Earth Coverage by Satellites in Circular Orbit

Alan R. Washburn

Department of Operations Research
Naval Postgraduate School

The purpose of many satellites is to observe or communicate with points on Earth's surface. Such functions require a line of sight that is neither too long nor too oblique, so only a certain segment of Earth will be covered by a satellite at any given time. Here the term "covered" is meant to include "can communicate with" if the purpose of the satellite is communications. The shape of this covered segment depends on circumstances — it might be a thin rectangle of width w for a satellite-borne side-looking radar, for example. The shape will always be taken to be a spherical cap here, but a rough equivalence with other shapes could be made by letting the largest dimension, (w, for the radar mentioned above) span the cap. The questions dealt with will be of the type "what fraction of Earth does a satellite system cover?" or "How long will it take a satellite to find something?" Both of those questions need to be made more precise before they can be answered.

Geometrical preliminaries

Figure 1 (top) shows an orbiting spacecraft sweeping a swath on Earth. The leading edge of that swath is a circular "cap" within the spacecraft horizon. The bottom part of Figure 1 shows four important quantities for determining the size of that cap, those being

R = radius of Earth (6378 km)
r = radius of orbit
α = cap angle
β = masking angle.
2γ = field of view

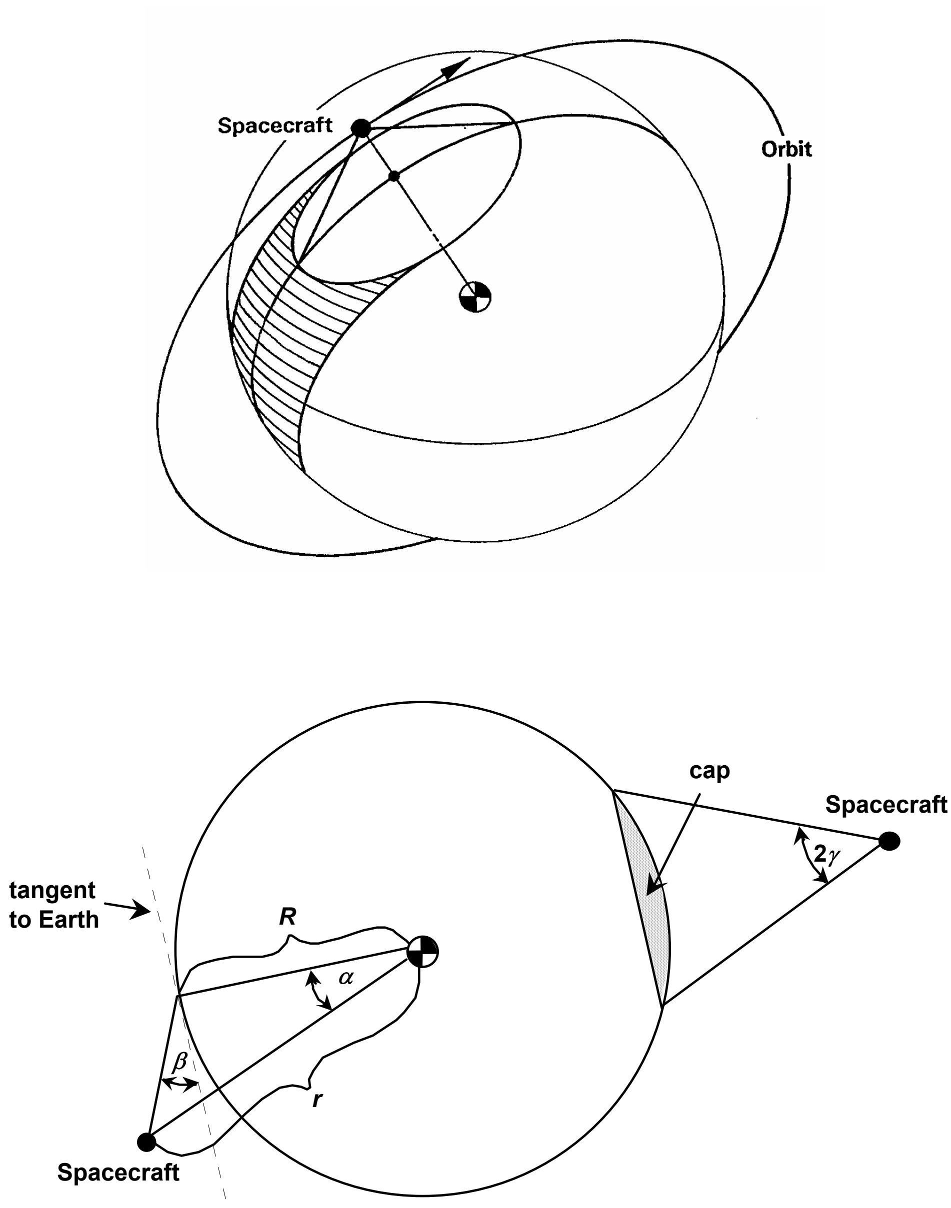

FIGURE 1. Illustrating the cap angle (α), the masking angle (β), and field of view (2γ).

