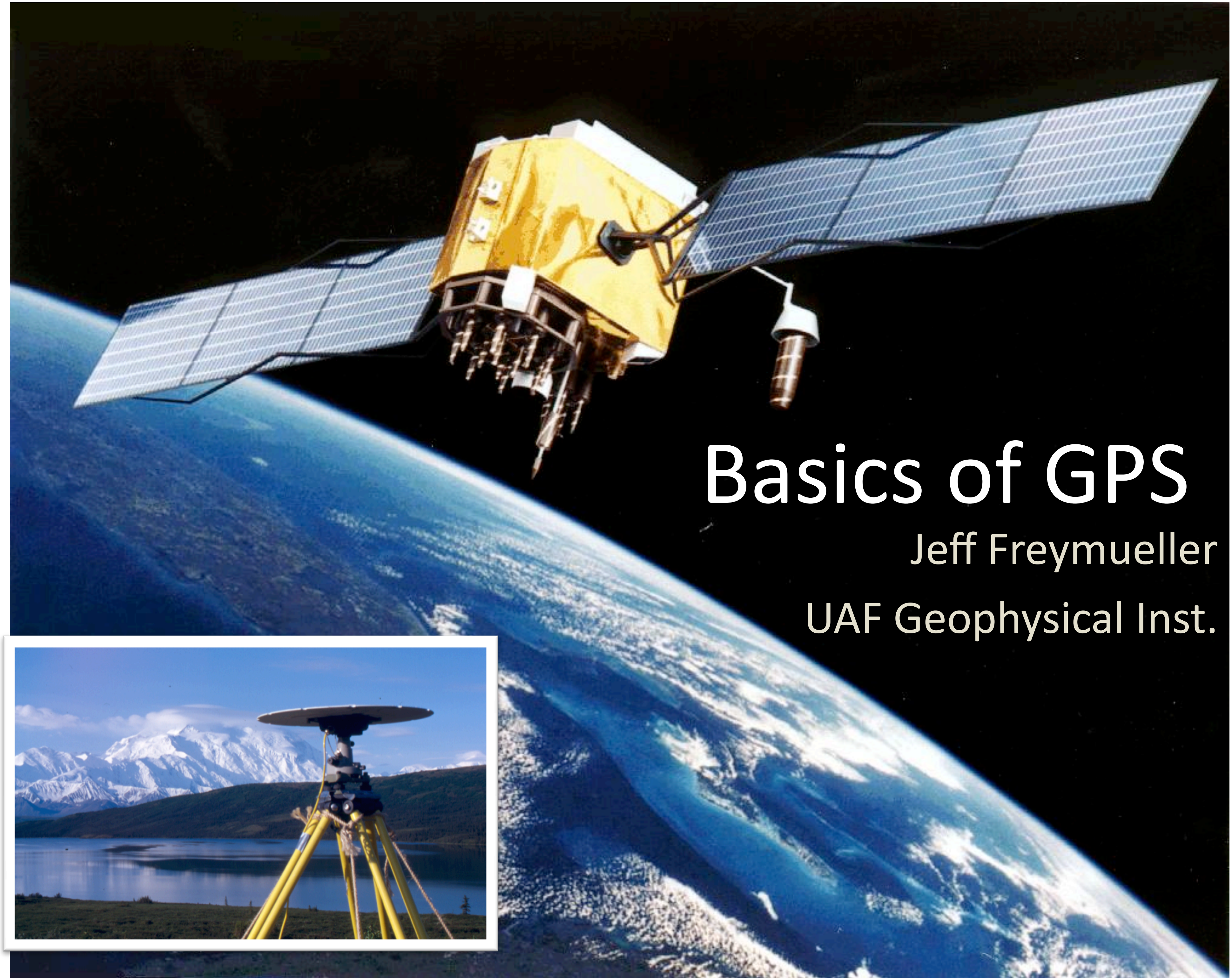
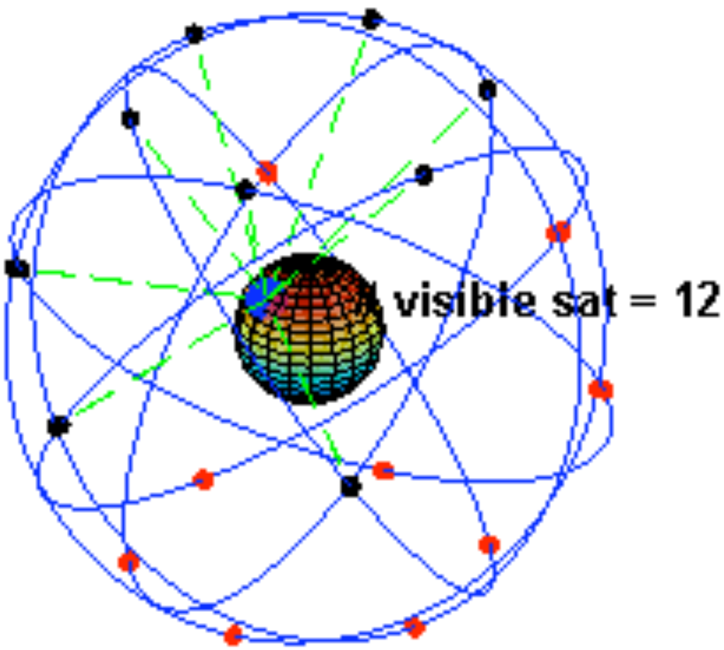


Basics of GPS

Jeff Freymueller

UAF Geophysical Inst.

GPS Design Timeline



- NAVSTAR = NAVigation System with Timing and Ranging
- Always-on, instant global positioning
- Development began in 1973
- First satellites launched 1978
- User equip tests 1980
- 1983: KLM 007 shot down by Soviet Union
  - Reagan mandates future civilian use of GPS

GPS User Hardware – Old!

GPS User Hardware – Modern!

Different Modes of Use

Navigation

- Instantaneous
- Single station
- Original intended use
- Accuracy
  - Few meters
  - Sub-meter w/differential corrections

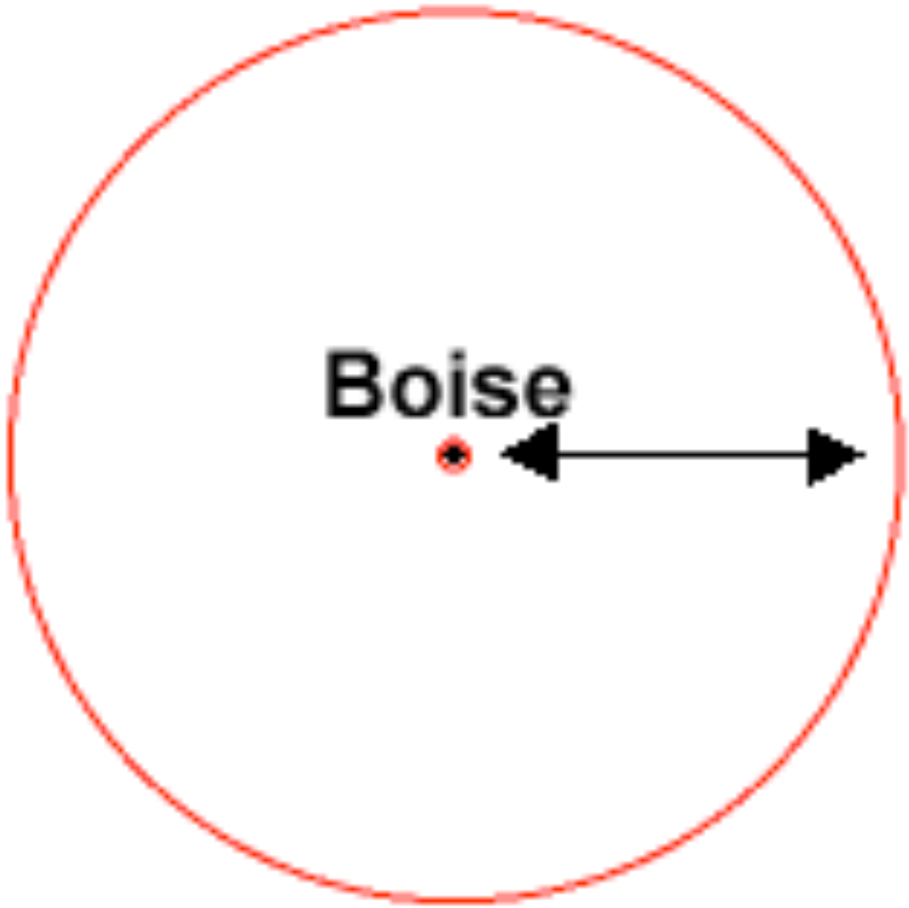
Surveying

- Usually post-process
- Usually multi-station
- Science or survey
- Accuracy
  - 1-2 cm at worst
  - 1-2 mm at best
- Also “seismology”

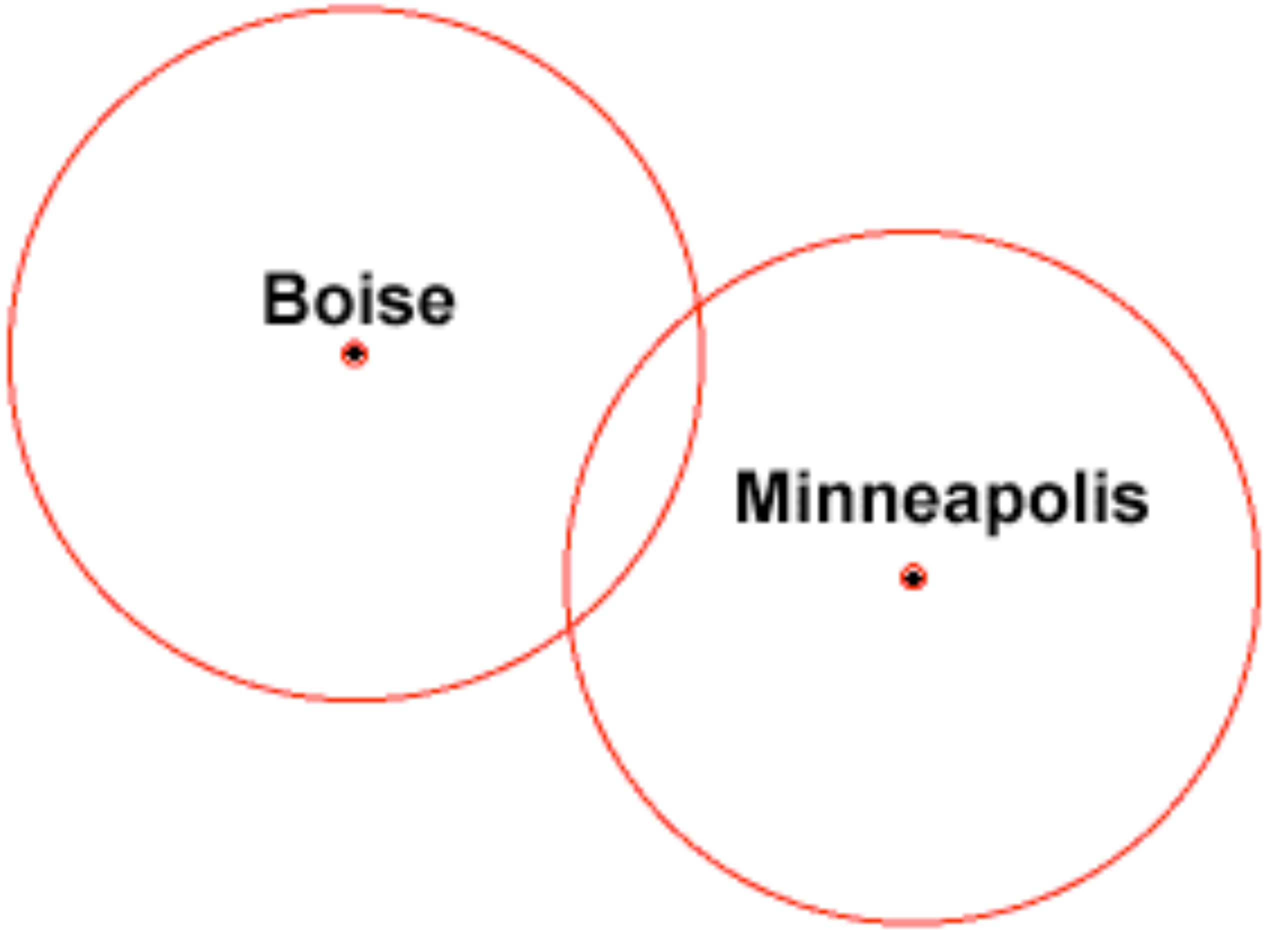
Basic Principles: Surveying

- Requires data from $n \geq 4$ satellites, $m \geq 2$ receivers
- Requires continuous tracking over time
- Post-processed using data from all receivers
- Use pseudorange and carrier phase measurements from each satellite to receiver
- Orbits of satellites fixed or estimated
- Clock error on satellites estimated
- Estimate receiver position (X,Y,Z) and clock error
- Model a wide variety of path delays and other effects

