Kalman Filtering: Theory and Practice Using MATLAB

Fourth Edition

Appendix C: Feasibility Analysis of Doppler-based Satellite Navigation

Mohinder S. Grewal

California State University at Fullerton

Angus P. Andrews

Retired from Rockwell Science Center, Thousand Oaks, California

Wiley

PREFACE

This appendix is intended to augment the examples used in [2] for demonstrating how one goes about assessing the accuracy of timing-based satellite navigation, by providing an example for its now-obsolete predecessor: Doppler-based satellite navigation. It was first analyzed before the Kalman filter, and it provides an example that does not need the Riccati equation for its performance analysis.

A simple model is used to demonstrate the feasibility of determining the longitude and latitude of a receiver antenna from measurements of Doppler frequency shifts during a single overhead pass of a satellite with known orbit. This was first demonstrated by William Guier and George Weiffenbach in 1958, leading to development of the Transit Navigation System, the world's first satellite navigation system. This example also demonstrates how the expected accuracy of Doppler-based satellite navigation depends on the relative location of the receiver antenna with respect to the satellite ground track, and on Doppler measurement accuracy.

This sort of "preliminary" analysis was originally used for determining whether Doppler-based satellite navigation was feasible, from the standpoint of whether it could provide a navigation solution for a fixed receiver location, and for assessing how accurate that solution might be. As such, the approach does not require the full capabilities of a Kalman filter for tracking a moving receiver with uncertain dynamics—as is now done for timing-based satellite navigation. This may have been fortunate at the time (late 1950s), because the Kalman filter [6] had not been introduced yet. However, these capabilities were soon added before Doppler-based satellite navigation became operational.

Appendix C

Feasibility Analysis of Doppler-based Satellite Navigation

If you want to find the secrets of the universe, think in terms of energy, frequency and vibration.
— ***Nikola Tesla*** *(1856–1943)*

C.1 Historical Background

C.1.1 The Doppler Effect

This effect is named for Austrian-born Christian Andreas Doppler (1803–1853), who developed a mathematical model for the observed phenomenon that the frequencies of acoustic signals are shifted upward when the source is coming toward the observer and downward when the source is going away from the observer. It is the same effect that—at radar frequencies—is used by traffic police to measure the speeds of automobiles on highways, and for Doppler-based satellite navigation.

C.1.2 Satellite Navigation

Earth's natural satellite—our moon—is thought to have been launched by an interplanetary collision around 4 billion years ago.

Sputnik I, the world's first artificial satellite, was launched by the former Soviet Union on October 4, 1957. This happened on a Friday evening on the East Coast of the United States, when nearly everyone had gone home for the weekend. When William Guier and George Weiffenbach showed up for work on the following Monday at the Applied Physics Laboratory (APL) of Johns Hopkins University, the Soviet satellite was the main topic of discussion at lunch in the cafeteria [5]. The two transmission frequencies around 20 MHz and 40 MHz had already been determined. Weiffenbach was in possession of a good 20 MHz receiver, so he and Guier decided to set up an antenna so they could "listen in" after lunch. The lower satellite transmission frequency was found to be offset about 1 kHz from 20 MHz, and APL was near enough to the 20 MHz standard broadcast by the National Bureau of Standards from station WWV that they could use it to remove the 20 MHz component to better observe the Doppler variations as the satellite passed from horizon to horizon. Weiffenbach had a spectrum analyzer he had acquired for his PhD thesis on microwave spectroscopy, so he was able to measure and record the Doppler variations from the nominal 1 kHz offset frequency relatively accurately. As the satellite passed overhead from horizon to horizon, it was observed to vary in frequency from about 500 Hz to 1500 Hz due to Doppler shift.

At that point, the selection of Doppler measurements had been determined by equipment availability, but it would turn out to be surprisingly fortuitous. Guier was curious to determine whether the satellite orbit could be determined from the Doppler variations alone, and was able to show that it could if the location of the receiver antenna was known (which it was).

Although the batteries aboard the Sputnik I satellite lasted only 23 days in orbit, Guier and Weiffenbach were able to continue their investigations for several months using the data recorded while it was transmitting. They also had the use of the Univac 1103A computer recently acquired by APL in their studies. In the process, they discovered

that the influence of low-altitude gravity anomalies on the Sputnik I trajectories could not be ignored, but that the same Doppler measurements could also be used to estimate the gravity anomalies and provide a relatively accurate ephemeris of the short-term satellite trajectory.

When these results were reviewed by Frank McClure at APL in March of 1958, he asked whether the problem could be inverted. That is, if the orbit were known, could the receiver location be determined from the Doppler measurements alone?

The answer turned out to be in the affirmative, and it was this discovery that led to the development of the first satellite navigation system [5].

A need for such a navigation system had been established a few years earlier, when the U.S. Navy commissioned a fleet of nuclear-powered, nuclear-missile-carrying submarines as a strategic deterrent in the Cold War against the USSR and its allies. However, the best inertial navigation technology at that time was not sufficiently accurate for launching ballistic missiles after many days at sea. Fortunately, the predicted accuracy of Doppler-based satellite navigation would solve that problem. Furthermore, the antenna exposure required for getting a position fix while submerged was deemed an acceptable risk if only Doppler frequency shift was to be used. Also, because the antenna would be essentially at sea level, only longitude and latitude had to be determined.