A target is visible to the satellite only if the satellite's elevation, as viewed by the target, is at least β. The relation between α and β is that, with $\rho = R/r < 1$, using the law of sines on the sides r and R, plus the fact that $\gamma + \alpha + \beta = \pi/2$,

$$\alpha = \cos^{-1}\left(\rho \cos\beta\right) - \beta; \qquad 0 \le \beta \le \pi/2. \tag{1}$$

Coverage may also be limited by a sensor's field of view, rather than a masking angle, in which case,

$$\alpha = \sin^{-1}\left(\left(\sin\gamma\right)/\rho\right) - \gamma; \qquad \sin\gamma \le \rho . \tag{2}$$

If sin $\gamma > \rho$, the sensor's field of view is not limiting and formula (1) applies with β=0.

A cap of size α covers a fraction f of Earth, where

$$f = \frac{1-\cos\alpha}{2}; \qquad 0 \le \alpha \le \pi/2 . \tag{3}$$

For example, consider a GPS satellite for which r = 26,561 km; this is the altitude corresponding to a 12 hour orbit. For a 10 degree masking angle, formula (1) with ρ = .24 produces sin α = .9158, so α = 1.16 radians, and f is .30. Each GPS satellite can see 30% of Earth. If GPS satellites were synchronous (24 hour orbit), r would be 42,164 km., and f would increase to .36. The largest possible value of f for fixed β is when ρ = 0 (satellite very far away), in which case $\alpha = \pi/2 - \beta$ and $f = (1 - \sin\beta)/2$. When $\beta = 10°$, this limiting fraction is .41. Of course, increasing f by increasing a satellite's altitude also has some disadvantages, increased insertion cost and on-board power requirements being among them.

Satellites intended for surveillance or remote sensing are often in orbits considerably below those of GPS, often only a few hundred kilometers above Earth's surface, in order to provide the required spatial resolution. Such satellites have relatively short rotational periods in accordance with Kepler's third law. The number of orbits completed by a satellite in the length of one sidereal day[1] is by definition the repetition factor Q. This factor is 2 for GPS, but is commonly in the range of 10-15 for satellites in low orbits.

[1] Earth's position with respect to the Sun is the same every day (24 hours). Earths's position with respect to the fixed stars is the same every sidereal day (23 hours and 56 minutes, the difference 4 minutes being 1/365 of a year)

These satellites have a much smaller coverage fraction f than does GPS, typically limited by a sensor's field of view.

Effects of inclination

A satellite's inclination (I) is the angle between Earth's equatorial plane and the plane that contains the satellite's orbit. A satellite's latitude (ψ) is the latitude of the point directly beneath the satellite on Earth. This latitude will never exceed I in absolute magnitude, and is determined by the anomaly θ according to

$$\sin\psi = \sin\theta \sin I \ . \tag{4}$$

The anomaly θ is an angular measure of the satellite's progress in moving around its orbit. For circular orbits, θ increases linearly with time. Formula (4), as well as formula (6) below, are special cases of Napier's rules for right spherical triangles, with the right angle being on the equator between latitude and longitude.

A satellite spends much of its time near the extreme latitudes $\pm I$. This effect is most easily quantified by regarding θ as a uniform random variable on $[0, 2\pi]$ and deriving the corresponding density of the random latitude ψ. Letting $f_I(x)$ be the density function of ψ, the result is that, in radian measure,

$$f_I(x) = \frac{1}{\pi} \frac{\cos x}{\sqrt{\sin^2 I - \sin^2 x}}; \qquad -I < x < I \ . \tag{5}$$

Figure 2 shows this density for inclinations of 30, 60, and 90°. Polar satellites ($I = 90°$) spend equal amounts of time above every latitude, but satellites with small inclinations have a strong tendency to appear near the latitude limits.

