National Aeronautics and Space Administration

An Introduction to Global Navigation Satellite Systems

Ben Ashman

Navigation and Mission Design Branch

Goddard Space Flight Center, Greenbelt, MD

IEEE International Frequency Control Symposium | Olympic Valley, CA | May 21, 2018

Navigation

- Navigation is the process of determining position and direction
- Generalization of the problem: estimate unknown parameters based on related observations

$$\mathbf{z} = \mathbf{h}(\theta) + \mathbf{v}$$

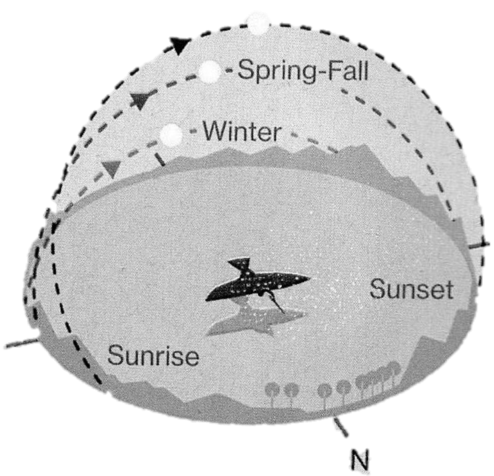
θ = parameter vector (e.g., Cartesian position and velocity, our “state”)

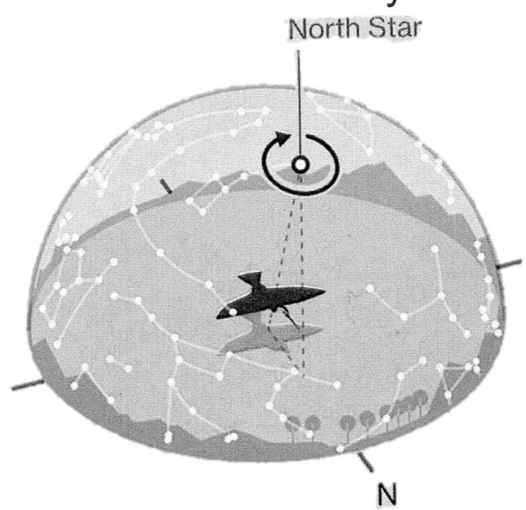
$\mathbf{z}$ = observation vector (i.e., set of measurements)

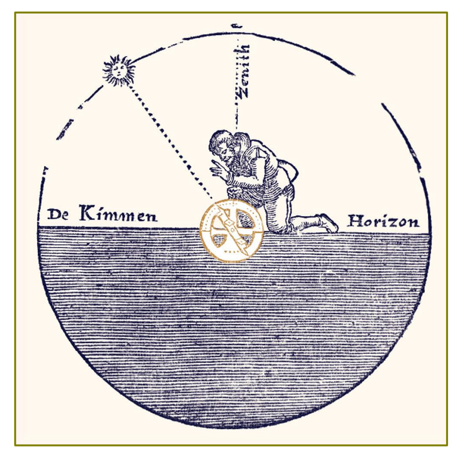
$\mathbf{v}$ = observation noise vector (i.e., measurement error)

$\mathbf{h}(\cdot)$ = relation between parameter set and observation set (i.e., measurement model)

- Given a parameter set, we seek an observation set, a relation between our parameters and observations, and an estimator $\hat{\theta}(\cdot)$, in order to form an estimate: $\hat{\theta} = \hat{\theta}(\mathbf{z})$
- Elegant and effective solutions have been devised by humans and other species for millennia







From left: day and night bird migration [20], astrolabe 1619 [21]

Navigation (continued)

- **Relative navigation**

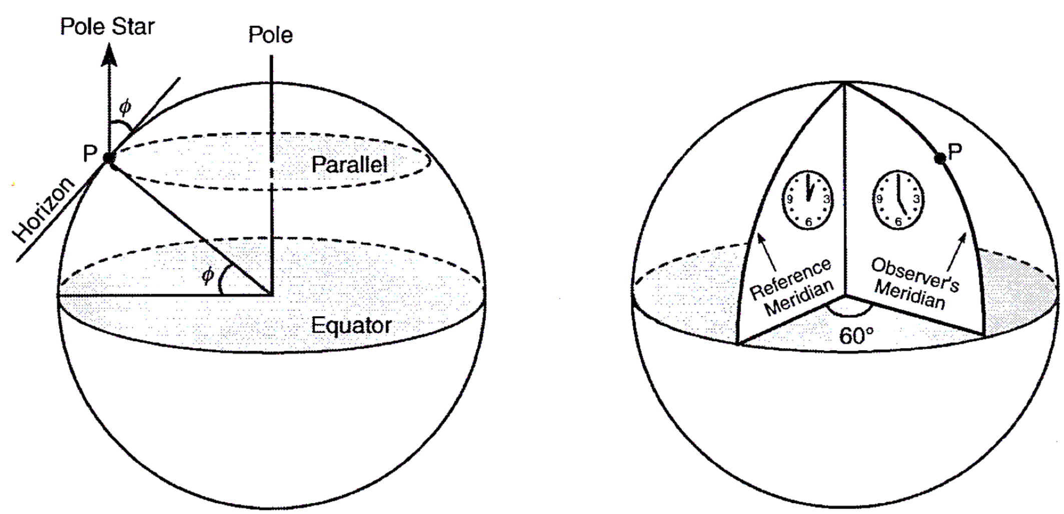
  Dead reckoning: monitor rate of travel and heading using a compass; prone to error, especially at sea

  Landmark bearings: angles to two known landmarks will constrain position in two dimensions

- **Absolute navigation: latitude and longitude (clocks vs. celestial)**

  Latitude: Measure the elevation of pole star above the horizon with a sextant or astrolabe

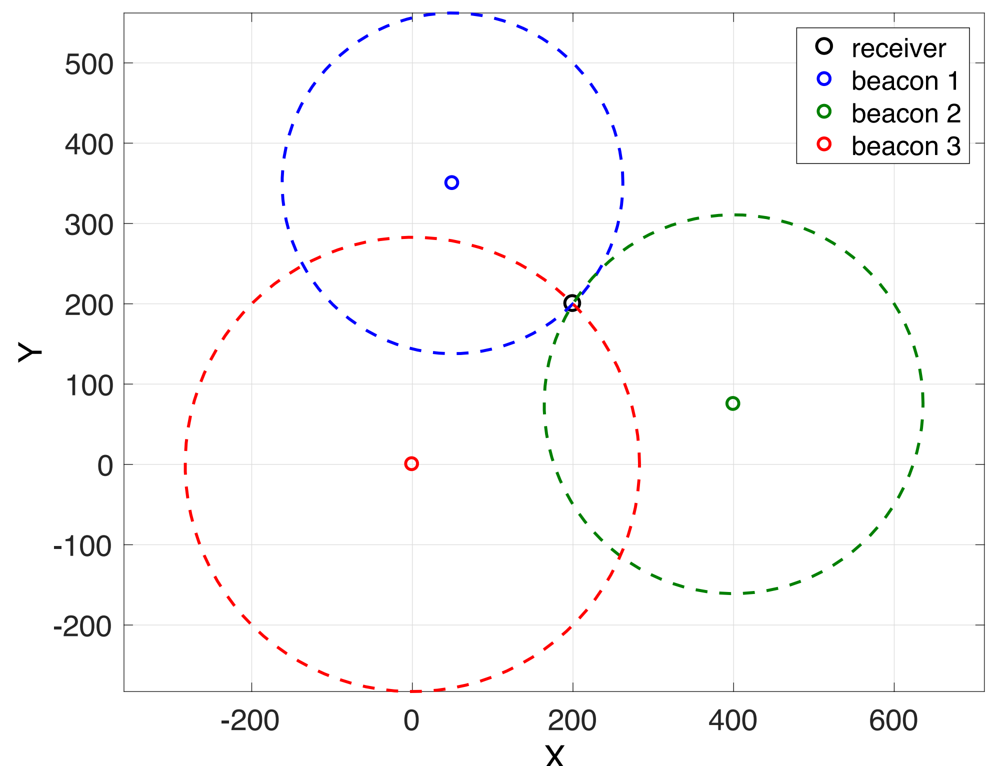
  Longitude: Very good clock or celestial (sextant for the elevation of celestial bodies above the horizon, accurate clock to determine the time of observations, almanac to find the predicted position of the body, magnetic compass to determine azimuth and maintain course continuity between celestial observations)



Latitude (left) and longitude (right) [1]

Radionavigation

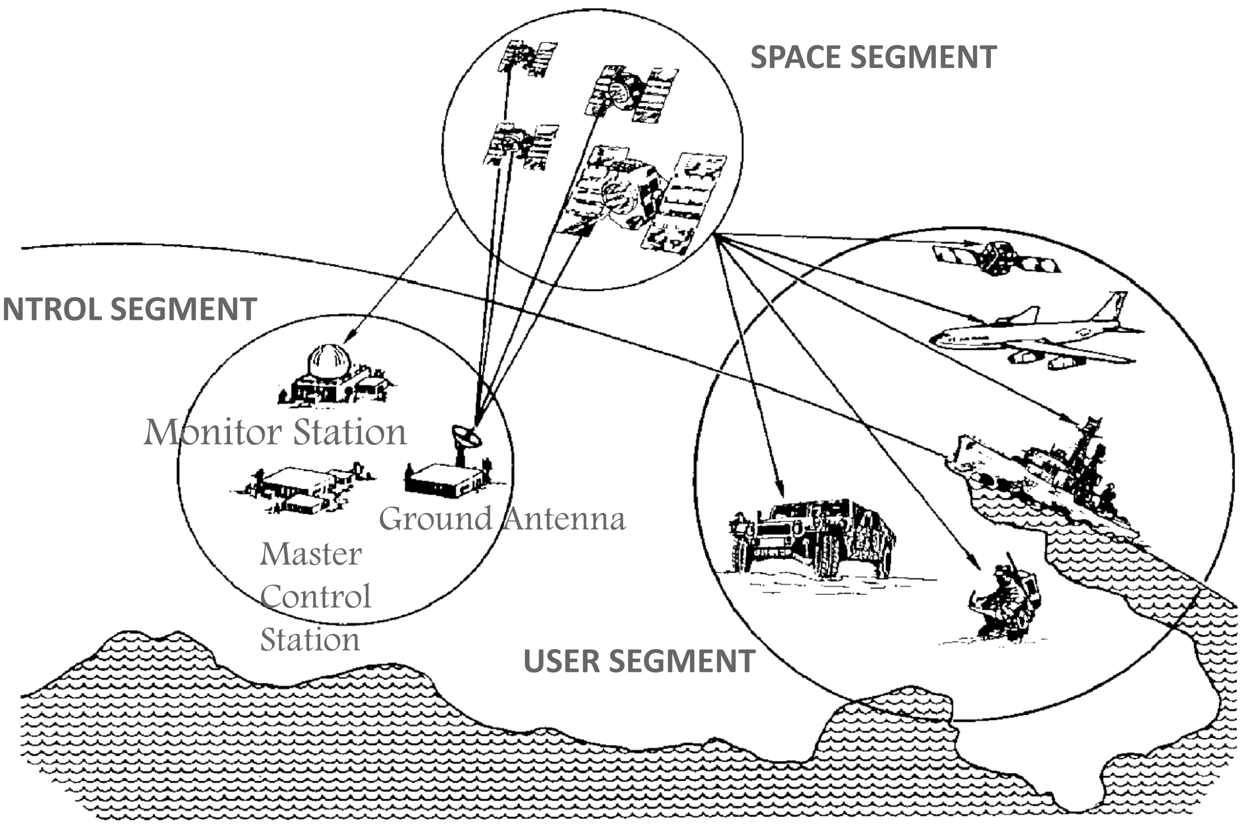
- Measurements: *distances* from known transmitter locations via the measurement of radio frequency signal transit time
- Solution to the estimation problem: *trilateration*, the determination of absolute or relative locations of points by measurement of distances using the geometry of circles, spheres, or triangles
- Ground based:
  - LORAN (1940s), Omega (1960s)
- Satellite-based:
  - Sputnik I (1957), Parus and Tskikada, Transit, MOSAIC, and SECOR (1960s)