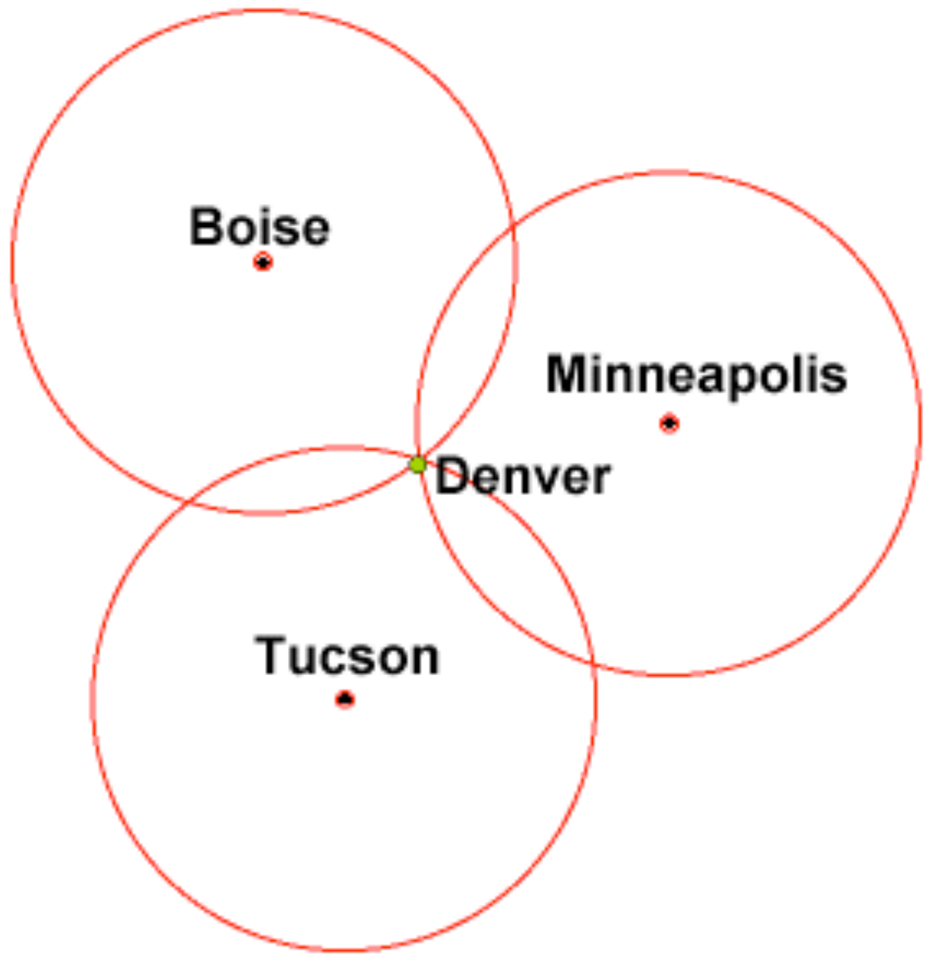
Positioning By Ranging 1



Positioning By Ranging 2



Positioning By Ranging 3



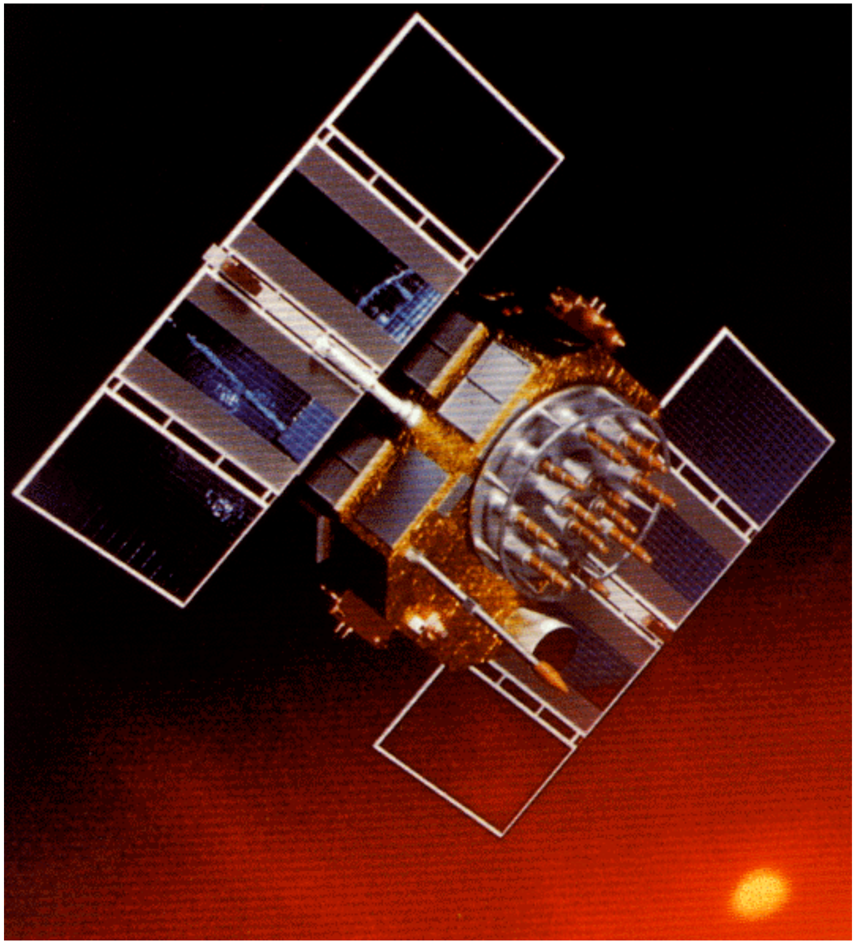
GPS Positioning

- Measure position by measuring ranges to satellites
  - A few satellites can serve an unlimited number of users on the ground, anywhere in the world
- How do we know where satellites are?
  - They broadcast their positions (orbits) in a ***navigation message***
  - (or) someone gives us precise orbits back in the lab
- Measured ranges are called ***pseudoranges***

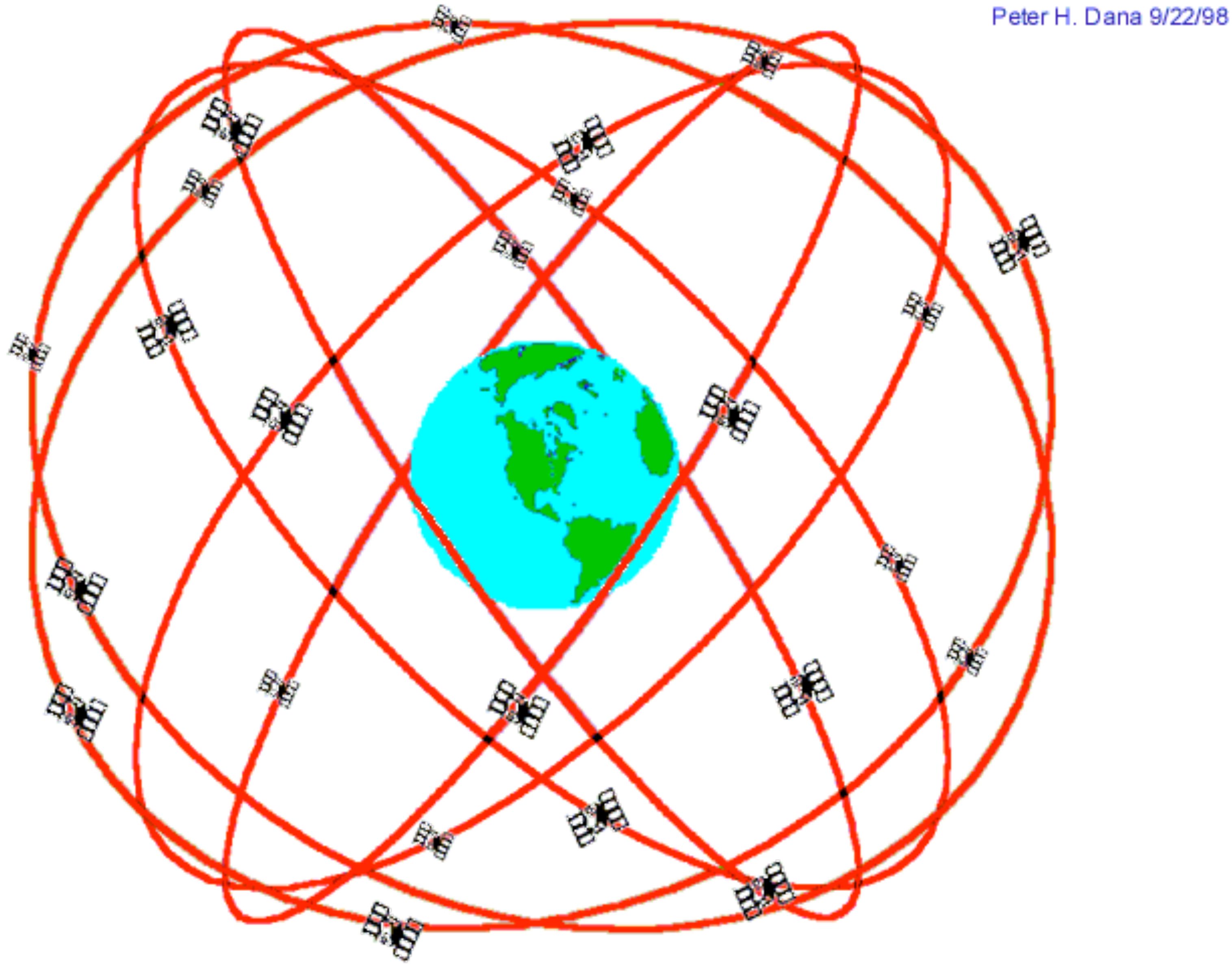
Why call it a “pseudorange”?