Relative distances

A satellite's ability to find (get within its covered cap) a specific target on Earth depends on the target's latitude. This is already evident from the results of the previous section, since it is clear that a target whose latitude exceeds $I + \alpha$ will never be seen by a satellite with positive inclination. Figure 2 implies that targets whose latitude is slightly

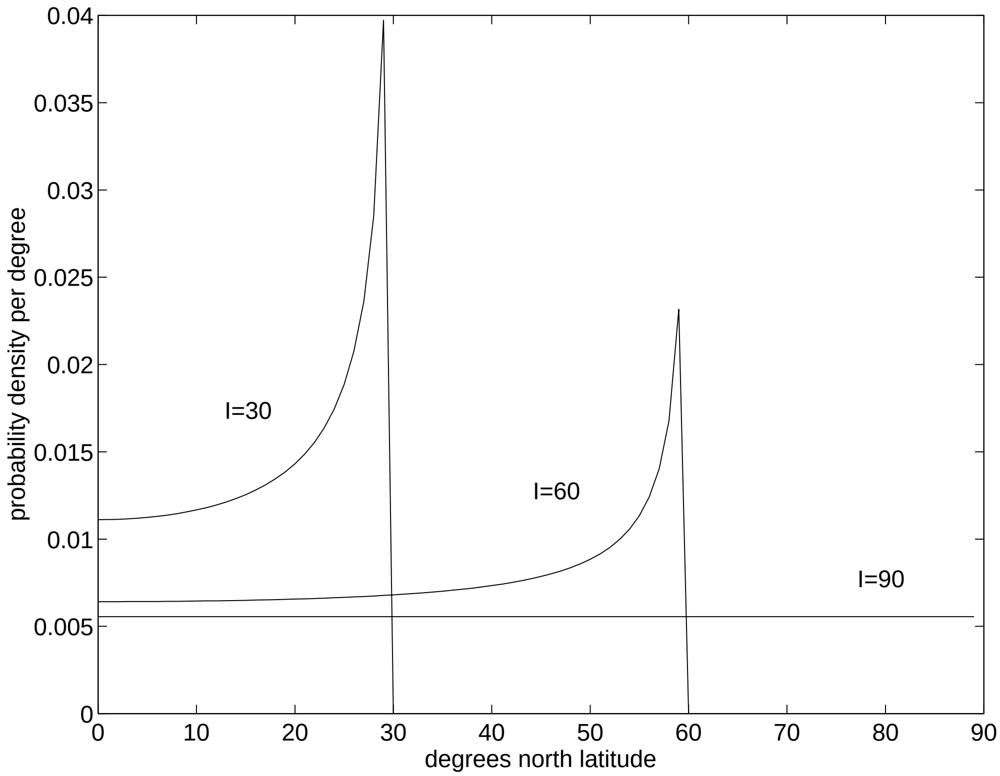

FIGURE 2. The density of a satellite's position over northern latitudes

smaller than I should be easiest to find, since the density of the satellite's latitude is greatest there. However, even targets within the satellite's latitude band may not be seen if the longitudes of target and satellite don't match. The relative longitude is influenced by Earth's rotation, as well as the satellite's. A careful analysis will require a reference system within which the locations of the satellite and its target can both be measured. The natural frame of reference is the celestial sphere, a sphere that is centered on Earth but which does not rotate. Relative to that sphere, the satellite revolves around a fixed, circular orbit. Latitude and longitude can be measured as usual if the celestial sphere is provided with an axis parallel to Earth's axis of rotation, except that longitude needs to be measured relative to some fixed direction in space, rather than the rotating longitude line that passes through Greenwich, England.

The latitude (ψ) and longitude (ϕ) of the satellite on the celestial sphere are determined by the inclination (I) of the orbit and the anomaly. Equation (4) applies to ψ, and the longitude is governed by

$$\tan \phi = \tan \theta \cos I. \tag{6}$$

The satellite's location is a periodic function of θ, and therefore also a periodic function of time. Note the implicit convention in (6) that longitude is measured relative to the direction to the satellite when it crosses the equator, since ψ and ϕ are both 0 when θ is 0.

Let ψ_E and ϕ_E be the latitude and longitude of the satellite's target on Earth. Since the target is assumed stationary,

$$\begin{aligned} \psi_E &= \psi_0 \\ \phi_E &= \phi_0 + \theta/Q \end{aligned} \tag{7}$$

where Q is the number of satellite orbits that correspond to one rotation of Earth (the repetition factor). ψ_0 and ϕ_0 are, by definition, the target's latitude and longitude when $\theta = 0$.