The Navy had studied the use of satellites as early as 1945 [1], but had found no justifiable application. Based on the findings at APL, a Navy project to develop a satellite navigation system based on Doppler measurements was initiated in 1958, and would achieve operational status in the 1960s [4]. The operational system would be called *Transit,* or *NAVSAT* (for *Navigational Satellite*) [7].

C.2 Methods for Proving Determinancy

Frank McClure's question to Guier and Weiffenbach was whether the location of a receiver antenna at sea level could be determined from the Doppler frequency variations in the signal of a satellite passing within view. The following is an account of a progression of mathematical methods for resolving such issues.

These derivations use a simple spherical-Earth model for the purpose of assessing the expected accuracy of Doppler-based satellite navigation, and for determining how the resulting navigational accuracy depends on the receiver antenna location relative to the satellite ground track, and on various noise sources. This model would not be good enough for the full Kalman filter implementation of the navigation solution, but it is adequate for statistical analysis of the resulting navigation errors of a more faithful Kalman filter model [6].

C.2.1 Least Squares Approach

The feasibility of satellite navigation using Doppler measurements was first demonstrated by Guier and Weiffenbach, using some of the same linearization methods used by Carl Friedrich Gauss in his 1801 determination of the orbit of the asteroid Ceres [2]. Gauss had far more observational data than unknowns needing to be determined. He was then able to pose this in the form of an overdetermined system of equations and use the method of least squares he had discovered earlier. In the case of the Doppler satellite navigation problem, Guier and Weiffenbach could use partial derivatives of the Doppler frequency measurements with respect to receiver location to model the antenna location problem as an overdetermined linear system [4]. The least-squares solution of an over-determined linear system $Hx = z$ for x, given H and z, is

$$\hat{x} = \left(H^T H\right)^{-1} H^T z, \tag{C.1}$$

if the **gramian** $\left(H^T H\right)$ is invertible. The determinant of this gramian matrix then becomes a discriminant for whether a unique solution is attainable.

C.2.2 Least Mean Squares Approach

By this time, however, least-mean-squares estimation had replaced the least-squares approach of Gauss, and the equivalent discriminant would then be the determinant of the 2×2 **information matrix** Y, defined as

$$Y \quad \stackrel{\text{def}}{=} \quad \sum_k H_k^T R_k^{-1} H_k \tag{C.2}$$

$$H_k \quad \stackrel{\text{def}}{=} \quad \left.\frac{\partial f_{k,\text{Doppler}}}{\partial E_{\text{ANT}}, N_{\text{ANT}}}\right|_{t_k} \tag{C.3}$$

where $f_{k,\text{Doppler}}$ is the k^{th} Doppler frequency measurement, the partial derivatives are with respect to the east (E) and north (N) location of the antenna evaluated under conditions of the Doppler measurement made at time t_k, and R_k is the variance of frequency measurement uncertainty on the k^{th} Doppler frequency measurement[1].

If the determinant of the information matrix Y is non-zero, then Y is nonsingular (i.e., invertible), and its inverse

$$P \stackrel{\text{def}}{=} Y^{-1} \tag{C.4}$$

is the 2×2 covariance matrix of mean-squared horizontal antenna location uncertainty after all the measurements from one or more satellite passes. In that case, the square root of its diagonal elements

$$\sqrt{p_{11} + p_{22}} = \sqrt{\text{E}\,\langle \varepsilon_E^2 + \varepsilon_N^2 \rangle} \tag{C.5}$$

is the RMS horizontal (2D) estimation error of antenna location.

The information matrix from the least-mean-squares approach then provides a quantitative assessment of the solution accuracy, whereas the determinant of the gramian had been only a qualitative indicator of the feasibility of a unique solution.

C.2.3 Dilution of Precision (DOP) Approach

Dilution of precision, is a term initially introduced to describe the impact of relative locations of radio beacons on ground-based radionavigation systems such as LORAN. Navigational accuracy of these phase-based or relative-timing-based systems can be different in different local horizontal directions, depending on the relative directions of arrival of signals from different beacons. The same is true for timing-based satellite navigation systems such as GPS and other global navigation satellite systems [3]. It is also true for Doppler-based satellite navigation, even though it may use a single satellite.

This general approach factors the RMS horizontal position uncertainty into the product of two factors:

1. The mean-squared measurement error R.
2. The inverse of the *unscaled information matrix* $Y^\star$, which represents the effect of the choice of the measurements to be used (but not their mean-squared uncertainties R_k). The measurements used in this application will be the Doppler frequencies from a passing satellite.

The unscaled information matrix characterizes performance metrics for the variables being estimated. The performance metric of interest in this case the **horizontal dilution of precision (HDOP)**, which is defined as[2]

$$\begin{aligned}
\textbf{HDOP} &\stackrel{\text{def}}{=} \sqrt{a_{11} + a_{22}} && \text{(C.6)}\\
A &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} && \text{(C.7)}\\
&\stackrel{\text{def}}{=} \{Y^\star\}^{-1} && \text{(C.8)}\\
Y^\star &\stackrel{\text{def}}{=} \sum_k H_k^T H_k && \text{(C.9)}\\
H_k &\stackrel{\text{def}}{=} \left. \frac{\partial f_{k,\text{Doppler}}}{\partial E_{\text{ANT}}, N_{\text{ANT}}} \right|_{t_k} && \text{(C.10)}
\end{aligned}$$

where, as before, $f_{k,\text{Doppler}}$ is the k^{th} Doppler frequency measurement, the partial derivatives are with respect to east and north displacements of antenna location, and the partial derivatives are evaluated under conditions of the Doppler measurement made at time t_k from a specified antenna location.