- Range is the distance from satellite to receiver, plus path delays.
- Pseudorange is distance plus effects of clock errors
- Geometric range ρ is true distance.
- $P = \rho + c^*(\text{clock errors}) + c^*(\text{path delays})$

Evolving Satellites



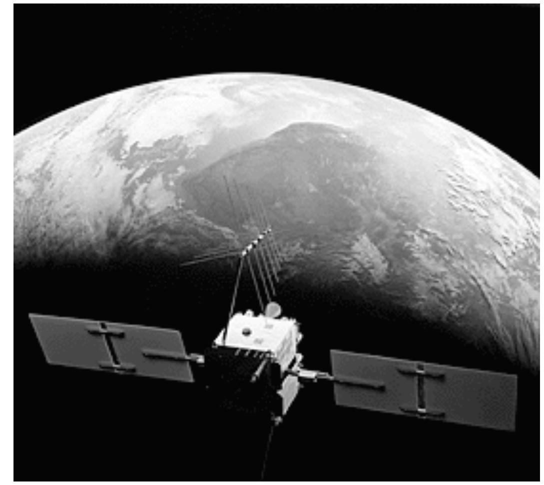
Satellite Constellation



Satellite Constellation Facts

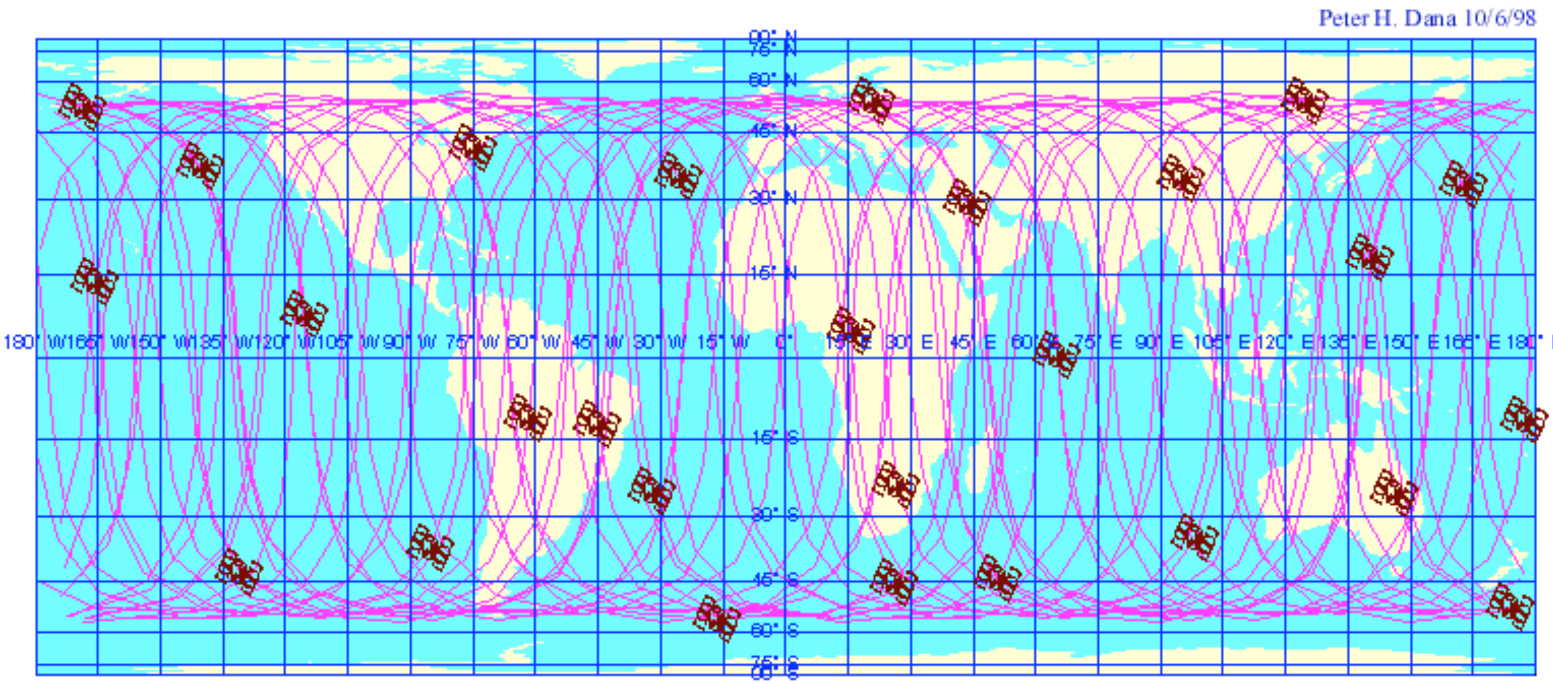
- Nominally 4 satellites (SVs) in each of 6 equally spaced orbital planes (now 5 in each plane).
- Orbital planes inclined 55° from equator.
- Nearly circular orbits R = 26,600 km ~ $4R_E$
- Orbital period is 11h 58m, two orbits per sidereal day
- Sidereal day is length of day defined by when stars appear in same place in sky
  - Differs from rotational day because of motion of earth around the sun.

Orbits



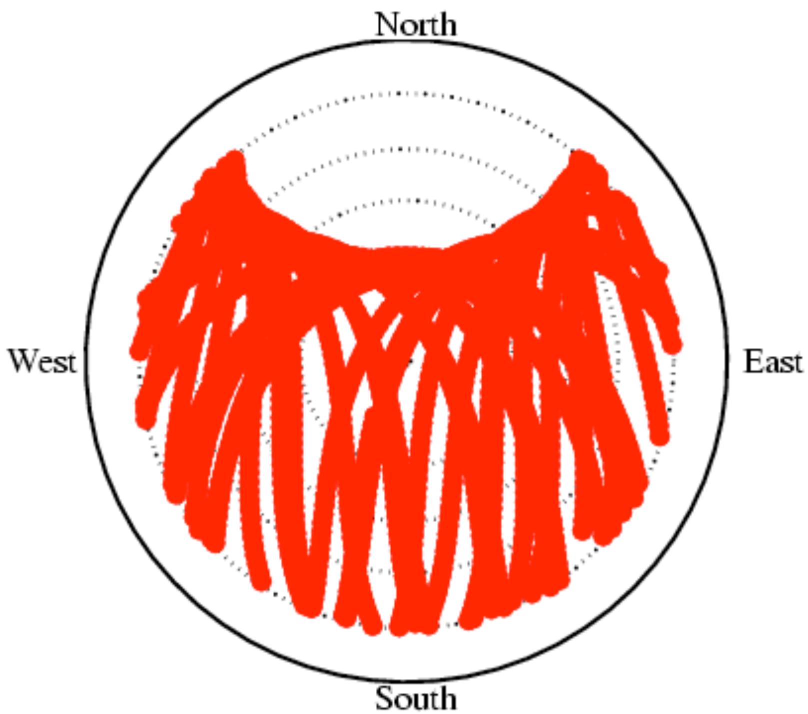
- Can estimate orbits or fix orbits to pre-determined values
- Representation of orbit
  - Broadcast: Keplerian elements
  - Tabular file of XYZ satellite positions
  - Trajectory: initial conditions + integrate equations of motion (needed to estimate orbits)
- In practice, highly precise orbits are available from the IGS
  - Ultra-Rapid: includes predict-ahead for real time use
  - Rapid: Available next day
  - Final: Available in <2 weeks

Satellite Ground Tracks

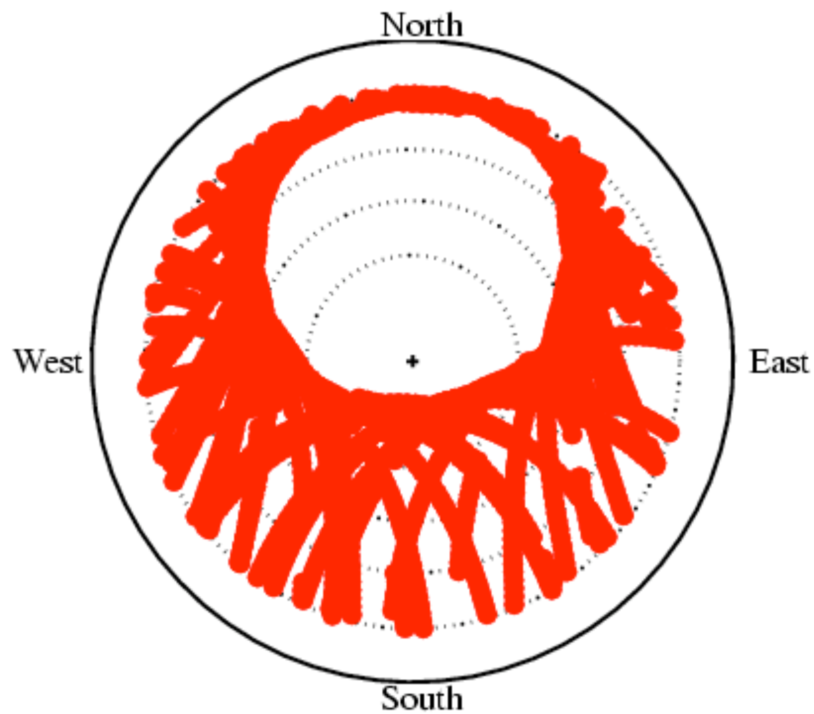


Global Positioning System Satellites and Orbits
for 27 Operational Satellites on September 29, 1998
Satellite Positions at 00:00:00 9/29/98 with 24 hours (2 orbits) of Ground Tracks to 00:00:00 9/30/98