Trilateration in 2 dimensions

GNSS

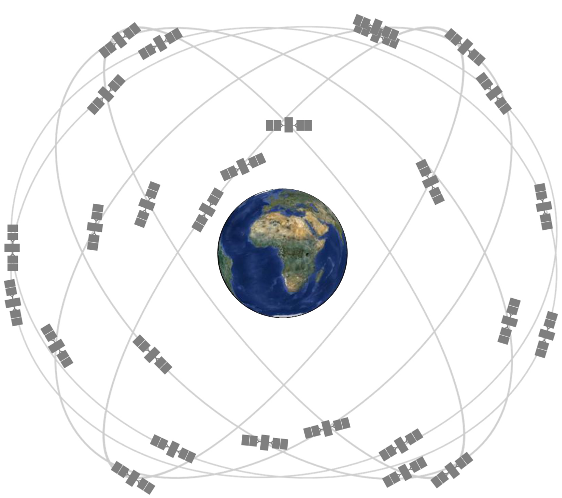
- Global Navigation Satellite Systems (GNSS): radionavigation perfected
- Features
  - Accuracy: 3D accuracies of a few meters and down to millimeters for users with specialized equipment and processing
  - Availability: signal availability anywhere on Earth with a clear view of the sky
  - Integrity: the assurance that expected performance will be realized



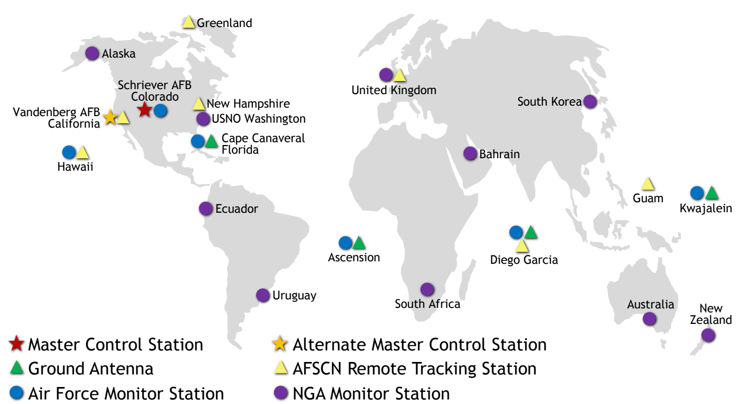
Components of Global Positioning System (GPS) [3]

GNSS (continued)

- Space segment
  - Constellation of satellites in near-circular, Medium Earth Orbits (~20,000 km), each satellite equipped with atomic clocks
- Control segment
  - Network of ground stations and antennas that perform monitoring of the constellation, check for anomalies, generate new orbit and clock predictions, build and send upload to spacecraft
- User segment
  - GNSS receivers—specialized radios that track GNSS signals and produce position and velocity solutions, typically with low-cost clocks



GPS Space Segment [4]

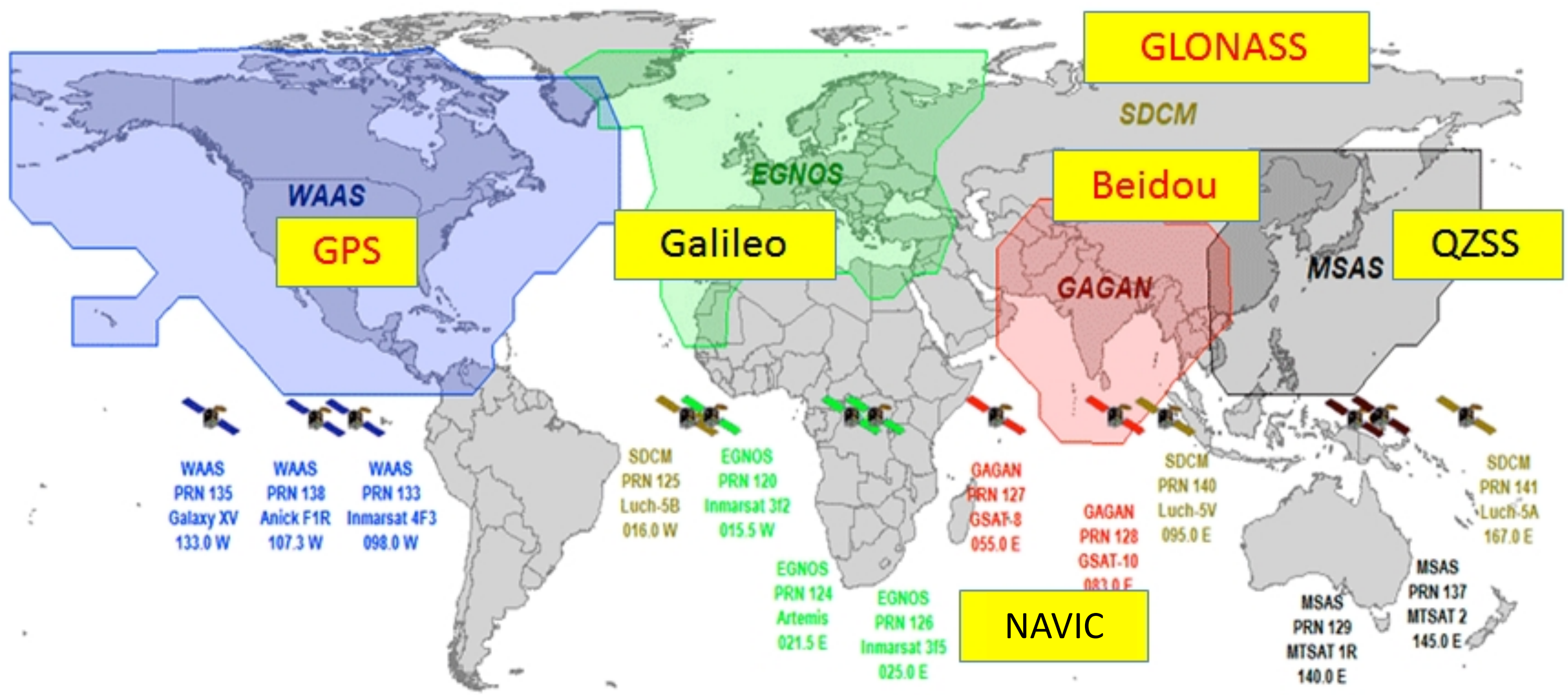


GPS Control Segment [5]

GPS Receivers [6]

GNSS constellations

- GNSS is an umbrella term for satellite constellations that broadcast signals from space for radionavigation
  - Systems with global coverage: GPS (United States), Galileo (European Union), GLONASS (Russia), BeiDou (China)
  - Systems with regional coverage: NAVIC (India), QZSS (Japan)
- This presentation uses GPS as a specific example, but generality is maintained where possible



GNSS constellations, augmentations, and regional constellations [7]

Outline

I. GNSS signals
II. GNSS receivers
III. Measurements
IV. Time reference, orbits, coordinate frames
V. Navigation solution
VI. Current status and future development

Block IIF GPS Satellite [2]

GPS signal structure

- What is required of a radionavigation signal?
  1. Propagation delay between transmitter and receiver can be measured
  2. Transmitters can be distinguished, enabling geometric diversity
  3. Modulation allowing the signal to propagate through space
- For any signal *p(t)* combined with Additive White Gaussian Noise (AWGN) *n(t)*,

$$p(t) + n(t)$$

correlation with a copy of p(t) maximizes the output signal to noise ratio (SNR) (i.e., optimal estimator in the Maximum Likelihood sense), so *p(t)* is designed to have a correlation shape that satisfies signal requirements 1 and 2

- Delay estimation

  Consider a known, continuous-time signal *p(t)* generated at the transmitter that arrives at the receiver with some delay tau: $p(t-\tau)$

  In order to estimate tau, a local replica of *p(t)* is formed at the receiver with test delay tau tilde. The delay estimate, tau hat, is the test delay that maximizes the average (over T_I) of the inner product:

$$\hat{\tau} = \arg\max_{\tilde{\tau}} \frac{1}{T_I} \int_{t-T_I}^{t} p(\alpha - \tilde{\tau}) p(\alpha - \tau) d\alpha$$

GPS signal structure: code

- Autocorrelation in terms of alignment error, $\epsilon = \tilde{\tau} - \tau$:

$$R(\epsilon) = \frac{1}{T_I} \int_{t-T_I}^{t} p(\alpha - \tilde{\tau}) p(\alpha - \tau) d\alpha$$

The ideal autocorrelation function would be

$$R(\epsilon) = \begin{cases} 1 & \text{for } \epsilon = 0 \\ 0 & \text{elsewhere} \end{cases}$$

- Multiple signals are required in order to form a position estimate, however. The trilateration problem relies on geometric diversity. One means of distinguishing transmitters is to minimize the cross correlation of signals from different transmitters:

$$R_{\mathrm{x}}(\tau) = \frac{1}{T_I} \int_{t-T_I}^{t} p^i(\alpha) p^j(\alpha - \tau) d\alpha \qquad\qquad R_{\mathrm{x}}(\epsilon) = 0 \quad \forall \epsilon$$

- These auto- and cross-correlation properties could be achieved with infinitely long random sequences of +1 and -1, known at the transmitter and receiver, and unique to each transmitter

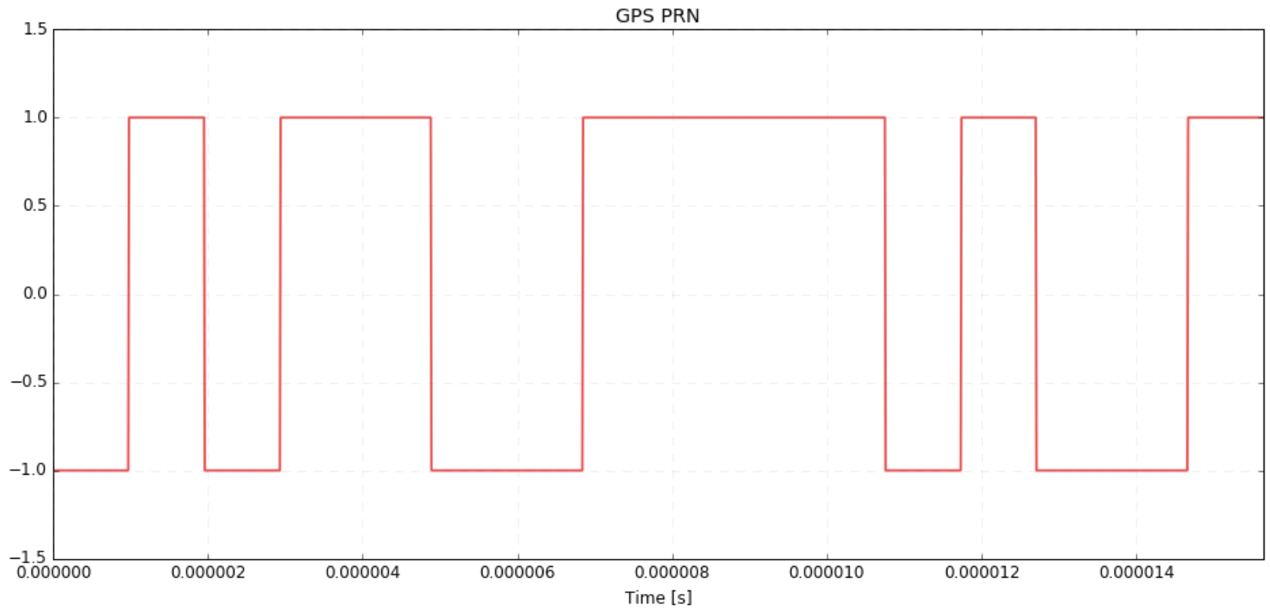
GPS signal structure: code

- This is accomplished using **Pseudorandom Noise (PRN) codes**

  Must be deterministic and finite for practical implementation, but sufficiently long and noise-like to approximate the desired autocorrelation and cross-correlation properties

- GPS Coarse Acquisition Code (C/A code) solution: Gold codes (modulo-2 sum of two linear feedback shift registers)

  Periodic sequence of {+1,-1} pulses called chips, unique to each GPS satellite, length 1023 with period of 1 ms (i.e., f_{chip} = 1.023 MHz)



Selection of GPS C/A code [8]

GPS signal structure: carrier

- Third navigation signal requirement: modulation allowing the signal to propagate through space

