AAS03-638

APPLICATIONS OF ERGODIC THEORY TO COVERAGE ANALYSIS

Martin W. Lo*

The study of differential equations, or dynamical systems in general, has two fundamentally different approaches. We are most familiar with the construction of solutions to differential equations. Another approach is to study the statistical behavior of the solutions. Ergodic Theory is one of the most developed methods to study the statistical behavior of the solutions of differential equations. In the theory of satellite orbits, the statistical behavior of the orbits is used to produce "Coverage Analysis" or how often a spacecraft is in view of a site on the ground. In this paper, we consider the use of Ergodic Theory for Coverage Analysis. This allows us to greatly simplify the computation of quantities such as the total time for which a ground station can see a satellite without ever integrating the trajectory, see Lo [1,2]. More over, for any quantity which is an integrable function of the ground track, its average may be computed similarly without the integration of the trajectory. For example, the data rate for a simple telecom system is a function of the distance between the satellite and the ground station. We show that such a function may be averaged using the Ergodic Theorem.

THE SATELLITE COVERAGE ANALYSIS PROBLEM

The study of differential equations, or dynamical systems in general, has two fundamentally different approaches. The more familiar approach is the construction and study of the solutions to differential equations. Another approach is the study of the statistical behavior of the solutions. Ergodic theory is one of the most developed methods for studying the statistical behavior of the solutions of differential equations. See Sinai[3] and Arnold[4] for references and an introduction to this field. In this paper, we apply these methods to the Satellite Coverage Analysis Problem.

The *Satellite Coverage Analysis Problem* is the study of the statistics of interactions between a satellite and other objects in space. The most common example of this problem is the analysis of the visibility of satellites to ground stations on Earth. A more complex problem is the analysis of the coverage between a rover on Mars and the Deep Space Network on Earth through a telecommunications spacecraft orbiting Mars such as the Mars Telecommunications Orbiter. In this paper we use ergodic theory to compute satellite coverage performance. This approach greatly simplifies the computation of quantities such as the total time for which a ground station can see a satellite without integrating the trajectory. This was shown by Lo [1,2]. More over, for any quantity which is an integrable function of the satellite position or ground track, its average may be

* Navigation and Mission Design Section, Jet Propulsion Laboratory, California Institute of Technology

computed similarly without the integration of the trajectory. For example, the data rate for a simple telecom system is a function of the distance between the satellite and the ground station. In this paper we show that such a function may be averaged using ergodic theory.

The Coverage Analysis Problem, at its simplest, is the study of the visibility properties of a satellite in orbit around the Earth from a point on Earth. In Figure 1, we depict the satellite ground track of a circular orbit and the circular region of visibility from a point P (at the center of the circle) on the Equator in the Pacific Ocean. Geometrically, whenever the satellite ground track enters this circle, it is in view from the station on the Equator. We define the following variables for this discussion. Let **D** be the circular region of visibility of a spacecraft from a ground station centered on the Equator in the Pacific Ocean. Let **A** be the annulus region defined by the ground tracks of the spacecraft. Let ℵ(D) denote the area function, it this example, the area of the region D on the sphere. See Figure 1.

Simplistically, one may think that the percentage of time **T** spent by the satellite in the circle D would be well approximated by the area of the intersection between the circle