24 hours of GPS Data



Southern California

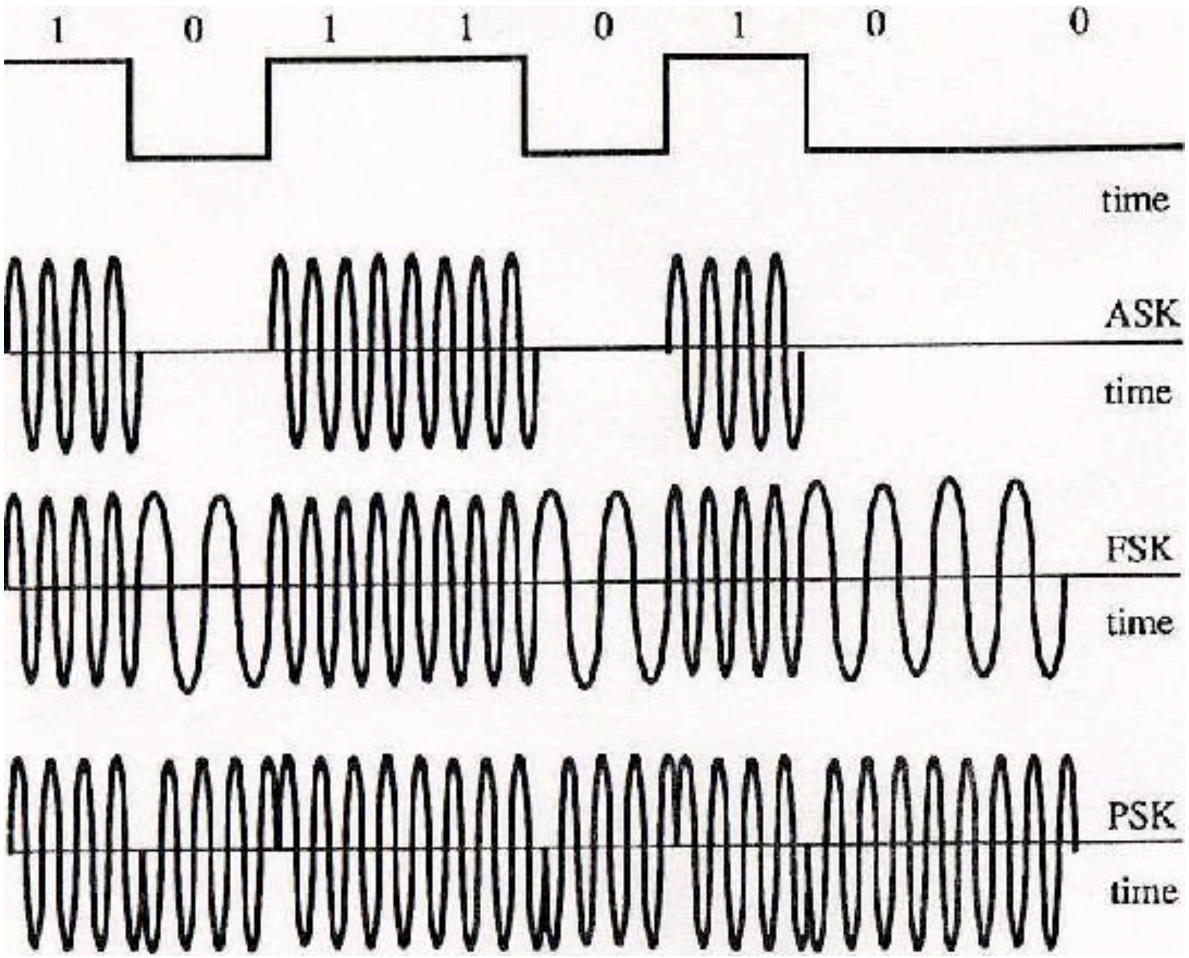


Fairbanks

GPS Signal Structure

- Two frequencies at L-band, L1 and L2
  - L1 at 154*10.23 MHz (~19 cm)
  - L2 at 120*10.23 MHz (~24 cm)
- Codes Modulated (phase modulation) onto each carrier
  - P-code at 10.23 MHz on L1 + L2
  - C/A (Coarse Acquisition) code at 1.023 MHz on L1 + L2 (new L2C)
  - Navigation message at 50 bits per second
- P and C/A codes are types of ***pseudo-random noise (PRN)*** codes

Types of signal modulation

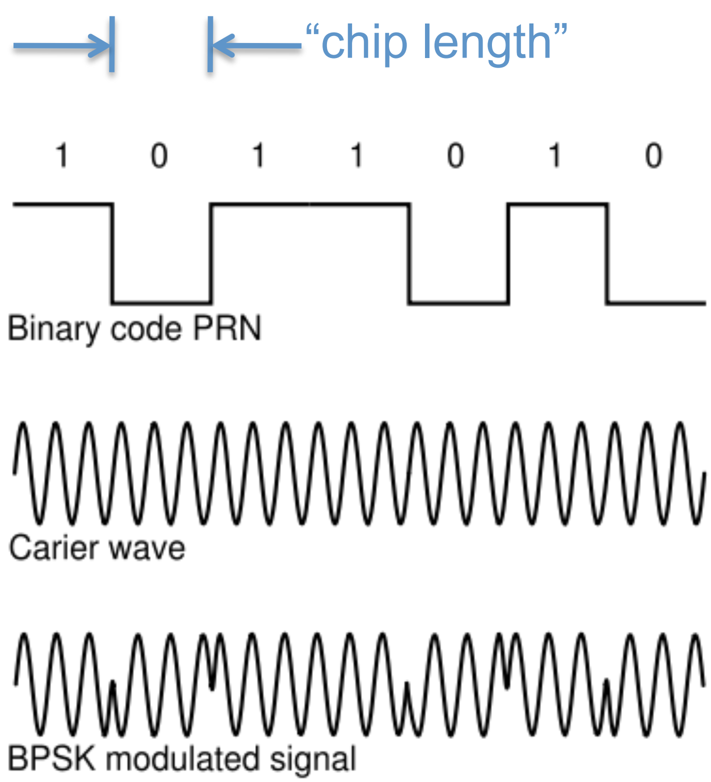


Amplitude Modulation (AM)

Frequency Modulation (FM)

Phase Modulation (PM)

Precision of Observations

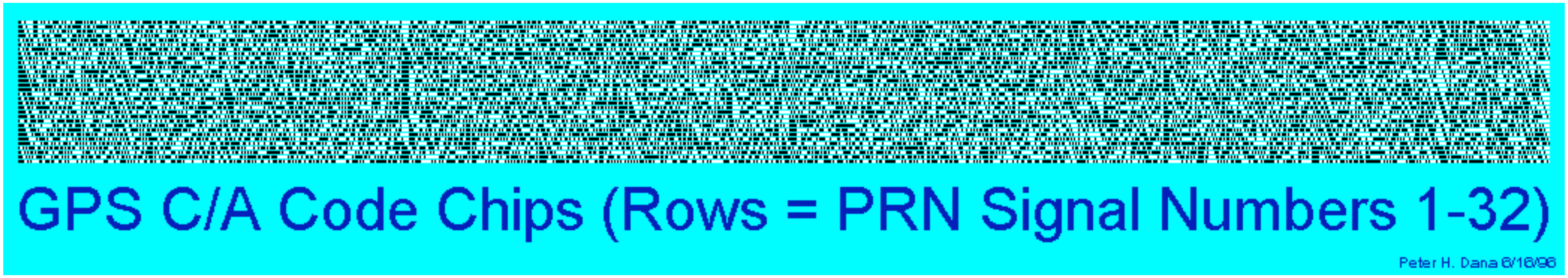


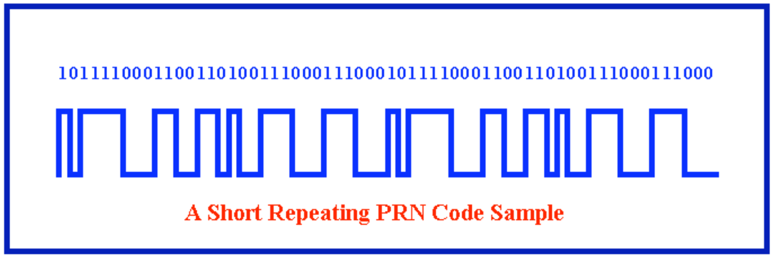
- Code “chip length” is the distance associated with each bit of the code.
  - C/A: 293 m
    - Repeats every ~300 km
  - P: 29.3 m
- Carrier wavelength is analogous to chip length == 2-3 orders of magnitude more precise

Pseudo-Random Noise

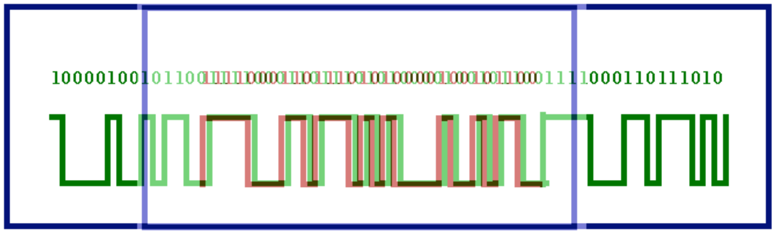
- Computers cannot generate true random numbers, but can generate a sequence of numbers with random statistical properties.
  - But the sequence can be repeated exactly
  - Begin with some starting value, then perform a series of operations
- C/A code has 1023 bits, repeats 1000 times per second
- P code has a lot of bits, repeats every 266.4 days; each SV gets a 7-day piece of code

Code Correlation for Ranging

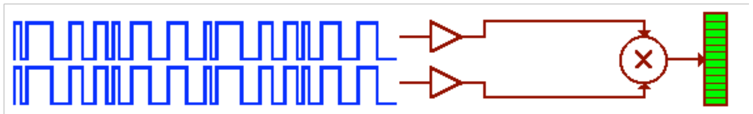


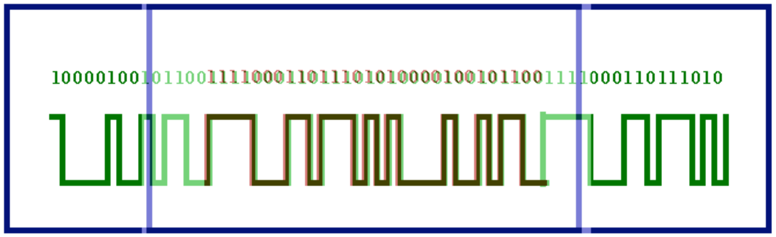


A Short Repeating PRN Code Sample



Partial Correlation of Identical Receiver and Satellite PRN Codes





Full Correlation (Code-Phase Lock) of Receiver and Satellite PRN Codes

Pseudorange Observation Model

- The correlation procedure produces a time shift, which is the fundamental pseudorange measurement.
  - Travel time = (time of reception) – (time of transmission)
- $P^S = (T - T^S)c$
  - T = receiver clock reading at reception
  - T^S = satellite clock reading at transmission
  - c = speed of light = 299792458 m/s

Observation Model with Clocks

- $P^S = (T - T^S)c$
  - $T = t + \tau$ $\quad |\tau| \leq 1$ millisecond
  - $T^S = t^S + \tau^S$ $\quad |\tau^S|$ is small (Cesium or Rubidium clocks)
  - t, t^S are true receive, transmit times
- Substituting
  - $P^S = [(t + \tau) - (t^S + \tau^S)]c$
  - $P^S = (t - t^S)c + (\tau - \tau^S)c$
  - $P^S = \rho^S(t,t^S) + (\tau - \tau^S)c$
- $\rho^S(t,t^S)$ is range from receiver at receive time to satellite at transmit time:

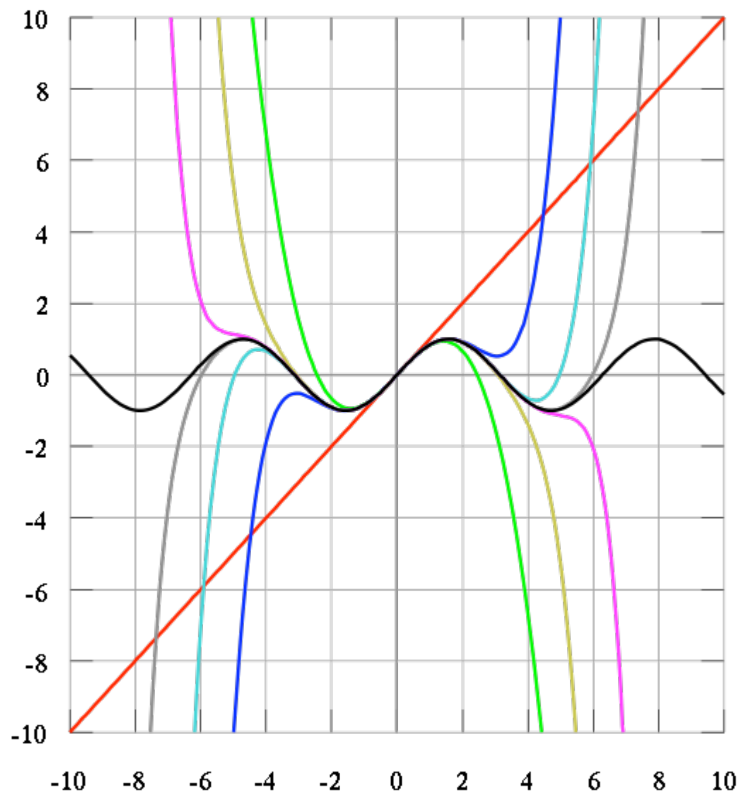
$$\rho^S(t,t^S) = \sqrt{\left(x^S(t^S) - x(t)\right)^2 + \left(y^S(t^S) - y(t)\right)^2 + \left(z^S(t^S) - z(t)\right)^2}$$