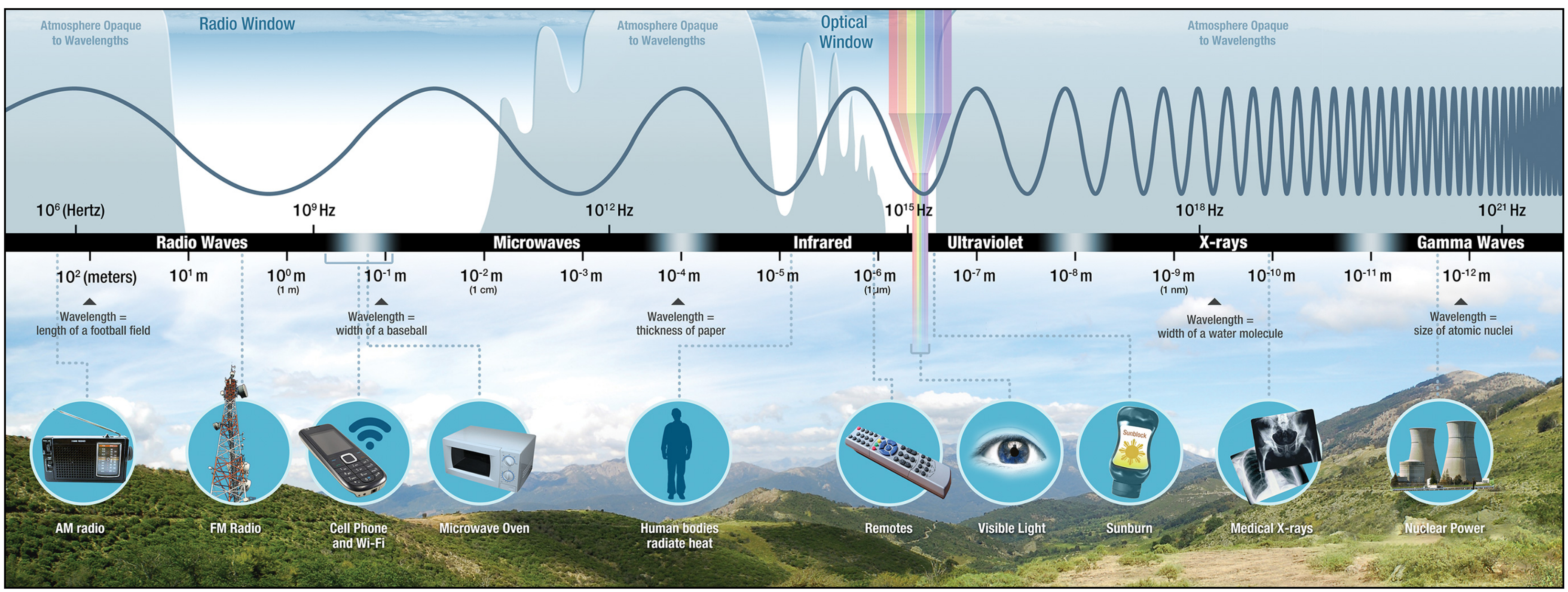
$$p(t)\cos(2\pi f_{carr} t)$$

- Radio frequencies used for satellite navigation—must penetrate atmosphere
- Apparent frequency at the receiver is Doppler shifted due to the relative motion of the transmitter and receiver

$$p(t-\tau(t))\cos(2\pi f_{carr}(t-\tau(t)))$$

$$p(t-\tau(t))\cos(2\pi(f_{carr}+f_D)t-\theta(t_0))$$

where $\tau(t)=\dot{\tau}t+\tau(t_0)$ and $f_D=-\dot{\tau}t=-\dot{r}(t)f_{carr}/c$

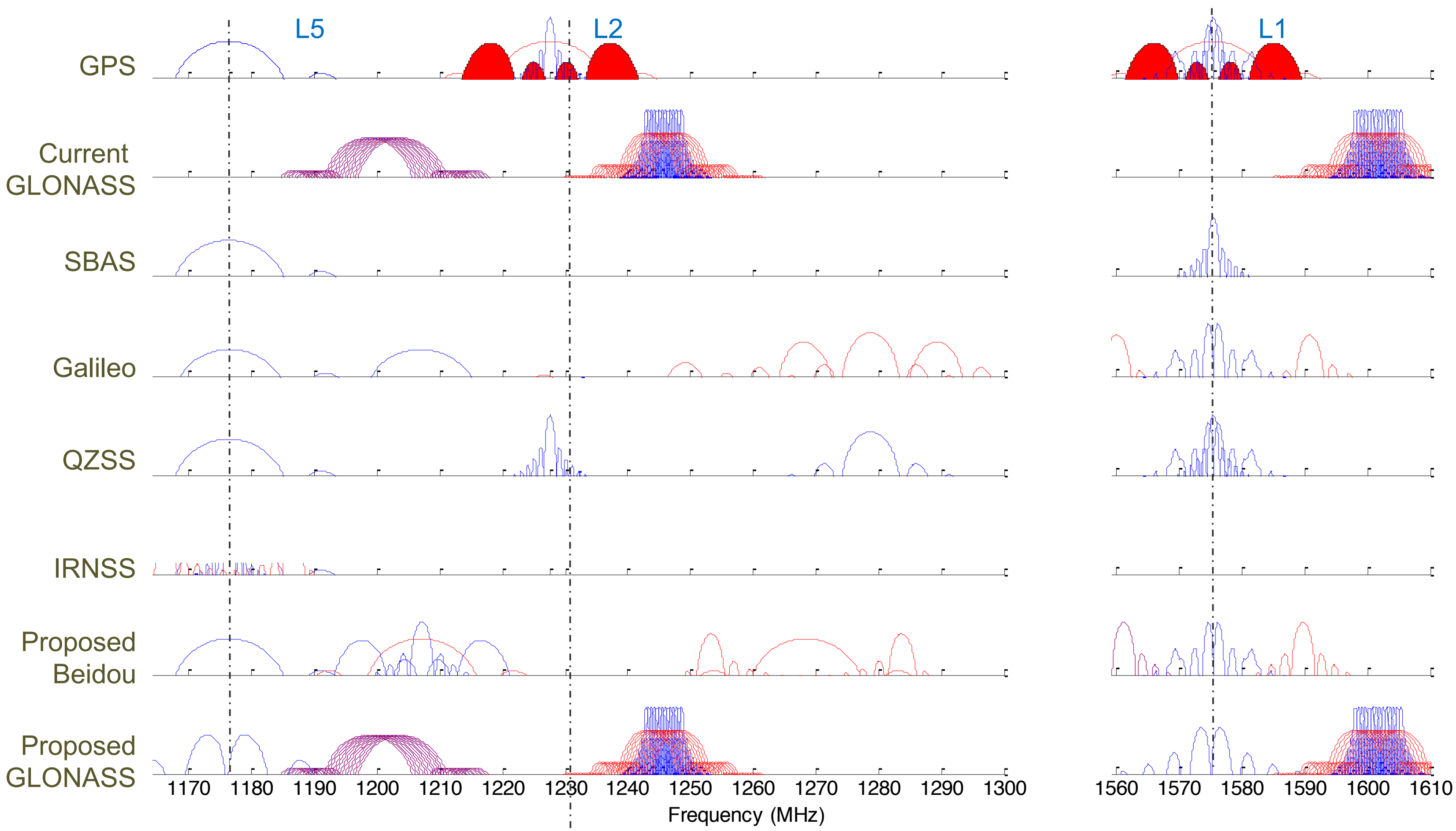


EM spectrum [9]

GNSS carriers

- A variety of carrier frequencies are used by GNSS providers

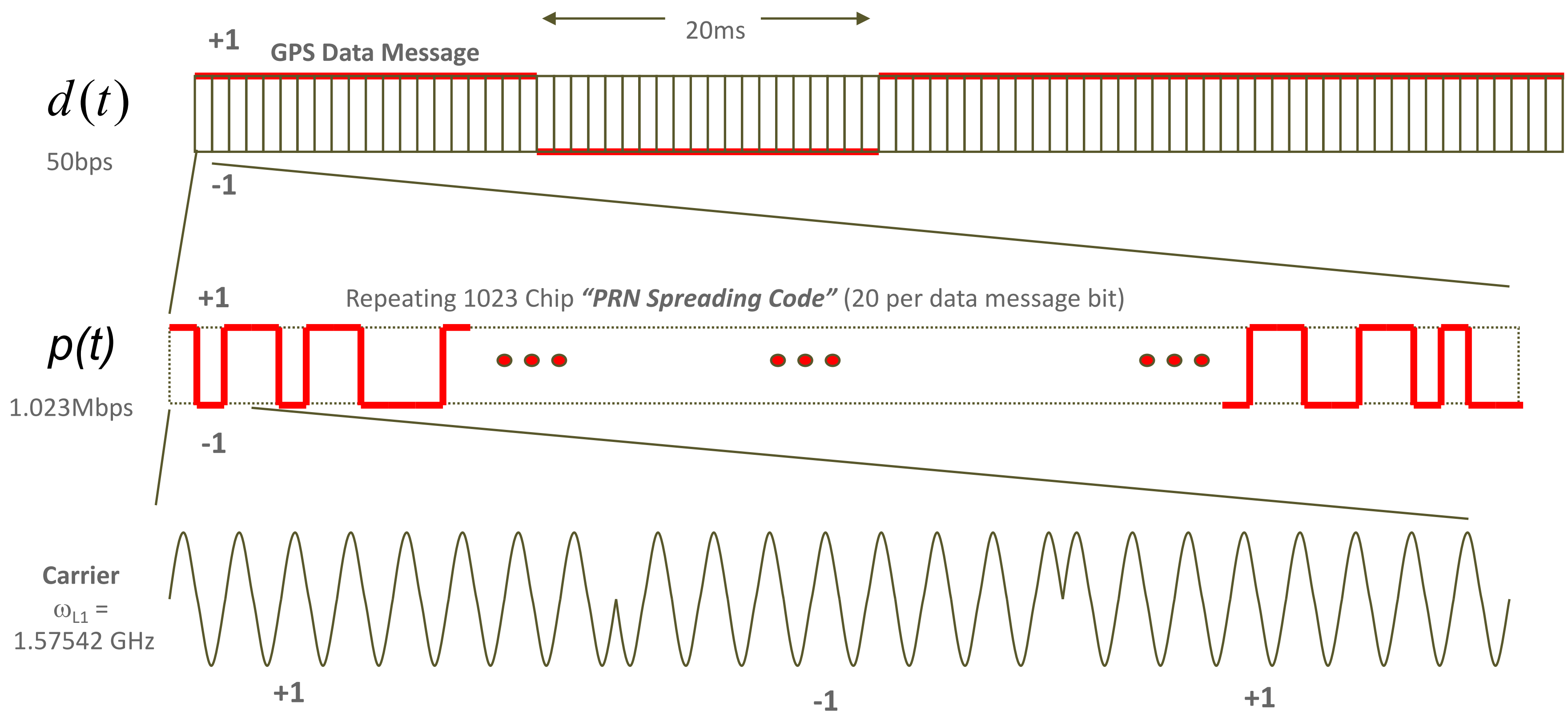
Color code: Blue—open signals Red—restricted or encrypted signals



- GPS L1 (f_{L1} = 1.57542 GHz) will be used as an example in this presentation

GPS signal structure (continued)

- Finally, signal is also modulated with 12.5 minute navigation message, a 50 bps binary sequence containing time tags, GPS satellite ephemerides (i.e., transmitter locations), etc.
- Time domain signal:



Composition of GPS signal in time domain [3]

- Received L1 frequency signal from the *i*-th satellite

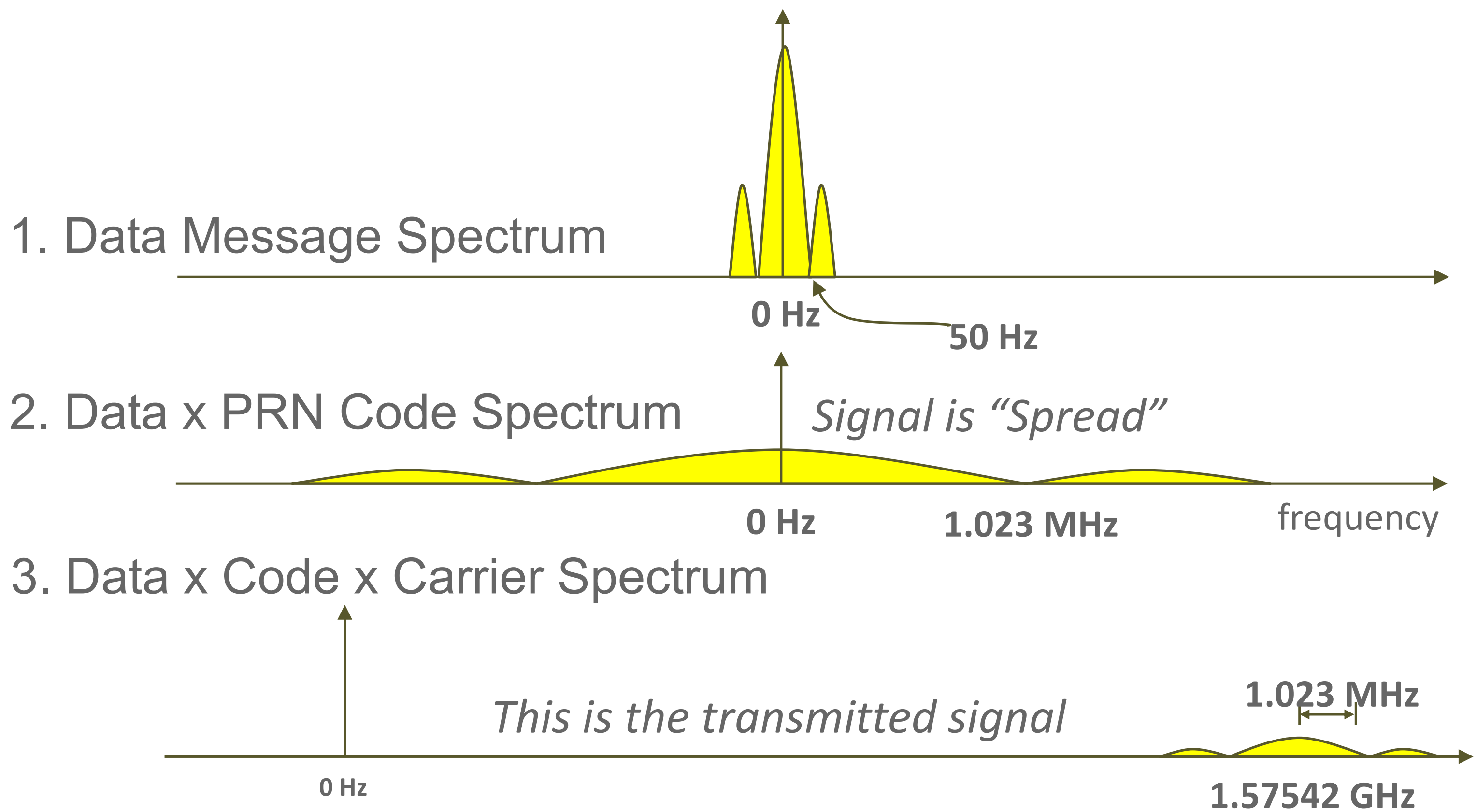
$$y^i(t) = \sqrt{2P_R} d^i(t - \tau^i(t)) p^i(t - \tau^i(t)) \cos(2\pi(f_{L1} + f_D^i)t + \theta^i(t_0)) + v^i(t)$$