The matrix $Y^\star$ is called the **unscaled** information matrix to distinguish it from the information matrix Y defined in Equation C.2. $Y^\star$ is not scaled by the inverse of the mean-squared measurement uncertainty.

[1] In this preliminary analysis, R_k will be considered time-invariant. However, some variation of measurement accuracy with time might be expected from the variation is satellite signal-to-noise ratio as the satellite passed overhead from horizon to horizon.

[2] Here, we follow the notation of [3] in defining HDOP.

If the mean-squared frequency measurement uncertainty R is constant over time, then the covariance matrix P of estimation uncertainty is the product of the scalar (mean-squared Doppler measurement error) factor R and the 2×2 matrix A:

$$\begin{bmatrix} \mathrm{E}\left\langle \varepsilon_E^2 \right\rangle & \mathrm{E}\left\langle \varepsilon_E \varepsilon_N \right\rangle \\ \mathrm{E}\left\langle \varepsilon_E \varepsilon_N \right\rangle & \mathrm{E}\left\langle \varepsilon_N^2 \right\rangle \end{bmatrix} \stackrel{\text{def}}{=} P \tag{C.11}$$

$$= Y^{-1} \tag{C.12}$$

$$= \left\{ Y^\star R^{-1} \right\}^{-1} \tag{C.13}$$

$$= R \left\{ Y^\star \right\}^{-1} \tag{C.14}$$

$$= RA \tag{C.15}$$

where ε_E is the east component of estimated antenna location error and ε_N is the north component of estimated antenna location error.

As a consequence, the horizontal dilution of precision

$$\mathbf{HDOP} = \sqrt{a_{11} + a_{22}} \tag{C.16}$$

$$= \frac{\sqrt{\mathrm{E}\left\langle \varepsilon_E^2 \right\rangle + \mathrm{E}\left\langle \varepsilon_N^2 \right\rangle}}{\sqrt{R}} \tag{C.17}$$

is essentially the multiplier of RMS measurement error that yields the resulting RMS horizontal antenna location estimation error.

In this case, this **HDOP** characterizes the efficacy of Doppler-based satellite navigation.

C.3 Model Derivations

In order to evaluate the efficacy of Doppler-based satellite navigation, one must first derive appropriate formulas for the partial derivatives of Doppler frequency with respect to horizontal displacements of the receiver antenna.

It is simpler to derive the Doppler frequency formulas in terms of the rate of change of the range between the satellite antenna and the receiver antenna, then use Leibniz's chain rule to get the partial derivatives of Doppler frequency with respect to horizontal components of receiver location.

C.3.1 Polar Orbit Model

These models will use the satellite trajectories of the Transit Navigation Satellite System, not the Sputnik I trajectories originally used in the initial proof-of-concept studies by Guier and Weiffenbach. In order to provide better world-wide coverage with a small number of satellites, Transit/NAVSAT satellites were placed in nearly circular polar orbits at altitudes in the order of 1000 kilometers, which is consistent with an orbital period of about 105 minutes.

Transit signal frequencies were at 150 and 400 MHz, so that ionospheric delays could be estimated from the relative phase delay between the two signals.

Satellite orbits are commonly specified in earth-centered inertial (ECI) coordinates, which have their origin at the center of mass of planet Earth, their z-axis in the direction of Earth's rotation axis, and their x-axis in the direction of the vernal equinox[3], as illustrated in Figure C.1, with parameters

$R_\oplus$, the radius of the(spherical) earth,

$\Omega_\oplus$, the rotation rate of Earth,

Ω_{SAT}, the orbital angular rate of the satellite,

h_{SAT}, the altitude of the satellite,

[3] All these "quasi-inertial" reference directions are not exactly constant, but change very slowly over time due to the influence of local gravitational fields. The inertial direction of the rotation axis of Earth, for example, moves about a tenth of a milliradian per year due to local gravity gradients. The Navigation and Ancillary Information Facility (NAIF) of the National Aeronautics and Space Administration keeps track of these variations and provides data and algorithms for users to adjust their data.

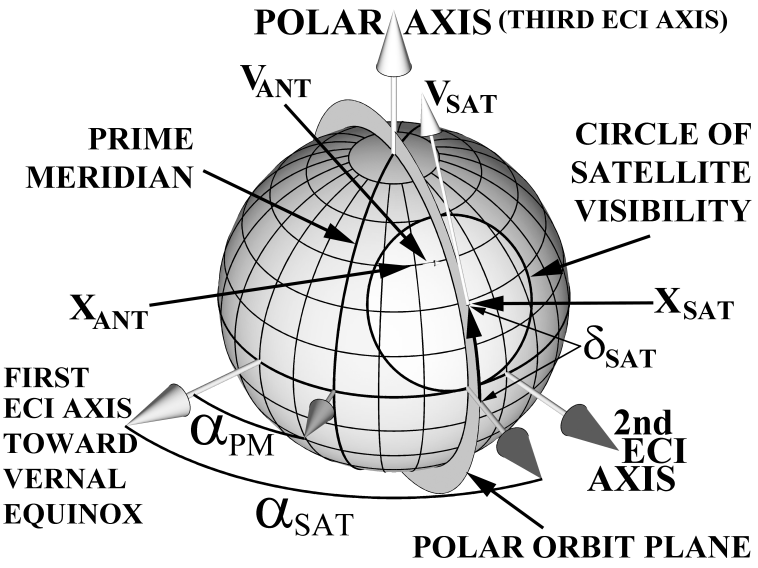

Figure C.1: Polar Transit Orbit in ECI coordinates.