Set of Simplified Observ. Equations

- Now, generalize to multiple satellites. We use a superscript to identify each satellite (don't confuse with an exponent). Later we will have to use a subscript to keep track of multiple receivers:
  - $P^{(1)} = [\,(x^{(1)} - x)^2 + (y^{(1)} - y)^2 + (z^{(1)} - z)^2\,]^{1/2} + c\tau - c\tau^{(1)}$
  - $P^{(2)} = [\,(x^{(2)} - x)^2 + (y^{(2)} - y)^2 + (z^{(2)} - z)^2\,]^{1/2} + c\tau - c\tau^{(2)}$
  - $P^{(3)} = [\,(x^{(3)} - x)^2 + (y^{(3)} - y)^2 + (z^{(3)} - z)^2\,]^{1/2} + c\tau - c\tau^{(3)}$
  - $P^{(4)} = [\,(x^{(4)} - x)^2 + (y^{(3)} - y)^2 + (z^{(3)} - z)^2\,]^{1/2} + c\tau - c\tau^{(4)}$

Linearizing Nonlinear Equations

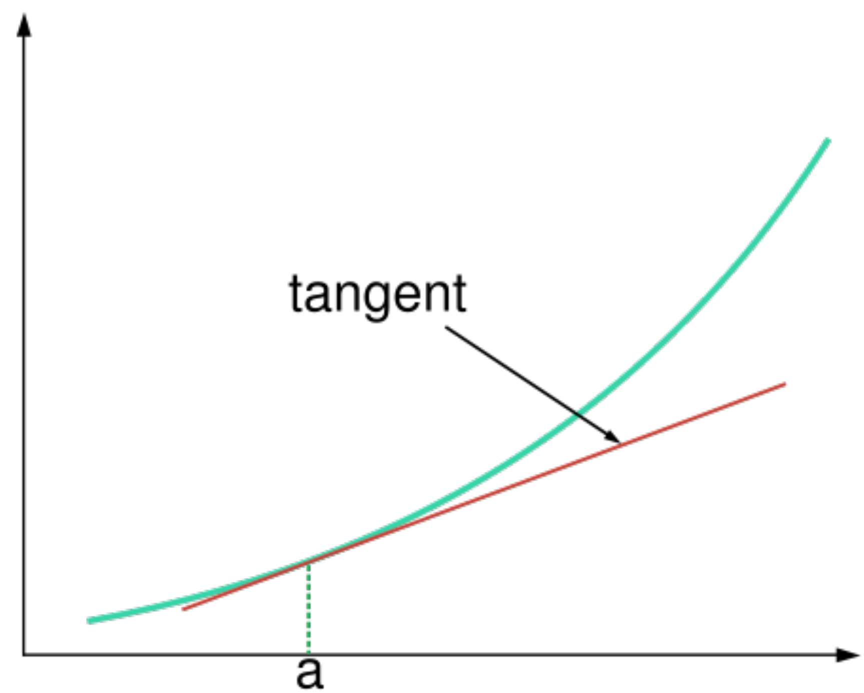
- There are simple ways to solve systems of linear equations, like matrix inversion or least squares. But we have a nonlinear problem. One approach is to linearize, or construct a linear approximation to the non-linear problem. We can do that with Taylor's theorem (Taylor Series)



$$f(x) = \boxed{f(a) + \frac{f'(a)}{1!}(x-a)} + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots,$$

Linearizing Part 2

- For problems with multiple variables, there is a simple extension of the Taylor Series.
- We linearize about approximate values (a,b)
- Partial derivatives are computed at (a,b)



$$f(x,y) \approx f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b).$$

Linearizing Our Equations

- We linearize our equations about approximate values (x_0, y_0, z_0, τ_0)

$$P(x,y,z,\tau) = P(x_0,y_0,z_0,\tau_0) + \frac{\partial P}{\partial x}\left(x - x_0\right) + \frac{\partial P}{\partial y}\left(y - y_0\right) + \frac{\partial P}{\partial z}\left(z - z_0\right) + \frac{\partial P}{\partial \tau}\left(\tau - \tau_0\right)$$

$$P(x,y,z,\tau) = P(x_0,y_0,z_0,\tau_0) + \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau$$

$$P(x,y,z,\tau) - P(x_0,y_0,z_0,\tau_0) = \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau$$

$$P_{observed} - P_{computed} = \frac{\partial P}{\partial x}\Delta x + \frac{\partial P}{\partial y}\Delta y + \frac{\partial P}{\partial z}\Delta z + \frac{\partial P}{\partial \tau}\Delta \tau$$

Matrix Equation

- It is easier to deal with this equation if we write it as a matrix equation:

$$\begin{pmatrix} \Delta P^{(1)} \\ \Delta P^{(2)} \\ \Delta P^{(3)} \\ \Delta P^{(4)} \end{pmatrix} = \begin{pmatrix} \frac{\partial P^{(1)}}{\partial x} & \frac{\partial P^{(1)}}{\partial y} & \frac{\partial P^{(1)}}{\partial z} & \frac{\partial P^{(1)}}{\partial \tau} \\ \frac{\partial P^{(2)}}{\partial x} & \frac{\partial P^{(2)}}{\partial y} & \frac{\partial P^{(2)}}{\partial z} & \frac{\partial P^{(2)}}{\partial \tau} \\ \frac{\partial P^{(3)}}{\partial x} & \frac{\partial P^{(3)}}{\partial y} & \frac{\partial P^{(3)}}{\partial z} & \frac{\partial P^{(3)}}{\partial \tau} \\ \frac{\partial P^{(4)}}{\partial x} & \frac{\partial P^{(4)}}{\partial y} & \frac{\partial P^{(4)}}{\partial z} & \frac{\partial P^{(4)}}{\partial \tau} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{pmatrix} + \begin{pmatrix} v^{(1)} \\ v^{(2)} \\ v^{(3)} \\ v^{(4)} \end{pmatrix}$$

Evaluate the Partial Derivatives

- This is often written in matrix form like
  - **b** = **Ax** + **v**   **A** is called the "Design matrix"
  - If $\rho^{(i)} = [\,(x_0 - x^{(i)})^2 + (y_0 - y^{(i)})^2 + (z_0 - z^{(i)})^2\,]^{1/2}$

$$A = \begin{pmatrix} \dfrac{x_0 - x^{(1)}}{\rho^{(1)}} & \dfrac{y_0 - y^{(1)}}{\rho^{(1)}} & \dfrac{z_0 - z^{(1)}}{\rho^{(1)}} & c \\ \dfrac{x_0 - x^{(2)}}{\rho^{(2)}} & \dfrac{y_0 - y^{(2)}}{\rho^{(2)}} & \dfrac{z_0 - z^{(2)}}{\rho^{(2)}} & c \\ \dfrac{x_0 - x^{(3)}}{\rho^{(3)}} & \dfrac{y_0 - y^{(3)}}{\rho^{(3)}} & \dfrac{z_0 - z^{(3)}}{\rho^{(3)}} & c \\ \dfrac{x_0 - x^{(4)}}{\rho^{(4)}} & \dfrac{y_0 - y^{(4)}}{\rho^{(4)}} & \dfrac{z_0 - z^{(4)}}{\rho^{(4)}} & c \end{pmatrix}$$

These have the form of trig functions, and can also be written in terms of the azimuth to the satellite and the inclination of the satellite above the horizon.

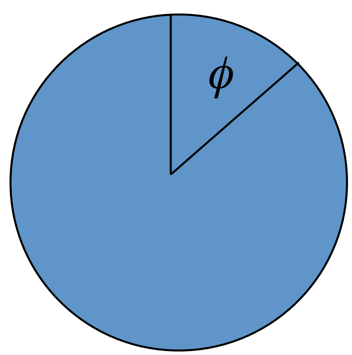
DOPs – Dilution of Precision

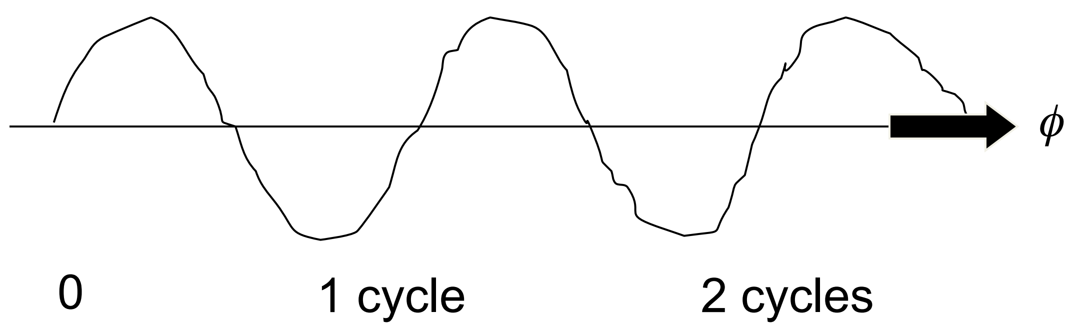
- Your handheld GPS probably reports a number called "PDOP", which stands for "Position Dilution of Precision". These are other DOPs as well, which all give measures of how the satellite geometry maps into position or time precision.
  - VDOP = σ_h
  - HDOP = $(\sigma_e^2 + \sigma_n^2)^{1/2}$
  - PDOP = $(\sigma_e^2 + \sigma_n^2 + \sigma_h^2)^{1/2}$
  - GDOP = $(\sigma_e^2 + \sigma_n^2 + \sigma_h^2 + c^2\sigma_\tau^2)^{1/2}$
  - TDOP = σ_τ
- Multiply PDOP by measurement precision to get uncertainty in 3D position.

PDOP > 5 considered poor

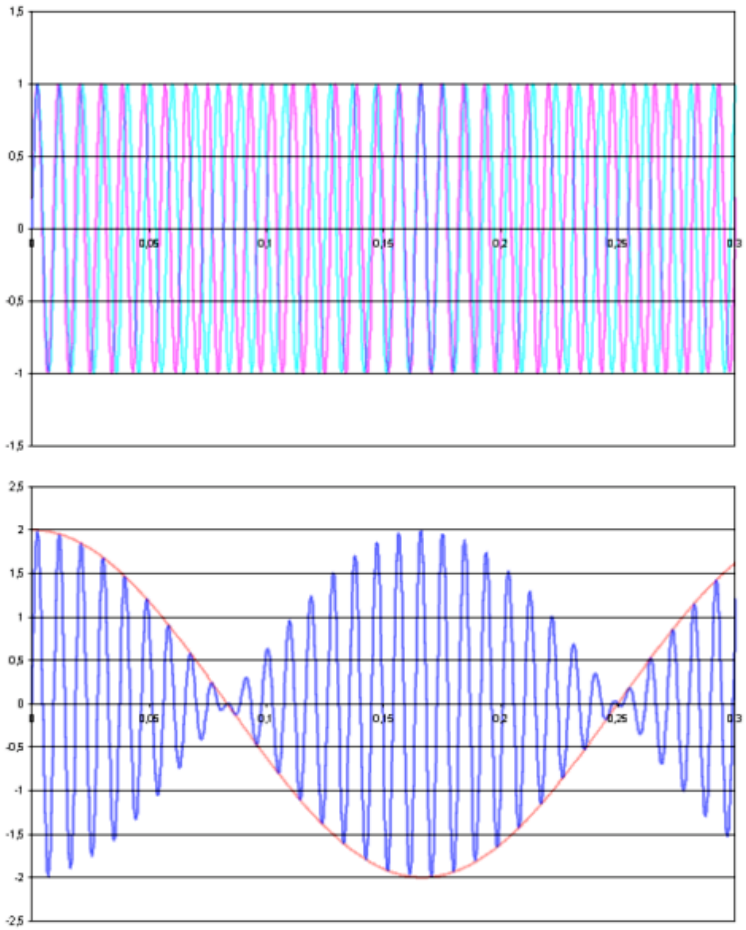
Phase Tracking

- Receiver measures changes in phase of carrier signal over time
  - First must remove codes to recover raw phase
  - Then track continuous phase, keeping record of the number of whole cycles
  - Phase has an integer ambiguity (initial value)
- Problems occur if receiver loses phase lock