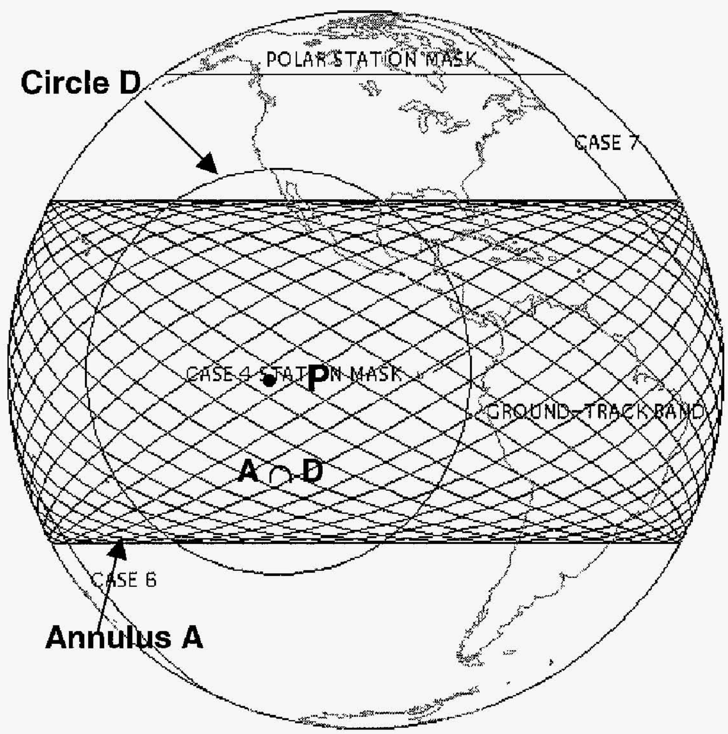

Figure 0 The ground tracks of a circular satellite forming the annulus A and the circular coverage region D of a point P on the Equator in the Pacific Ocean. The orbital radius is 7714.14 km and the inclination is 28.5 deg. The circle centered at P on the equator labeled CASE 4 is the station maks of a fictitious station on an ocean platform with longitude equator to that of the DSN Goldstone Station.

and the annulus defined by the ground track, divided by the area of the annulus defined by the ground track, i.e. T would be equal to the expression: $\aleph (\mathbf{D} \cap \mathbf{A}) / \aleph (\mathbf{A})$. In fact, this is a very bad approximation. A clue as to why this is a bad approximation is given by the density of the ground tracks which depends on the latitude and is quite uneven. Moreover, the speed of the nadir of the satellite along the ground track is not constant because the Earth is rotating. Also, the orbital plane is precessing due to the J_2 gravity harmonic. But all is not lost.

The heuristics of this reasoning is intuitively correct. But, instead of the ratio of the geometric areas of the two regions mentioned earlier, we need to weigh the area depending on the ground velocity and some how account for the expansion and contraction of the ground tracks. This new weighted area function is technically called an "invariant measure**" usually denoted by "μ". As a weighted area element following the satellite nadir along the ground track, the area of the element is preserved. Hence the weighted area element is invariant under the motion of the satellite ground track. When such a measure of the area is available, then indeed the percentage of time spent by the satellite in the circular region is given by the measure of the intersection of the circular region with the annulus, divided by the measure of the annulus, i.e. T is equal to the expression: $\mu (\mathbf{A} \cap \mathbf{D}) / \mu (\mathbf{A})$.

Such a measure was constructed in Lo [1,2]. However, for this to work, it is necessary that the ground tracks not be periodic. But it is shown that even when the ground track is periodic, provided the repeat cycle is not too small, this approximation is fairly good. This means that instead of finding the view periods from a propagated trajectory to compute the amount of time a satellite is in view of a ground station, also called the "time average", we can replace this by a simple area integral with a weighted area. This weigted average is called the "space average". This in essence is the Ergodic Theorm that we can replace time averages by space average. Typically, time averages are more difficult to compute since it requires the solution of differential equations. Where as the space average is much easier to solve as it requires only a single area integral. Note, for the space average, J_2 is only used for the verification that the orbit ground track is not periodic. It never enters into the area integral.

We should note here that we are assuming that the satellite orbit is being maintained so that effects of the Earth gravity's higher harmonics, the luni-solar perturbation, solar radiation pressure, and drag are being compensated. The maneuvers will perturb the orbit node, but the orbital elements such as semimajor axis, eccentricity are essentially preserved. Assuming the maneuvers are sufficiently random without a bias that would cause the ground tracks to become periodic in some fashion, this theory applies to the coverage problem.