α_{SAT}, the right ascension of the satellite polar orbit,

$\delta_{\mathrm{SAT}}(0)$, the declination of the satellite above the equatorial plane at time $t = 0$,

$\alpha_{\mathrm{PM}}(0)$, the right ascension of the prime meridian at reference time $t = 0$,

$\mathbf{LAT}_{\mathrm{ANT}}$, the latitude of the antenna location,

$\mathbf{LON}_{\mathrm{ANT}}$, the longitude of the antenna location.

In Figure C.1, the area on the surface within which the satellite can be seen above the local horizon is indicated by the more lightly shaded circle. Outside that circle, no direct Doppler signal is available as the satellite passes by beyond the horizon.

Figure C.1 shows only a single polar satellite orbit, although the full operational system would include five operational satellites and as many "spares" parked in the five separate polar orbits to provide backup if any operational satellite failed. However, for the purpose of showing the feasibility of using only Doppler measurements for satellite navigation, it suffices to consider just one satellite on a single pass within view of the receiver antenna.

C.3.2 Satellite and Antenna Position Models

C.3.2.1 Satellite Position

The satellite position in ECI coordinates depends on the **right ascension** of the orbit, which is the angle measured in the equatorial plane from the vernal equinox to the ECI longitude at which the satellite passes from the southern hemisphere into the northern hemisphere. If α_{SAT} is the right ascension of the polar orbit, then the position of the satellite in ECI coordinates can be represented as

$$x_{\mathrm{SAT}}(t) = (R_{\oplus} + h_{\mathrm{SAT}}) \begin{bmatrix} \cos(t\Omega_{\mathrm{SAT}} + \delta_{\mathrm{SAT}}(0)) \cos(\alpha_{\mathrm{SAT}}) \\ \cos(t\Omega_{\mathrm{SAT}} + \delta_{\mathrm{SAT}}(0)) \sin(\alpha_{\mathrm{SAT}}) \\ \sin(t\Omega_{\mathrm{SAT}} + \delta_{\mathrm{SAT}}(0)) \end{bmatrix}, \tag{C.18}$$

where $\delta_{\mathrm{SAT}}(0)$ is the declination of the satellite in its orbit when $t = 0$.

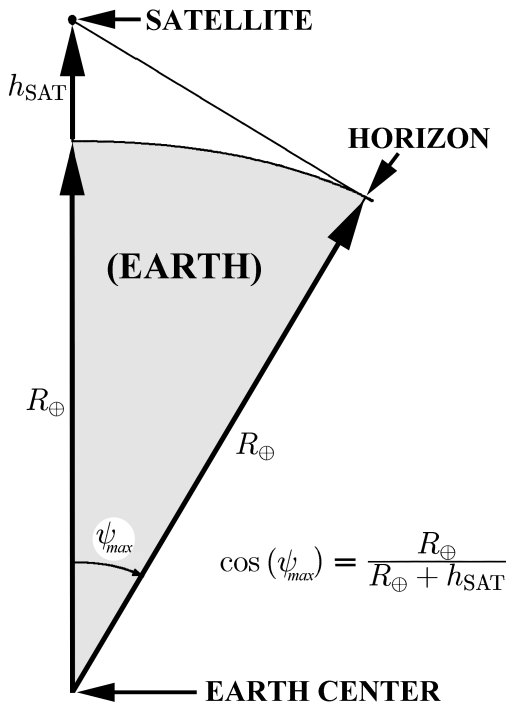

Figure C.2: Satellite horizon geometry

C.3.3 Receiver Antenna Position

The receiver antenna is assumed to be located at sea level at a fixed longitude $\mathrm{LON_{ANT}}$ and latitude $\mathrm{LAT_{ANT}}$ in Earth-centered Earth-fixed (ECEF) coordinates, which will be rotating in ECI coordinates. Unlike longitude, latitude has the same value in ECEF coordinates as right ascension (α_{ANT}) in ECI coordinates.

The receiver antenna right ascension in ECI coordinates will then be

$$\alpha_{\mathrm{ANT}}(t) \quad = \quad \alpha_{\mathrm{PM}}(0) + \mathrm{LON_{ANT}} + \Omega_\oplus t, \tag{C.19}$$

where $\alpha_{\mathrm{PM}}(0)$ is the right ascension of the prime meridian at time $t = 0$, $\mathrm{LON_{ANT}}$ is the longitude of the receiver antenna with respect to the prime meridian, and $\Omega_\oplus$ is the inertial rotation rate of Earth.