Measurement Trick: Beat Phase



- Remove PRN code modulation by multiplying signal by code – removes phase shifts and recovers original carrier signal
- Mix received phase with reference phase signal
- Filter high frequency beats and measure phase of low frequency beat

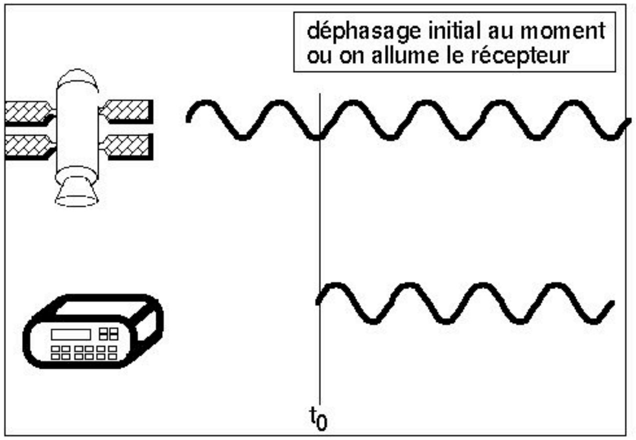
$$\sin(2\pi f_1 t) + \sin(2\pi f_2 t) = 2\cos\left(2\pi \frac{f_1 - f_2}{2} t\right) \sin\left(2\pi \frac{f_1 + f_2}{2} t\right)$$

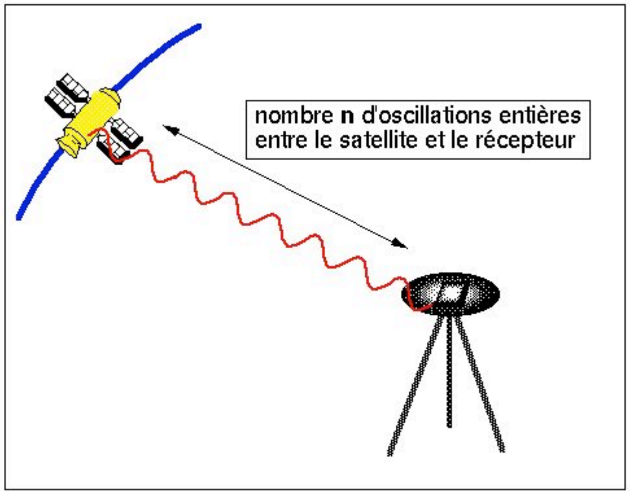
Measuring Beat Phase

- Doppler Shift shifts each SVs frequency slightly
- Receiver generates reference signal at nominal GPS frequency
- Beat phase and beat frequency are
  - $\phi_B(t) = \phi_R(t) - \phi_G(t)$
  - $f_B = f_R - f_G$
- Beat frequency is much lower than nominal, easier to measure beat phase, but we can recover all variations in phase of the transmitted carrier signal from the beat phase

Phase Ambiguity

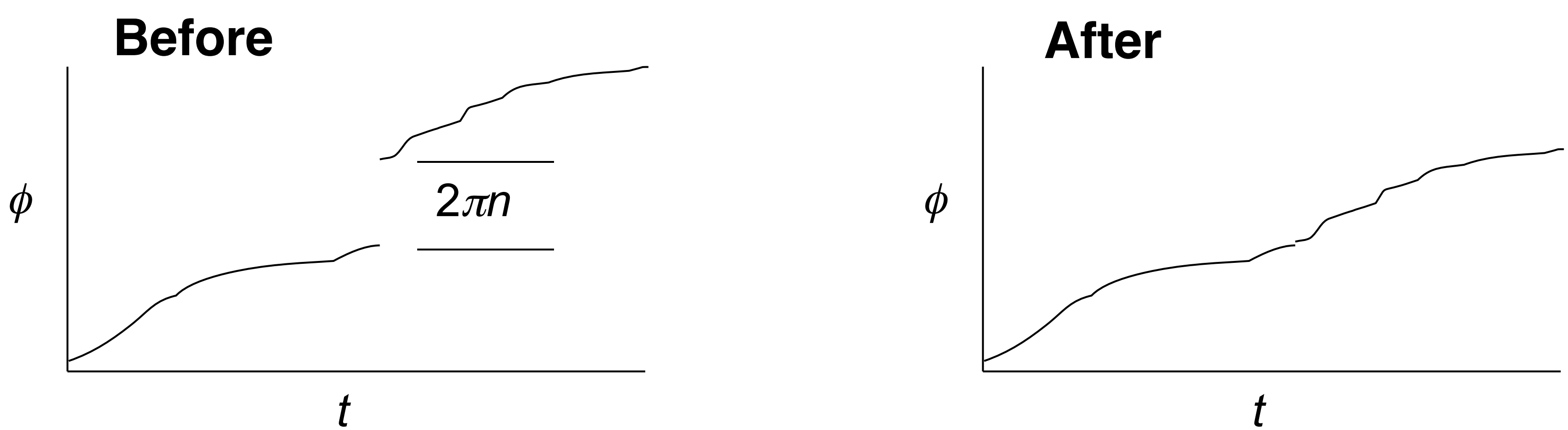
- One drawback of the beat phase is that we can add an arbitrary constant number of cycles to the transmitted carrier signal, and we would get exactly the same beat phase:
  - $\Phi + N = \phi_R - \phi_G$
  - Actual recorded phase is Φ
- Also, we must track the phase continuously. If we lose track of the phase over time, and start over, we get a different N.
  - Losing track is called a "cycle slip"





Loss of Lock (Cycle Slips)

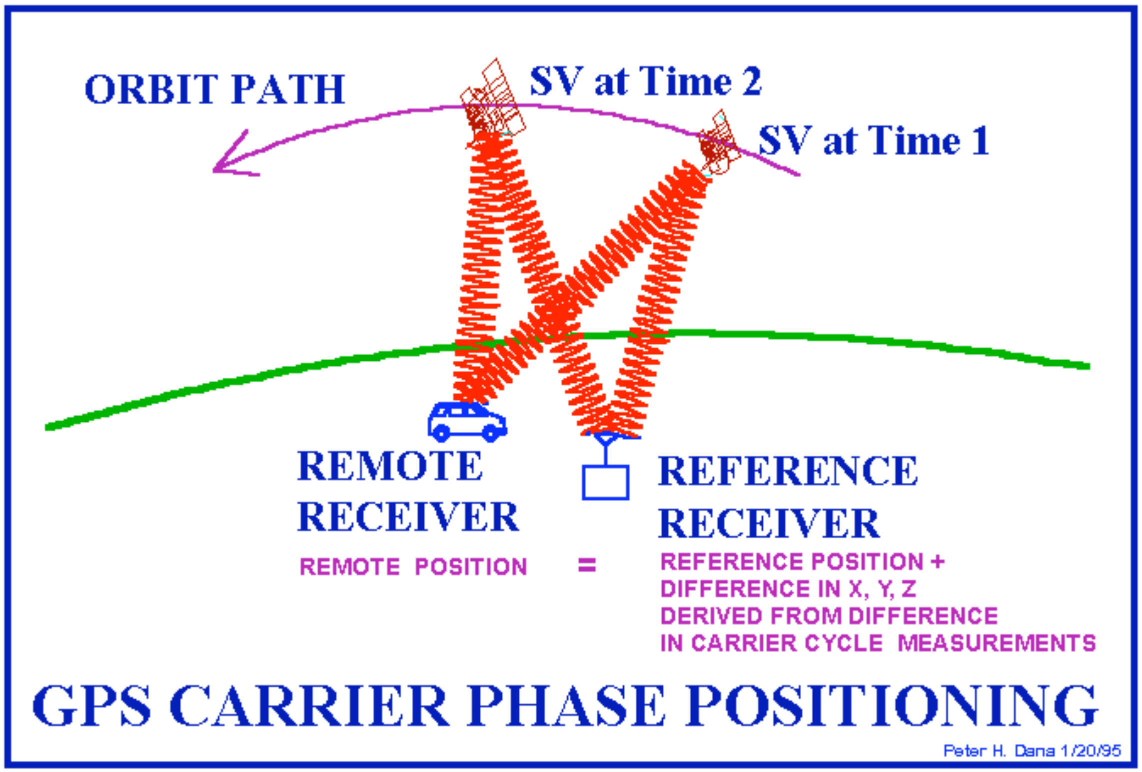
- If receiver loses phase lock, there will be a jump of an integer number of cycles in the phase data
- This must be detected and repaired by the analysis software
- Slightly different procedures are usually applied multiple times to find all of the cycle slips



Observation Equations

- Compare the phase observation model with the pseudorange model:
  - $L_A^j(T_A) = c(T_A - T^j) + B_a^j$
  - $P_A^j(T_A) = c(T_A - T^j)$
  - Exactly the same except for the phase bias!
- We need to add more terms to deal with the clock errors and with path delay terms
  - $L_A^j(T_A) = \rho_A^j(t_A, t^j) + c\tau_A - c\tau^j + Z_a^j - I_A^j + B_A^j$
  - $P_A^j(T_A) = \rho_A^j(t_A, t^j) + c\tau_A - c\tau^j + Z_a^j + I_A^j$
  - Path delay terms are Z for the troposphere, and I for ionosphere. We'll come back to these later on.

Differencing Techniques

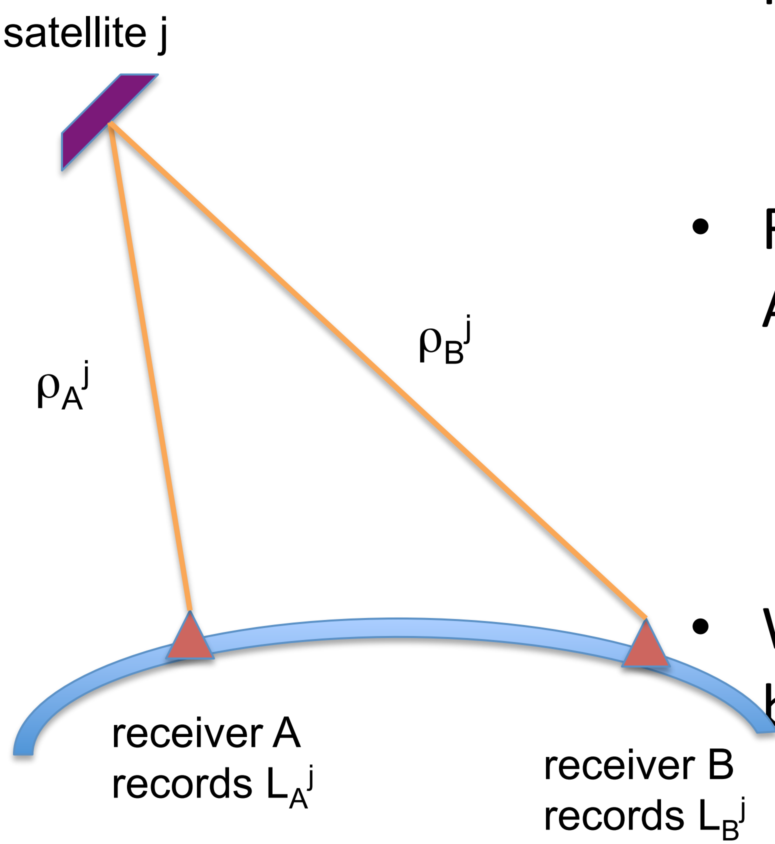


- Receiver and satellite clock biases can be removed by differencing data from mulitple satellites and/or receivers.
  - Difference between receivers ("single difference") removes satellite clock
  - Difference between satellites ("single difference") removes receiver clock
  - Difference of differences ("double difference") removes both clocks