GPS signal structure (continued)

- Received L1 signal from the *i*-th satellite

$$y^i(t) = \sqrt{2P_R}d^i(t-\tau^i(t))p^i(t-\tau^i(t))\cos(2\pi(f_{L1}+f_D^i)t+\theta^i(t_0))+v^i(t)$$

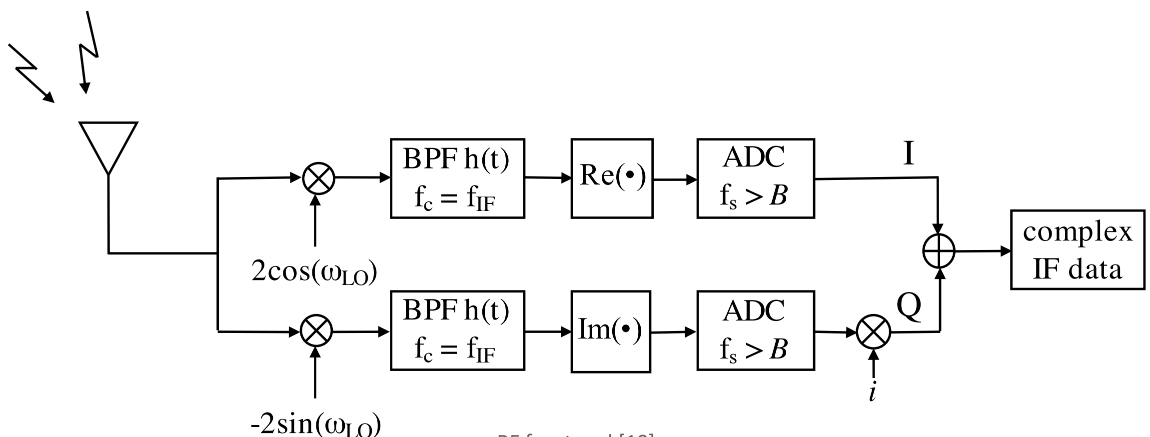
- Frequency domain signal:



Composition of GPS signal in frequency domain [3]

GNSS receivers

- Receiver has three main tasks:
  1. **Acquisition:** Determine which satellites are visible and estimate the propagation delay and Doppler associated with each
  2. **Tracking:** Refine the delay and Doppler estimates and track these features as they change over time
  3. **Navigation:** Use measurements from all visible signals to estimate the receiver's position and velocity
- First the radio frequency signal is downconverted to an intermediate frequency (IF) for processing

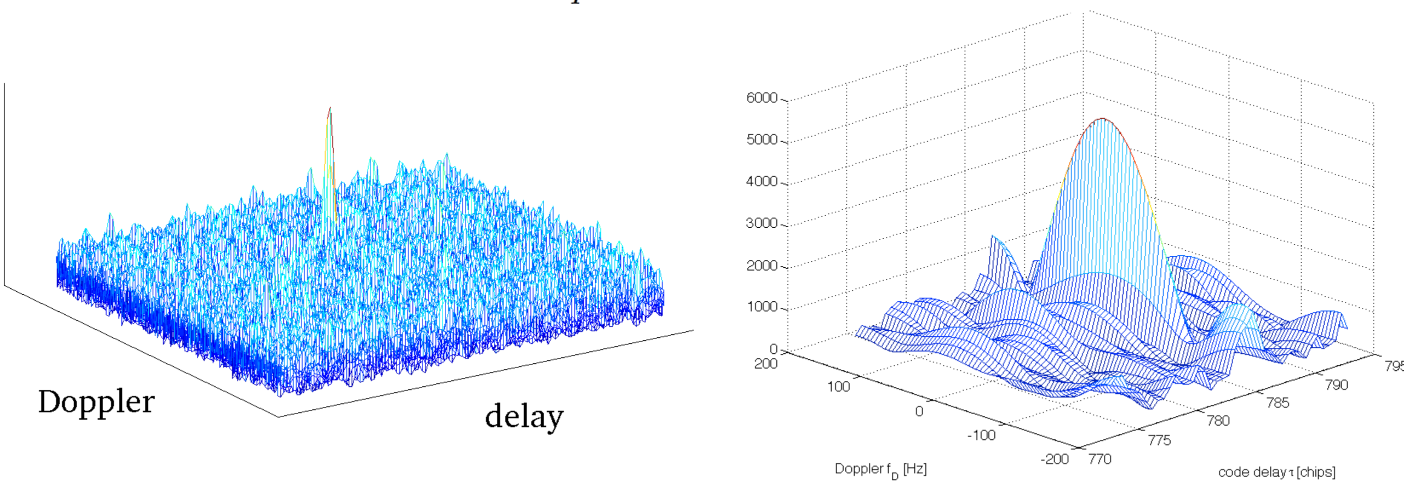


RF front end [10]

GNSS receivers: acquisition

- Acquisition seeks to determine whether a particular satellite is visible (via its unique PRN) and estimate its delay (modulo one code period, 1 ms) and Doppler
- Correlation of an incoming signal with a local replica, mismatched in frequency and delay, forms what is known as an asymmetric ambiguity function:

$$\chi(\tau, f_D) = \int_{t-T_I}^{t} p(\alpha)p^*(\alpha - \tau)e^{i2\pi\Delta f_D \alpha}\, d\alpha$$

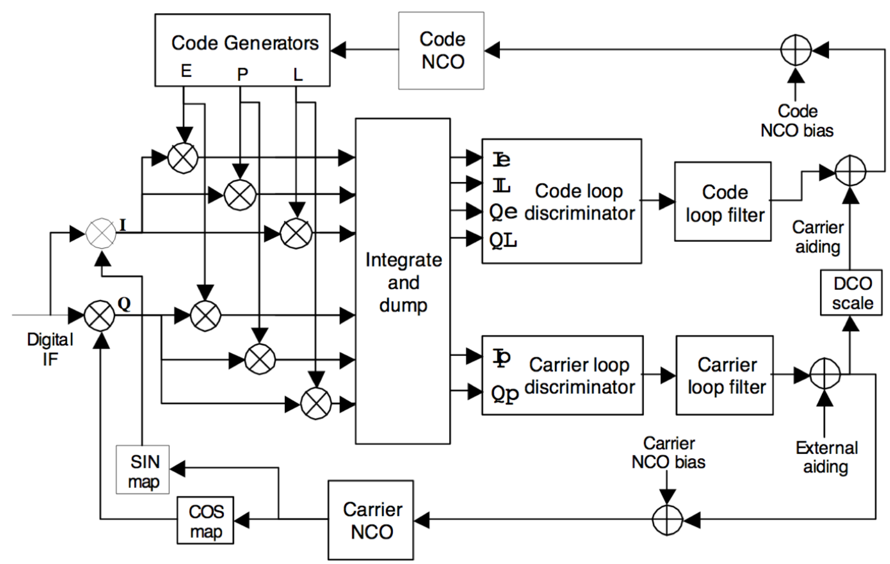


Ambiguity function magnitude complete [3] (left) and zoomed in [11] (right)

- Delay and Doppler values are tried over a search space. Correlation magnitudes are compared to the noise floor—if the carrier to noise spectral density exceeds a threshold, the signal is determined visible and the delay and Doppler at the correlation peak are used to seed tracking.
- Pre-detection integration time, T_I, is an important parameter in detecting weak signals.

GNSS receivers: tracking

- Tracking seeks to refine the delay and Doppler estimates produced by acquisition
  1. Input signal is correlated with a local replica
  2. Correlation result is filtered to produce error terms that quantify the difference between the input and local signal
  3. A feedback process makes adjustments to the local signal replica according to the error terms
- In addition to converging on the input signal delay and Doppler parameters, the tracking of a dynamic signal allows for measurements of changing signal features and more accurate estimates of the signal to noise ratio
- Most receivers compute three correlations per signal: Early, Prompt, and Late
  - Phase of prompt corr. gives error signal for carrier tracking
  - Comparing size of Early and Late corr. gives error signal for PRN code tracking
- Coupled feedback loops DLL and PLL maintain lock on code and carrier signal parameters



Code and carrier tracking [3] from [11]

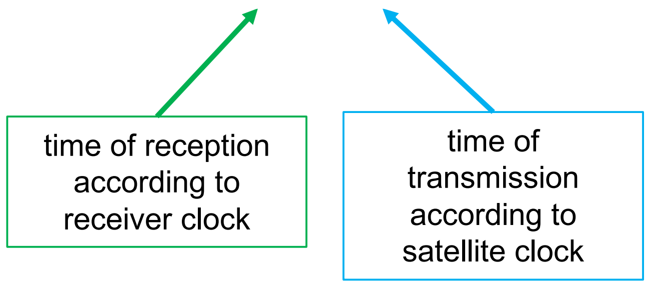
GNSS measurements

- GNSS observables (i.e., receiver outputs)
  1. **Pseudorange:** propagation delay plus receiver clock bias (measured from the PRN code to a fraction of a chip: ~meter level accuracy)
  2. **Doppler:** measured frequency shift of the received carrier
  3. **Carrier phase:** measured fractional and accumulated whole cycle phase of the carrier (measured to small fraction of 19 cm cycle: ~mm precision)
  4. **C/N_0:** carrier to noise spectral density estimate in dB-Hz

GNSS measurements: pseudorange

- Pseudorange measured from the *i*-th satellite ("pseudo" because of receiver clock bias):

$$\rho^i = c(\tilde{t}_r - \tilde{t}_t^i)$$



GNSS measurements: pseudorange

- Pseudorange measured from the *i*-th satellite ("pseudo" because of receiver clock bias):

$$
\begin{aligned}
\rho^i &= c(\tilde{t}_r - \tilde{t}_t^i) \\
&= c((t_r + \delta t_r) - (t_t^i + \delta t_t^i)) + \epsilon
\end{aligned}
$$

time of reception according to receiver clock

time of transmission according to satellite clock

- Transmission and receive times each expressed as a sum of the "true" time (i.e., the time according to a common time standard, such as GPST) plus an unknown bias

GNSS measurements: pseudorange

- Pseudorange measured from the *i*-th satellite (“pseudo” because of receiver clock bias):

$$
\begin{aligned}
\rho^i &= c(\tilde{t}_r - \tilde{t}_t^i) \\
&= c((t_r + \delta t_r) - (t_t^i + \delta t_t^i)) + \epsilon \\
&= c\delta t_r - c\delta t_t^i + |x - x_t^i| + Q^i + I_{L1}^i + T^i + \epsilon
\end{aligned}
$$

true receiver position (x)

broadcast satellite position (x_t^i)

- Propagation delay:

$$\tau^i = (r^i + Q^i + I_{L1}^i + T^i)/c$$

Q^i is the satellite orbit error

r^i is the geometric range between the i-th transmitting satellite and the receiver, $|x - x_t^i|$

I_{L1}^i is the delay due to the ionosphere, a region of ionized gas in the upper atmosphere where the time varying density of free electrons and ions introduces a dispersive (frequency dependent) delay

T^i is the delay due to the troposphere, the lowest region of the atmosphere, a non-dispersive medium consisting of dry gases and water vapor

GNSS measurements: pseudorange (cont.)

