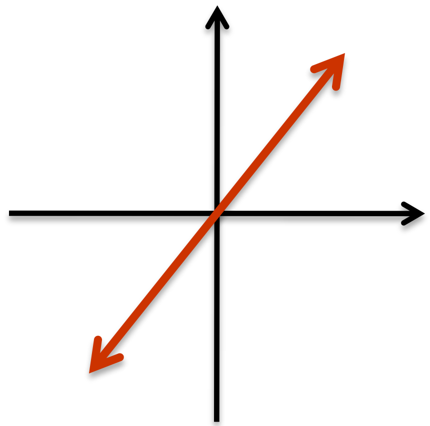
Chapter 4
EM Propagation, GPS Signal Structure, User Equipment

- Review of EM Propagation
- Overview of GNSS Signals
- GPS Signal Structure
- How a GNSS Receiver Works
- Benefits of GNSS Signal Structure (Spread Spectrum)
- Code Multipath
- GNSS User Equipment
- Time Scales

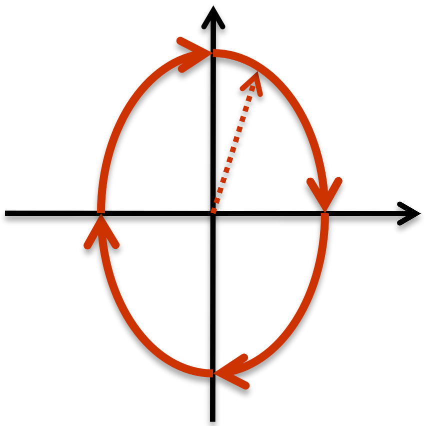
Review of EM Propagation

Wave Polarization

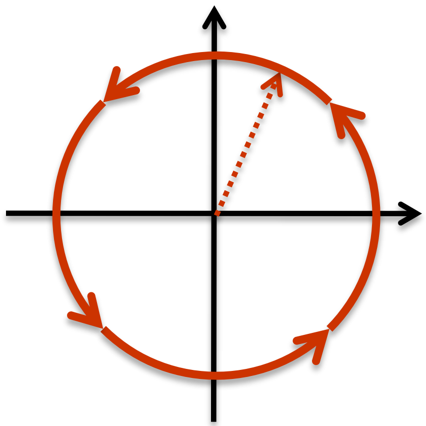
- The polarization of an electromagnetic wave polarization is defined by its E-field plane
- For a wave propagating out of the plane of the page/slide



Linear Polarization



Left-Handed Elliptical Polarization

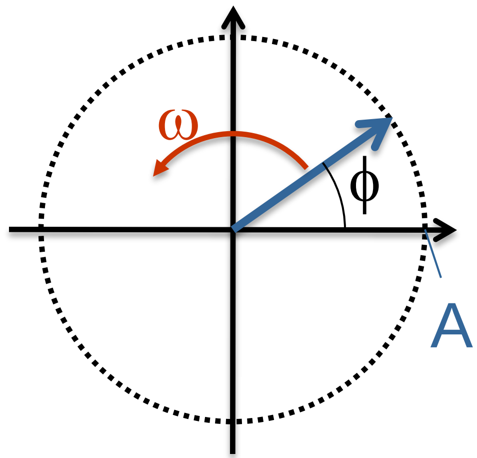


Right-Handed Circular Polarization

- GPS signals are right-hand circularly polarized (RHCP) to deal with the Faraday effect (polarization rotation of an EM wave propagating through ions in the ionosphere)

Phasors

- An EM wave can be interpreted as a phasor



A Amplitude

ω Frequency [rad/s]

ϕ Phase [rad]

- The phasor interpretation is ***independent*** of the wave's polarization and is useful for explaining how a GNSS receiver works

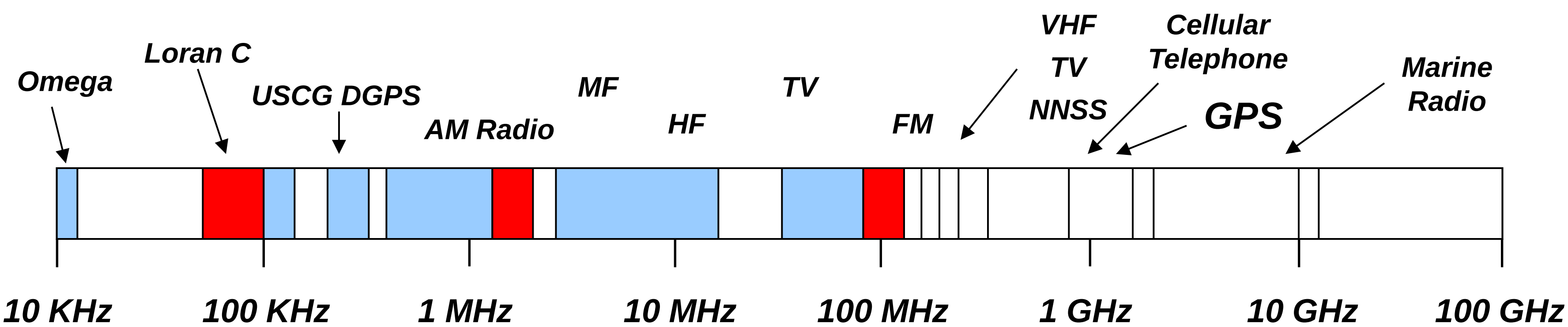
Radio Frequency (RF)

- RF frequencies range from 3 Hz to 300 GHz; grouped into several bands
  - Each band represents an order of magnitude increase of frequency
- Some of the frequency bands are shown in the table below

Designation	Abbreviation	Frequencies	Wavelengths
Very Low Frequency	VLF	3 kHz - 30 kHz	100 km - 10 km
Low Frequency	LF	30 kHz - 300 kHz	10 km - 1 km
Medium Frequency	MF	300 kHz - 3 MHz	1 km - 100 m
High Frequency	HF	3 MHz - 30 MHz	100 m - 10 m
Very High Frequency	VHF	30 MHz - 300 MHz	10 m - 1 m
Ultra High Frequency	UHF	300 MHz - 3 GHz	1 m - 100 mm
Super High Frequency	SHF	3 GHz - 30 GHz	100 mm - 10 mm
Extremely High Frequency	EHF	30 GHz - 300 GHz	10 mm - 1 mm

RF Spectrum Assignment

- The global use of the radio spectrum is governed by International Telecommunications Union (ITU)
  - Use of RF spectrum is governed by a given country’s internal Federal regulations (in Canada the governing body is Industry Canada, and in the U.S. it is the FCC)
- Other parts of the EM spectrum (such as optical and infrared) are also commonly used for range measurements



- Full spectrum allocation is available at

  http://www.ic.gc.ca/eic/site/smt-gst.nsf/vwapj/spectallocation-08.pdf/$FILE/spectallocation-08.pdf

Some Common RF Effects

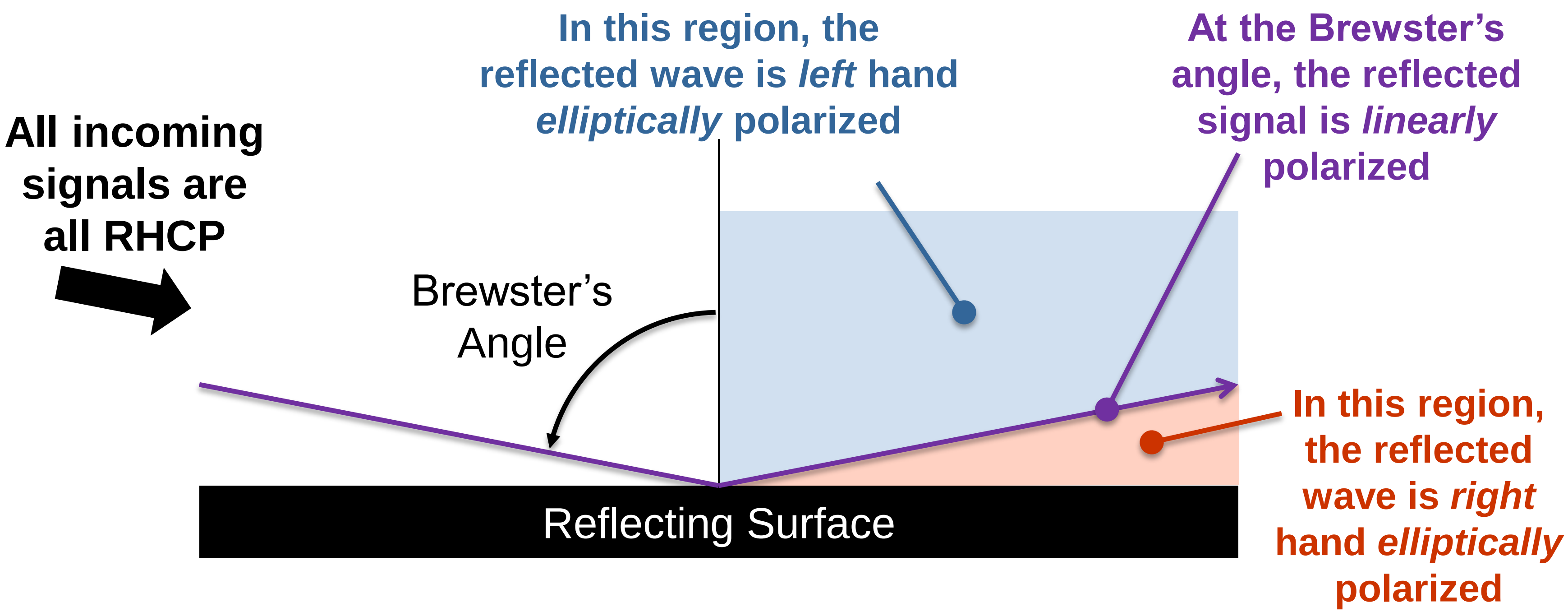
- ***Absorption*** refers to the conversion of EM wave energy into another form of energy (e.g., heat) as a result of its interaction with matter
- ***Attenuation*** is the decrease in the intensity of an EM wave as a result of absorption of energy, but not including the reduction due to geometric spreading (discussed below)
- ***Refraction*** is the change in speed and possibly direction of a wave as it propagates from one medium to another

$$n = \frac{v_{vacuum}}{v_{medium}} \qquad\qquad n_a \sin\left(q_a\right) = n_b \sin\left(q_b\right)$$

- ***Diffraction*** is a process in which the direction of energy flow of a wave deviates when it passes an obstacle such as the Earth's curvature, mountains, man-made structures, etc.

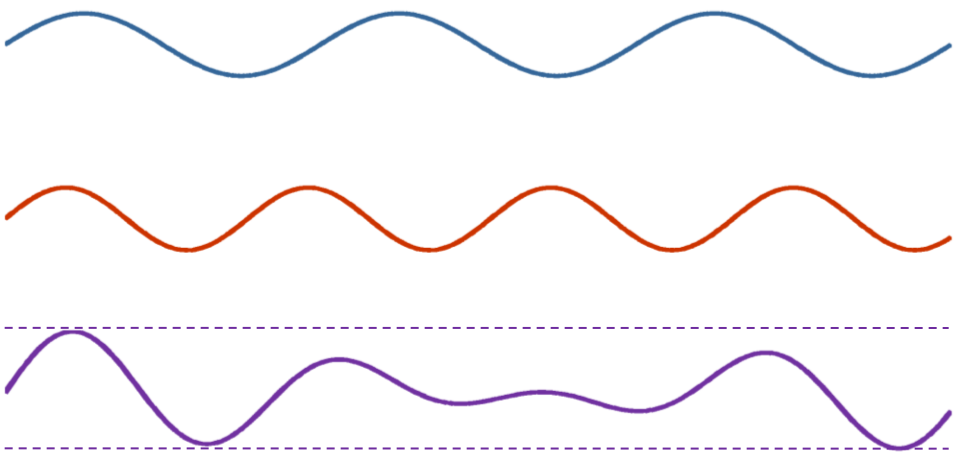
Reflection

- An EM wave may be partially or completely reflected by the boundary between two media. If the reflected wave also reaches the user (along with the direct signal) we call this ***multipath*** and it leads to fading and measurement errors. The wave's polarization will also change upon reflection.