Advantages/Disadvantages of Differencing

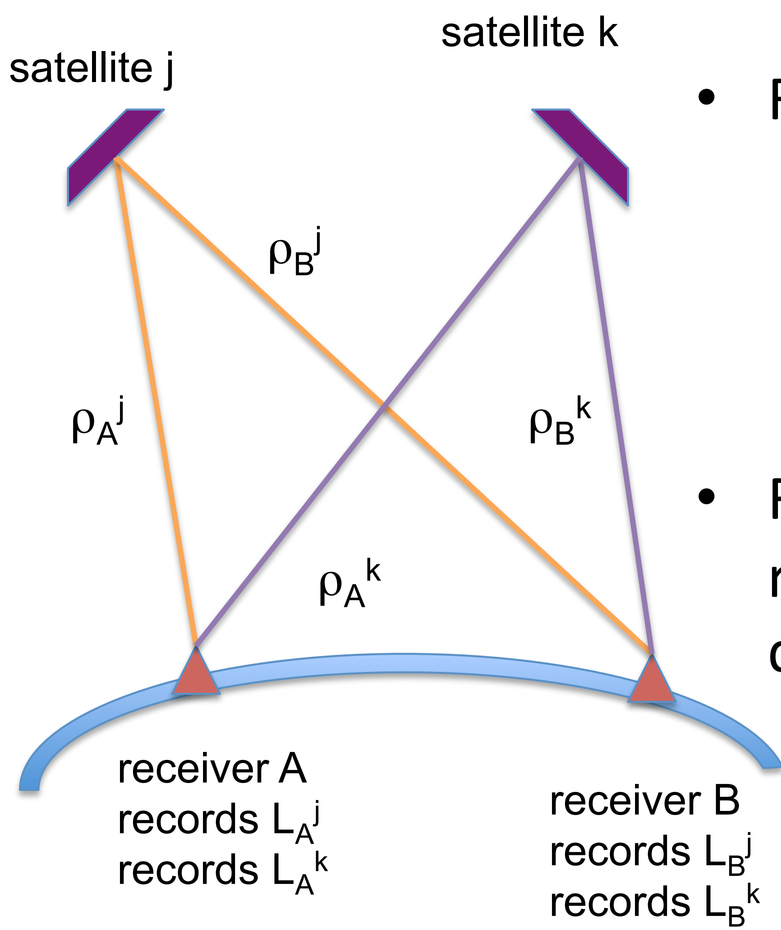
- Advantage
  - Removes clock errors, which are a pain
  - Makes for faster estimation (fewer parameters if you do not need to estimate clock error at every observation time)
  - Phase bias parameters reduce to integer values
- Disadvantages
  - Requires a method to select differences
  - Requires additional bookkeeping
  - Notation gets messy

Single Difference



- Receivers A and B observe:
  - $L_A^j = \rho_A^j + c\tau_A - c\tau^j + B_A^j$
  - $L_B^j = \rho_B^j + c\tau_B - c\tau^j + B_B^j$
- Form a difference between receivers A and B
  - $\Delta L_{AB}^j = L_A^j - L_B^j$
  - $\Delta L_{AB}^j = (\rho_A^j - \rho_B^j) + (c\tau_A - c\tau_B) + (B_A^j - B_B^j)$
  - $\Delta L_{AB}^j = \Delta\rho_{AB}^j + c\Delta\tau_{AB} + \Delta B_{AB}^j$
- We use Δ to indicate a difference between ground receivers.

Double Difference

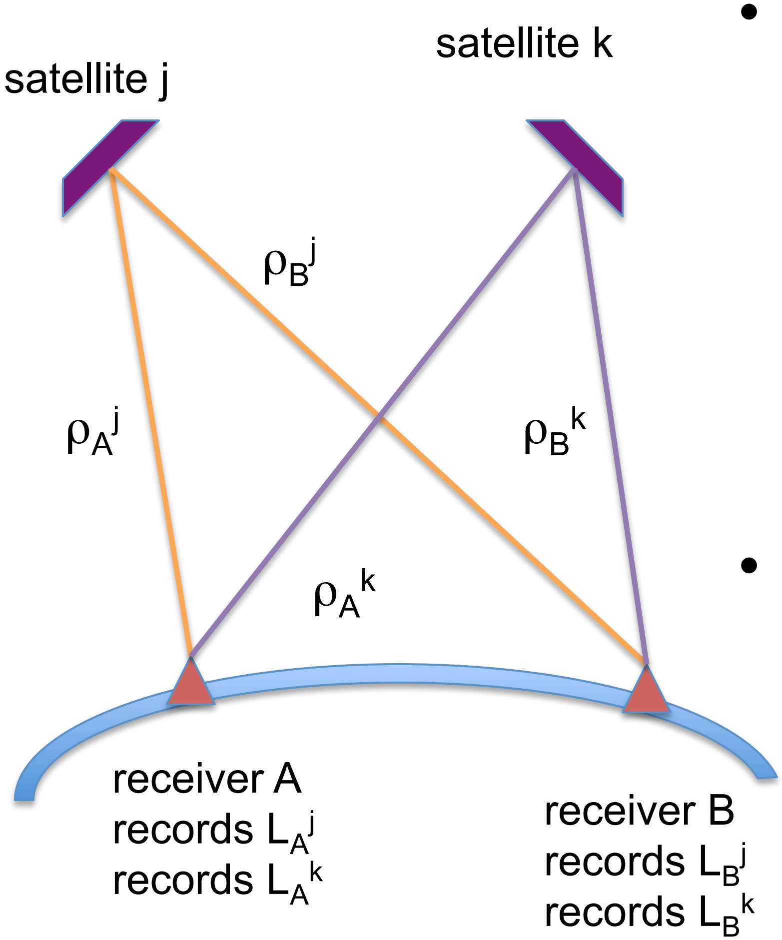


- Receivers A and B observe:
  - $L_A^j = \rho_A^j + c\tau_A - c\tau^j + B_A^j$
  - $L_A^k = \rho_A^k + c\tau_A - c\tau^k + B_A^k$
  - $L_B^j = \rho_B^j + c\tau_B - c\tau^j + B_B^j$
  - $L_B^k = \rho_B^k + c\tau_B - c\tau^k + B_B^k$
- Form the single difference between receivers A and B, and then difference between satellites j and k:
  - $\Delta\Delta L_{AB}^{jk} = \Delta L_{AB}^j - \Delta L_{AB}^k$
  - $\Delta\Delta L_{AB}^{jk} = (\Delta\rho_{AB}^j - \Delta\rho_{AB}^k) + (\Delta B_{AB}^j - \Delta B_{AB}^k)$
  - $\Delta\Delta L_{AB}^{jk} = \Delta\Delta\rho_{AB}^{jk} + \Delta\Delta B_{AB}^{jk}$
- We use $\Delta\Delta$ to indicate a double difference.

Double-Differenced Ambiguity

- The double-differenced phase ambiguities become exactly integers:
  - $\Delta\Delta B_{AB}{}^{jk} = \Delta B_{AB}{}^{j} - \Delta B_{AB}{}^{k} = \lambda_0 \Delta N_{AB}{}^{jk}$
- Each B has three parts:
  - $B_A{}^j = \lambda_0 (N_A{}^j + \phi_{0A} - \phi_0{}^j)$
  - The receiver bias ϕ_{0A} is common to all satellites, and differences out like the receiver clock.
  - The satellite bias $\phi_0{}^j$ is common to all receivers, and differences out like the satellite clock.
- There are some clever techniques to remove the phase ambiguity completely, if you can resolve it to the correct integer.

Triple Difference



- The triple difference adds a difference in time. If you difference the double-differenced observations from one epoch in time to those of the previous epoch, you get a triple difference:
  - $\Delta\Delta\Delta L_{AB}^{jk}(i+l, i) = \Delta\Delta\Delta\rho_{AB}^{jk}(i+l, i)$
- We use $\Delta\Delta\Delta$ to indicate a double difference. The triple difference also removes most of the geometric strength from GPS, so it produces only weakly determined positions. But the triple difference could be applied to kinematic problems.

Final Notes on Differencing

- When you difference between receivers, then in effect you are now estimating the baseline vector between the two receivers, rather than the two positions.
- You have to take some care in choosing which differences to use
  - Cannot use linearly dependent observations
  - Must be careful in choosing baselines to difference, satellites to difference between.
- Each software does it differently
- Some softwares do not difference at all, but estimate clock errors instead.

Ionospheric Calibration

- To a very good approximation, the path delay due to ionospheric refraction is proportional to $1/f^2$
- The phase is ***advanced***, while the pseudorange data are ***delayed***
  - Information travels at group velocity
- Specifically, the path delay is $(40.3/f^2)$TEC, where TEC is the total electron content. This path delay can be as large as meters.
  - The delay term $I_a^j = 40.3TEC/f^2$; $f = f_1$ for L1, f_2 for L2

Ionosphere-free Combination

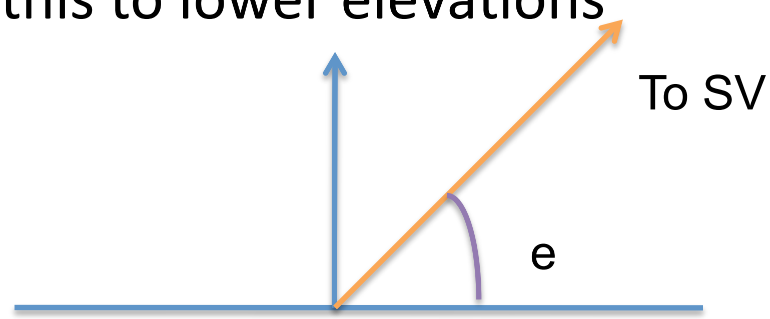
- We can remove the effects of the ionosphere by forming a linear combination of the data at the two frequencies
  - $L_C = -f_1^2/(f_2^2 - f_1^2)L_1 + f_2^2/(f_2^2 - f_1^2)L_2$
  - $P_C = -f_1^2/(f_2^2 - f_1^2)P_1 + f_2^2/(f_2^2 - f_1^2)P_2$
- Try it: For L1 and L2, the biases are $(40.3TEC/f_1^2, 40.3TEC/f_2^2)$
- Note that the two coefficients sum to 1. They have values of approximately (-1.54, 2.54)
- This removes all ionospheric effects except for a $1/f^4$ dependence. There are now "second-order ionosphere" models coming into use, which have an impact on positions at the few mm level or less.