For the assumed spherical Earth model, the position of the receiver antenna in ECI coordinates is then

$$x_{\mathrm{ANT}}(t) \quad = \quad R_\oplus \begin{bmatrix} \cos(\mathrm{LAT_{ANT}})\cos(t\Omega_\oplus + \theta_{\mathrm{ANT}} + \alpha_{\mathrm{PM}}(0)) \\ \cos(\mathrm{LAT_{ANT}})\sin(t\Omega_\oplus + \theta_{\mathrm{ANT}} + \alpha_{\mathrm{PM}}(0)) \\ \sin(\mathrm{LAT_{ANT}}) \end{bmatrix}. \tag{C.20}$$

C.3.4 Signal Models

C.3.4.1 Availability of Doppler Signals at Receiver Antenna

This is the issue of whether the satellite appears above the horizon at the receiver location, as depicted in cross-section in Figure C.2.

The resulting triangle formed by the satellite, the receiver antenna, and the center of Earth is a right triangle when the receiver antenna is on the horizon, as depicted, with right angle then at the receiver antenna location. As a consequence, the limiting great circle angle ψ from the sub-point of the satellite to the receiver on the horizon will

be

$$\psi_{\max} = \cos^{-1}\left(\frac{R_\oplus}{R_\oplus + h_{\mathrm{SAT}}}\right), \tag{C.21}$$

where h_{SAT} is the altitude of the satellite and $R_\oplus$ is the radius of the (assumed) spherical earth.

Consequently, the satellite is visible above the horizon at the antenna if and only if the dot product

$$x_{\mathrm{ANT}}^T x_{\mathrm{SAT}} \geq \cos(\psi_{\max}) R_\oplus (R_\oplus + h_{\mathrm{SAT}}), \tag{C.22}$$

the formula used for identifying the circle of satellite visibility in Figure C.1.

C.3.4.2 Range Model

Using the formulas derived above, the range between the satellite antenna and receiver antenna can be derived as the vector norm

$$\rho(t) = |x_{\mathrm{SAT}}(t) - x_{\mathrm{ANT}}(t)|. \tag{C.23}$$

The signal frequency at the receiver antenna is then

$$f_{\mathrm{ANT}} = \frac{c f_{\mathrm{SAT}}}{c + \dot{\rho}} \tag{C.24}$$

$$\dot{\rho} \stackrel{\mathrm{def}}{=} \frac{d}{dt}\rho(t), \tag{C.25}$$

where c is the speed of light and f_{SAT} is the frequency transmitted by the satellite.

C.3.5 Doppler Frequency Model

Doppler shift is due to the time-rate-of change of the range between the transmitter antenna (i.e., the satellite antenna) and the receiver antenna.

The Doppler frequency is then

$$f_{\mathrm{Doppler}} = f_{\mathrm{ANT}} - f_{\mathrm{SAT}} \tag{C.26}$$

$$= -\frac{f_{\mathrm{SAT}}\dot{\rho}}{c + \dot{\rho}} \tag{C.27}$$

$$\approx -\frac{f_{\mathrm{SAT}}}{c}\dot{\rho} \tag{C.28}$$

for

$$|\dot{\rho}| \ll c. \tag{C.29}$$

The sensitivity of Doppler frequency to variations in the range rate is then

$$\frac{\partial f_{\mathrm{Doppler}}}{\partial \dot{\rho}} \approx -\frac{f_{\mathrm{SAT}}}{c}. \tag{C.30}$$

C.3.6 Sensitivities to North and East Antenna Position

The modeling above is in terms of antenna position in longitude and latitude, although sensitivities with respect to position in meters would be more useful in practice. The effect of the change of variables on the necessary partial derivatives can be modeled using Leibniz's chain rule.

If we let

$\delta_{E,\mathrm{ANT}}$ represent variations is EASTING [m],

$\delta_{N,\mathrm{ANT}}$ represent variations is NORTHING [m],

of the receiver antenna location, then

$$\frac{\partial f_{\text{Doppler}}}{\partial \delta_{E,\text{ANT}}} = \frac{\partial f_{\text{Doppler}}}{\partial \text{LON}_{\text{ANT}}} \frac{\partial \text{LON}_{\text{ANT}}}{\partial \delta_{E,\text{ANT}}} \tag{C.31}$$

$$\frac{\partial \text{LON}_{\text{ANT}}}{\partial \delta_{E,\text{ANT}}} = \frac{1}{\cos\left(\text{LAT}_{\text{ANT}}\right) R_{\oplus}}, \tag{C.32}$$

$$\frac{\partial f_{\text{Doppler}}}{\partial \delta_{N,\text{ANT}}} = \frac{\partial f_{\text{Doppler}}}{\partial \text{LAT}_{\text{ANT}}} \frac{\partial \text{LAT}_{\text{ANT}}}{\partial \delta_{N,\text{ANT}}} \tag{C.33}$$

$$\frac{\partial \text{LAT}_{\text{ANT}}}{\partial \delta_{N,\text{ANT}}} = \frac{1}{R_{\oplus}}, \tag{C.34}$$

where the longitude and latitude units are radians.

If the unit of f_{Doppler} is Hertz and the unit of horizontal displacement is meters, then the units of the alternative 2×2 unscaled information matrix

$$Y^{\star} \stackrel{\text{def}}{=} \sum_k H_k^T H_k \tag{C.35}$$

$$H_k \stackrel{\text{def}}{=} \left[\left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{E,\text{ANT}}} \right|_{t_k} \quad \left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{N,\text{ANT}}} \right|_{t_k} \right] \tag{C.36}$$

will be $\left[(\text{Hz/m})^2\right]$, and the units of the resulting HDOP will be [m/Hz].