Signal Fading

- ***Fading*** represents variation of RF field strength at a particular point caused by changes in the transmission path with time
  - More specifically, fading results when signals from two or more paths combine, thus causing constructive or destructive interference
  - For GNSS, this results primarily from reflection of signal from nearby objects (“multipath”) but can also occur because of atmospheric effects known as scintillation



Desired (direct) signal

+

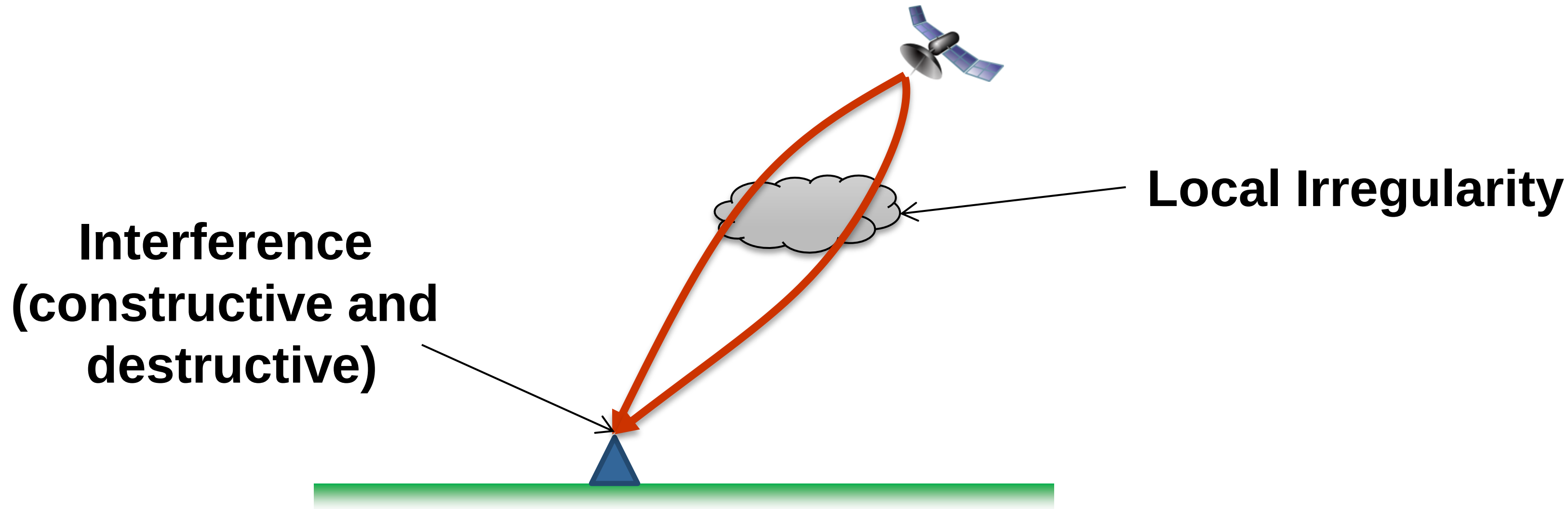
Second (reflected) signal

=

Composite signal
(Note amplitude variations which can cause the receiver to be unable to track the signal)

Scintillation

- ***Scintillation*** is the rapid fluctuation of the amplitude and phase of a wave passing through a medium with small-scale irregularities.



- For GNSS, the primary source of scintillation – when it occurs – is the ionosphere (atmosphere above ~50 km). Troposphere (atmosphere below ~11 km) scintillation can also occur but is much less frequent.

Propagation Velocity at GNSS Frequencies

- When a media is said to be ***non-dispersive***, it means the index of refraction is independent of the frequency of waves passing through it. Conversely, the index of refraction for ***dispersive media*** is a function of the wave's frequency.

- In the context of GNSS, the atmosphere can be divided into two parts that affect the GNSS signals differently:
  - The ***troposphere*** is non-dispersive below 20 GHz. The index of refraction is a function of temperature, pressure, and water vapor only.
  - The ***ionosphere*** is dispersive in the GNSS band and the index of fraction is thus a function of the wave's frequency.

- When the refractive index is dispersive, it can be accurately determined using measurements on two frequencies. This was the main reason for having two frequencies in GNSS.

Signal Strength

Decibels

- Units of decibels (dB) are often used to quote gain and attenuation

$$[\text{dB}] = 10\log_{10}[\text{ratio}]$$

Ratio	dB
1	$10 \log_{10}(1) = 0$ dB
2	$10 \log_{10}(2) = 3$ dB
4	$10 \log_{10}(4) = 6$ dB
8	$10 \log_{10}(8) = 9$ dB
10	$10 \log_{10}(10) = 10$ dB

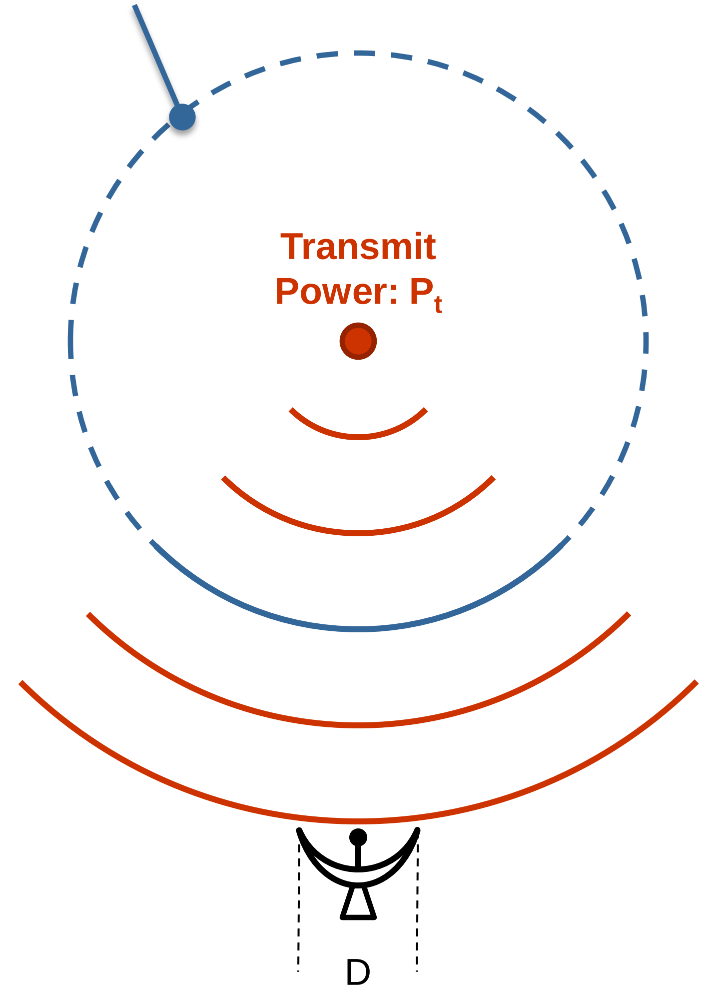
- Following Ohm's Law ($P = I^2R = V^2/R$), we can write the following expressions for power ratios

$$[\text{dB}]_{\text{Gain}} = 10\log_{10}\frac{P_t}{P_r} = 20\log_{10}\frac{V_t}{V_r} = 20\log_{10}\frac{I_t}{I_r}$$

Geometrical Spreading and Free-Space Loss (1/3)

- Geometrical spreading, is the decrease in EM field strength as wave energy radiates over an ever increasing area. It is independent of reflections, refraction, diffraction, absorption or scattering.
- Consider the 2D case with an isotropic transmitter emitting a single wave front with total power P_t. As the wave front expands, the power is spread over the circumference of the circle, thus decreasing is power density (i.e., power per unit length of circumference). The amount of power received at the receiver is a function of the size (more specifically, aperture) of the antenna

$$\text{Power Density at a Given Distance} = P_d = \frac{P_t}{\text{Circumference}}$$



$$\text{Received Power} = P_r = P_d \cdot D$$

Geometrical Spreading and Free-Space Loss (2/3)

- Power density for an isotropic, or omni-directional, transmitter in 3D

$$P_d = \frac{P_t}{4\pi R^2}$$

- Power density for a non-isotropic transmitter with gain G_t relative to the isotropic case

$$P_d = \frac{P_t}{4\pi R^2} \cdot G_t$$

- Aperture of an omni-directional receiving antenna

$$A = \frac{\lambda^2}{4\pi}$$

- Aperture of receiver antenna with a gain of G_r

$$A = \frac{\lambda^2}{4\pi} \cdot G_r$$

- Received power is thus given by

$$P_r = P_d A = \frac{P_t G_t G_r \lambda^2}{(4\pi R)^2}$$

Geometrical Spreading and Free-Space Loss (3/3)

- Finally, the free-space loss is given by the ratio of transmit to the received power without considering the antenna gains (i.e., assume they are unity)

$$FSL = 10\log_{10}\left(\frac{P_t}{P_r}\right) = 10\log_{10}\left(\frac{(4\pi R)^2}{\lambda^2}\right) = 20\log_{10}\left(\frac{4\pi R}{\lambda}\right)$$

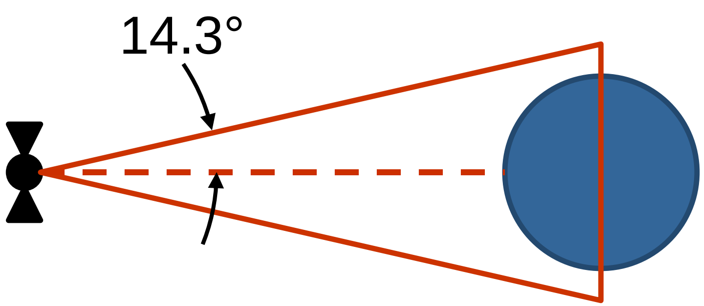
- If the antenna gains are considered we can write

$$FSL = 20\log_{10}\left(\frac{4\pi R}{\lambda}\right) - G_r^{dB} - G_t^{dB}$$

NOTE: In general, G_t and G_r are expressed as ratios, whereas the "dB" superscript indicates the corresponding decibel values (which are more commonly used to define antenna gains)

GPS Free-Space Loss

- The farthest a GPS satellite can be from a receiver on the surface of the Earth is approximately 25,225 km. At L1, the corresponding free-space loss is 184.4 dB. As a ratio, this means the signal is reduced by a factor of $10^{18.4}$ relative to the transmit power!
- To overcome this, the antenna gain pattern is directed toward the Earth, which provides a gain of 13.4 dB.
  - If you had to design the beam pattern, what would it look like and why?



Reference: Czopek, F.M. and S. Shollenberger (1993), **Description and performance of the GPS Block I and II L-Band antenna and link budget**. Proc. GPS93, ION, pp. 37-43.

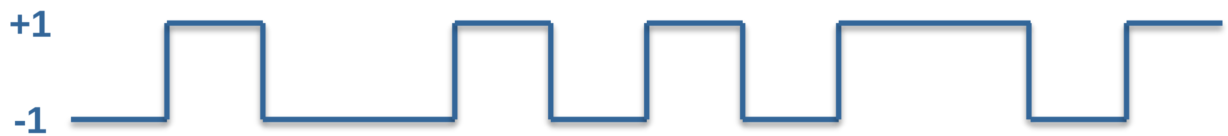
Overview of GNSS Signals

Desired GNSS Signal Properties

Characteristic	How this is Accomplished
Tolerance to other GNSS signals in the same frequency band (i.e., each satellite be able to be uniquely identified)	• Multiple Access capability is built into the signals (details to follow)
Tolerance to multipath (i.e., reception of same source from multiple directions, usually due to reflection)	• Increasing the bandwidth (chipping rate) of the signals improves multipath tolerance
Tolerance to jamming and interference (intentional or not) and to spoofing (i.e., trying to fool a receiver using a similar signal)	• System is based on spread spectrum techniques • Increasing the bandwidth of the signals improves multipath tolerance
Ability to remove ionospheric effects	• Broadcast on two (or more) frequencies (e.g., L1 and L2)

Basic GNSS Signal Structure

- In general, any GNSS signal can be thought of as being comprised of
  - A sinusoidal wave ("carrier wave", or just "carrier")
  - A ranging code that is used to measure the distance to the satellites. This also provides robustness to interference. The ranging code is a square wave consisting of a sequence of ±1's.



  - A navigation message (optional) than contains information on ephemeris, almanac, clock parameters, etc. It consists of a sequence of bits represented as ±1's (i.e., it is also a square wave).
- The ranging code and navigation message are modulated onto the carrier wave using phase modulation (PM). Other modulation approaches include amplitude modulation (AM), frequency modulation (FM) and time modulation (TM), but these are not used by GNSS signals.