- Propagation delay:

$$\tau^i = (r^i + Q^i + I^i_{L1} + T^i)/c$$

Acquisition and tracking measure code phase, i.e., ambiguous time of transmission modulo one code period (1 ms for GPS C/A code, or approximately 300 km): $p(t - \tau)$

The navigation message must be decoded to form a pseudorange

- Navigation message is organized into six second subframes, each beginning with an 8-bit Telemetry word (TLM) and Hand-over word (HOW), the latter of which contains the satellite time the subframe was transmitted

svtime_of_transmission =

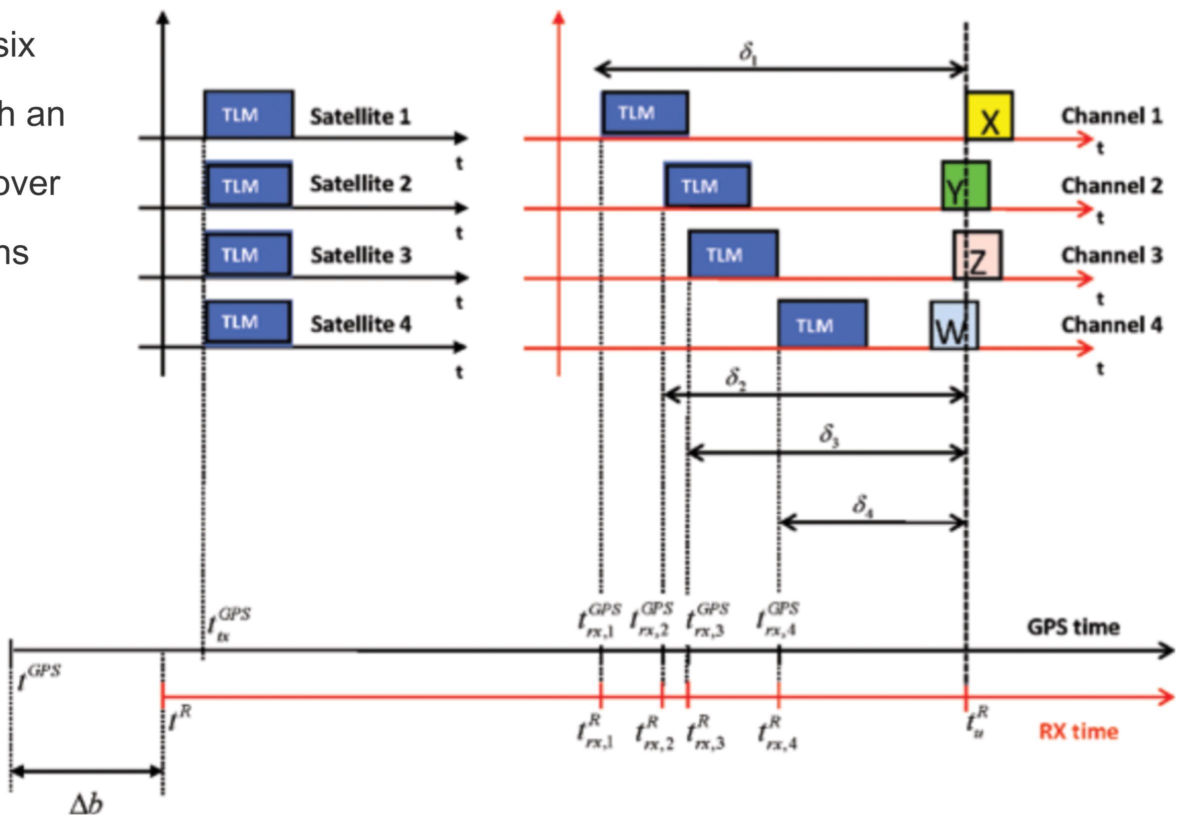
svtime_of_subframe

*+bits_since_subframe_start*0.02*

*+whole_cacodes_since_bit_start*0.001*

+whole_chips_into_current_cacode/1.023e6

+frac_code_phase_chips/1.023e6;



Calculation of pseudorange using four satellites [12]

GNSS measurements: pseudorange (cont.)

- Propagation delay:

$$\tau^i = (r^i + Q^i + I^i_{L1} + T^i)/c$$

- Orbit error

  Maintained to within ~1 m RSS by Control Segment

- Ionosphere

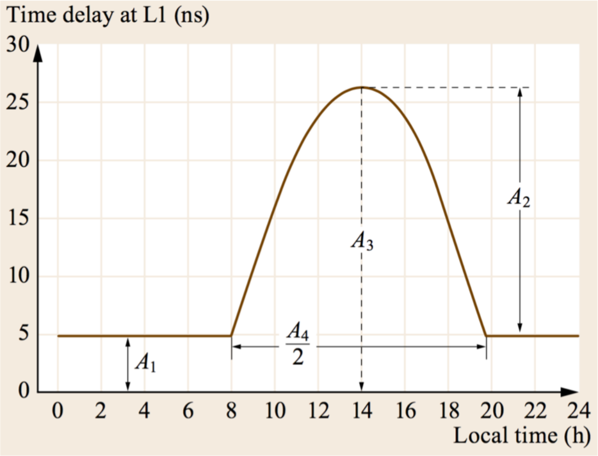
  Group delay for pseudorange due to the ionosphere:

$$I^i_{L1} = \frac{40.3}{f^2_{L1}}\text{TEC}$$

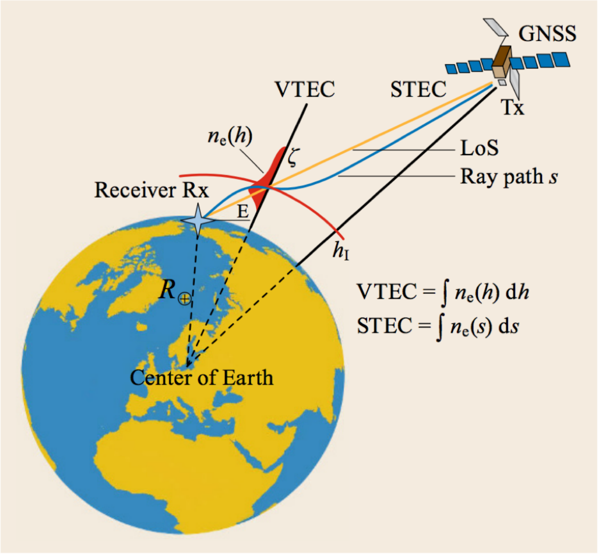
  where TEC is the Total Electron Content in a 1 m^2 column from the receiver the transmitter.

  Ionospheric delay can be corrected by using measurements from two frequencies (note frequency dependence, here we use GPS L1) or through a model that predicts TEC

  GPS uses the Klobuchar model, in which four parameters (defined by eight numbers in the navigation message) are used to define the daily zenith variation at the ionospheric pierce point



GPS ionospheric delay model: Klobuchar [13]



Geometry of zenith TEC (vertical TEC, VTEC) and slant TEC (STEC) [13]

GNSS measurements: pseudorange (cont.)

- Propagation delay:

$$\tau^i = (r^i + Q^i + I^i_{L1} + T^i)/c$$

- Troposphere
  - Not frequency dependent, wet (< 0.25 m, large variation) and dry (~2 m, small variation) components
  - Corrected using models (e.g., Hopfield) that incorporate empirical corrections—typically average meteorological parameters for latitude, longitude, and season

- Complete pseudorange expression:

$$\rho^i = r^i + ct_{b,r} - ct^i_{b,s} + Q^i + I^i_{L1} + T^i + \epsilon^i$$

- Relativity
  - Second-order Doppler shift: a clock moving in an inertial frame runs slower than a clock at rest
  - Gravitational frequency shift: a clock at rest in a lower gravitational potential runs slower than a clock at rest in a higher gravitational potential
  - GNSS space segment atomic clocks are offset to compensate for these effects—without correction satellite clocks would gain almost 40 microseconds per day (~10 km range error)

GNSS measurements: pseudorange (cont.)

- Propagation delay:

$$\rho^i = r^i + ct_{b,r} - ct^i_{b,s} + Q^i + I^i_{L1} + T^i + \epsilon^i$$

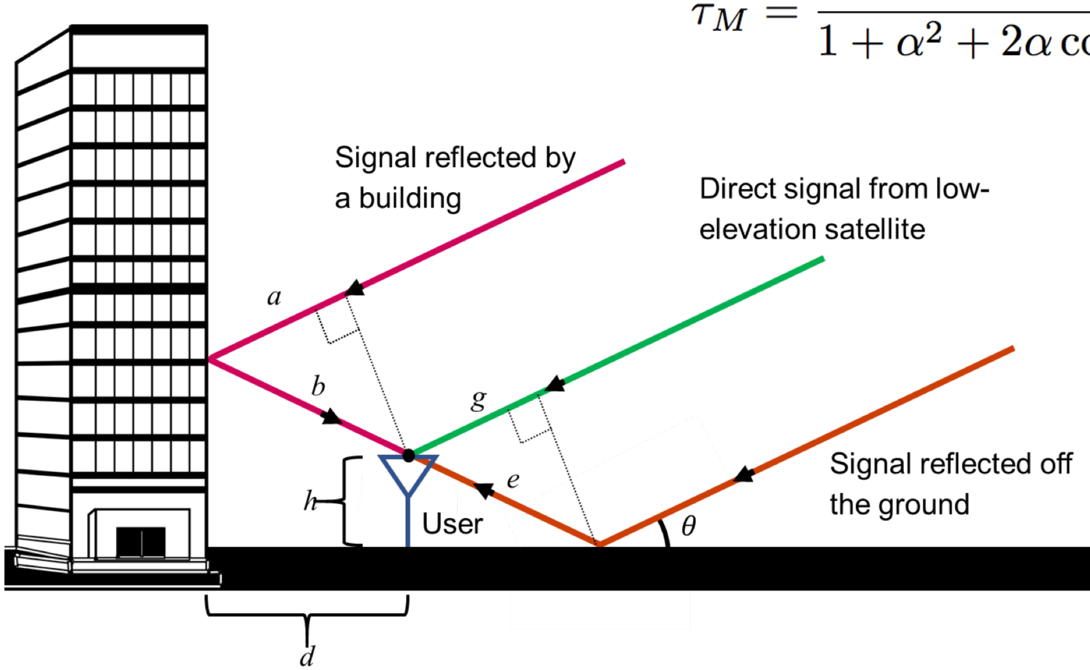
- Multipath

  Reflected signals are received as delayed, attenuated replicas of the direct signal

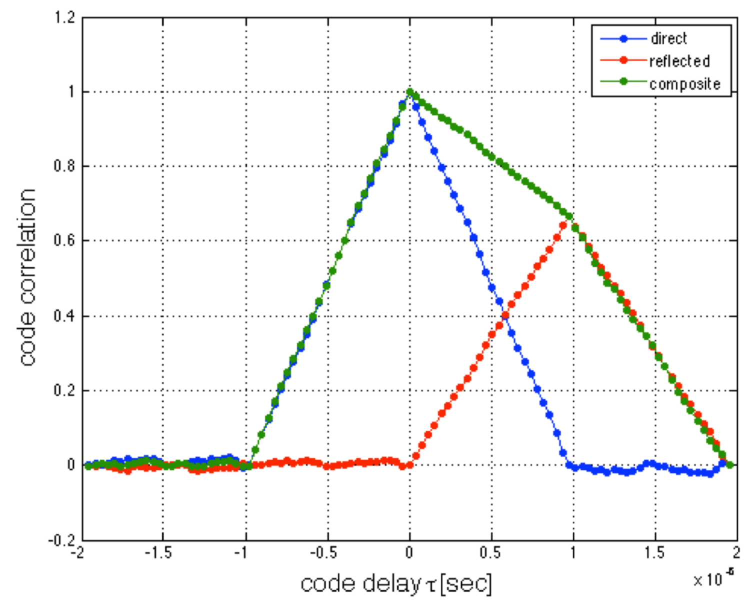
  Correlation shape of the combined signal causes an error in the code tracking loop that depends on geometry, number and strength of reflections, and tracking loop design

  E.g., one signal, noncoherent DLL:

$$\tau_M = \frac{\alpha^2\delta + \alpha\delta\cos\psi}{1 + \alpha^2 + 2\alpha\cos\psi}$$



Direct, building-reflected, and ground-reflected signal paths in multipath interference scenario [19]



Direct, reflected, and composite correlations [10]

GNSS measurements (cont.)