ERGODIC THEORY

** Measure is a generalization of the concept of area and volume for sets of arbitrary dimensions.

Ergodic theory has its origins in the study of statistical mechanics in the 19^{th} century. Maxwell, Boltzmann, Gibbs, and Poincaré were the first to propose a statistical approach to study differential equations. A classical problem is the following: Given a particle moving randomly within a closed and bounded box B; at time 0 the particle is known to be in the subset C of our box B; how frequently will this particle visit the subset C within our box as the time goes to infinity? Poicaré's Recurrence Theorem tells us that the particle will repeatedly visit the set C infinitely often (see Sinai[5]). In fact, the probability that the particle can be found in C is given by the volume of the set C divided by the volume of the box B. This is geometrically intuitive.

The fact that the probability the particle can be found in the set C is given by the quantity Volume(C) / Volume (B) is a profound result. Our original question is about the time average of the particle visiting the set C; our answer is that it is given by the space average of the set C, i.e. its volume normalized by the total volume of the box B. This equivalence of "time average" with "space average" is at the heart of ergodic theory. The reason this is so powerful is because we can replace knowledge of the time history of a particle (its trajectory) in a dynamical system (a set of differential equations) by a definite integral over subsets within the phase space (such as the 6 dimensional state space of position and velocity for a satellite). This means that without integrating the differential equations, we can obtain valuable statistical information about the dynamical systems by computing definite integrals which are often much easier to do.

In order to apply ergodic theory to a dynamical system described by a set of differential equations, one must first obtain a volume function on the phase space which is invariant under the trajectory flow $\varphi_t(x)^\dagger$ prescribed by the differential equations. This is known as an "invariant measure"; it is just a volume element weighted by a function to compensate for the contractions and expansions of the trajectories in the phase space. The construction of the invariant measure is the hard part of the problem. Fortunately for our problem, this has been done in Lo [1]. We denote this measure, or the volume function by μ, which is normalized to give a total volume equal to 1. We indicate the differential volume element by $d\mu$. We now define more precisely what we mean by time and space averages.

The *time mean* $\langle f \rangle$ of a function f is defined by:

$$\langle f \rangle = \lim_{t\to\infty} \frac{1}{T} \int_0^T f(\varphi_t(x))\, dt \ , \qquad x \in M, t \in R. \tag{1}$$

The space mean of a function $\bar{f}$ is defined by:

$$\bar{f} = \int_M f(x)\, d\mu \ . \tag{2}$$

---

† The flow $\varphi_t(x)$ is the solution to the differential equation at time "t" with initial condition "x" in the phase space. It describes the complete set of solutions to the differential equation. The analogy is to streamlines in a fluid flow.

Here x is a point in the phase space M, and R is the set of real numbers. The fundamental theorem of ergodic theory is the Birkhoff-Khinchin theorem (Theorem 6.4 in Arnold [4]) which simply states that the time mean (1) is equal to the space mean (2) for dynamical systems. In other words, that $< f > = \overline{f}$. We will not get into the details of the necessary conditions for this theorem to hold; suffice it to say that for satellite motion under the gravitation of an oblate planet, these conditions are satisfied.

THE SATELLITE VIEW PERIOD PRBLEM

We first review the results from Lo [1] for the view period problem. Given a ground station located at "x" on Earth, we want the amount of time the ground station is in view of the satellite. We assume the satellite is moving in a circular orbit about an oblate planet with J_2 perturbation resulting in the drifting of the ascending node of the orbit. However, to compute the station visibility, we treat the planet as a sphere. In order for the Birkhoff-Khinchin Theorem to apply, we must further assume that the orbital ground track is not repeating. Although this excludes some of the most important satellite orbits, all is not lost. For those orbits with short ground track repeat cycles, the statistics may be quickly computed using the standard trajectory integration approach. For those orbits with long repeat cycles, the ergodic theory provides a reasonable approximation for quick analyses as noted in Lo [1]. The interesting thing is that the only place where the value of J_2 is needed is in the verification that the satellite ground track is non-repeating. The