C.4 Implementation

C.4.1 Improving Implementation Efficiency

The sensitivities of Doppler frequency to $\delta_{E,\text{ANT}}$ and $\delta_{N,\text{ANT}}$, the East and North components of receiver antenna location, can be represented by the chain rule as factors of partial derivatives:

$$H_k = \left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{E,\text{ANT}},\, \delta_{N,\text{ANT}}} \right|_{t_k} \tag{C.37}$$

$$= \left[\left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{E,\text{ANT}}} \right|_{t_k} \quad \left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{N,\text{ANT}}} \right|_{t_k} \right] \tag{C.38}$$

$$\left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{E,\text{ANT}}} \right|_{t_k} = \left[\frac{\partial f_{k,\text{Doppler}}}{\partial \dot{\rho}} \right] \underbrace{\left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LON}_{\text{ANT}}} \right|_{t_k} \right]}_{\text{time-varying}} \left[\left. \frac{\partial \mathbf{LON}_{\text{ANT}}}{\partial \delta_{E,\text{ANT}}} \right|_{\text{LAT}_{\text{ANT}}} \right] \tag{C.39}$$

$$\left. \frac{\partial f_{k,\text{Doppler}}}{\partial \delta_{N,\text{ANT}}} \right|_{t_k} = \left[\frac{\partial f_{k,\text{Doppler}}}{\partial \dot{\rho}} \right] \underbrace{\left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LAT}_{\text{ANT}}} \right|_{t_k} \right]}_{\text{time-varying}} \left[\frac{\partial \mathbf{LAT}_{\text{ANT}}}{\partial \delta_{N,\text{ANT}}} \right]. \tag{C.40}$$

However, only the middle factor of each product of partial derivatives depends on the summation index used in

computing the unscaled information matrix $Y^\star$:

$$
\begin{aligned}
Y^\star &\stackrel{\text{def}}{=} \sum_k H_k^T H_k && \text{(C.41)}\\
&= \left[\frac{\partial f_{k,\text{Doppler}}}{\partial \dot\rho}\right]^2 \times \begin{bmatrix} \frac{\partial \mathbf{LON}_{\text{ANT}}}{\partial \delta_{E,\text{ANT}}} & 0 \\ 0 & \frac{\partial \mathbf{LAT}_{\text{ANT}}}{\partial \delta_{N,\text{ANT}}} \end{bmatrix} \\
&\quad \times \left\{ \sum_k \begin{bmatrix} \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]^2 & \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right] \\ \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right] & \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]^2 \end{bmatrix} \right\} \\
&\quad \times \begin{bmatrix} \frac{\partial \mathbf{LON}_{\text{ANT}}}{\partial \delta_{E,\text{ANT}}} & 0 \\ 0 & \frac{\partial \mathbf{LAT}_{\text{ANT}}}{\partial \delta_{N,\text{ANT}}} \end{bmatrix} && \text{(C.42)}\\
&= \left[-\frac{f_{\text{SAT}}}{c}\right]^2 \times \begin{bmatrix} \frac{1}{\cos(\text{LAT}_{\text{ANT}})\, R_\oplus} & 0 \\ 0 & \frac{1}{R_\oplus} \end{bmatrix} \\
&\quad \times \left\{ \begin{bmatrix} \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]^2 & \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right] \\ \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right] & \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]^2 \end{bmatrix} \right\} \\
&\quad \times \begin{bmatrix} \frac{1}{\cos(\text{LAT}_{\text{ANT}})\, R_\oplus} & 0 \\ 0 & \frac{1}{R_\oplus} \end{bmatrix}, && \text{(C.43)}
\end{aligned}
$$

where we have used Equations C.30, C.32, and C.34 for substitutions.

As a consequence, the matrix

$$
\begin{aligned}
A &\stackrel{\text{def}}{=} \{Y^\star\}^{-1} && \text{(C.44)}\\
&= \frac{c^2 R_\oplus^2}{f_{\text{SAT}}^2} \begin{bmatrix} \cos(\text{LAT}_{\text{ANT}}) & 0 \\ 0 & 1 \end{bmatrix} \\
&\quad \times \begin{bmatrix} \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]^2 & \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right] \\ \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]\left[\frac{\partial \dot\rho}{\partial \mathbf{LON}_{\text{ANT}}}\Big|_{t_k}\right] & \sum_k \left[\frac{\partial \dot\rho}{\partial \mathbf{LAT}_{\text{ANT}}}\Big|_{t_k}\right]^2 \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} \cos(\text{LAT}_{\text{ANT}}) & 0 \\ 0 & 1 \end{bmatrix}, && \text{(C.45)}
\end{aligned}
$$

where all but one of the matrices involved can be inverted in closed form, and the remaining 2×2 matrix involves only the partial derivatives of range rate ($\dot\rho$) with respect to longitude and latitude. Being a 2×2 matrix, it can also

be inverted in closed form, in terms of the sums γ_{ij}, where

$$\Gamma \stackrel{\text{def}}{=} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \tag{C.46}$$

$$\gamma_{11} = \sum_k \left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LON}_{\mathrm{ANT}}} \right|_{t_k} \right]^2 \tag{C.47}$$