- Measured **Doppler** shift is a combination of the changing geometric line of sight and the unknown receiver clock drift. Time derivative of the pseudorange:

$$\dot{\rho}^i(t) = \dot{r}^i(t) + c\dot{t}^i_{b,r} - c\dot{t}^i_{b,s}$$

recall: $f_D = \dot{r}(t) f_{carr}/c$

unknown receiver clock drift

transmitter clock drift approximately known

- **Carrier phase**

$$\phi^i(t) = -2\pi f_{carr} \rho^i(t)/c + 2\pi M$$

Can be measured with much higher precision than code phase (i.e., pseudorange), ~cm for GPS, but ambiguous on the order of carrier cycles, 19 cm for GPS. Combine with code measurements or use for precise measurement of change (Accumulated Delta Range)

Ionosphere also induces a delay, but opposite in sign relative to pseudorange, leading to a code/carrier divergence

Multipath also introduces an error in carrier phase measurements, as the geometry changes and the received reflections cycle through constructive and destructive interference with the direct signal. For a single reflection:

$$\phi^i_M(t) = \left(2\pi\Delta^i/\lambda_{carr} + \phi^i_R\right) \mathrm{MOD} 2\pi$$

GNSS measurements (cont.)

- **Carrier phase (cont.)**

$$\phi^i_M(t) = \left(2\pi\Delta^i/\lambda_{carr} + \phi^i_R\right) \mathrm{MOD} 2\pi$$

Ground reflection

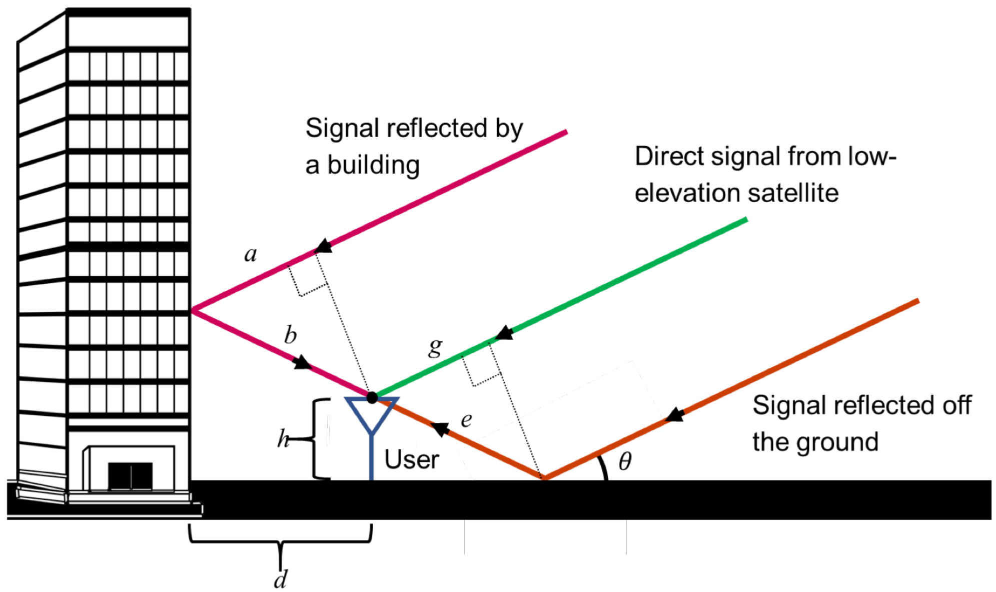
$$\Delta = a + b = 2d\cos\theta$$

$$\frac{\partial \phi^i_M}{\partial t} = -\left(\frac{4\pi d}{\lambda_{carr}}\sin\theta\right)\frac{\partial\theta}{\partial t}$$

Wall reflection

$$\Delta = e - g = 2h\sin\theta$$

$$\frac{\partial \phi^i_M}{\partial t} = \left(\frac{4\pi h}{\lambda_{carr}}\cos\theta\right)\frac{\partial\theta}{\partial t}$$



Direct, building-reflected, and ground-reflected signal paths in multipath interference scenario [19]

- **Carrier to noise spectral density (C/N_0)** is the signal power divided by the measurement noise power density. Unlike signal to noise ratio, SNR, this is independent of the receiver bandwidth *B*. It is a power to noise density per unit frequency, expressed in units dB-Hz:

$$\frac{C}{N_0} = SNR + B = (P_R - B - N_0) + B$$

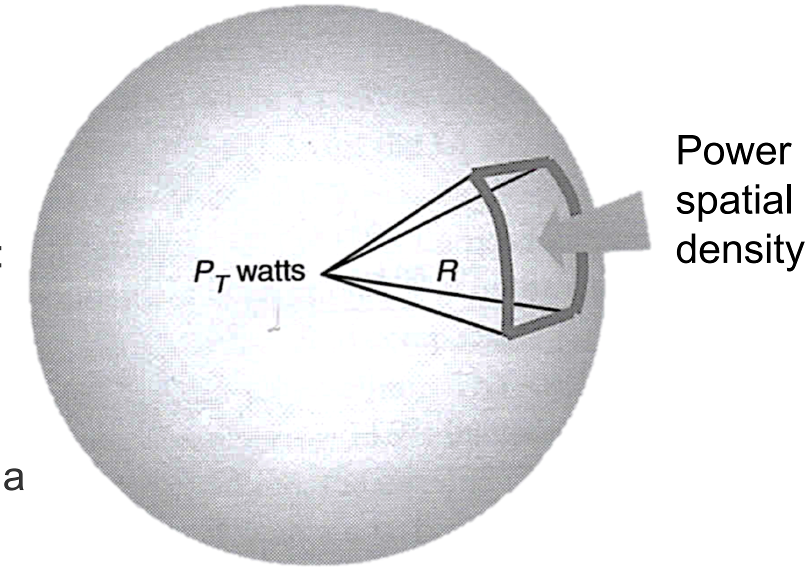
GNSS measurements: link budget

- GNSS link budget

  The received signal power is a combination of power spatial density produced by the transmitter at the receiver and the effective area of the receive antenna:

$$P_R = \mathcal{P}_{T,rcvr} A_R$$

  Effective area is a measure of an antenna's ability to capture power in an electric field on the antenna from a certain direction, defined as: $A_{eff} = G\lambda^2/4\pi$



Power spatial density [1]

An isotropic antenna radiates power equally in all directions. At a given distance from the transmitter, *R*, the power density is simply the transmitted power divided by the surface area of the sphere: $P_T/(4\pi R^2)$ This accounts for spreading loss.

Spreading loss can be offset by focusing the transmitted power in a particular direction, a property described by the transmit antenna gain, G_T. The power density at the receiver is:

$$\mathcal{P}_{T,rcvr} = \frac{P_T G_T}{4\pi R^2}$$

GNSS measurements: link budget

- GNSS received power

  The power density at the receiver is:

$$\mathcal{P}_{T,rcvr} = \frac{P_T G_T}{4\pi R^2}$$

  Thus for receive antenna gain G_R the received power is given by the Friis transmission formula:

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi)^2}$$

- Noise power per frequency unit: $N_0 = kT_{Eff}$

  Effective temperature used to characterize all noise, not just thermal

- Carrier to Noise Spectral Density

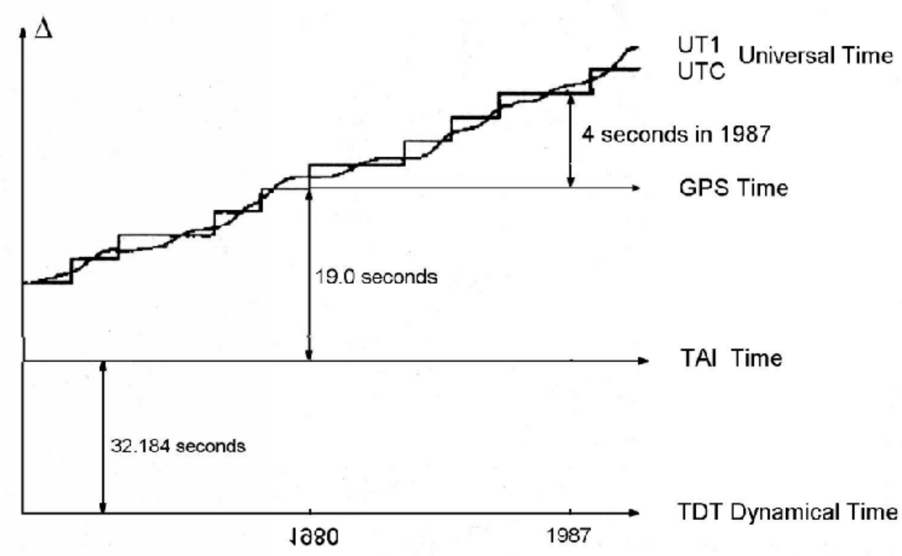
$$\frac{C}{N_0} = \frac{P_R}{N_0} = \frac{P_R}{P_N}\frac{1}{B}$$

  Typical values for GPS:

  T_E = 290 K → N_0 = -201 dBW-Hz, P_R = -156 dBW → C/N_0 = 45 dB-Hz

GNSS time reference

- GNSS requires a common time scale for computing ranges
- GPS Time (GPST) is the operational time scale of GPS
- To keep satellites on GPST adequately, atomic clocks are required
  - Corrections in the navigation message are used to synchronize satellites to GPST
  - For example, to limit clock error to 1 m over 12 hrs requires drift < 8 x 10^{-14} s/s
- GPST coarsely steered to align with Universal Consolidated Time (UTC) as maintained by the US Naval Observatory via corrections in the navigation message
- **Traceability to UTC USNO enables precise time and frequency transfer on a global scale**

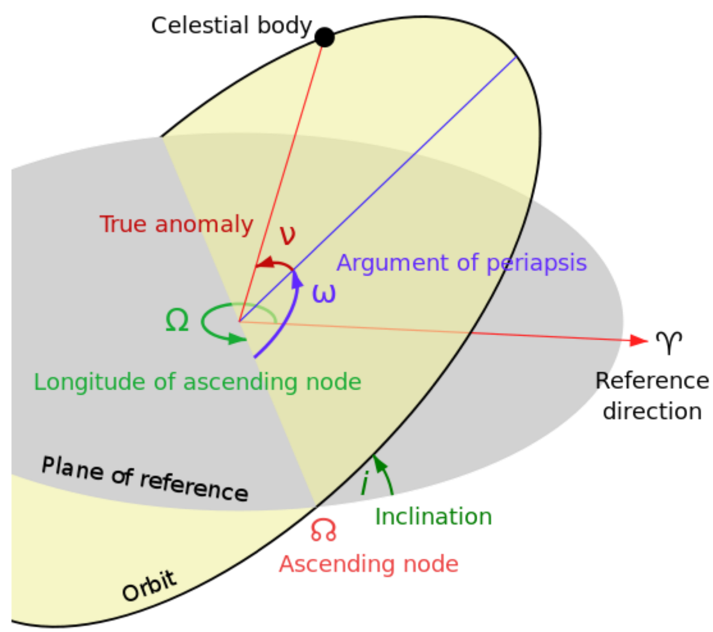


Time systems [14]

- Tidal friction and other processes that cause a significant redistribution of mass are slowing the Earth's rotation, lengthening the solar day by ~2 ms / century
- UTC incorporates leap seconds to maintain alignment with sidereal time (UT1), but GPST does not. This difference is a persistent challenge for receiver designers and users.