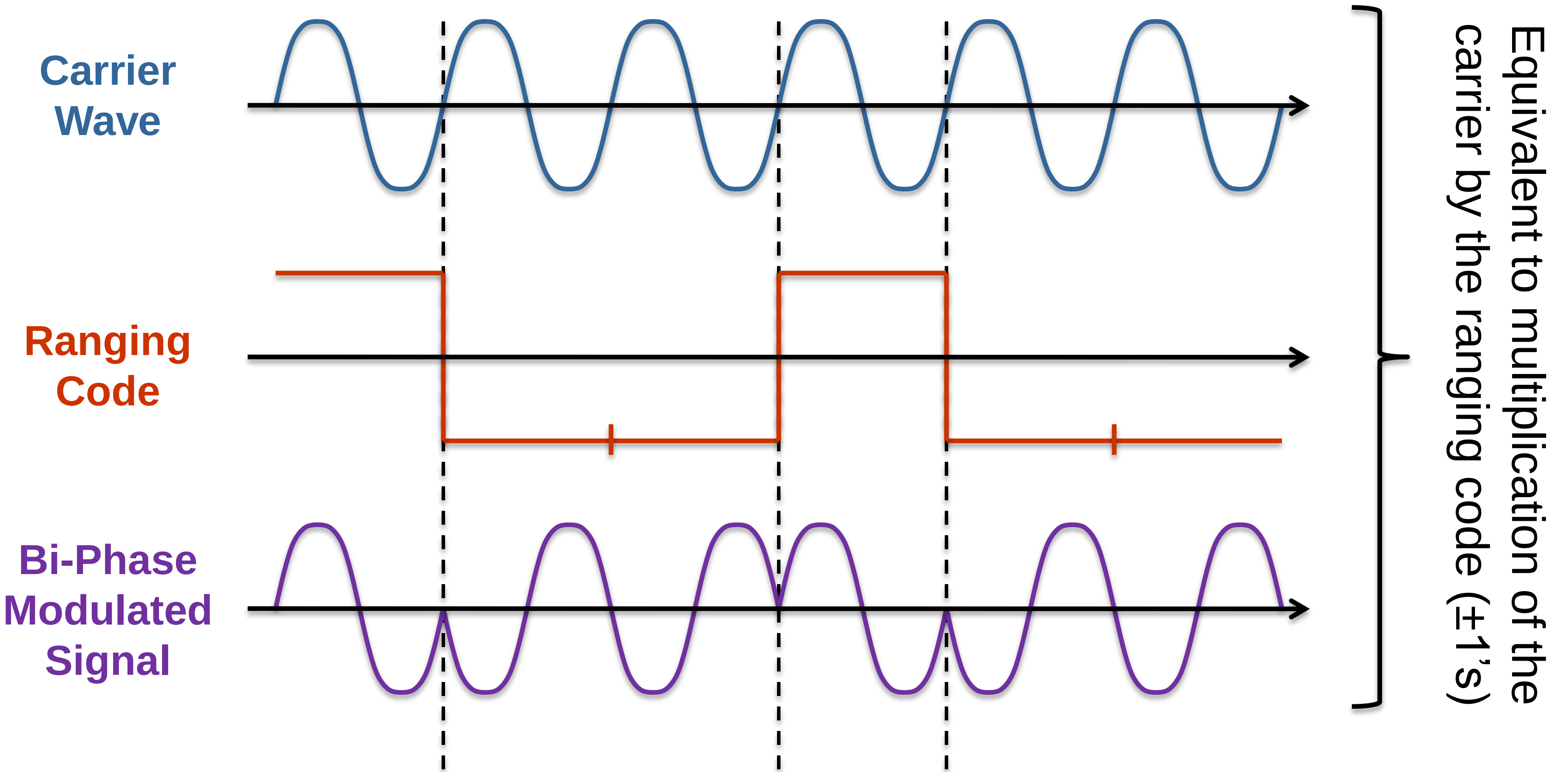
Multiple Access Capability

- Multiple access capability means that a receiver is able to "access" signals from multiple satellites at the same time. There are several ways of realizing this in GNSS:
  - Code Division Multiple Access (CDMA) is when all satellites have the same carrier frequency but different ranging codes. This is used by GPS, Galileo, Compass and future GLONASS satellites.
  - Frequency Division Multiple Access (FDMA) is when all satellites have the same ranging codes but different carrier frequencies. This approach is currently used by GLONASS (although GLONASS will shift to CDMA too).

- Other multiple access techniques include time division multiple access (TDMA) and space division multiple access (SDMA), although these are not used in GNSS

Bi-Phase Modulation

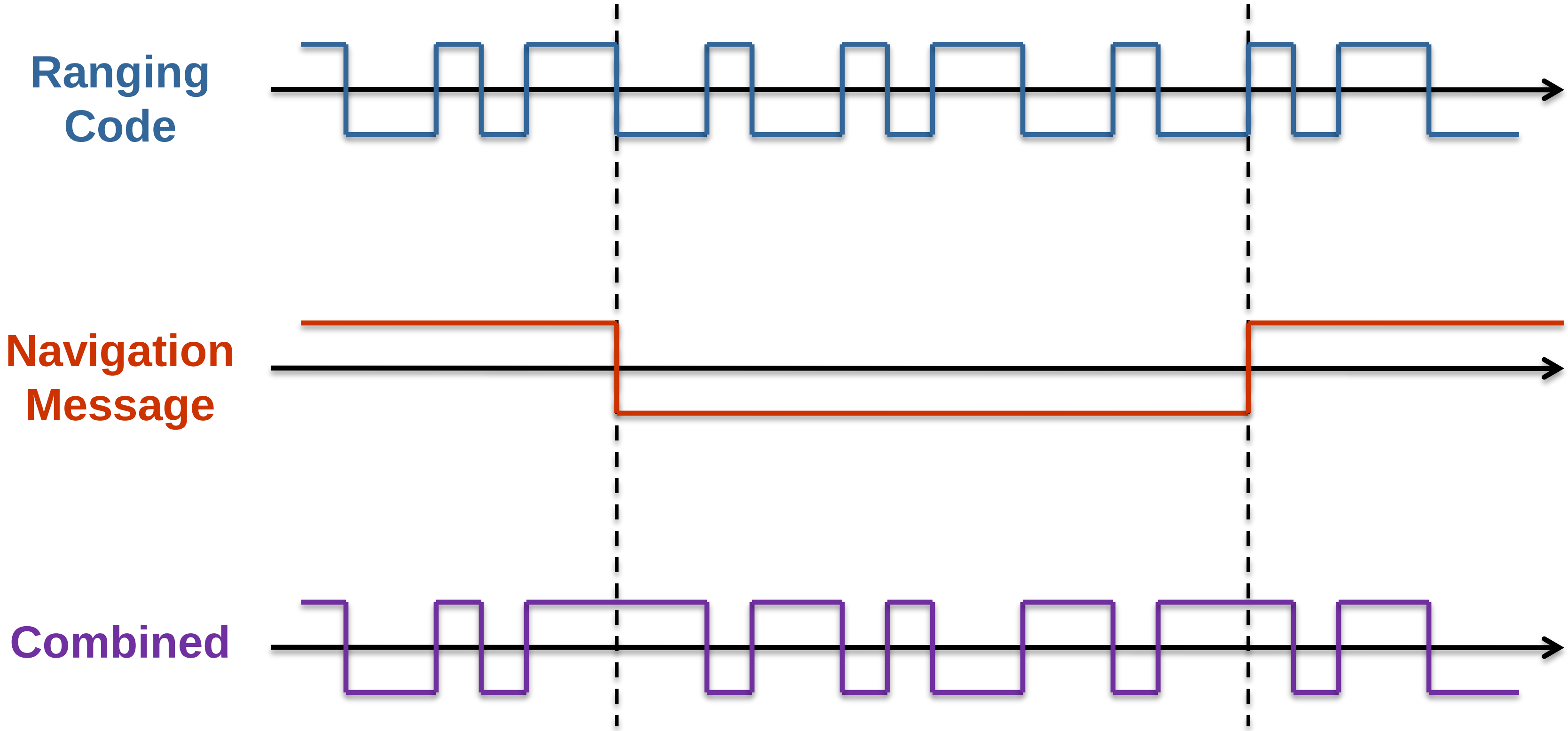
- We only consider bi-phase modulation which introduces a 180° phase shift in the carrier when the modulation changes



Not to scale

Including Navigation Message

- The navigation message is further modulated onto the carrier. Combination of ranging code and navigation message is done using multiplication of sequences of ±1's, which is equivalent to modulo-2 addition of sequences of zeros and ones.

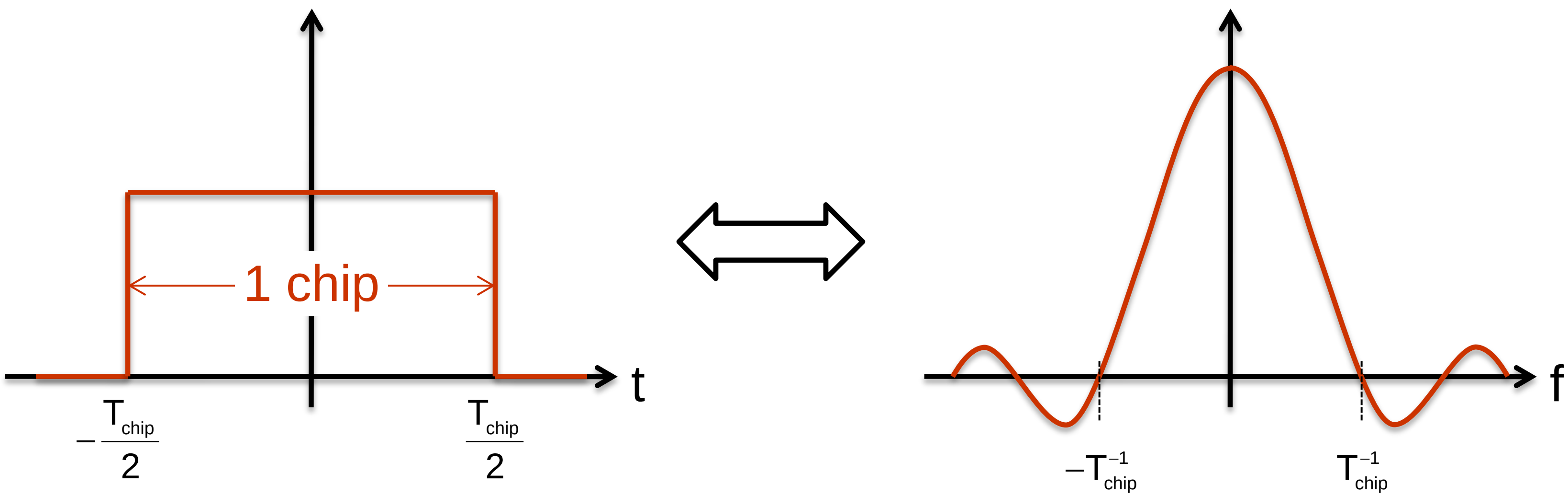


Not to scale

Spectrum of Square Waves (1/2)

- As already discussed, ranging codes and the navigation message are square waves. Assuming a period of T_{chip} (chipping rate of f_{chip}) their Fourier transform is given by.

$$\mathrm{rect}\left(\frac{t}{T_{chip}}\right) \Longleftrightarrow T_{chip}\,\mathrm{sinc}\left(\pi f T_{chip}\right)$$

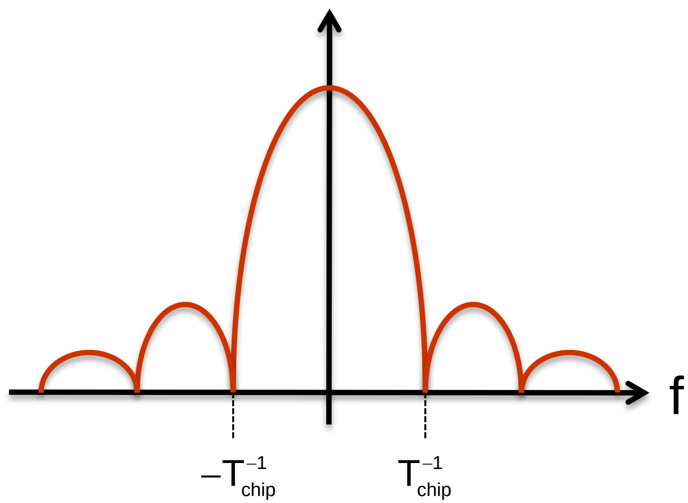


Note: Higher chipping rates give higher bandwidths, and bandwidth plays a very important role in terms of system performance

Spectrum of Ranging Code (2/2)

- The power spectrum is the square of the Fourier transform

$$T_{chip}^{2}\,\mathrm{sinc}^{2}\left(\pi f T_{chip}\right)$$



- This is the common way of showing a GNSS signal in the frequency domain

GPS Signal Structure

GNSS Signal Information

- This course looks very briefly at the GPS signal structure and the format of the navigation message

- The complete definition of any GNSS signal is given by that system's Interface Control Document (ICD)
  - For GPS, it can be downloaded from http://www.gps.gov/technical/icwg/

- Another good source of information for GPS is the "Global Positioning System Standard Positioning Service Performance Standard"
  - Available online at: http://www.gps.gov/technical/ps/

GPS Signal Structure Overview

- The satellite has an internal "fundamental frequency", f_0, of 10.23 MHz. All *carrier waves* are integer multiples of f_0
    - $f_{L1} = 154 \times f_0 = 1575.42$ MHz
    - $f_{L2} = 120 \times f_0 = 1227.60$ MHz
    - $f_{L5} = 115 \times f_0 = 1176.45$ MHz

- The following *ranging codes* are used in GPS (all of which are periodic)
  - Coarse acquisition code: C/A-code
    - Available on L1
  - Precise code: P-code
    - Available on L1 and L2
    - Encrypted by U.S. military (using a classified Y-code) to yield the P(Y)-code
  - Modernization
    - L2C – Available on 9 satellites (as of Jan 2012)
    - L5 – Available on 2 satellites (as of Jan 2012)
    - L1C – First launch ~2013

GPS C/A and P Codes

Property	C/A-Code	P-Code
Chipping Rate	1.023 MHz	10.23 MHz
Chip Length (analogous to carrier wavelength)	293.1 m	29.31 m
Code Length	1,023 chips	2.3547 x 10^{14} chips
Code Period	1 ms	266.4 days (split into 7-day segments; one per satellite)
Minimum Received Power	-158.5 dBW	-161.5 dBW (IIR-M satellites or newer)

- As we will see, the chip length (and thus chipping rate) is important in terms of a receiver's sensitivity to multipath
- The higher chipping rate for the P-code also makes it less sensitive to jamming and interference
- The short period of the C/A makes it easy/fast to acquire
- 'dBW' denotes power relative to one Watt (1 W)