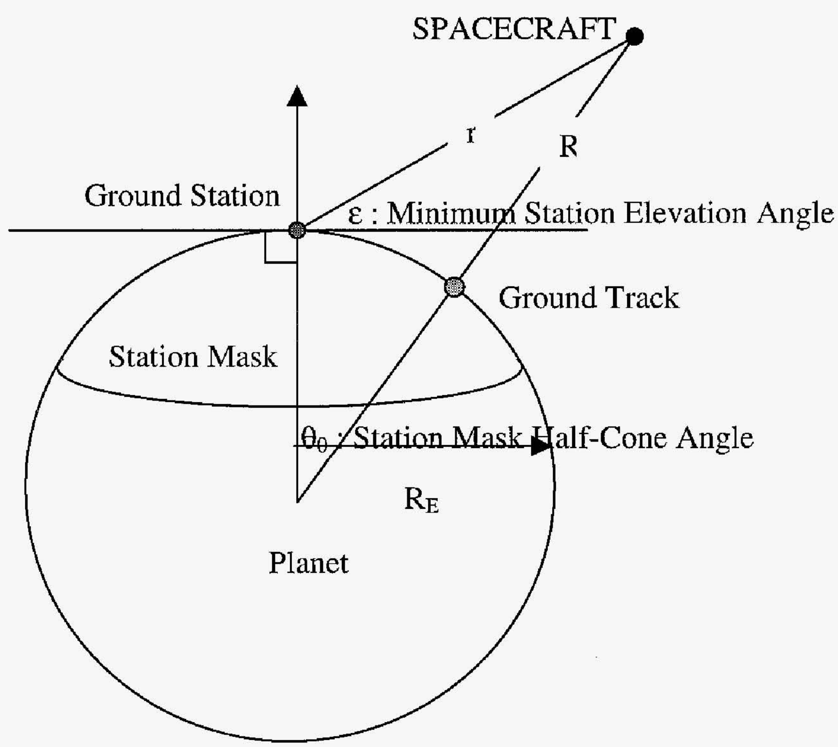

Figure 2. This shows the geometry of the station mask which is determined by the altitude of the spacecraft.

expression for the invariant measure itself, $d\mu$, does not include the J_2 coefficient. The long-term station view period ρ is defined by Lo [1] as

$$\rho = \lim_{T\to\infty} \frac{P(T)}{T} , \tag{3}$$

where P(T) is the total time the satellite is in view of the ground station from time 0 to time T. In other words, ρ is the fraction of time the satellite can see the ground station; and ρT gives a good approximation for value of P(T) for sufficiently large T.

In this case, the function f (x) is the characteristic function of the station mask, A (e.g. see Figures 1 and 2). In other words, f (x) is 1 when x is in A and 0 otherwise. The set A is the circular region on the planet centered around the ground station defined by the minimum elevation angle ε of the ground station. The resulting space average integral is given by

$$\rho = \int_{\varphi_1}^{\varphi_2} g(\varphi)h(\varphi)d\varphi , \tag{4}$$

where

$$g(\varphi) = 2 \ \mathrm{acos}\,[(\cos\beta_0 - \sin\varphi \sin\varphi_0)/\cos\varphi_0 \cos\varphi], \tag{5}$$

$$h(\varphi) = \cos\varphi/(\,2\pi^2\sqrt{\sin^2 i - \sin^2\varphi}),$$

and

φ_0 = station latitude, (6)
θ_0 = station longitude,
β_0 = station mask angular radius = $90° - \varepsilon - \mathrm{asin}(R_E \cos(\varepsilon) / R)$,
R_E = planet radius,
R = spacecraft circular orbit radius,
ε = minimum station elevation angle,
i = spacecraft orbit inclination, assume always ≥ 0,
Li = i (for i