Let A be the angle between (ψ, ϕ) and (ψ_E, ϕ_E). This is the angle that needs to be smaller than the cap angle α if the satellite is to detect the target. Then, using spherical trigonometry,

$$\cos A = \cos\psi_E \cos\psi \cos\left(\phi_E - \phi\right) + \sin\psi_E \sin\psi \ . \tag{8}$$

By expanding $\cos(\phi_E - \phi)$ and using (4), equation (8) can be rewritten as

$$\cos A = B\cos\theta + C\sin\theta \ , \tag{9}$$

where $B = \cos \psi_E \cos\phi_E$, and $C = \cos \psi_E \sin \phi_E \cos I + \sin \psi_E \sin I.$ (10)

Figure 3 shows a graph of A versus time as measured by satellite revolutions ($\theta/2\pi$) when $Q = 15$, $\phi_0 = 1$ radian, $I = 1$ radian, and for $\psi_0 =$ 1.2, 1.0, and .8 radians. The separation A is in all cases a sinusoidal function modulated by a slowly varying envelope that is shown supporting it. In the bottom graph the target is located outside of the latitude band covered by the satellite, so the minimal separation is $\psi_0 - I = .2$ radians. In the top graph the target is located inside the covered latitude band, and the minimal separation is almost 0. Note that the envelope is zero twice during every Earth rotation (15 satellite revolutions) in this case. The center graph has the target exactly on the edge of the covered band.

The features shown in Figure 3 are typical, as can be shown by rewriting cos A in yet another way. Equation (9) can be written

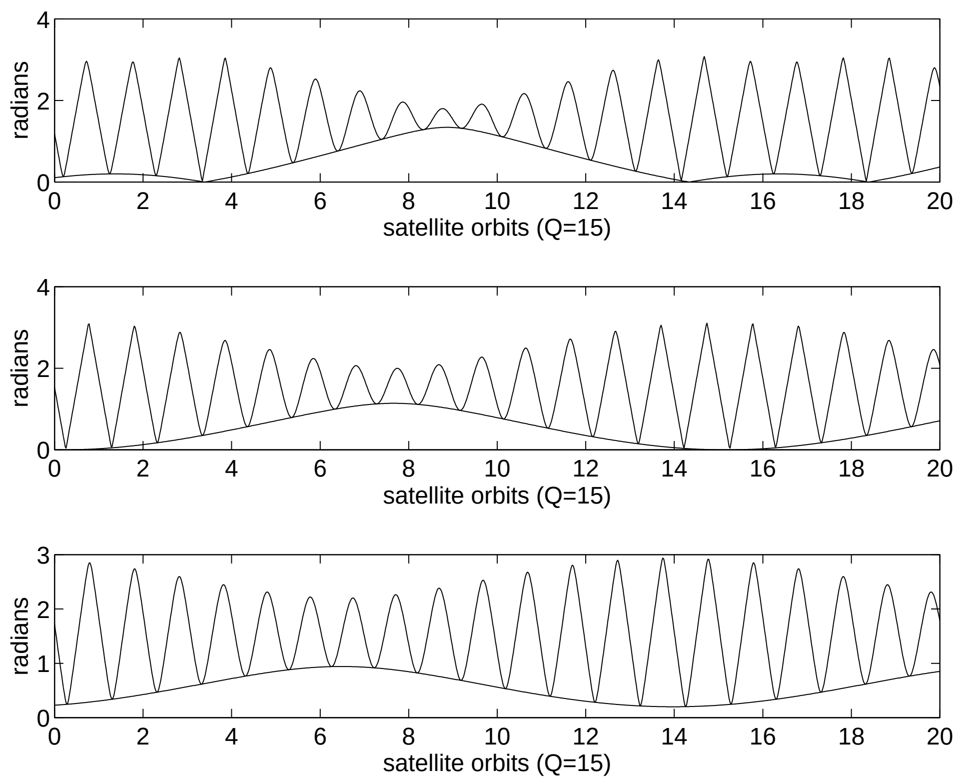

FIGURE 3. The angle separating a satellite from its target versus time

$$\cos A = \cos E \sin(\theta + D), \tag{11}$$

where $\cos^2 E = B^2 + C^2$ and $\tan D = B/A$. Since neither B nor C involves θ, D and E have period Q when time is measured in satellite orbits; that is, D and E are slowly varying functions of time. E is the angle separating the target from the satellite's *orbit*, necessarily smaller than the angle A separating the target from the satellite itself. E is the envelope shown in Figure 3. E is zero two times per Earth rotation in the top graph of Figure 3. This is because the satellite's orbit covers the target once ascending and once descending, as must always happen for any target within the latitude band covered by a satellite.

It can be shown (see appendix) that

$$\sin E = \left|\sin\psi_E \cos I - \sin\phi_E \sin I \cos\psi_E\right|. \tag{12}$$

The only time dependent factor in (12) is ϕ_E. If $\left|\sin\psi_E\right| < \left|\sin I\right|$, there will always be some value of ϕ_E that makes $E = 0$; otherwise, the smallest possible value for E is $\left|\psi_E - I\right|$ or $\left|\psi_E + I\right|$, whichever is smaller. All of these features are evident in Figure 3.

