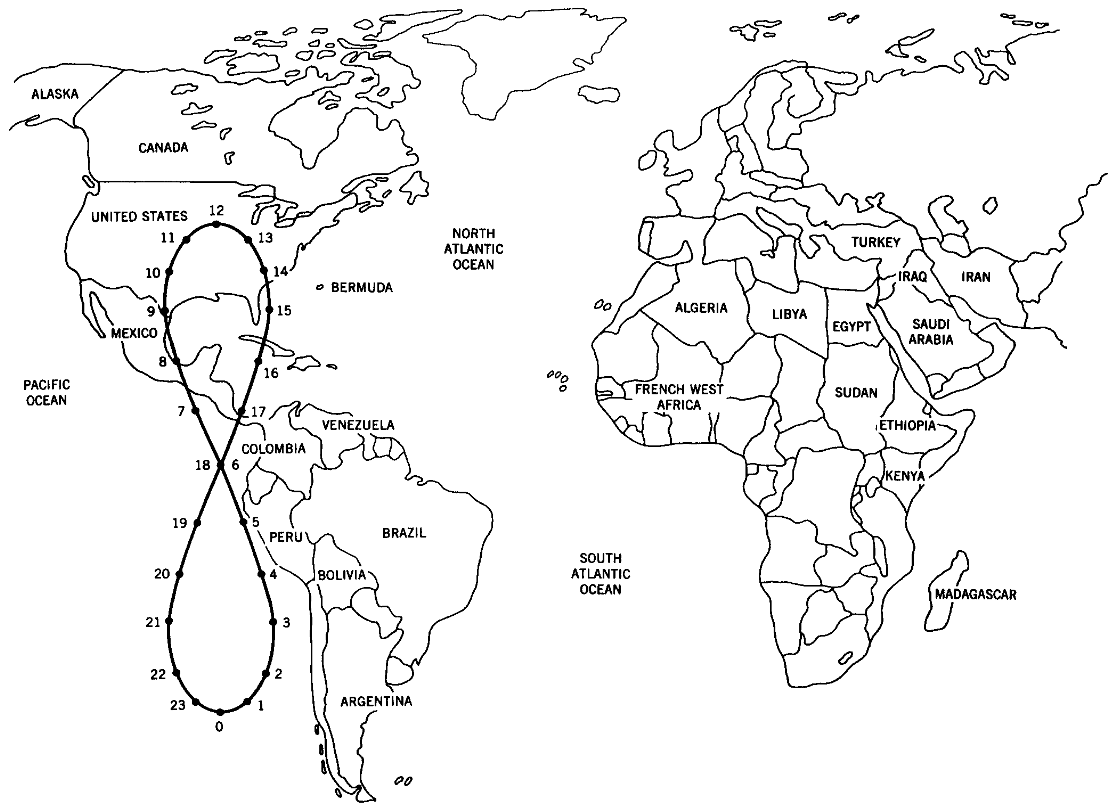

Figure 3. Ground track of a circular synchronous orbit with IN = 45°

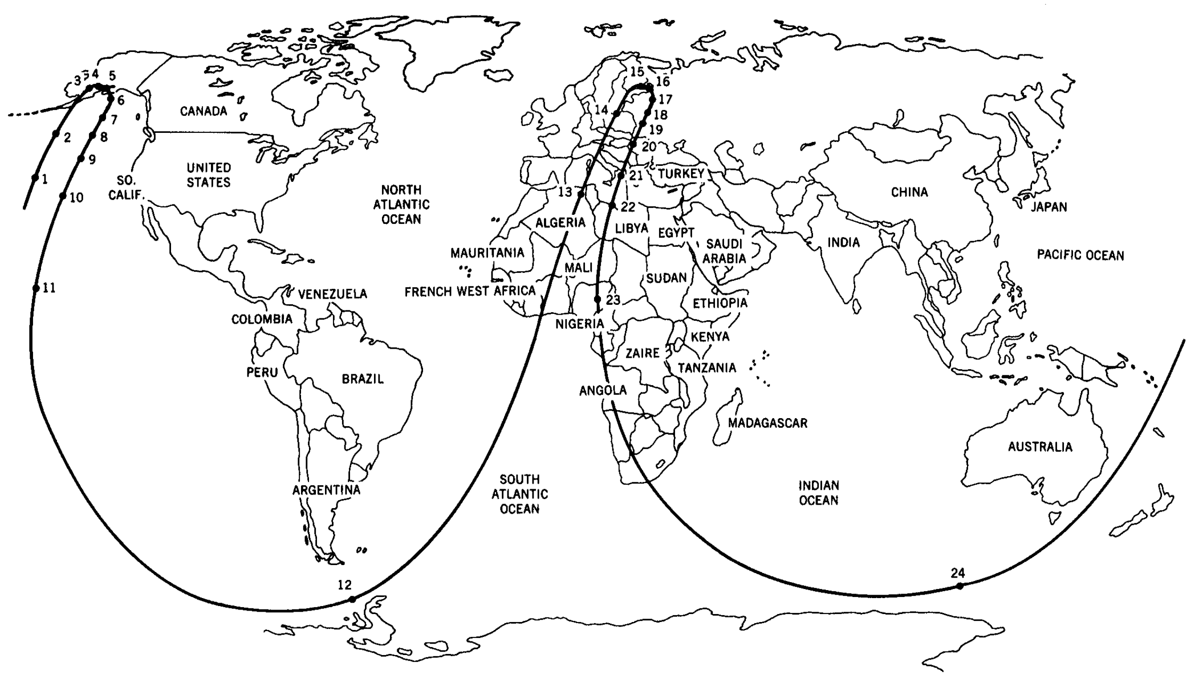

Figure 4. Ground track of an elliptical orbit with repetition factor Q=2

5. Perturbations

Due to its rotation, Earth is a bit fat around the Equator. As a result, the motion of a satellite is perturbed from what it would be if Earth were a perfect sphere. The effect is the same one that applies to a spinning top when Earth's gravity tries to make the axis of spin horizontal. Instead of falling down, the axis of the top's rotation precesses about the vertical. Now, a satellite is basically just a big top, and the bulge at the Equator continually tries to move the satellite's axis of rotation into the Equatorial plane. The reaction of the satellite is that its axis of rotation precesses about the North Pole; the inclination of the orbit remains constant, but the Right ascension (RA) changes slowly with time. RA decreases for prograde orbits and increases for retrograde orbits, with polar orbits being unaffected. The magnitude of this drift in RA is on the order of several degrees per day, depending on the altitude and inclination of the satellite, so the effect on Earth coverage can be substantial over only a few days. Satellite tracking programs invariably account for it.

Orbit precession is an analytic annoyance, but a clever orbit designer can put even this phenomenon to use. Recall that a synchronous orbit passes over the same point on Earth once every sidereal day. If that point is in sunlight at one time, it will be in darkness six months later because of the Sun's orbit around Earth. This can be awkward if photography is involved – it would be better if the satellite could pass