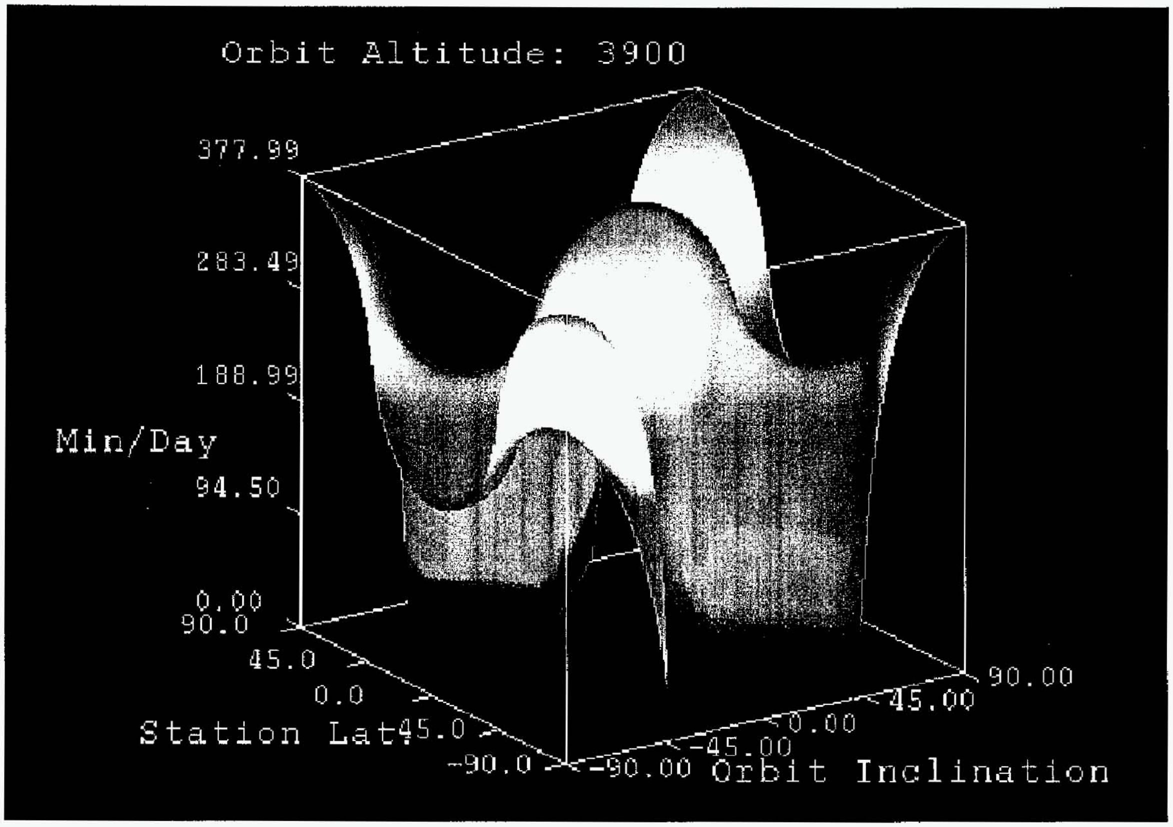

Figure 3. The total visibility for any satellite at 3900 km altitude to every point on the Earth. The X-axis is orbital inclination, the Y-axis is station latitude, and the Z-axis is the total visibility period in minutes per day.

Figure 3 above shows the power of this approach. We are able to provide the global coverage properties of all circular orbits at 3900 km altitude to every point on the Earth using the visibility ratio provided by equation (7). This required a few seconds to compute. Where as if we were to compute this using the integration of trajectories, it would require many hours, perhaps even days of computation to provide such a result. The simplicity of the equations allows analysts to use them in tools such as Excel or Matlab for quick studies of coverage analysis.Since this is a first-order approximation, the results around critical inclination are inaccurate.