GNSS orbits

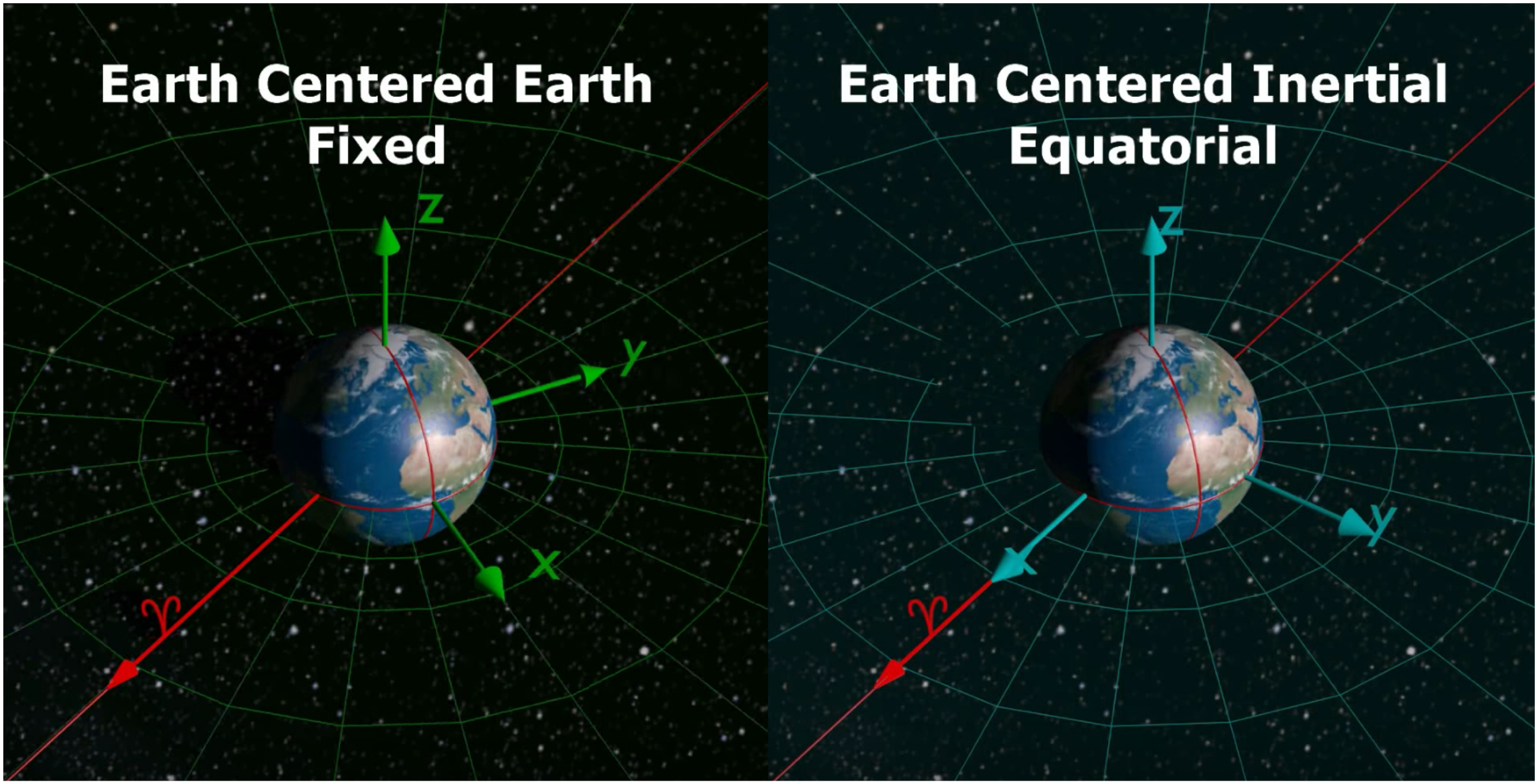
- Satellites make great reference points
  - Small number can provide global coverage
  - They can be precisely located
- Orbital mechanics are well understood and satellite orbit determination is a refined science; for GPS, for example, the MCS estimates and predicts satellite orbits to less than 1 m
- Dual frequency observables from a network of monitor stations used to estimate orbits and satellite clock biases
- Each GPS satellite broadcasts its ephemeris (valid for 2-4 hours) and an almanac (subset of ephemeris parameters for every satellite in the constellation—not accurate enough for navigation, but accurate enough for a satellite search)
- International GNSS Service (IGS) and others maintain large networks of monitor stations, use advanced techniques to locate satellites with cm-level accuracy



Keplerian elements [16]

GNSS Coordinate Frames

- Earth Centered Earth Fixed (rotating reference frame) versus Earth Centered Inertial (ECI)



ECEF and ECI reference frames animation [15]

- Explanatory video: https://youtu.be/DbYapFLJsPA

GNSS Coordinate Frames

- GNSS orbit determination is performed in an inertial (non-rotating) frame
  - Example: Earth Mean Equator and Mean Equinox of the J2000 epoch (January 1, 2000 at 12:00 TT), x-axis is aligned with the mean equinox, z-axis aligned with the Earth's spin axis or celestial north Pole
- Terrestrial navigation is performed in an Earth-fixed frame (rotating with the Earth) for convenience to users
  - Example: GPS uses the WGS84 frame, a 3-dimensional coordinate reference frame for establishing geodetic latitude, longitude, and heights for navigation. Defined by the US National Geospatial Intelligence Agency.

Navigation solution

- Position estimation with pseudorange

  Want to estimate receiver position and clock bias at some instant in time:

$$\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \text{ and } b = c\delta t_{b,r}$$

  Given N > 4 pseudorange measurements (corrected for transmitter clock bias):

$$\rho^i = \left|\mathbf{x} - \mathbf{x}_t^i\right| + b + v^i$$

  Standard approach is to solve as a non-linear least squares (NLLS) problem by Gauss-Newton method:

$$\text{minimize} \qquad J(\hat{\mathbf{x}}, \hat{b}) = \sum_{i=1}^{N} \left(\rho^i - (\left|\hat{\mathbf{x}} - \mathbf{x}_t^i\right| + \hat{b})\right)^2$$

1. Linearize about initial guess $(\hat{\mathbf{x}}_0, \hat{b}_0)$
2. Solve linear least squares problem for $(d\hat{\mathbf{x}}, d\hat{b})$
3. Set $\hat{\mathbf{x}}_1 = \hat{\mathbf{x}}_0 + d\hat{\mathbf{x}},\ \hat{b}_1 = \hat{b}_0 + d\hat{b}$
4. Iterate

Navigation solution: DOP

- *In general*, when solving the linear least squares problem

$$z = H\theta + v\ ,\ Cov\{z\} = \sigma_z^2 I$$

- The covariance of the least squares solution θ^* is

$$\sigma_\theta^2 = Cov\{\theta^*\} = \sigma_z^2 (H^T H)^{-1} = \sigma_z^2 W$$

- *W* (the inverse Gramian matrix) transforms measurement noise into solution noise

- In GPS, the *i*-th row of *H* is

$$\mathbf{h}_i = [\mathbf{u}_i^T, 1] \text{ with } \mathbf{u}_i = \frac{\mathbf{x} - \mathbf{x}_t^i}{|\mathbf{x} - \mathbf{x}_t^i|} \text{ (unit vector from transmitter to receiver)}$$

- Thus, *W* is determined by the geometry of the visible transmitters. Dilution of Precision (DOP):

$$GDOP := \sqrt{\sum_{i=1}^{4} W_{ii}} \quad PDOP := \sqrt{\sum_{i=1}^{3} W_{ii}} \quad TDOP := \sqrt{W_{44}}$$

- Examples

  If transmitters are in a plane, *H* is rank deficient and $GDOP = \infty$

  If transmitters are located at corners of a tetrahedron $GDOP = \sqrt{3}$ (minimum for N = 4)

Navigation solution: typical GPS error budget

Error Source	Basic single freq	Precise dual-freq, assisted
Ionosphere (< 1000 km)	~3 m (single frequency, using broadcast model)	Dual frequency <1 cm
Troposphere (< 20 km)	0.1-1 m	1 cm level using estimators, advanced models
GPS orbits	<2.0 m (broadcast ephem)	1 cm, Int. GNSS service (IGS)
GPS clocks	<2.0 m (broadcast clock)	1 cm (IGS)
Multipath (“clean” environment)	0.5-1 m code	0.5-1 cm carrier
Receiver Noise	0.25-0.5 m code	1-2 mm carrier
RSS range error	4 m	2 cm
Typical GDOP	2	2
RSS solution error	8 m	4 cm

- Disclaimer: for illustration purposes only

Current status and future development

- Status of current GNSS constellations

  GPS (US) - fully operational with global coverage since 1995, 31 satellites in orbit, issued a request for proposals in February 2018 for the next block of satellites, GPS 3F

  GLONASS (Russia) – full operational capacity / global coverage achieved in 1995, lost and then regained in 2011, 24 satellites in orbit, next satellite block, K2, to enter service soon

  Galileo (European Union) – first launch in 2005, full operational capacity expected in 2019, 30 satellites in orbit, will provide global search and rescue (SAR) functionality

  BeiDou (China) – currently providing service in Asia-Pacific, expected to reach global coverage in 2020, 22 satellites in orbit, may be more accurate than GPS, includes GEO sats

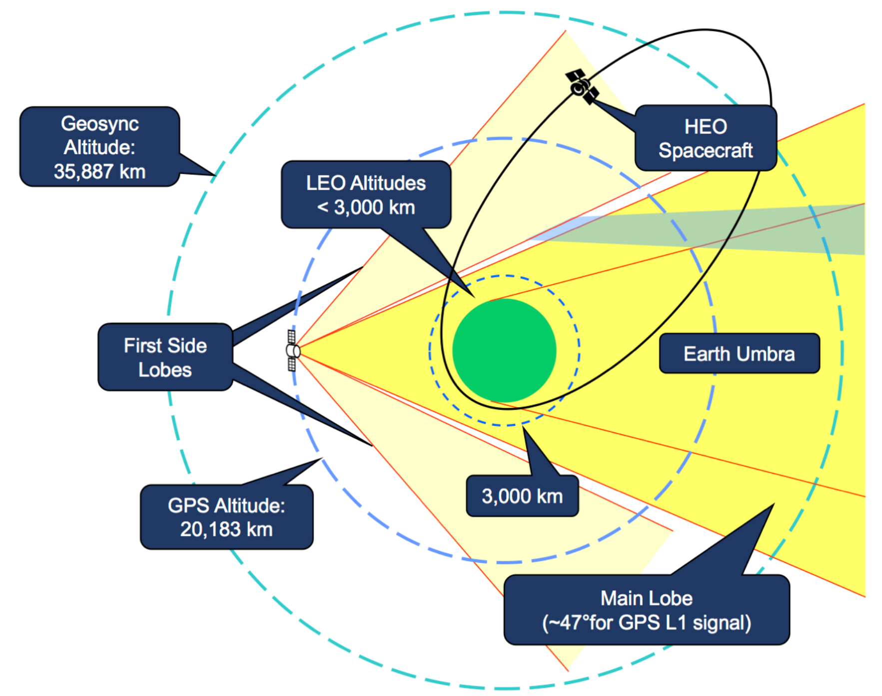
  NAVIC – provides service to India region, expected to be fully operational now, 7 satellites

  QZSS – provides service to Japan region, preliminary service available now, 4 satellites

- International collaboration is facilitated through the International Committee on GNSS (ICG) and other forums with the objective of inter-operability among the different constellations

Space applications: space service volume

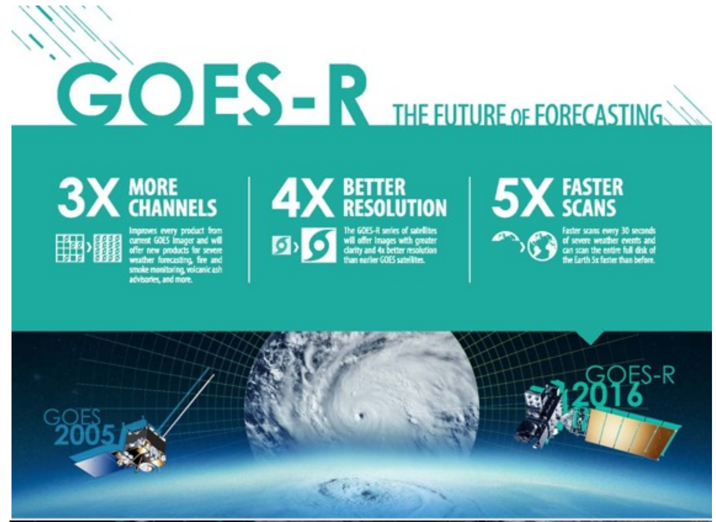
- The Space Service Volume (SSV) is defined as the volume of space surrounding the Earth from the edge of LEO to GEO, i.e., 3,000 km to 36,000 km altitude
- The SSV overlaps and extends beyond the GNSS constellations, so use of signals in this region often requires signal reception from satellites on the opposite side of the Earth – main lobes and sidelobes
- Signal availability constrained by poor geometry, Earth occultation, and weak signal strength
- Formal altitude limit of GNSS usage in space is 36,000 km, but the practical limit is known to extend well beyond this.