GPS Signals Leaving the Satellite

- For this course, we only concern ourselves with the "legacy" L1 and L2 signals

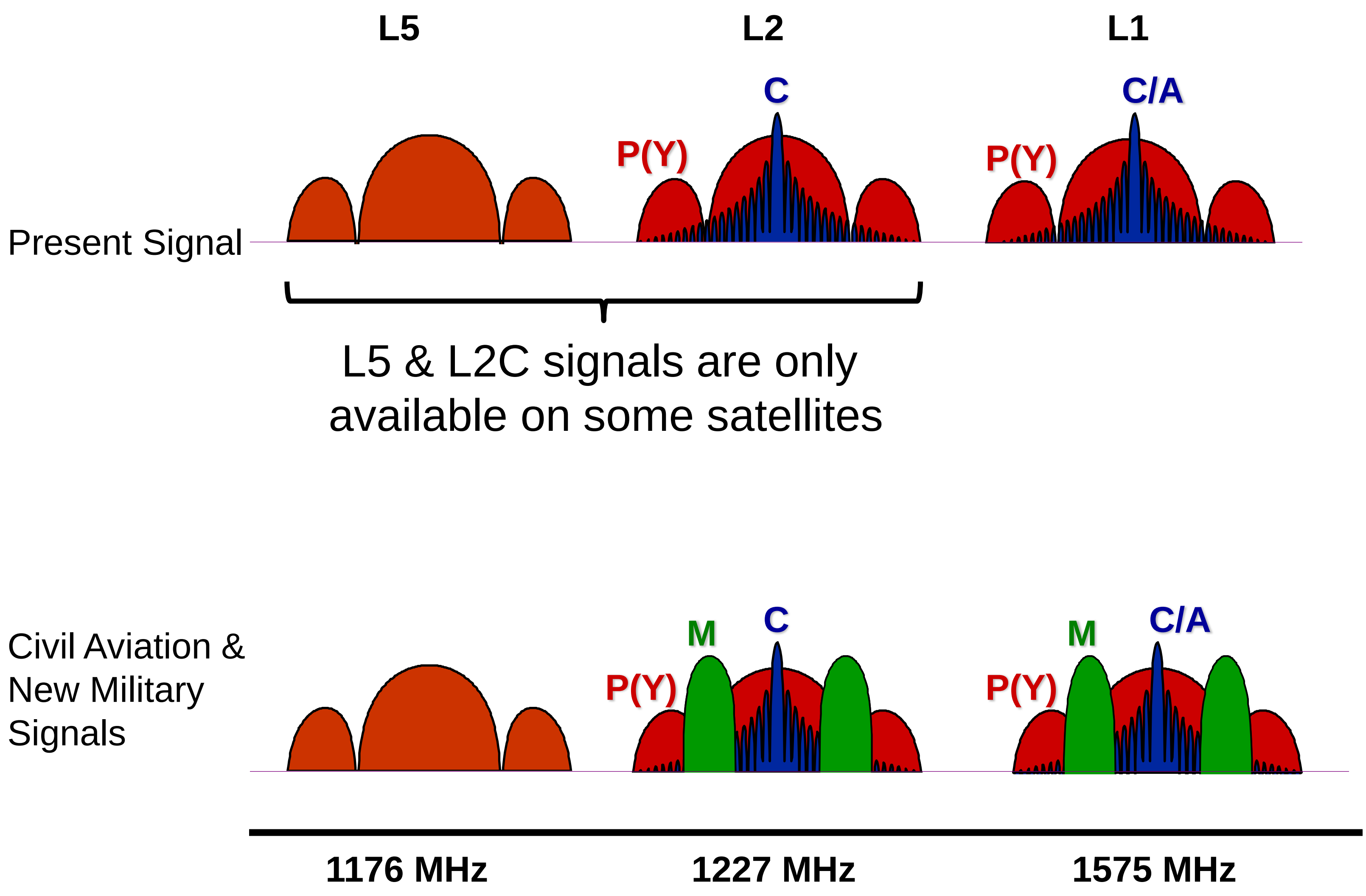
$$L1(t) = AP(t)N(t)\cos(2\pi f_{L1}t) + \sqrt{2}AC(t)N(t)\sin(2\pi f_{L1}t)$$

$$L2(t) = AP(t)N(t)\cos(2\pi f_{L2}t)$$

where

A	is the amplitude
P	is the P-code ranging code
C	is the C/A-code ranging code
N	is the navigation message (data bits)
f_{L1}	is the L1 carrier frequency
f_{L2}	is the L2 carrier frequency

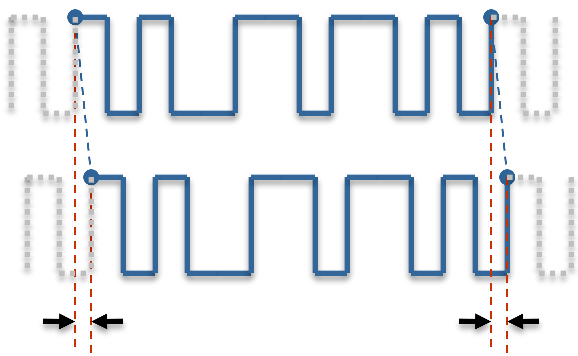
Current and Future GPS Signals



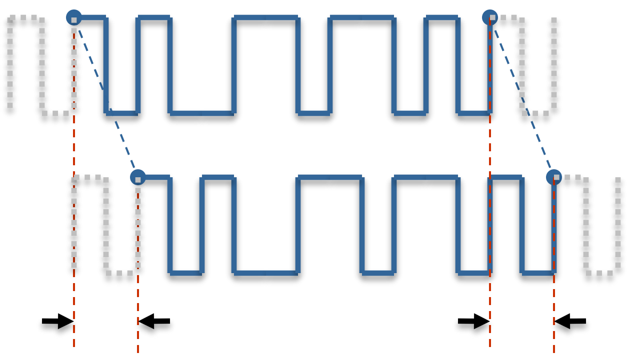
Reference: Julien, O. and C. Macabiau (2006) **What are the major differences between Galileo and GPS current and forthcoming frequencies?**, GNSS Solutions, Inside GNSS Magazine, Vol 1, No 4, pp. 22-25.
Available at: http://www.insidegnss.com/auto/MayJune06GNSSSolutions.pdf.

Correlation of Ranging Codes

- GNSS ranging codes are carefully designed to have special auto- and cross-correlation properties. You will recall from statistics that cross-correlation refers to how related two parameters are to each other (e.g., number of layers of clothes is correlated with the outside temperature). Similarly, auto-correlation refers to how related something is with a (time) shifted version of itself.



For small shifts, the two signals still look quite similar (high correlation)



For large shifts, the two signals look quite different (low correlation)

Correlation Functions

- An ***auto-correlation function (ACF)***, $R_x(\tau)$, mathematically defines the correlation (similarity) of a signal relative to a version of itself shifted in time by τ. This is important for ranging purposes

$$R_x(\tau) = \frac{1}{T}\int_0^T x(t)\,x(t-\tau)\,dt \qquad\qquad R_x[k] = \frac{1}{N}\sum_{i=1}^{N} x_i x_{i-k}$$

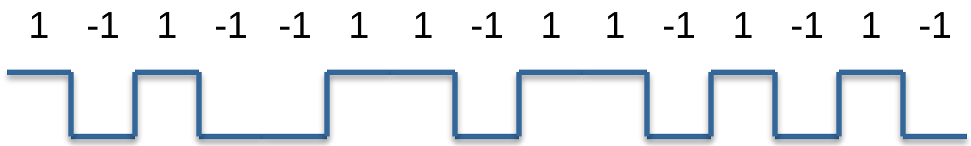
  where N is the number of chips and T is the period

- A ***cross-correlation function (CCF)***, $R_{xy}(\tau)$, describes the correlation (similarity) of a signal relative to another signal shifted in time by τ. This is important for CDMA implementation

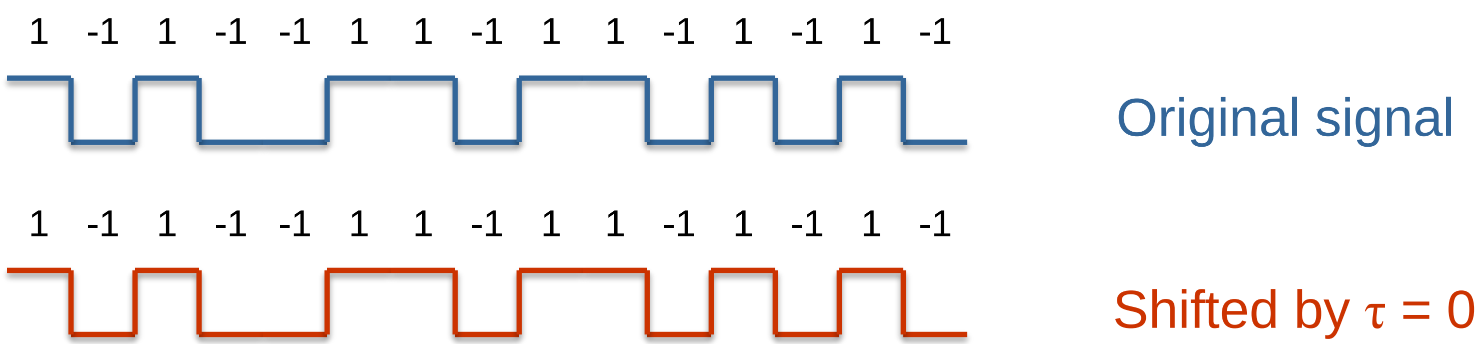
$$R_{xy}(\tau) = \frac{1}{T}\int_0^T x(t)\,y(t-\tau)\,dt \qquad\qquad R_{xy}[k] = \frac{1}{N}\sum_{i=1}^{N} x_i y_{i-k}$$

Example ACF for Periodic Sequences (1/2)

- Consider the following sequence, which is assumed to be periodic (as are the ranging codes)



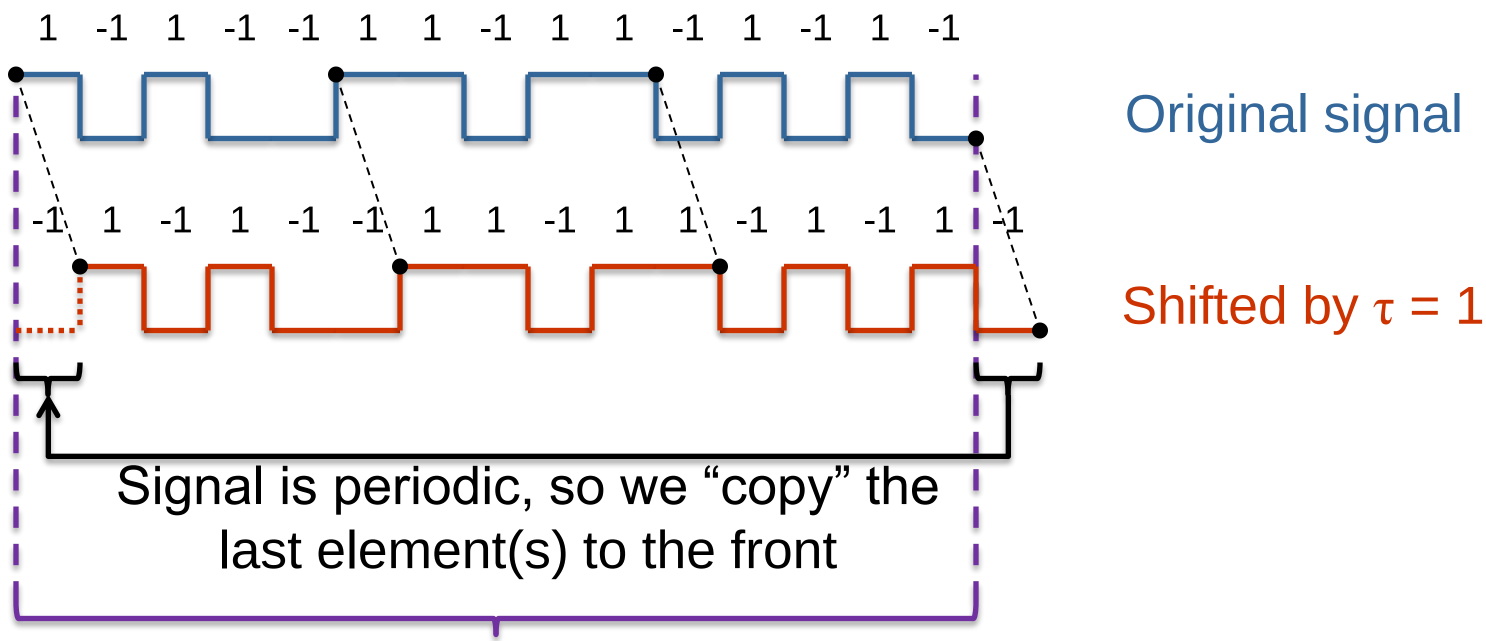
- For $\tau = 0$



- Evaluating R(0) (i.e., multiplying and adding and dividing by the length of the sequence) gives a value of unity

Example ACF for Periodic Sequences (2/2)

- For $\tau = 1$



Multiply and add over the period of the code to get a value of -9. This value is then divided by the number of elements in the sequence to get R(1) = -0.6. This is less than unity, that is, R(0).