THE DATA TRANSMISSION PROBLEM

The *Data Transmission Problem* refers to the problem of estimating the total volume of data transmitted between the satellite and the ground. For this application of the ergodic integral (6), we assume the transmission is from the ground to the satellite; for example, a Mars rover transmitting data to a telecommunications relay satellite around Mars for eventual transmission to the Earth. Here we examine the data transmission from the ground to the satellite only.

We assume the data transmission rate is f (r), where "r" is the distance between the spacecraft and the ground station. Then according to the Birkhoff-Khinchin Theorem, the average data rate is given by the following integral

$$\delta \;=\; \int_{-g(\varphi)/2}^{g(\varphi)/2} \int_{\varphi_1}^{\varphi_2} f(r) h(\varphi)\, d\theta\, d\varphi\,, \tag{8}$$

and the average volume of data transmitted for the period T is just $T\delta$.

Suppose we assume the data rate is the following simple expression

$$f(r) \;=\; \kappa/r^2 \text{ MBPS (megabits per second)}, \tag{9}$$

where κ is a constant characterizing the link. We need to express r in terms of the integration variables φ and θ (latitude and longitude). This is simple using spherical coordinates. Suppose the satellite ground track is at the latitude and longitude (φ, θ), this means the satellite is at the location $S = (R, \varphi, \theta)$ in spherical coordinates rotating with the Earth. Similarly, the ground station is located at $G = (R_E, \varphi_0, \theta_0)$. The distance r between S and G is just $|S - G|$ which may be easily computed by converting to Cartesian coordinates.

For more complex data transmission rate functions, one simply follows the recipe outlined above in equation (8) and proceed accordingly.

CONCLUSION

The Satellite Coverage Problem is a challenging problem even for simple circular orbits. This is because the coverage is both dynamical and combinatorial in nature. To provide statistical analysis using conventional methods requires the computation of the trajectory geometry for long durations. The use of ergodic theory allows one to estimate some of these statistical parameters without trajectory computations at all. One simply replaces the time average process by a space averaging process which results in definite integrals that can be easily and quickly computed numerically. The time averaging process is tedious and consumes a tremendous amount of resources being computationally and time-wise, requiring serious software development. Whereas the space averaging process is easily implemented in Matlab or Excel spreadsheets. For architectural studies, parametric what-if scenario analyses, the ergodic approach is extremely attractive.

In this paper, we presented the measure $\int_C d\mu$ as a surface integral in equation (7) which is suitable for computing the space average of various quantities f (x), where x is the satellite state. By the ergodic theorem, this definite integral is equal to the time average of the same quantity f (x(t)). As an application, we used the data transmission rate from a rover on Mars to a telecommunications satellite orbiting Mars in a circular orbit with non-repeating ground track given in equation (8).

For elliptical orbits, the problem is much more challenging. In addition to the variation in the node, the argument of periapsis is also evolving. Hence the averaging process must take both effects into consideration. This increases the dimensions of the problem. Once again, the construction of the invariant measure on the proper parameter space is the hard part of the problem. The resulting multidimensional integral over space may be easily implemented using montecarlo integration techniques.

ACKNOWLEDGEMENTS

This work was carried out at the Jet Propulsion Laboratory of the California Institute of Technology under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. Lo, M.W., "The Long-Term Forecase of Station View Period", *JPL TDA Progress Report 42-118*, pp. 1-14, April-June, 1994.
2. Lo, M.W., "The Invariant Measure for the Satellite Ground Station View Period Problem," C. Simó (ed.), *Hamiltonian Systems with Three or More Degrees of Freedom*, NATO ASI Series, Kluwer Academic Publishers, 1999.
3. Sinai, Ya.G., *Introduction to Ergodic Theory*, Princeton University Press, Princeton, N.J., 1976.
4. Arnold, V.I., A. Avez, *Ergodic Problems of Classical Mechanics*, Addison Wesley, 1989, New York.
5. Sinai, Ya.G., Topics in Ergodic Theory, Princeton University Press, Princeton, N.J., 1994.