Probability of detection and random search

The initial value of the anomaly θ was taken to be 0 in Figure 3, but this value θ_0 could in principle be any number between 0 and 2π. Given θ_0, ϕ_0, ψ_0, Q, I, and α, the amount of time until the satellite first covers the target (if ever) is deterministic, and can be easily found by inspecting diagrams like those in Figure 3. However, there are so many parameters involved that it is difficult to get a grasp on how well a given satellite covers the various latitudes within its covered band. Figure 2 implies that coverage is best near the band edges, but the coverage situation is complicated by the effects of Earth's rotation. As a first approach to quantifying Earth coverage, consider a latitude ψ_E within the satellite's covered band. As the satellite moves from its southernmost to northernmost latitude (a "pass"), it will include a certain fraction g of latitude ψ_E within its covered cap. If the target's longitude is random, this fraction is a probability of detection. The first order of business is therefore to determine g.

It will be assumed in the following that $\alpha << \pi$, in which case the local coverage pattern as the satellite passes over latitude ψ_E is as if Earth were flat. Figure 4 shows the covered region and the covered length W of the latitude line.

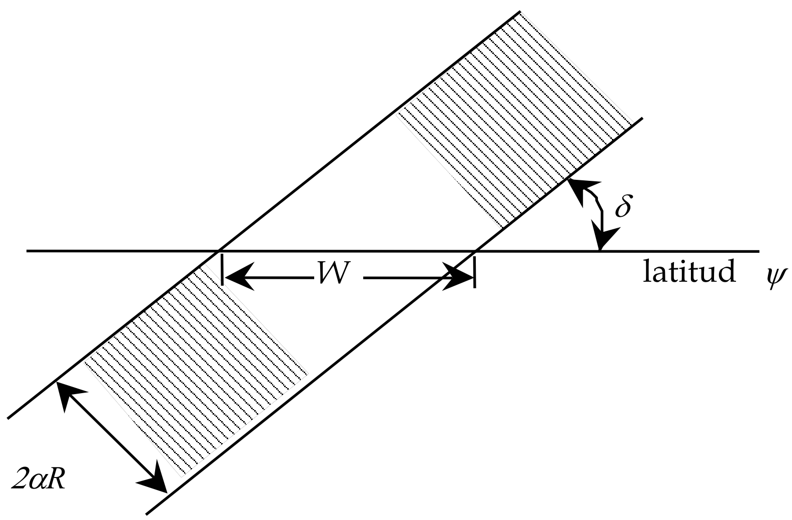

FIGURE 4. Satellite coverage for small α.

The width of the covered strip is $2\alpha R$, where R is Earth's radius, but W is larger than $2\alpha R$ because the satellite does not ascend vertically through the latitude line. The angle δ is determined by the velocity of the satellite relative to Earth, the horizontal and vertical components of which are $R\left(\frac{d\phi}{d\theta} - \omega\right)\cos\psi$ and $R\frac{d\psi}{d\theta}$, where time is measured by the anomaly θ and $\omega = 1/Q$. ω is Earth's rotation rate in terms of radians of rotation per radian of anomaly. By differentiating (4), one discovers that $\frac{d\psi}{d\theta} = \cos\theta \sin I/\cos\psi$. By differentiating (6) and using trigonometric identities, one discovers that $\frac{d\phi}{d\theta} = \cos I/\cos^2\psi$. Finally,

$$\tan\delta = \frac{\frac{d\psi}{d\theta}}{\left(\frac{d\phi}{d\theta} - \omega\right)\cos\psi_E} = \frac{\cos\theta \sin I}{\cos I - \omega \cos^2 \psi_E}. \quad (13)$$

The cos θ factor can be eliminated from (13) by using (4), after which

$$\tan\delta = \frac{\sqrt{\sin^2 I - \sin^2 \psi_E}}{\cos I - \omega \cos^2 \psi_E}. \quad (14)$$

The length of the latitude line is $2\pi R \cos \psi_E$, so the fraction of the line covered per pass is

$$g(\psi_E, \omega, I, \alpha) = \frac{W}{2\pi R \cos\psi_E} = \frac{2\alpha R/\sin\delta}{2\pi R \cos\psi_E} = \frac{\alpha}{\pi \sin\beta \cos\psi_E}. \quad (15)$$

Using the fact that $1/\sin^2\delta = 1 + 1/\tan^2\delta$, equation (15) can be put in the form

$$g(\psi_E, \omega, I, \alpha) = \frac{\alpha v(\psi_E, I, \omega) f_I(\psi_E)}{\cos\psi_E}, \quad (16)$$

where $f_I(\psi_E)$ is the density function given by (5), and where

$$v(\psi_E, I, \omega) = \sqrt{1 - 2\omega\cos I + \omega^2 \cos^2 \psi_E}\ . \quad (17)$$