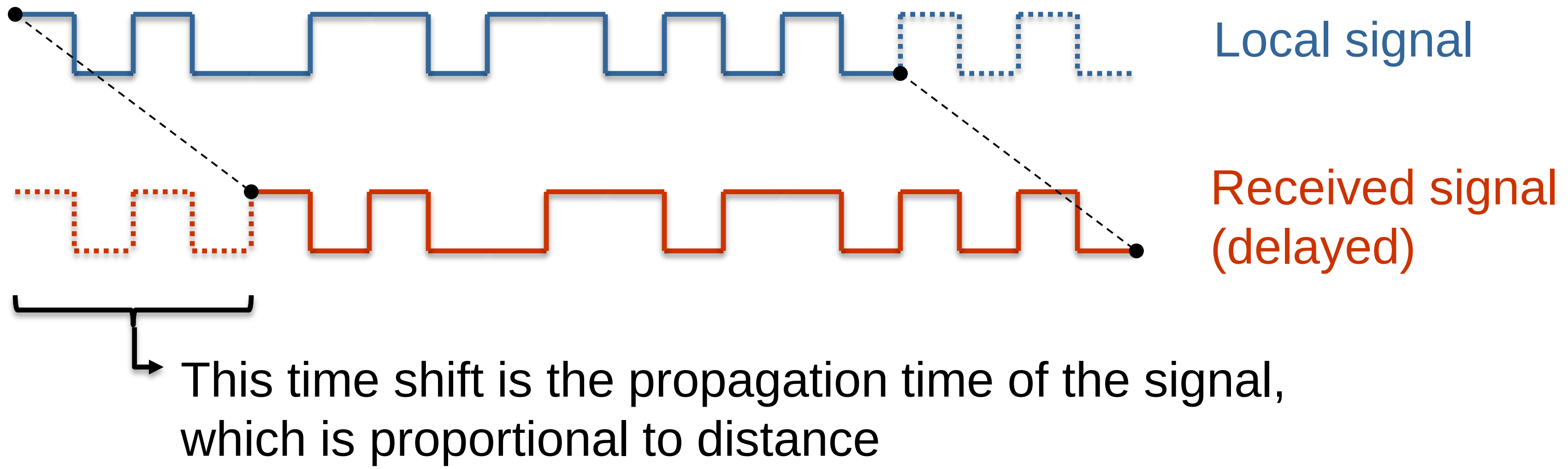
Ideal Correlation Properties (1/2)

- Ideally, the ACF should be unity at $\tau = 0$ (i.e., perfect correlation) and zero for all other values of τ (i.e., no correlation). This would allow for the unambiguous determination of the peak value.

- Conversely, the CCF between different satellites' ranging codes should ideally always be zero (i.e., no correlation). This would ensure that signal from different satellites are not confused with each other. This is the basis of code division multiple access (CDMA), which is used in nearly all GNSS.

Ideal Correlation Properties (2/2)

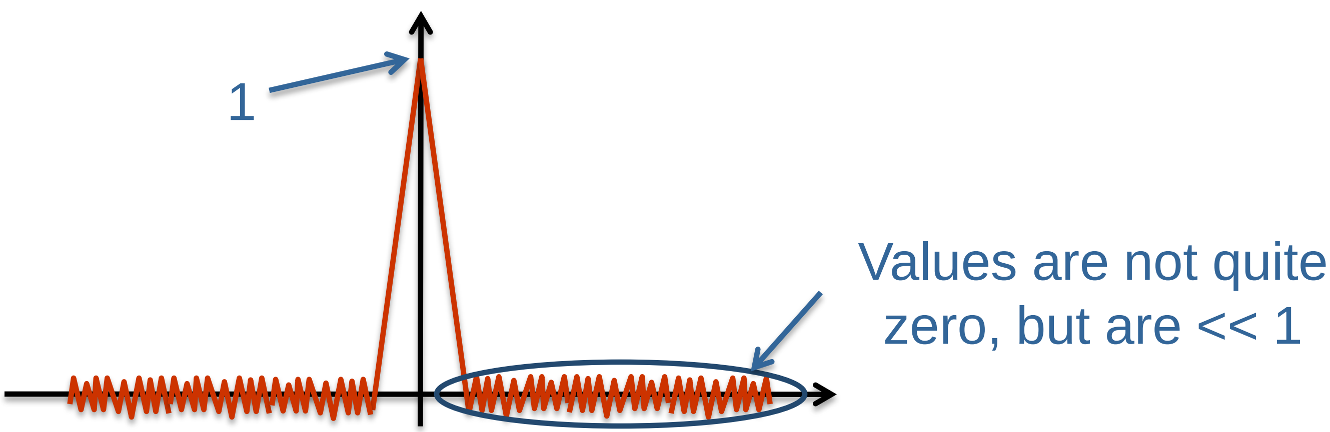
- Why is any of this important?
  - To answer this, we have to briefly explain what happens inside a GNSS receiver assuming the receiver is perfectly synchronized to GNSS time. In this case, a GNSS receiver generates the ranging code internally and determines the time shift required to align the local code with the code received from the satellite. The time shift that maximizes the correlation between the local and incoming signal is the time it took the signal to propagate from the satellite to the receiver (which can then be scaled to a distance).



Pseudo-Random Noise (PRN) Codes

- The only sequences that generate the ideal ACF and CCF properties are random noise sequences. However, since these are completely random, they cannot be reproduced in the receiver. Instead, the ranging codes are ***pseudo-random noise (PRN)*** sequences, meaning they appear nearly random, but are actually reproducible (deterministic).

- In GPS, each satellite is identified by a separate PRN code. For the C/A code, the ACF is shown below



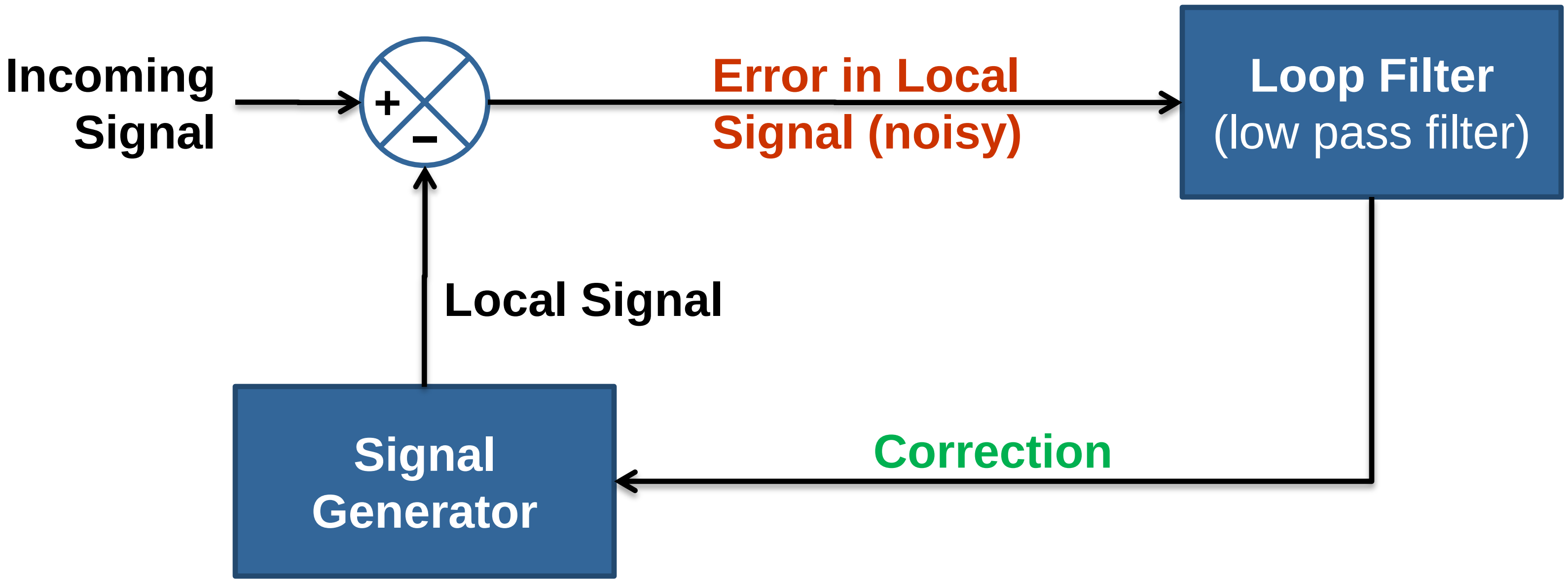
How a GNSS Receiver Works
(at a very high level)

Basic Concept

- Overall objective of signal tracking is to generate the incoming satellite signals locally. This consists of ***both*** the ***code part*** (code tracking) and the ***carrier part*** (phase tracking).
- The incoming and locally generated signals are then compared to determine the error in the local signal. Using this error, the local signal is adjusted to more closely match the incoming signal. This process continues on and is called ***signal tracking***, or simply ***tracking***, and is performed in a ***tracking loop***.
- While tracking, since the local signal closely matches the received signal, GPS measurements ***conceptually*** can be made using knowledge of only the locally generated signal.
- If a receiver is no longer able to track a signal the receiver is said to have **"lost lock"** on that signal

Code and Carrier Tracking Loops

- For the purpose of this course, we treat code and carrier tracking separately but in reality they must be done together and each will impact the other (i.e., bad carrier tracking will negatively affect code tracking and vice versa). The basic tracking loop implementation is the same for both and is shown below.



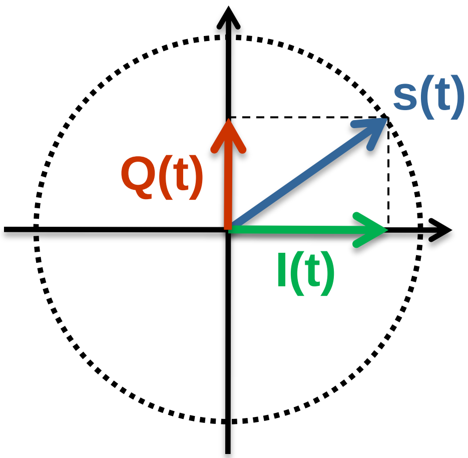
Signal Tracking

- The following slides give a very general overview of how a GNSS receiver tracks signals. Although the code and carrier tracking are presented separately, they are actually done in parallel and you cannot have one without the other.

- Code tracking is performed in a ***Delay Lock Loop (DLL)*** and phase tracking is performed in a ***Costas Phase Lock Loop (PLL)***. Generally, carrier tracking is more difficult because the wavelength of the carrier is shorter than the code chip length (and thus is harder to track).

Carrier Tracking (1/3)

- Recall that the carrier phase can be represented as a phasor



$$s(t) = I(t) + jQ(t)$$
$$= A\cos(\phi(t)) + jA\sin(\phi(t))$$

$$\phi(t) = 2\pi ft$$

Carrier Tracking (2/3)

- Assume we generate a local signal as

$$\hat{s}(t) = \cos\left(\hat{\phi}(t)\right) + j\sin\left(\hat{\phi}(t)\right) \qquad \hat{\phi}(t) = 2\pi\hat{f}t$$

- We can then multiply to get

$$\begin{aligned} r(t) &= s(t) \times \hat{s}(t) \\ &= \cos\left(\delta\phi(t)\right) + j\sin\left(\delta\phi(t)\right) \\ &= \hat{I}(t) + j\hat{Q}(t) \end{aligned}$$

where

$$\delta\phi(t) = \phi(t) - \hat{\phi}(t)$$

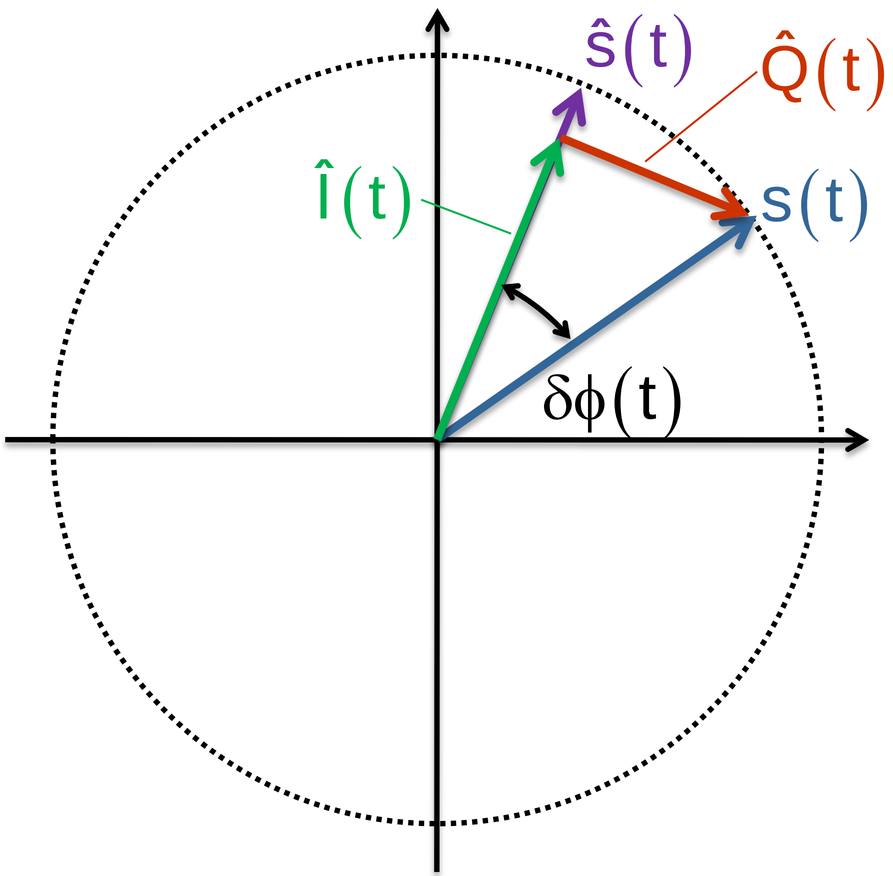
Carrier Tracking (3/3)

- The mathematics on the previous slide is equivalent to projecting the incoming signal onto the local signal.