over a given point at the same time(s) every day. Precession can be used to make that happen. If the satellite is in a retrograde orbit, then its axis of rotation will move counterclockwise around the North Pole, just like the Sun. If the axis makes exactly one rotation per year, then the orbit will *Sun-synchronous*, the desired effect. Satellites in near-polar retrograde orbits are usually put there with this effect in mind. They are likely to be in low orbits, since the precession effect dies out quickly with altitude.

Another perturbation that must be allowed for is the inexorable effect of friction, particularly for satellites in low orbit. Sooner or later, every such satellite will spiral inward until it finally crashes on the surface of Earth, if it hasn't burned up already in the atmosphere. In the meantime, the six parameters of the orbit that approximates the satellite's actual track must be periodically adjusted.

6. Statistics about satellites

Satellite databases always include RA, IN, three elliptical parameters, one parameter to locate the satellite on the ellipse, and possibly other information. The three elliptical parameters are:

eccentricity (EC) This is half of the ratio of the distance between foci to the semimajor axis. Necessarily $0 \le EC < 1$. EC is zero for a circular orbit. As EC approaches 1, a small perturbation could turn the ellipse into a parabola, in which case the satellite would never come back.

argument of perigee (AP) This is the angle from the ascending node to the perigee, measured in degrees within the plane of rotation. Thus $0 \le AP \le 360$.

mean motion (MM) This is the number of revolutions per day. MM is 1.0027 for a synchronous satellite. The reciprocal of MM is the period T in days.