The function $v(\psi_E, I, \omega)$ is the surface speed of the subsatellite point (the point on Earth's surface between the satellite and Earth's center) in units of radians per radian of anomaly. Note that

- $v(\psi_E, I, 0) = 1$ (no Earth rotation)
- $v(0, 0, 1) = 0$ (geostationary case)
- $v(0, \pi, 1) = 2$ (a retrograde case)

Since a satellite makes two passes per revolution over each latitude line and Q revolutions per day, a target on latitude ψ_E will be detected $2Qg(\psi_E, \omega, I, \alpha)$ times per day, on the average. If there is a collection of N satellites, each with a small cap angle, then the total average rate at which the collection detects a target on latitude ψ_E is (recall $Q_i = 1/\omega_i$)

$$TR = \sum_{i=1}^{N} 2g\left(\psi_E, \omega_i, I_i, \alpha_i\right) / \omega_i \ . \tag{18}$$

In computing *TR*, it must be borne in mind that $g(\psi_E, \omega_i, I_i, \alpha_i) = 0$ unless $\sin^2 \psi_E <$ $\sin^2 I_i$. If α_i is a constant α, then *TR* is $(180\alpha/\pi)\, H(\psi_E)$, where

$$H\left(\psi_E\right) = \frac{1}{90\cos\psi_E} \sum_{i=1}^{N} v\left(\psi_E, I_i, \omega_i\right) f_{I_i}\left(\psi_E\right) / \omega_i \ . \tag{19}$$

The function $H(\psi_E)$ is "average detections per degree of cap angle per day on latitude ψ_E". Figure 5 shows $H(\psi_E)$ as a function of latitude for the collection of 2000 satellites in low Earth orbit in early 1996, treating all orbits as if they were circular. The maximum value of the sum is about 280 detections per degree per day at 82 degrees. If you want to *avoid* satellites, don't go to a pole (too many polar orbiters) or the Equator (every satellite crosses the Equator). The best place is at latitude 85°, only three degrees above

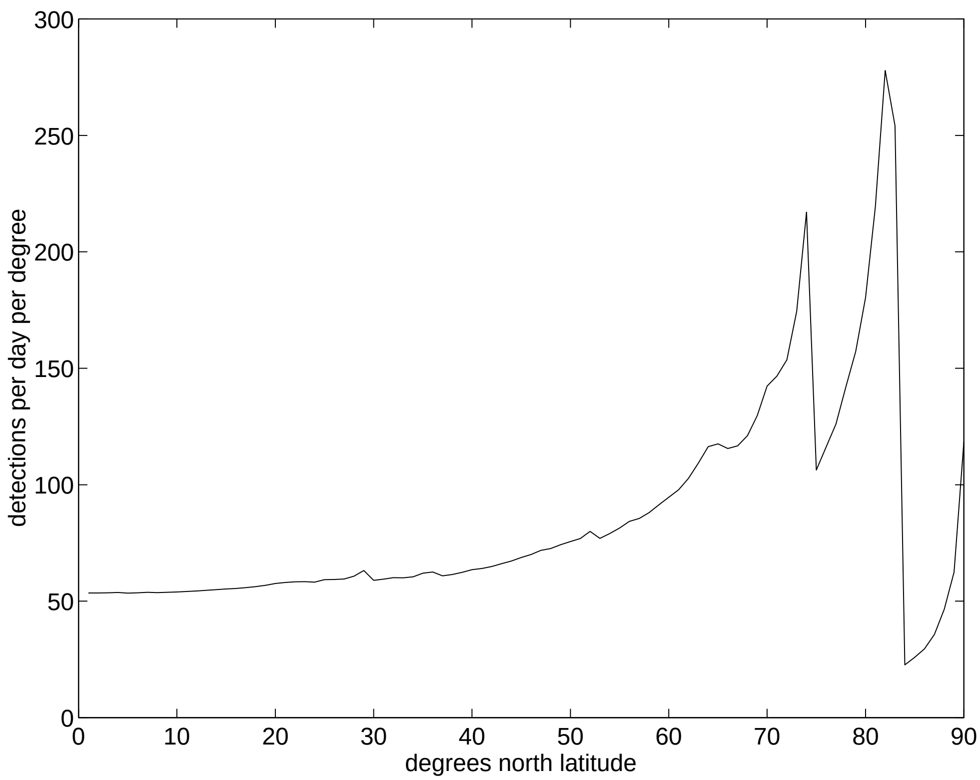

FIGURE 5. Intensity of satellite coverage versus latitude

the worst place (the 85 degree latitude avoids a large number of Cosmos satellites with inclination 83 degrees). Even at 85°, a .1 degree circle (radius 10 km) would still include the passes of about three satellites (the subsatellite points, that is) per day, on the average. The graph for southern latitudes would of course be identical to Figure 5.