- If the local signal matches the incoming signal Q-hat will be zero

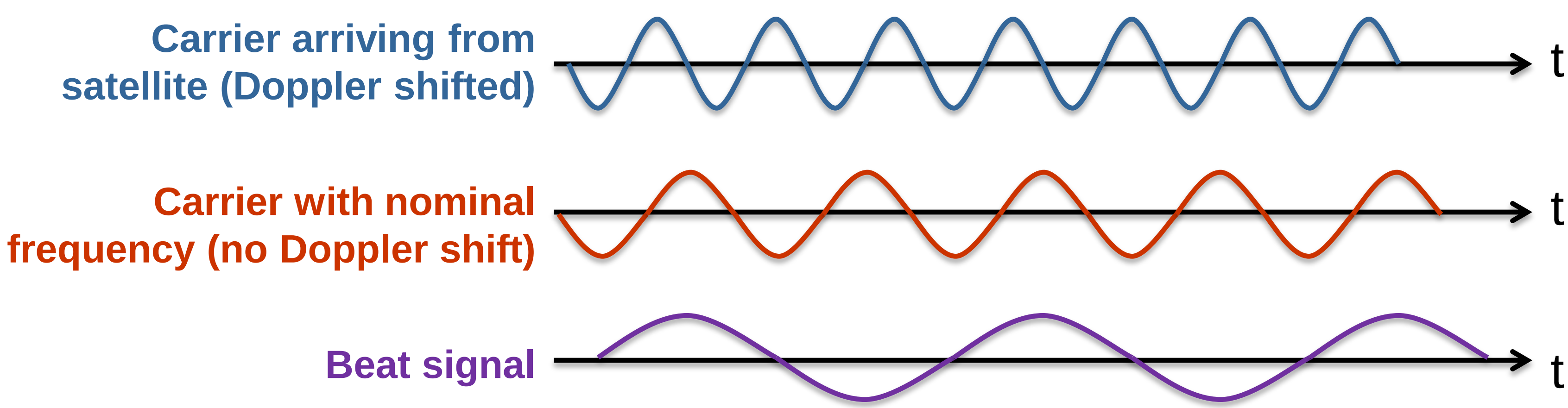
- The phase error can then be estimated as

$$\delta\hat{\phi}(t) = \tan^{-1}\left(\frac{\hat{Q}(t)}{\hat{I}(t)}\right)$$



Carrier Phase Measurements

- GNSS receivers can measure the phase of the incoming signal with an accuracy of a few degrees, which implies the receiver is also tracking the frequency of the signal. With the actual (received) frequency known, the Doppler frequency can be computed by subtracting the nominal signal frequency (e.g., at L1).
- Once the Doppler is available, the carrier phase measurement is actually by integrating the Doppler (and is sometimes called the accumulated Doppler range).
- This can be interpreted as measuring the phase of a signal whose frequency is the difference between incoming and nominal signals (called a beat frequency in physics)



Code Tracking (1/3)

- Code tracking is based on the auto-correlation function. Assuming the incoming signal has a code phase (chip number) of τ, the incoming signal can be written as

$$C(\tau)$$

- Within the receiver, an “early” (E) and “late” (L) version the signal are generated as follows ($\Delta\tau$ is usually 0.5 chips or less)

$$\hat{E} = C(\hat{\tau} + \Delta\tau)$$

$$\hat{L} = C(\hat{\tau} - \Delta\tau)$$

- Sometimes a “prompt” (P) version is also generated

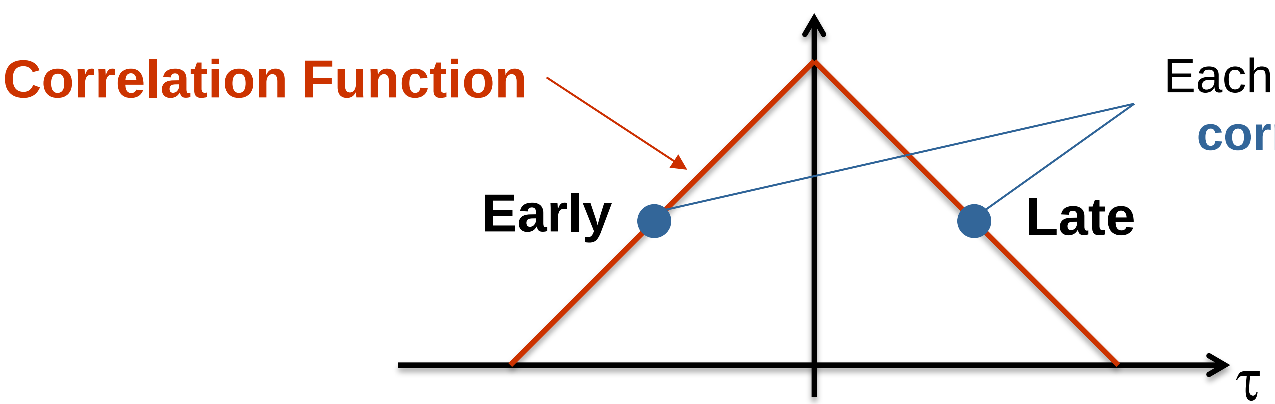
$$\hat{P} = C(\hat{\tau})$$

Code Tracking (2/3)

- Each of the local codes (early, prompt and late) is multiplied by the incoming signal and added together. Considering only the early and late, we can write

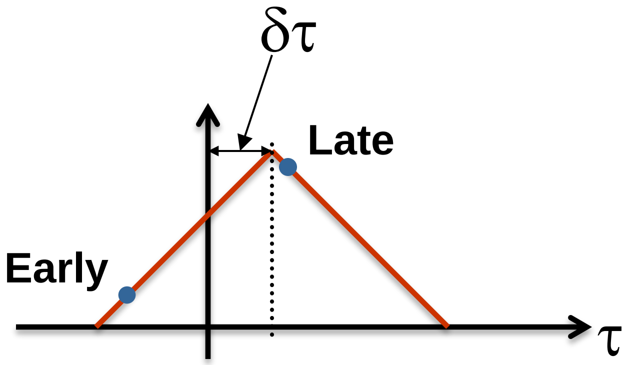
$$\begin{aligned} E &= \sum C(\tau)C(\hat{\tau}+\Delta\tau) \\ &= R(\tau-\hat{\tau}-\Delta\tau) \\ &= R(\delta\tau-\Delta\tau) \end{aligned} \qquad \begin{aligned} L &= \sum C(\tau)C(\hat{\tau}-\Delta\tau) \\ &= R(\tau-\hat{\tau}+\Delta\tau) \\ &= R(\delta\tau+\Delta\tau) \end{aligned}$$

- Each value is a point on the auto-correlation function and is called a ***correlator output***.

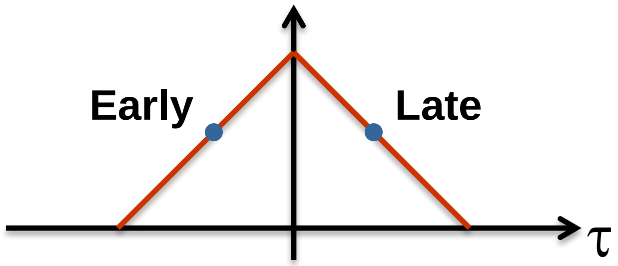


Code Tracking (3/3)

- Differencing the early and late correlator outputs indicates which correlator (early or late) contains more energy and therefore whether the receiver needs to advance or delay the locally generated version of the code (to align it with the incoming signal).
- Ideally the two paths are "balanced", and the resulting difference is zero. The midpoint between the early and late correlators is called the ***tracking point***.



Early & Late Correlator Outputs
<u>with</u> Tracking Error



Early & Late Correlator Outputs
<u>without</u> Tracking Error

Pseudorange Measurements

- When the local code is aligned to the incoming signal, the time it took the signal to propagate from the satellite to the receiver can be determined (details are beyond the scope of this course but the principle is exactly as described). Scaling this time by the speed of light gives a measure of the range to the satellite.

- However, the receiver cannot generate a perfect replica of the signal because its internal clock has errors and we therefore call the resulting range a ***pseudorange***. We have already seen how to deal with this clock error, namely by estimating it in our least-squares solution.

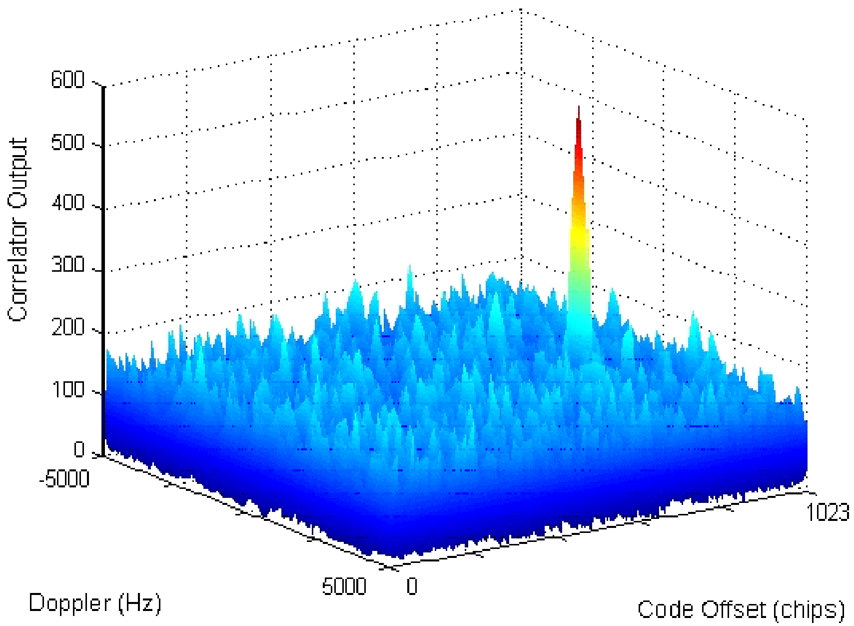
Signal Acquisition (1/2)

- The previous slides have described the signal tracking process in which the local signal is steered to align with the incoming signal as closely as possible. However, tracking must be initialized with some coarse values of the signal parameters.

- ***Signal acquisition*** is the process of obtaining approximate values for the *code phase* and the *carrier Doppler* (frequency). It is performed by searching all possible combinations for the code phase and Doppler. For each combination, the local signal is correlated with the incoming signal. If the signal has the assumed signal parameters, the correlation output will be large and the signal is said to be acquired. If the signal does not have the assumed parameters, the correlation output will be low and the signal will not be detected/acquired.

Signal Acquisition (2/2)

- Results form a typical acquisition search is shown to the right. As can be seen, there is only one very strong correlation result; the code phase and Doppler shift at this peak are used to initialize the tracking process.
  - Also of note is that there are other smaller peaks. What is this caused by?
- Finally, acquisition is generally more difficult than tracking.



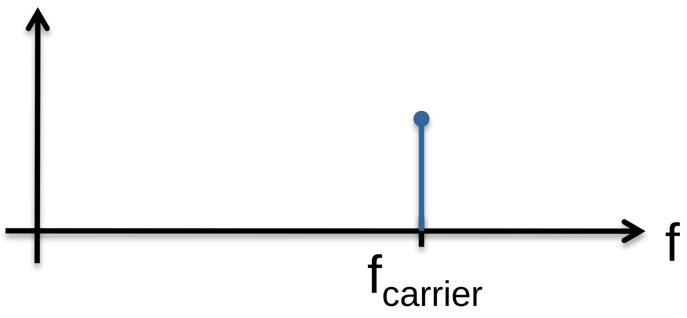
Reference: Weill, L.R. (2011) **Why is acquisition of GNSS signals generally more difficult than tracking and what are the limiting factors?**, Inside GNSS, 6(1), pp. 20-23

Benefits of GNSS Signal Structure (Spread Spectrum)

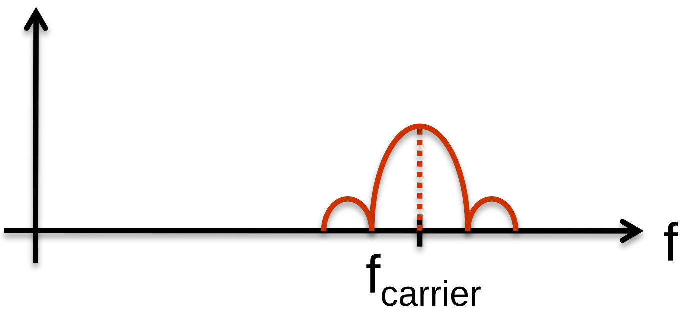
Spread Spectrum Systems

- Spread spectrum systems take a signal and deliberately spread it across a wider bandwidth in the frequency domain. Assuming the carrier phase is the “signal”, modulation (which is effectively multiplication) of the carrier by the ranging code spreads the signal, as shown below.

Spectrum of Carrier Only

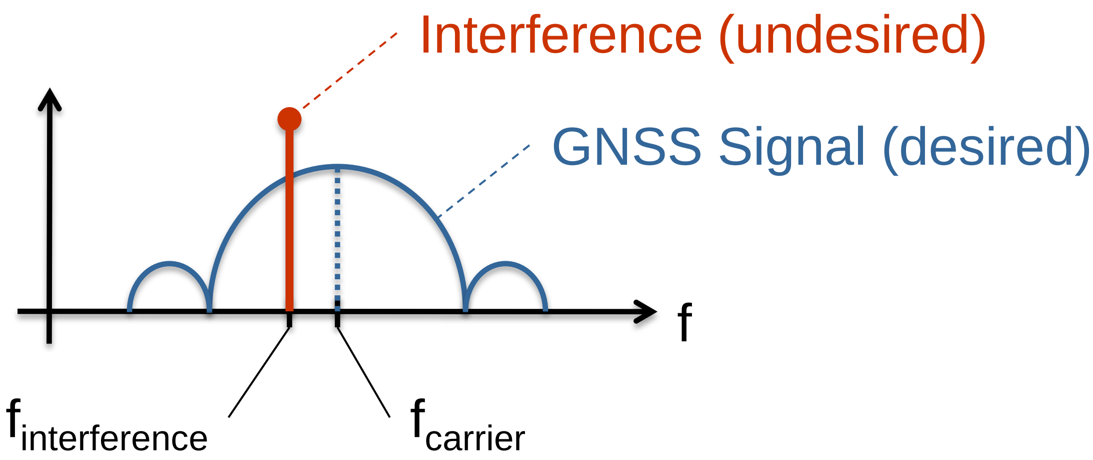


Spectrum of Carrier Modulated with (multiplied by) the Ranging Code



Resistance to Interference (1/3)

- In the context of GNSS, the spread spectrum signal structure provides some resistance to interference and jamming. Specifically, assume the incoming signal is corrupted by an interference signal, as below.