$$\gamma_{12} = \sum_k \left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LON}_{\mathrm{ANT}}} \right|_{t_k} \right] \left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LAT}_{\mathrm{ANT}}} \right|_{t_k} \right] \tag{C.48}$$

$$= \gamma_{21} \tag{C.49}$$

$$\gamma_{22} = \sum_k \left[\left. \frac{\partial \dot{\rho}}{\partial \mathbf{LAT}_{\mathrm{ANT}}} \right|_{t_k} \right]^2 \tag{C.50}$$

$$\Psi \stackrel{\text{def}}{=} \{\Gamma\}^{-1} \tag{C.51}$$

$$= \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} \tag{C.52}$$

$$\psi_{11} = \frac{\gamma_{22}}{\det \Gamma} \tag{C.53}$$

$$\psi_{12} = -\frac{\gamma_{12}}{\det \Gamma} \tag{C.54}$$

$$= \psi_{21} \tag{C.55}$$

$$\psi_{22} = \frac{\gamma_{11}}{\det \Gamma} \tag{C.56}$$

$$\det \Gamma = \gamma_{11}\gamma_{22} - {\gamma_{12}}^2 \tag{C.57}$$

Consequently,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{C.58}$$

$$= \frac{c^2 {R_\oplus}^2}{{f_{\mathrm{SAT}}}^2} \begin{bmatrix} \dfrac{\left(\cos\left(LAT_{ANT}\right)\right)^2 \gamma_{22}}{\gamma_{11}\gamma_{22} - {\gamma_{12}}^2} & -\dfrac{\cos\left(LAT_{ANT}\right)\gamma_{12}}{\gamma_{11}\gamma_{22} - {\gamma_{12}}^2} \\ -\dfrac{\cos\left(LAT_{ANT}\right)\gamma_{12}}{\gamma_{11}\gamma_{22} - {\gamma_{12}}^2} & \dfrac{\gamma_{11}}{\gamma_{11}\gamma_{22} - {\gamma_{12}}^2} \end{bmatrix} \tag{C.59}$$

and the horizontal dilution of precision

$$\mathbf{HDOP} = \sqrt{a_{11} + a_{22}} \tag{C.60}$$

$$= \frac{cR_\oplus}{f_{\mathrm{SAT}}} \sqrt{\frac{\cos\left(\mathrm{LAT}_{\mathrm{ANT}}\right)^2 \gamma_{22} + \gamma_{11}}{\gamma_{11}\gamma_{22} - {\gamma_{12}}^2}} \tag{C.61}$$

in terms of the summations of partial derivatives or $\dot{\rho}$ with respect to longitude and latitude.

This implementation eliminates a lot of multiplies inside summing loops.

C.5 MATLAB Implementation

Equations for the model were derived using the Maplesoft symbolic math environment, where they could be translated into LaTeX for publication and into MATLAB scripts for computation. The essential MATLAB scripts are on the Wiley website, using the following set of variable and parameter names and units [in square brackets]:

`alphaPM0` is the right ascension of the prime meridian at time $t = 0$ [rad]

`alphaSAT` is the right ascension of the polar satellite orbit [rad]

`delta0` is the declination of the satellite at time $t = 0$ [rad]

`fSAT` is the satellite transmission frequency [Hz]

`hSAT` is the altitude of the circular satellite orbit [m]

`LatAnt` is the latitude of the receiver antenna [rad]

`LonAnt` is the longitude of the receiver antenna [rad]

`OmegaEarth` is the Earth rotation rate [rad/s]

`OmegaSAT` is the satellite inertial rotation rate around Earth [rad/s]

`Rearth` is the radius of Earth [m]

`t` is the time t at which Doppler-based range rate is measured [s]

The measurement sensitivity derivations were done in terms of $\dot{\rho}$ to reduce the run time. Key m-files are:

`AntLoc.m` computes the antenna location in ECI coordinates as a function of time.

`SatLoc.m` computes the satellite location in ECI coordinates as a function of time.

`dRRdLAT.m` computes the partial derivative of range rate with respect to antenna latitude.

`dRRdLON.m` computes the partial derivative of range rate with respect to antenna longitude.

`Transit4.m` computes HDOP as a function of antenna longitude and latitude on a $120^o \times 160^o$ grid during a single satellite pass. Outputs include histograms of the condition numbers of the 2×2 information matrices (some of which are infinite), and the HDOP values for those locations with acceptable condition numbers for computation of HDOP. These histograms were used for setting contour levels.

C.6 Results and Conclusions

The resulting contour plot of horizontal dilution of precision (HDOP) from a single satellite pass is plotted versus longitude and latitude in Figure C.3. Features of interest in this plot include the following:

1. As might be expected, **Guier and Weiffenbach were right**, and they were able to prove it using a vacuum-tube computer with multiplication times in the order of a few hundred microseconds.

2. For the Transit orbits, position is observable from Doppler frequency measurements made within a swath of width ≈ 60 degrees centered along the ground track of the satellite, and unobservable outside that swath. That swath is where the satellite is observable during part of its pass-by.

3. The width of that swath is a function of satellite altitude. The offset center of the longitude range is due to the geometry shown in Figure C.1, with the prime meridian at $\approx 30^o$ right ascension at $t = 0$ and the right ascension of the satellite orbit at $\approx 75^o$.