The word "average" used above in describing detection rates refers to an average over the target's longitude. When Q is an integer, it could be that targets on most longitudes will never be seen, and that the average rate is achieved by including a high rate of detection on whatever set of longitudes the satellite happens to include in its ground track. In such circumstances, knowledge of the detection rate alone will not be of much use in predicting the time to detection or the intervals between detections. However, there are several phenomena that should make detections tend toward a Poisson process, among which are

- the target's longitude may vary from pass to pass
- the total rate may be achieved by summing the rates of several satellites
- "detection" may only be a necessary condition for some other stochastic event
- Q may not be an integer or even a rational number.

In such circumstances, it may be reasonable to refer to the detection rate as λ, the conventional symbol for the rate of a Poisson process, to make predictions accordingly, and to speak of the search process as "random". On account of the large number (2000) of uncoordinated satellites involved in computing Figure 5, for example, it is reasonable to refer to the rate being plotted as λ, and to argue that incidents where some satellite comes within 10 km of a point fixed at 85 degrees are a Poisson process with rate 3/day.

For targets that are only intermittently visible, the amount of coverage time per pass may be important. This time depends on the lateral range X in radians from the satellite's track to the target, being 0 if $|X| \geq \alpha$. If α is small and $|X| < \alpha$, the amount of track length in radians over which the target is visible is $Y = 2\sqrt{\alpha^2 - X^2}$. This can be converted to a coverage time by dividing by the satellite's ground speed measured in radians per orbit, $2\pi v(\psi_E, I, \omega)$.

When α is small, X will be uniformly distributed over the interval $[-\alpha, \alpha]$, given that $|X| < \alpha$. Under the same condition, the average value of Y is therefore $\pi\alpha/2$. Letting τ be the orbital period, the average contact time per detection is thus

$$(\pi\alpha/2)(\tau)/(2\pi v(\psi_E, I, \omega))$$

or

$$T(\psi_E, \omega, I, \alpha) = \frac{\alpha\tau}{4v(\psi_E, I, \omega)}. \tag{20}$$

Since there are two chances at detection per orbit, the fraction of the time that the target is within view of the satellite is $2g(\psi_E, \omega, I, \alpha)\, T(\psi_E, \omega, I, \alpha)/\tau$, or

$$c(\psi_E, I) = \frac{\alpha^2 f(\psi_E, I)}{2\cos\psi_E}. \tag{21}$$

Note that $c(\psi_E, I)$ does not depend on Earth's rotational rate ω.

As an example of the kind of calculations that are possible, consider a submarine patrolling in the vicinity of $\psi_E = 45$ degrees. The submarine is invisible unless it uses its periscope, which it does occasionally in a Poisson process with rate $r = .1$/day. If the periscope emerges at a time when a satellite is within $\alpha = .1$ radians of it, the submarine will be detected. A single satellite with $I = 60$ degrees and $Q = 10$ searches for periscope emergences. How long will it take for the satellite to detect the submarine? One way to answer the question is to reason that each satellite pass "covers" the submarine with probability $g(45°, .1, 60°, .1\ \text{rad}) = .0857$, but that each coverage is unlikely to detect a periscope emergence because the average length of coverage is only $T(45°, .1, 60°, .1\ \text{rad}) = .063$ hours (using $\tau = 2.4$ hours in (20)). The detection probability per pass is only $(.0857)\,(.063r) = .000540$, so the average number of passes required is

1/.00540 = 1852. Since there are 20 passes per day, the time required is 92.6 days. An alternate method is to argue that the Poisson detection rate is $\lambda = r\, c(\psi_E, I) =$ (.1/hour) (.0045) = .00045/hr., the reciprocal of which is 92.5 days. The alternate method is simpler but perhaps less intuitive. If the periscope stayed up for a non-negligible amount of time, then a pass-based analysis would be required.

In calculating detection probabilities when the target wishes to avoid detection, it should be borne in mind that one of the weaknesses of satellite surveillance is that the locations of satellites are quite predictable. If the target's policy is to act only when no satellite is present, then of course the detection probability will be 0. There is some literature on the design of constellations of satellites where certain latitudes, at least, are *always* covered. For example, see Rider (1986), Hanson and Higgins (1990), or Wilkinson (1994).

References

Hanson, J.M. and Higgins, W.B., "Designing Good Geosynchronous Constellations," *Journal of the Astronautical Sciences*, Vol. 38, No. 2, April-June 1990, pp. 143-159.