NOTE: In the above diagram, the interference is intentionally drawn to be "larger" than the desire signal.

Resistance to Interference (2/3)

- Recall that within the receiver, the local ranging code is multiplied by the incoming signal

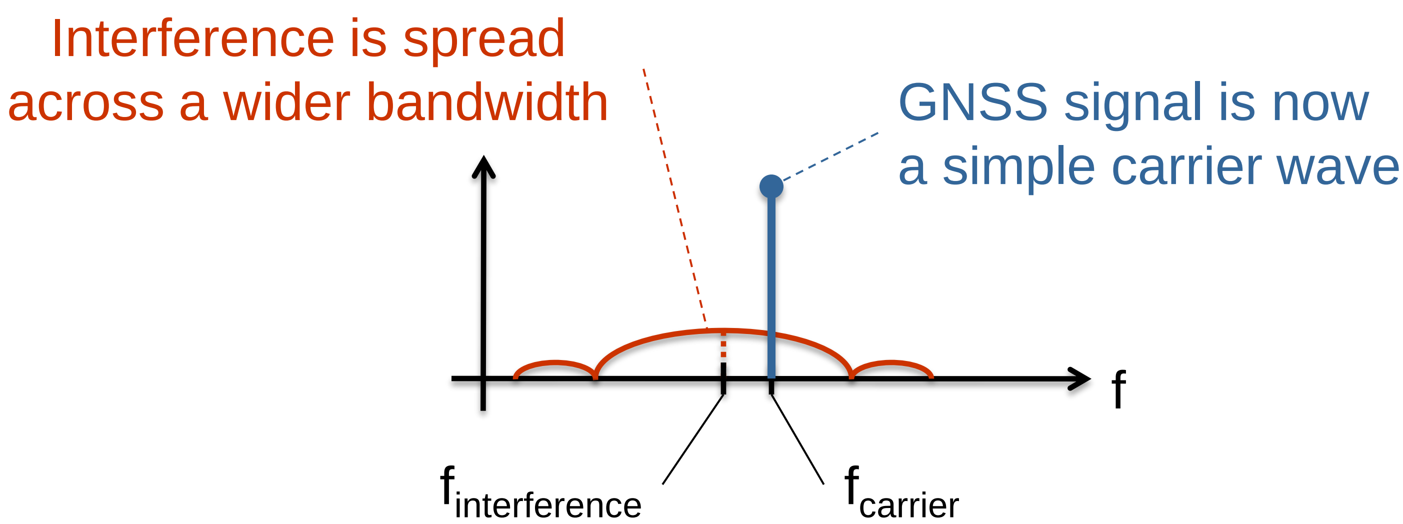
$$s(t)\times\hat{C}(t)=\left[C(t)\sin\left(f_{carrier}\cdot t\right)+\sin\left(f_{interference}\cdot t\right)\right]\times C(\hat{\tau})$$
$$=C(t)\sin\left(f_{carrier}\cdot t\right)\times C(\hat{\tau})+\sin\left(f_{interference}\cdot t\right)\times C(\hat{\tau})$$

- If the local and incoming codes are aligned, the two “cancel” (the product is unity) leaving only the desired carrier

- Second term is the multiplication of a sine wave by the ranging code. However, we already know that multiplying a carrier wave by a ranging code spreads the carrier wave across a wider bandwidth (this is the principle of spread spectrum systems).

Resistance to Interference (3/3)

- The result is that the desired carrier is recovered and the energy/power in the interferer is spread out over a wider bandwidth



NOTE: In the above diagram, the desired signal is now more pronounced than the interfering signal, thus illustrating how spread spectrum systems provide some resistance to interference and jamming.

Code Multipath

Code Multipath

- Multipath occurs when the signal arrives at the receiver via multiple paths, usually because of reflection. For a static receiver, the received signal can thus be represented as

$$s(t) = C(t)\sin(ft) + aC(t - Dt)\sin(ft - f\,Dt)$$

α represents a loss of power due to reflection ($\alpha < 1$)

Δt is extra time it takes the reflected signal to reach the receiver

- Assuming ‘f Δt’ is an integer number of cycles (i.e., the direct and reflected signals have the same phase), correlation of the incoming and local signals gives

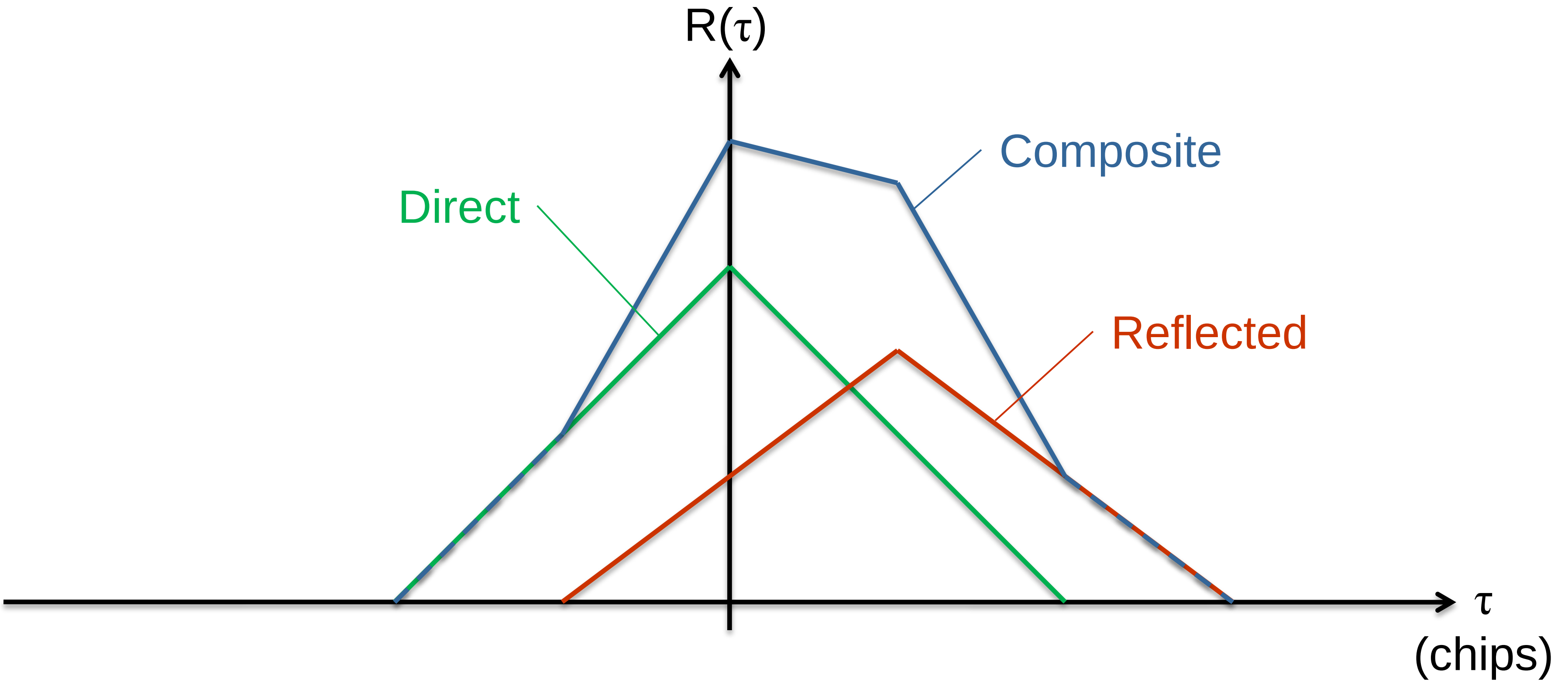
$$\frac{1}{T}\int_0^T \left[C(t)C(\hat{\tau}) + \alpha C(t-\Delta t)C(\hat{\tau})\right]dt = R(t-\hat{\tau}) + \alpha R(t-\Delta t-\hat{\tau})$$

Direct (desired) signal

Reflected (undesired) signal

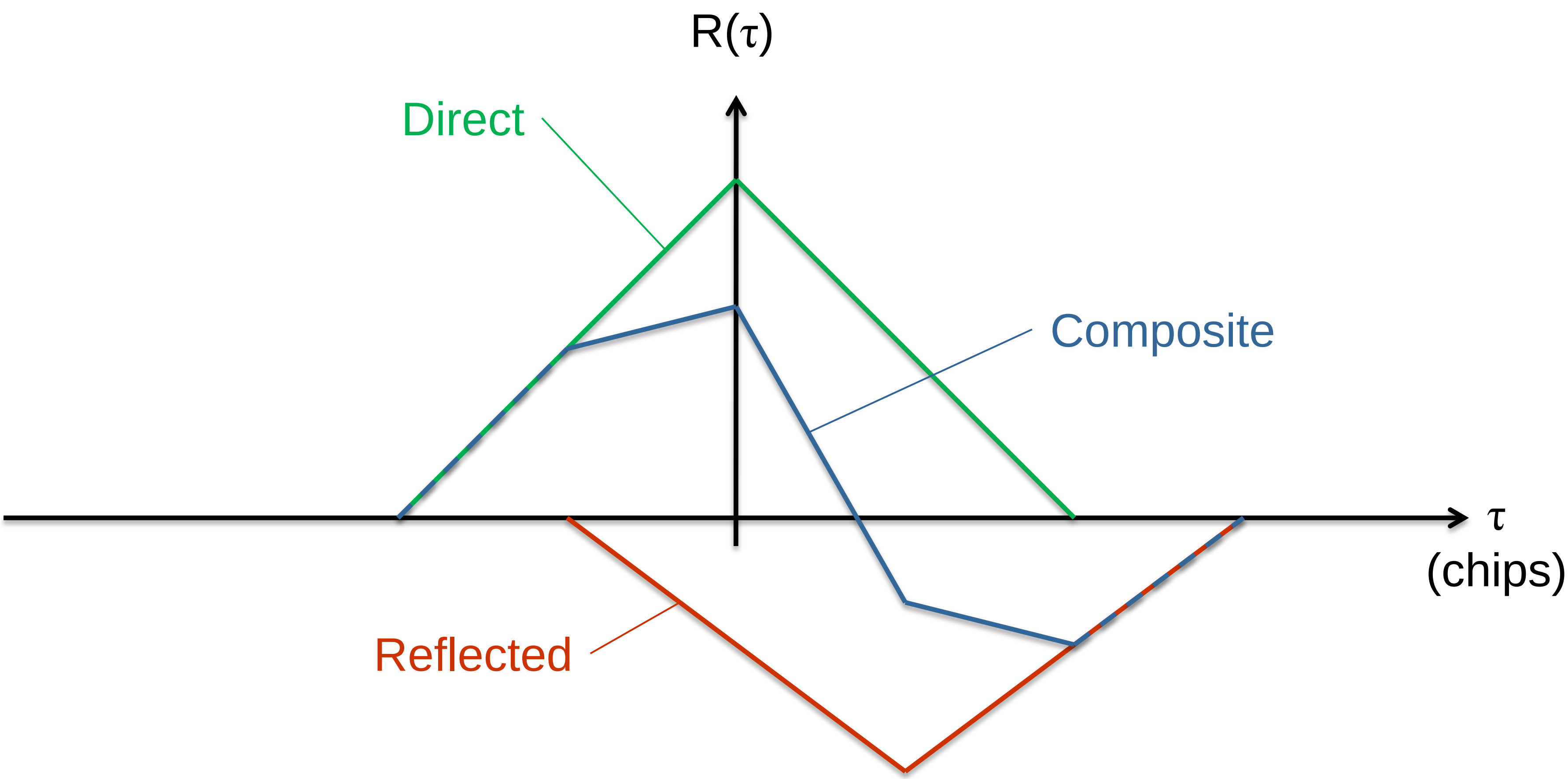
In-Phase Code Multipath

- ***In-phase*** multipath results when the direct and reflected signals have the same phase. Regardless, the receiver only sees the sum of the two signals, that is, the composite signal below.