The sixth parameter locates the satellite on the ellipse at some reference time. The *true anomaly* is the angle from perigee to the satellite (satellite terminology has inherited terms like "argument" and "anomaly" from the ancient science of astronomy — both of these are names for angles). The *mean anomaly* (MA) is the angle from perigee to where the satellite would be if it moved around its orbit at a constant angular rate, which it really doesn't. MA is essentially a measure of time since perigee, so it is used in preference to true anomaly in spite of its odd definition.

In theory a single reference time (1957, say) would do for all satellites, but in practice each satellite has its own *epoch* (EP), a time in the hopefully recent past at which the other six parameters are asserted to have been accurate. Of the six parameters, only MA should depend on EP in theory. In practice the other five also have a slight dependence on account of perturbations.

Data for various collections of satellites can be downloaded from the Internet. A good starting location is `http://emerald.feldberg.brandeis.edu/~progrmer/satellite/index.html`. The most common format is NASA format where each satellite gets three lines. In that format EP is in the second line, along with other data that includes the last two digits of the year of launch (YR). The third line includes IN, RA, EC, AP, MA, and MM, in that order, except that EC is given as an integer that should be multiplied by 10^{-7}. NASA format is invariably called two-line format, apparently with the idea that the first line, which contains only the satellite's name, should not be counted. The six lines below describe OSCAR 10 and ANIK C2 (TELESAT-7):

```
OSCAR 10
1 14129U 83058B   96046.56186910 -.00000124  00000-0  10000-3 0  4034
2 14129  26.3643 223.9075 5983094 351.0706   1.5442  2.05879267 67343
ANIK C2 (TELESAT-7)
1 14133U 83059  B 96056.22874657 -.00000269  00000-0  10000-3 0  6892
2 14133   3.6497  66.1128 0001290 248.7237 206.0228  1.00270375 16217
```

Lines labelled 1 and 2 have the following format (elements defined above are shown in [ ]):

Column	Description of element in line labelled 1
01-01	Line Number of Element Data
03-07	Satellite Number
10-11	International Designator (Launch year) [YR]
12-14	International Designator (Launch number of the year)
15-17	International Designator (Piece of launch)
19-20	Epoch Year (Last two digits of year) [EP]
21-32	Epoch (Julian Day and fractional portion) [EP]
34-43	First Time Derivative of the Mean Motion divided by 2 or Ballistic Coefficient (Depending of ephemeris type)
45-52	Second Time Derivative of Mean Motion divided by 6.

54-61	BSTAR drag term if GP4 perturbation theory was used. Otherwise, radiation pressure coefficient.
63-63	Ephemeris type
65-68	Element number (counts the number of updates)
69-69	Check Sum (Modulo 10)

Column	**Description of element in line labelled 2**
01-01	Line Number of Element Data
03-07	Satellite Number
09-16	Inclination [IN]
18-25	Right Ascension of the Ascending Node [RA]
27-33	Eccentricity (times 1E7) [EC]
35-42	Argument of Perigee [AP]
44-51	Mean Anomaly [MA]
53-63	Mean Motion [MM]
64-68	Revolution number at epoch
69-69	Check Sum (Modulo 10)

NASA format does not include a field for the satellite's purpose, owner, or launcher, but such information can be found in other databases. The key field is the satellite number that is shown on both lines in the two-line format. Every satellite has a unique number, and it never changes.

You may have access to a file SAT.DAT that was created by downloading the 3530 satellites that were orbiting in March, 1996, and processing the data so that each satellite has a single line in SAT.DAT. The eight columns show (NUM,YR,IN,RA,EC,AP, MA,MM) for each satellite, where NUM is the satellite number. For example, the line corresponding to OSCAR 10 is

14129 83   26.3643 223.9075 .5983 351.0706   1.5442 2.05879267

SAT.DAT is an ASCII file that should be readable by most statistical software, so statistical measures of various kinds can be derived from it.