Rider, L., "Analytic Design of Satellite Constellations for Zonal Earth Coverage Using Inclined Circular Orbits," *Journal of the Astronautical Sciences*, Vol. 34, No. 1, January-March 1986, pp. 31-64.

Wilkinson, C.K., "Coverage Regions: How They are Computed and Used," *Journal of the Astronautical Sciences*, Vol. 42, No. 1, January-March 1994, pp. 47-70.

APPENDIX

The Separation Angle A

From (8), the angular separation A between satellite (ϕ, ψ) and target (ϕ_E, ψ_E) is

$$\cos A = \cos\psi_E \cos\psi \cos\left(\phi_E - \phi\right) + \sin\psi_E \sin\psi \ . \tag{A1}$$

If the point (ϕ, ψ) is governed by equations (4) and (7), there will be some time (anomaly) θ at which this separation is minimal. The object in this appendix is to investigate the dynamics of A, especially the minimal value.

Since $\cos\left(\phi_E - \phi\right) = \cos\phi_E \cos\phi + \sin\phi_E \sin\phi$, and since (from (4)) $\sin\phi = \sin\theta \cos I/\cos\psi$ and $\cos\phi = \cos\theta/\cos\psi$, a simple substitution and cancellation shows that (9) is true, and (9) can be put in the form (11) using a trigonometric identity. Let $R = \cos E$, so that $R^2 = B^2 + C^2$, and let $\mu = \sin\phi_E \tan I/\tan\psi_E$. Then $\cos^2\phi_E = 1 - \mu^2\tan^2\psi_E/\tan^2 I$, so

$$B^2 = \left(\cos\psi_E \cos\phi_E\right)^2 = \cos^2\psi_E - \mu^2 \sin^2\psi_E \big/ \tan^2 I \ . \tag{A2}$$

Also, since $\cos\psi_E \sin\phi_E \cos I = \mu \sin\psi_E \cos^2 I/\sin I$,

$$C^2 = \sin^2\psi_E \left(\mu \cos^2 I/\sin I + \sin I\right)^2 . \tag{A3}$$

Now find the coefficients of μ^0, μ^1, and μ^2 in R^2. The result is

$$R^2 = 1 - \sin^2\psi_E \cos^2 I + 2\mu\left(\sin^2\psi_E \cos^2 I\right) - \mu^2 \sin^2\psi_E \cos^2 I$$

or

$$R^2 = 1 - \sin^2\psi_E \cos^2 I \left(1-\mu\right)^2 . \tag{A4}$$

Since $R = \cos E$, (A4) can also be written

$$\sin^2 E = \sin^2\psi_E \cos^2 I\left(1-\mu\right)^2 , \tag{A5}$$

which is equivalent to (11). The minimal value of $\sin^2 E$ is achieved when $\mu = 1$, as will be feasible if $\left|\tan\psi_E\right| < \left|\tan I\right|$. As a curiosity, since (A5) is a concave function of μ, the *maximal* value will be achieved at an extreme value of μ, which will be when ϕ_E is either $\pi/2$ or $-\pi/2$. It will be characteristic of the maximum angle E_{max} that $0 \leq E_{max} \leq \pi/2$, and that

$$\sin E_{max} = \left|\sin\left(\psi_E \pm I\right)\right|. \tag{A6}$$

For example, if $I = 2/3\pi$ (retrograde orbit) and $\psi_E = -\pi/2$ (South Pole), the angle between target and orbit is $\pi/6$.

Exercises

1. Obtaining a terrestrial fix with GPS satellites requires that at least 4 be simultaneously in view. The 24 GPS satellites are actually placed in carefully selected orbits to make this event be as likely as possible over the mid-latitudes where most users are found, but what if they were not? Supposing each satellite can view 30% of Earth's surface,

   a) What is the average number of satellites that can see a randomly selected point on Earth's surface?

   b) If all 24 satellites had independent, identically distributed, uniform locations on the celestial sphere, what would be the probability that at least 4 would be able to see a user in Monterey, California?

2. Derive equation (5).

3. Use equation (3) to derive and sketch the density of ψ, the latitude of a point selected randomly on Earth's surface.

4. Use the answer to exercise 3 and equation (21) to show that the average fraction of the time that a randomly selected point on Earth's surface is within view of a satellite is $\alpha^2/4$, where α is a small cap angle, regardless of the inclination I. Explain why this result is "obvious".

5. It is claimed on p. 13 that the average value of Y is $\pi\alpha/2$. Prove it.

6. A missile is launched at latitude 45°. If a satellite is within a masking elevation of 5°, the launch will be witnessed by a thermal sensor carried on board. There are 10 such satellites, each in a circular orbit with $Q = 40/\pi$, (an irrational number), and $I = 60°$. What is the probability that the launch will be witnessed? State any assumptions.