Some Other Important Models

- Tropospheric delay (estimated)
  - Hopfield, Saastamoinen, Lanyi, **<u>Niell</u>**
- Earth tides (well known)
  - Up to ~70 cm amplitude
- Ocean Tidal Loading
  - Response of solid earth to changing load of ocean tides
- Antenna Phase Center variations with elevation
  - Phase center is the point on the antenna that we actually measure distances to
  - It is an imaginary point in space, not a physical point

Troposphere

- Tropospheric path delay affects both frequencies identically. It has two components.
  - "Dry" delay: due to air mass (~ proportional to pressure)
  - "Wet" delay: due to integrated water vapor along path
- Delay in both cases is largest at low elevations above the horizon, because the path length through atmosphere is longer there.
- In practice, we estimate a zenith delay, and use a "mapping function" of elevation angle to map this to lower elevations
  - Mapping function is ~1/sin(i)



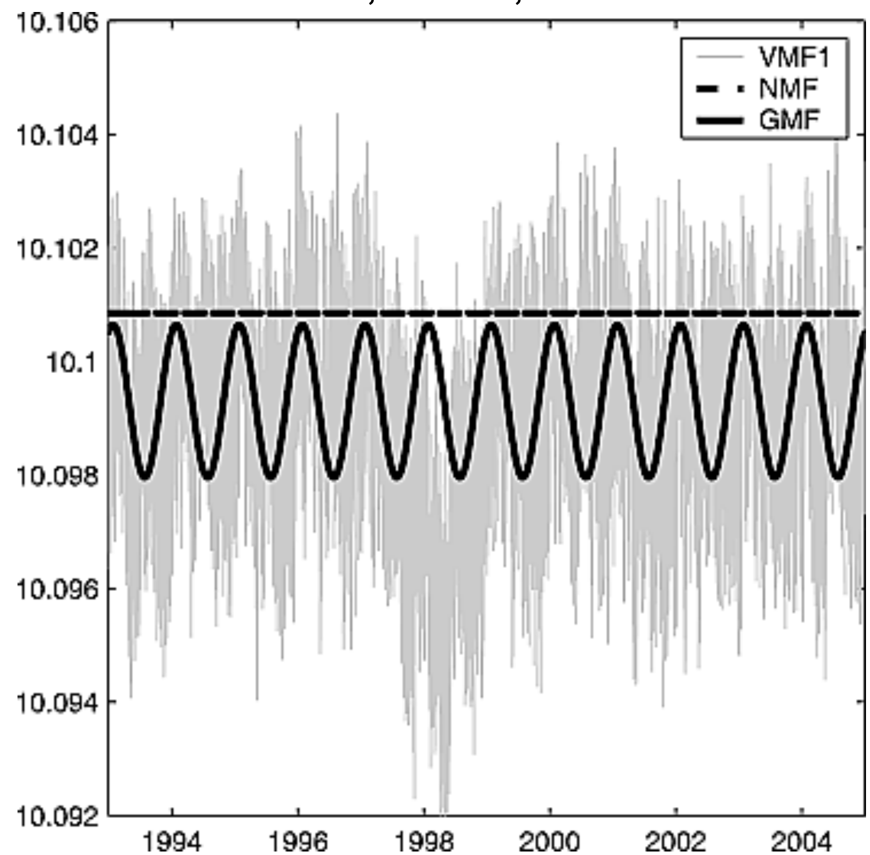
Wet Tropospheric Mapping Function

- Detailed form of mapping function is a continued fraction:

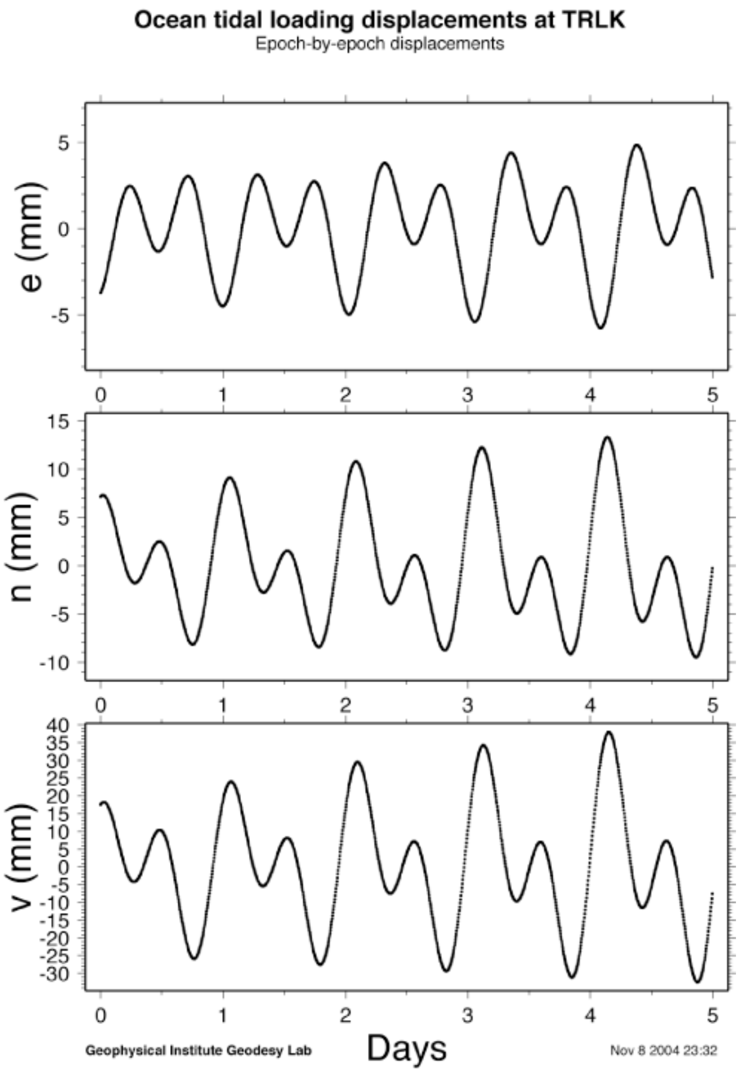
$$\mathrm{mf}(e) = \frac{1 + \dfrac{a}{1 + \dfrac{b}{1 + c}}}{\sin e + \dfrac{a}{\sin e + \dfrac{b}{\sin e + c}}}$$

  - Approx. for layered atmosphere
- Path delay at some elevation angle e is ZTD*mf(e)
- Mapping functions vary with space and time based on distribution of water vapor.

Fortaleza, Brazil, 5° elevation

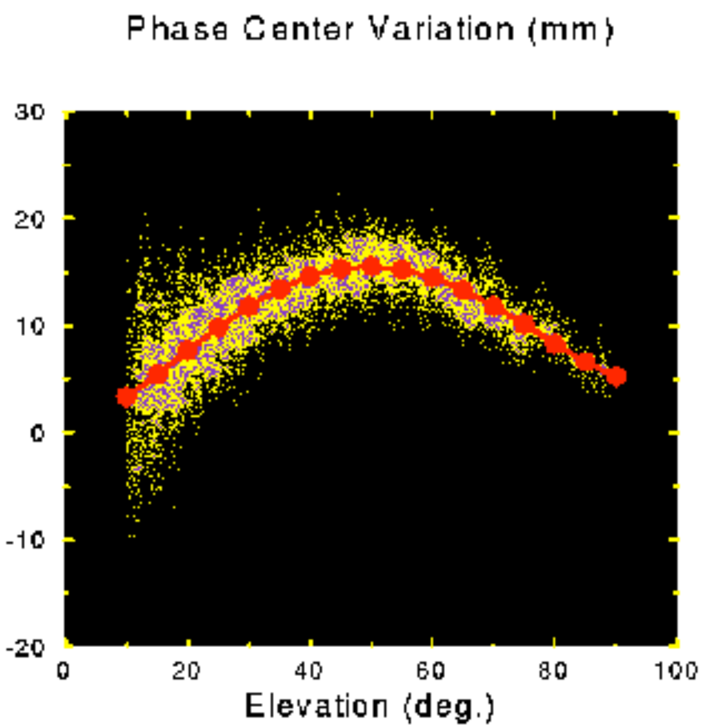


Ocean Tidal Loading



- Solid earth responds to changing load of ocean tides
- Displacements large near coast, where tidal range is large
- Details depend on ocean tides, coastline
- Accurate removal depends on good tidal models

Antenna Phase Center Models

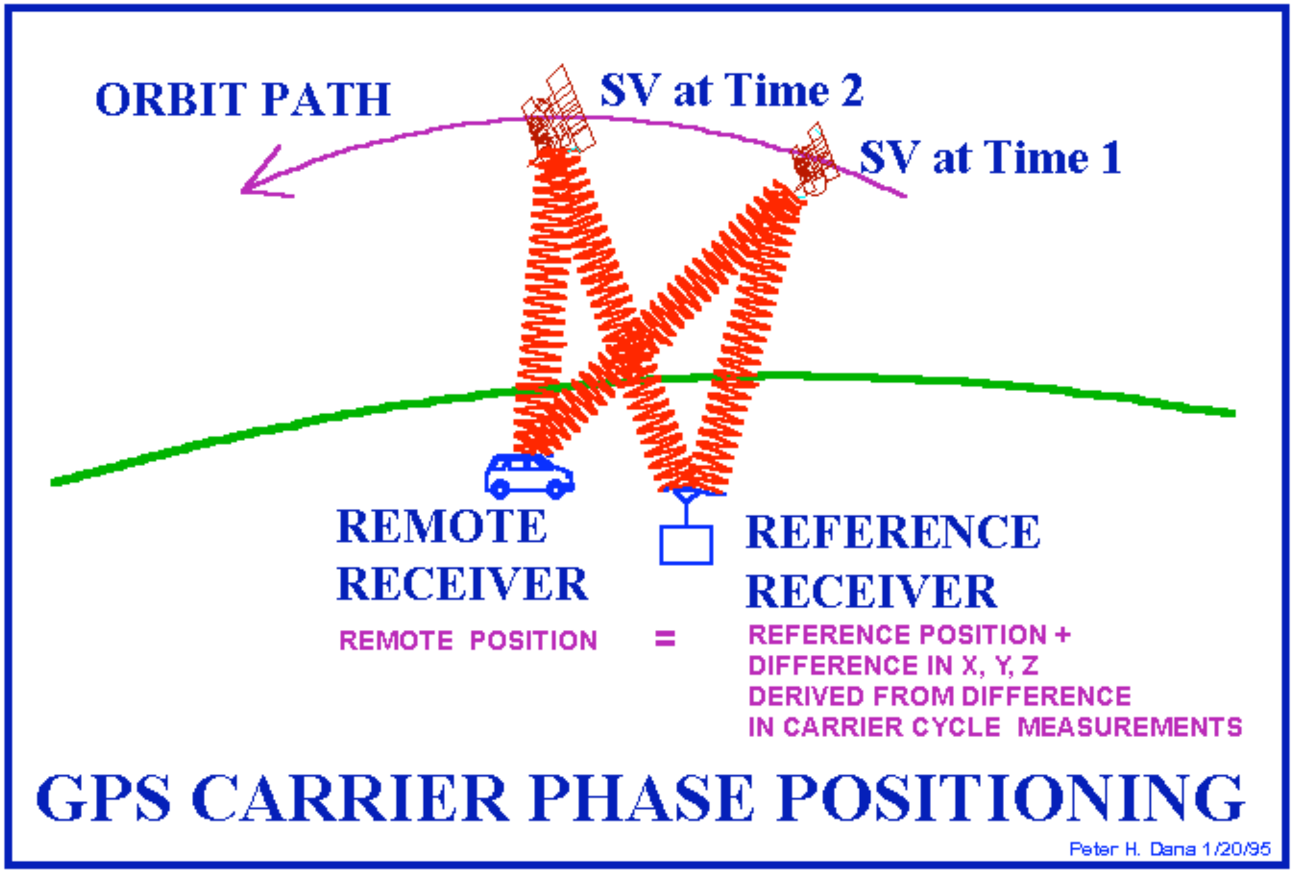


- Ideally, phase center is a point in space.
- Different for every type of antenna
- In reality, the phase center depends on the azimuth and elevation of incoming signal.
- Models assume azimuthal symmetry and fit elevation-dependence

Ambiguity Resolution

- Ambiguity resolution is a trick that can dramatically improve position quality for short surveys or kinematic positioning.
- If you know the ambiguity is an integer, and can determine which integer, then you can fix the ambiguity to that integer value.
  - Removing the ambiguity parameter dramatically improves the strength of the data to be used for determining the position
- We'll talk about this more in the kinematic discussion

Product is Differential Position



Or a set of relative positions (all sites in network relative to each other)