4. That swath appears to head ever-so-slightly west of north, due to Earth rotation.

5. The apparent widening of the swath away from the equator in this longitude-latitude plot is due to the meridians squeezing together toward the poles. The ≈ 60-degree width of the "observability" swath covers the full range of longitudes at latitudes within 30 degrees of either pole.

6. There is a narrow, steep "ridge" of degraded HDOP down the middle of that swath. Figure C.4 is an HDOP cross-section plot along the equator of Figure C.3, showing the sharpness and that ridge. This semi-log plot shows the sharp crest of that ridge as having HDOP values about 50 time greater than the minimum values in the "HDOP canyon," and very steep canyon walls at the limits of observability where the satellite passes just beyond the horizon from the receiver antenna. This ridge is most likely due to diminished sensitivity of Doppler frequency to cross-track antenna location when the satellite passes nearly overhead. The width of the ridge

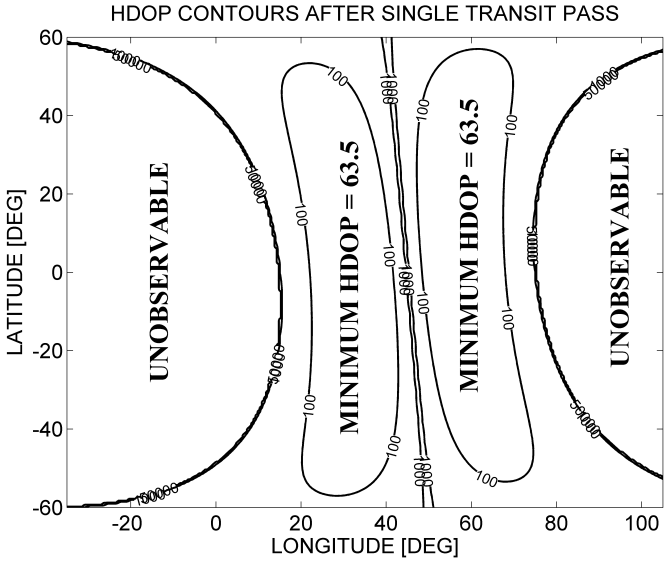

Figure C.3: Single-pass horizontal dilution of position versus longitude and latitude.

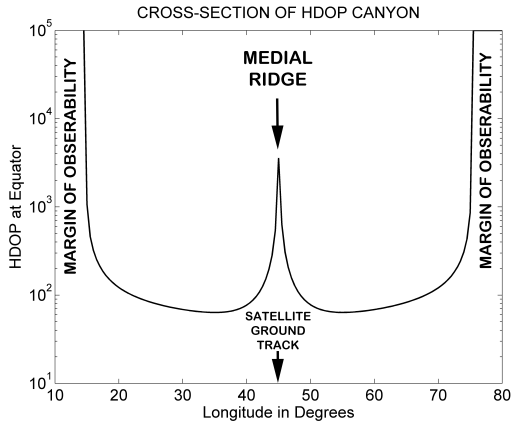

Figure C.4: Cross-section of "HDOP canyon" along equator.

increases noticeably near the poles, another result of the "meridian crowding problem" for longitude-latitude plots Except at the poles, this effect can be mitigated by waiting another 105 minutes for the next satellite pass, when the satellite ground track will have moved westward about 26 degrees of latitude.

7. The reason for the closing-out of the contours at the 100-level near the poles is not known, but is possibly an artifact of the coding implementation. The satellite trajectory begins 30 degrees south of the latitude range shown, and ends 30 degrees north of the latitude range plotted, so it should not be caused by any missing information. Except for the satellite-track "ridge" of HDOP, we would expect navigation near the poles to be as accurate as at lower latitudes.

8. Units of HDOP are RMS radial position error in meters per Hertz of RMS Doppler frequency measurement error. For a transmitter frequency of 150MHz (the low-frequency Trident channel), a Doppler frequency error of 1 Hz would require transmitter and receiver clock accuracies in the order of 10^{-8} parts per part, which is in the same order of accuracy as space-based clocks in that time era (1960–1990). Consequently, RMS position location accuracy would probably be in the order of 100 meters.

These results assumed stationary receiver locations. Later implementations (using the Kalman filter) allowed some degree of uncertainty in the dynamics of antenna location.

Bibliography

[1] P. Dickson, *Sputnik: The Shock of the Century,* Walker Publishing, New York, 2001, p. 69.

[2] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB,* 4th edition, Wiley, Hoboken, NJ, 2015.

[3] M. S. Grewal, A. P. Andrews, and C. G. Bartone, *Global Navigation Satellite Systems, Inertial Navigation, and Integration,* 3rd edition, Wiley, Hoboken, NJ, 2013.

[4] W. H. Guier and G. C. Weiffenbach, "A satellite Doppler navigation system," *Proceedings of the IRE,* Vol. 48, No. 4, 1960.

[5] W. H. Guier and G. C. Weiffenbach, "Genesis of satellite navigation," *Johns Hopkins APL Technical Digest,* Vol. 18, No. 2, pp. 178–181, 1997.

[6] R. E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Journal of Basic Engineering*, Vol. 82, pp. 34–45, 1960.

[7] T. A. Stansell, *The Transit Navigation System: Status, Theory, Performance, Applications,* Magnavox, 1978.