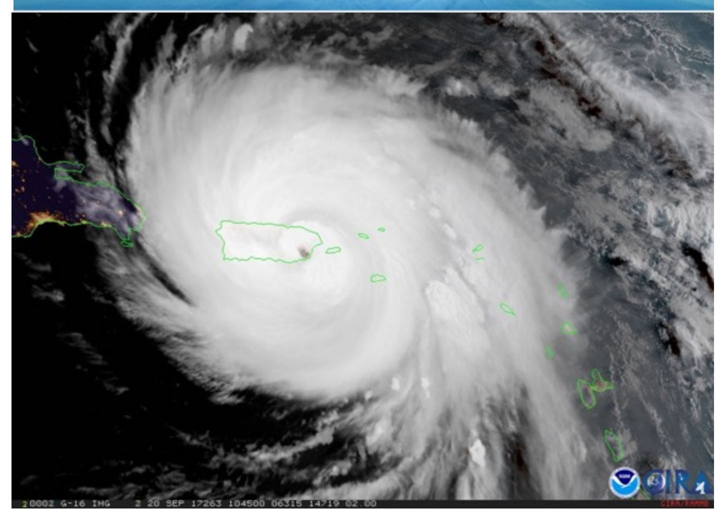


Space Service Volume [17]

Recent flight experiences: GOES-16

- Geostationary Operational Environment Satellite 16 (GOES-16)
- GOES-R, -S, -T, -U: 4th generation NOAA operational weather satellites
- GOES-R/GOES-16 Launch: 19 Nov 2016
- 15 year life, series operational through mid-2030s
- Employs GPS at GEO to meet stringent navigation requirements
- Relies on beyond-spec GPS sidelobe signals

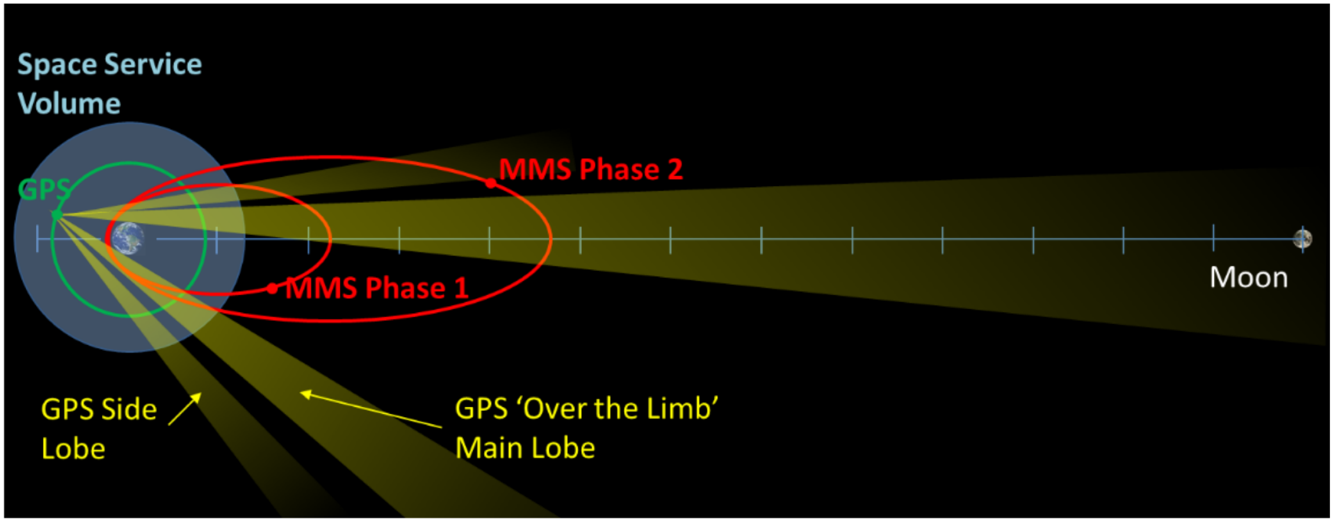




GOES-16 image of Hurricane Maria making landfall over Puerto Rico [18]

Recent flight experiences: MMS

- Magnetospheric Multiscale (MMS) mission
- Launched March 12, 2015
- Four spacecraft form a tetrahedron near apogee for performing magnetospheric science measurements (space weather)
- Four spacecraft in highly eccentric orbits
- Phase 1: 1.2 x 12 Earth Radii (Re) Orbit (7,600 km x 76,000 km)
- Phase 2B: Extends apogee to 25 Re (~150,000 km) (40% of way to Moon)



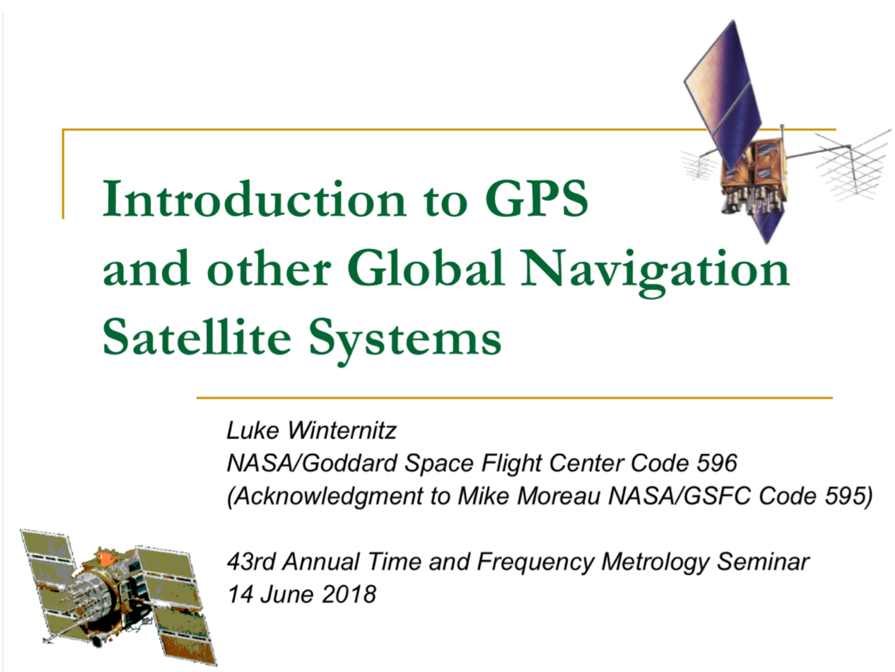
MMS phase 1 and phase 2 [18]

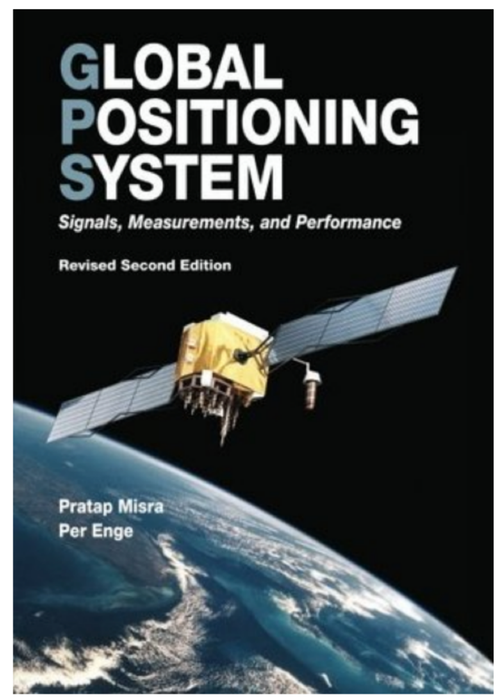
Active areas of research

- Reflectometry: measurement of wind speeds by determining ocean surface roughness from reflected GNSS signals, measurement of soil moisture from reflected GNSS signals
- Radio occultation: detection and characterization of seismic events through analysis of GNSS signals propagating over the horizon through the upper atmosphere
- Autonomous navigation: Enables formation flying, provides robustness to signal outages, significantly reduces ground station tracking and ground-based orbit determination costs
- High-altitude / lunar GNSS: US is planning several human spaceflight missions (Exploration Missions 1-3) in the next few years, as well as a permanent way-station in the vicinity of the moon (a lunar “gateway”). Initial studies have shown GNSS signals to be available and usable at lunar distances; use of GNSS would enable periods of autonomy and provide a stable and accurate timing source for hosted science and technology payloads.

References

- Luke Winternitz, “Introduction to GPS and other Global Navigation Satellite Systems,” in *Proceedings of the 43rd Annual Time and Frequency Metrology Seminar*, Boulder, CO, 14 June 2018.
- James Garrison, *AAE575: Introduction to Satellite Navigation and Positioning*, Purdue University, Fall 2011.
- Pratap Misra and Per Enge, *Global Positioning System*, 2nd ed. Lincoln, MA: Ganga-Jamuna Press, 2005.





Benjamin Ashman, Ph.D.

Navigation and Mission Design Branch

Code 595, NASA Goddard Space Flight Center

Greenbelt, MD 20771

benjamin.w.ashman@nasa.gov

Image references

1. M. Pratap and P. Enge, Global Positioning System, 2nd ed. Lincoln, MA: Ganga-Jamuna Press, 2005.
2. https://mk0spaceflightnoa02a.kinstacdn.com/wp-content/uploads/2014/12/GPS_IIF.jpg
3. Luke Winternitz, “Introduction to GPS and other Global Navigation Satellite Systems,” in Proceedings of the 43rd Annual Time and Frequency Metrology Seminar, Boulder, CO, 14 June 2018.
4. https://www.gps.gov/multimedia/images/constellation.jpg
5. https://www.gps.gov/multimedia/images/GPS-control-segment-map.pdf
6. http://www8.garmin.com/aboutGPS/
7. http://www.elenageosystems.com/GNSS.aspx
8. https://natronics.github.io/blag/2014/gps-viz-1/
9. https://smd-prod.s3.amazonaws.com/science-blue/s3fs-public/thumbnails/image/EMS-Introduction.jpeg
10. B. Ashman, “Incorporation of GNSS Multipath to Improve Autonomous Rendezvous, Docking, and Proximity Operations,” Ph.D. dissertation, Purdue University, 2016.
11. M. Moreau, “GPS Receiver Architecture for Autonomous Navigation in High Earth Orbits,” Ph.D. dissertation, University of Colorado, 2001.
12. http://insidegnss.com/wp-content/uploads/2018/01/IGM_janfeb12-Solutions.pdf
13. https://link.springer.com/chapter/10.1007/978-3-319-42928-1_6
14. http://www.navipedia.net/index.php/Transformations_between_Time_Systems
15. https://youtu.be/DbYapFLJsPA
16. https://en.wikipedia.org/wiki/Orbital_elements#/media/File:Orbit1.svg
17. B. Ashman, J. Parker, F. Bauer, M. Esswein, “Exploring the Limits of High Altitude GPS for Lunar Missions,” AAS GN&C Conference, Breckenridge, CO, American Astronautical Society, February 2018.
18. J. Miller and J. Parker, “NASA GNSS Activities,” International Committee on GNSS 12, Kyoto, Japan, December 2017.
19. G. McGraw, P. Groves, and B. Ashman, "Robust Positioning in the Presence of Multipath and NLOS GNSS Signals," Chapter 21 in 21st Century PNT, Jade Morton editor, 2018.
20. National Geographic March 2018
21. https://timeandnavigation.si.edu/navigating-at-sea/navigating-without-a-clock/celestial-navigation

Optimality of correlation with respect to SNR

An optimal decision statistic is one that maximizes the signal to noise ratio (SNR). This means that for an input signal $x(t) = s(t) + n(t)$, with signal component $s(t)$ and AWGN $n(t)$, a linear time invariant (LTI) process, $h(t)$, is needed that maximizes the SNR of the output $z(t) = z_s(t) + z_n(t)$. In terms of the input and $h(t)$, the output is

$$z(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\,d\tau + \int_{-\infty}^{\infty} h(\tau)n(t-\tau)\,d\tau,$$

and the SNR of the output sampled at time T_I is

$$SNR \equiv \frac{|z_s(T_I)|^2}{E\left[|z_n(T_I)|^2\right]},$$

so the problem is to find

$$h_M(t) = \arg\max_{h(t)} \frac{|z_s(T_I)|^2}{E\left[|z_n(T_I)|^2\right]}.$$

By Parseval's theorem,

$$|z_s(T_I)|^2 = \left|\int_{-\infty}^{\infty} h(\tau)s(T_I-\tau)\right|^2 d\tau = \left|\int_{-\infty}^{\infty} H(f)S(f)e^{i2\pi T_I f}\,df\right|^2$$

and

$$E[|z_n(T_I)|^2] = R_{z_n,z_n}(0) = \int_{-\infty}^{\infty} S_{z_n,z_n}(f)\,df = \int_{-\infty}^{\infty} |H(f)|^2 S_{nn}(f)\,df.$$

Optimality of correlation (cont.)

For white noise, $S(f) = N_0/2$, so

$$SNR = \frac{\left| \int_{-\infty}^{\infty} H(f) S(f) e^{i2\pi T_I f}\, df \right|^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2\, df}.$$

By the Schwartz inequality, for any two square-integrable functions f and g

$$\left| \int f(t) g(t)\, dt \right|^2 \leq \int |f(t)|^2\, dt \cdot \int |g(t)|^2\, dt,$$

with equality if and only if $f(t) = \lambda g^*(t)$, where λ is a scalar and the asterisk indicates complex conjugate. Signal to noise ratio is therefore bounded:

$$SNR \leq \frac{\int_{-\infty}^{\infty} |H(f)|^2\, df \int_{-\infty}^{\infty} |S(f)|^2\, df}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2\, df} = \frac{2}{N_0} \int_{-\infty}^{\infty} |S(f)|^2\, df.$$

Equality holds when $H(f) = \lambda S^*(f) e^{-2\pi f T_I}$. This corresponds to

$$h(t) = \mathcal{F}^{-1}\left\{ H(f) \right\} = \lambda s(T_I - t).$$

Thus, SNR is maximized when $h(t)$ is a scaled, flipped, time-delayed copy of the input signal.

- See *Introduction to Digital Communications* by Michael B. Pursley