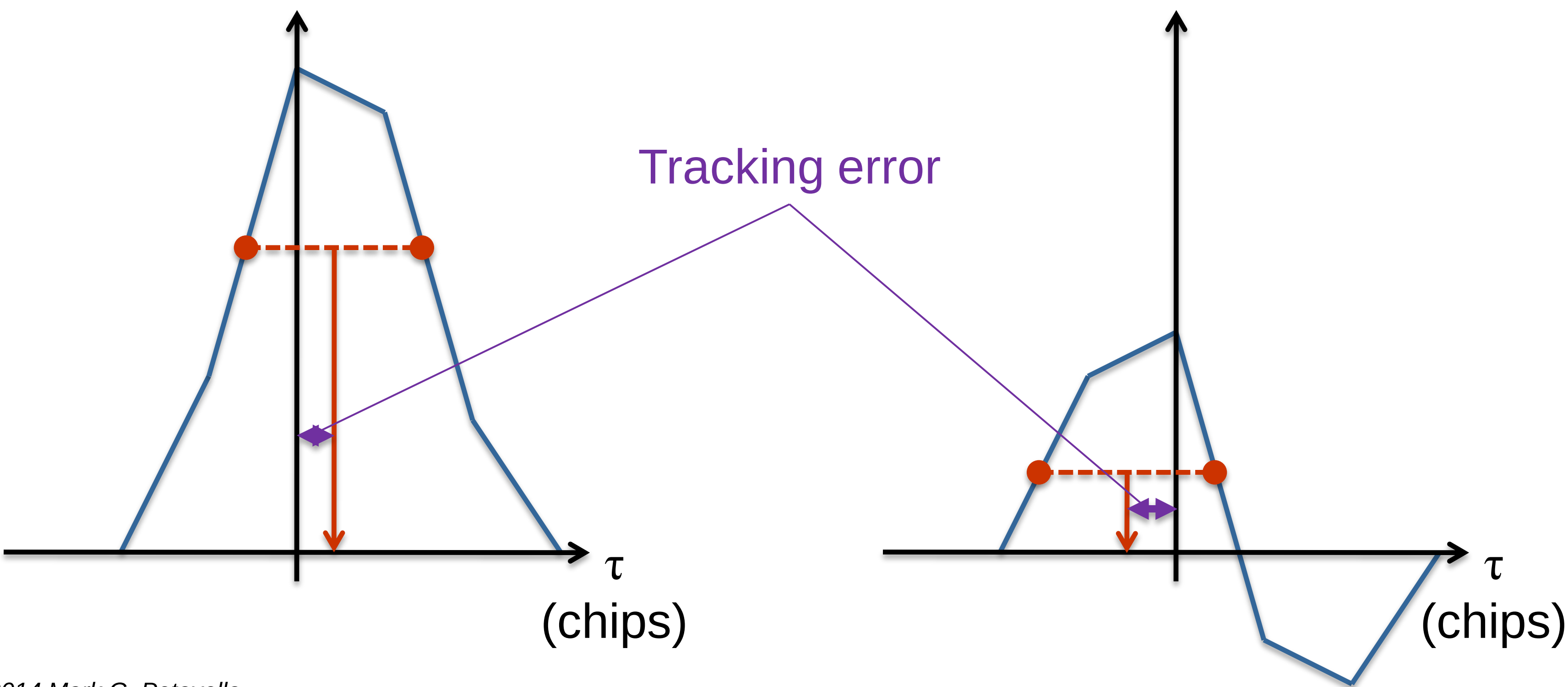
Out of Phase Code Multipath

- ***Out of phase*** multipath is when the direct and reflected signals have different phases. For a worst case of 180° phase difference, the following result is obtained.



Code Tracking with Multipath

- Recall that GNSS receivers try to balance/equalize the output of the early and late correlators and that the midpoint between the correlators is the "tracking point", which ideally should be zero.
- This approach introduces tracking errors in the presence of multipath, as shown below. Note that the diagrams only show the composite correlation functions.



Notes on Code Multipath

- As shown on the previous slide, multipath can make the signal appear longer or shorter. Multipath does not influence receiver for reflections longer than 1.5 chips.

- Finally, note that the x-axis on the auto-correlation plots is in units of code phase (or chips). This can be scaled to units of length by multiplying by the chip length. In other words, a shorter chip gives less multipath.
  - What does this say about the desired chipping rate of the signal?
  - What does this say about the relative performance of the P-code and the C/A-code in the presence of multipath?

GNSS User Equipment

GNSS User Equipment

- In terms of equipment, you usually need to select an ***antenna*** and a ***receiver***. Some newer receivers integrate both together, but for the purpose of this course, we will consider them separately.

- Selection of the equipment will vary greatly depending on your application. Some of the considerations include
  - Cost
  - Size & weight
  - Power consumption
  - Desired accuracy
  - Desired measurements (in some cases)
  - Operational environment (weather such as snow, sea water, etc.)

GNSS Receiver Categories

- In general, there are two categories of GNSS receivers

- Geodetic-grade receivers
  - Typically used for surveying or high-performance applications
  - Cost is one to several thousand dollars depending on capabilities
  - May include original equipment manufacturer (OEM) models (i.e., may not be a prepackaged product)

- Consumer-grade receivers
  - Much lower cost (<$1 to a few hundred dollars)
  - Wider range of applications (mostly due to cost)
  - *Chipsets* are a subset of this category where you only get the chip, not a prepackaged product

Geodetic-Grade Receivers

- Highest quality receivers and antennas
- Provide more accurate measurements
  - Provide advanced algorithms to mitigate multipath (especially pseudorange multipath)
  - Pseudorange measurement noise is typically on the order of cm-dm
- Receiver outputs carrier phase measurements, which are needed for high-accuracy positioning (i.e., cm-level)
  - May or may not provide carrier phase processing onboard in real-time (using float or fixed ambiguities; see Chapter 7 for details on this)
- May or may not track more than one frequency
  - L1 only
  - L1+L2 (see slides below for tracking P(Y)-code)
  - L1+L2+L5
- May be able to track satellites from multiple GNSS
- Generally high power consumption

Consumer-Grade Receivers

- Typically only uses/provides pseudorange measurements
  - Carrier phase measurements are not usually available
  - Position accuracy is typically dm to m-level at best
- Usually do not provide advanced algorithms to mitigate multipath or noise
  - Pseudorange noise may be as high as several metres
- May be able to track satellites from multiple GNSS but usually only a one frequency
- Lower power consumption

- Most consumer-grade receivers are also ***high-sensitivity*** receivers

High-Sensitivity GNSS Receivers (1/2)

- Recall that within the tracking loop, the incoming and local signals are correlated. Assuming this is done perfectly, the correlator output would be (using 1 ms of data)

$$\begin{aligned} \text{Prompt Correlator} &= \sum_{1\,\text{ms}} C(\tau)C(\tau) + \text{noise} \\ &= R(0) + \text{noise} \\ &= A + \text{noise} \end{aligned}$$

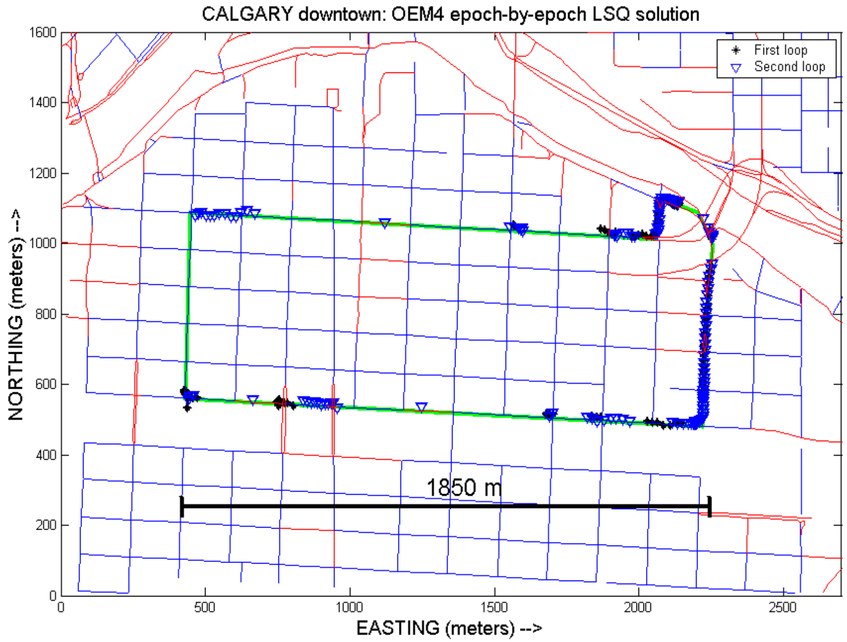
- It turns out that the power of the signal (i.e., A) increases linearly with the amount of data used to compute the correlation (e.g., 1 ms, 2 ms, 5 ms, etc.) whereas the noise power increases with the square-root of the amount of data

High-Sensitivity GNSS Receivers (2/2)

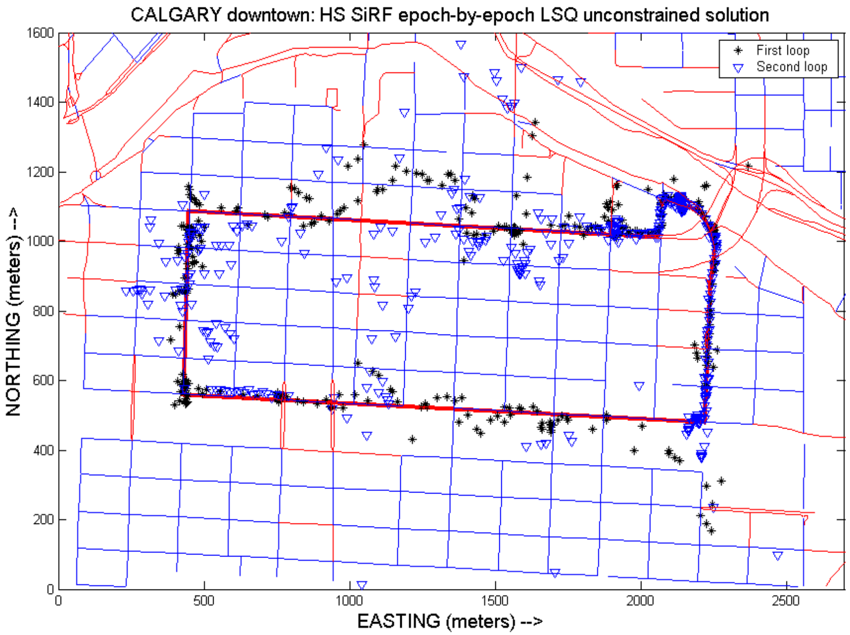
- Given the result of the previous slide, the more data used for the correlation, the weaker the signal you can track
  - There are several reasons why you cannot use an indefinite amount of data (not discussed here) and thus the ***sensitivity*** of a receiver – that is, the power of the weakest signal it can track – is ultimately limited
- As their name suggest, high-sensitivity receivers use advanced signal processing techniques in order to provide better sensitivity, that is, to track weaker signals
  - These types of receivers are need to compute your position indoors, for example
  - This increases the ***availability*** of the receiver (i.e., the amount of time it can provide a solution), but at the cost of ***reduced accuracy***
- Are there applications where you think this type of technology is useful? Justify your answer.

Example of High-Sensitivity GNSS Performance

"Standard" Receiver



High-Sensitivity Receiver



- Standard receiver gives few, but accurate, solutions
- High-sensitivity receiver given many, less accurate, solutions

Reference: Mezentsev, O., Y. Lu, G. Lachapelle and R. Klukas (2002) ***Vehicular Navigation in Urban Canyons Using a High Sensitivity GPS Receiver Augmented with a Low Cost Rate Gyro***, Proceedings of ION GPS 2002.

Software-Based GNSS Receivers

- Since the turn of the millennium, there has been considerable development in the area of software-based GNSS receivers wherein – after sampling the signal – all processing is done in software

- This provides a very flexible environment for improving receiver performance
  - Easily add new GNSS signals and systems
  - Design receivers for specific applications
  - Test new signal tracking algorithms

- Although this is largely a research approach, some commercial receivers provide some re-configurability

P(Y)-Code Signal Tracking

- Encryption of the P-code with the Y-code means that civil user cannot directly track the signal. However, some civilian receivers are able to track the L2 signal using advanced technology
  - All methods introduce a loss in signal-to-noise (SNR) ratio relative to direct P-code tracking; the *best case* loss is 14 dB (~25-fold!)
  - These technologies are very complicated and receiver manufacturers charge a premium for them (dual-frequency receivers are therefore much more expensive than L1-only receivers)
- Practical impact is poorer quality measurements on L2
  - More likely to experience cycle slips on L2 (more on cycle slips later)
  - Receiver is more likely to lose lock on L2 signals
  - L2 measurements are noisier than L1 measurements (mostly for pseudoranges)
- The above does <u>**not**</u> apply to the new L2C signal, since this signal is freely available to civilian users

P-Code Generation

- The P-code is 266.4 days long and is divided into 38 different 7-day sequences; one per satellite. The P-code itself is unclassified

- Before being transmitted, the P-code is encrypted with the classified Y-code to yield the P(Y)-code
  - Ideally, users need to know the Y-code to track the signal
  - Military receivers have special chips/keys that allow them to track the P-code
  - Civil users have to use special techniques to track the P-code

- The denial of the P-code is called ***anti-spoofing*** (AS)
  - Intended to limit accuracy and improve robustness to jamming

GNSS Receiver Characteristics

- Selection of the receiver primarily involves selecting from amongst the different technologies/capabilities available. These include
  - **Number of frequencies**: Single frequency receivers are lower in cost but are not typically used for cm work.
  - **Number of systems**: GPS+GLONASS integrated receivers are becoming more prominent and some multi-system receivers are already for when other GNSS come online. Tracking satellites from more systems improves accuracy and reliability, but generally costs more (especially for high accuracy applications)
  - **Accuracy**: What is the quality of measurements and the corresponding position? Of importance here is whether or not some form of advanced correlator strategies are employed to reduce multipath.
  - **Re-acquisition time**: How well does the receiver perform in environments where views of the sky are periodic or short lived?
  - **Time to first fix**: How long does it take the receiver to compute a solution once it is turned on?
  - **Flexibility**: Is the receiver a "turn key" system or is it intended to be integrated in custom products? What is the size and power consumption? What type of case is available?

GNSS Antenna Characteristics

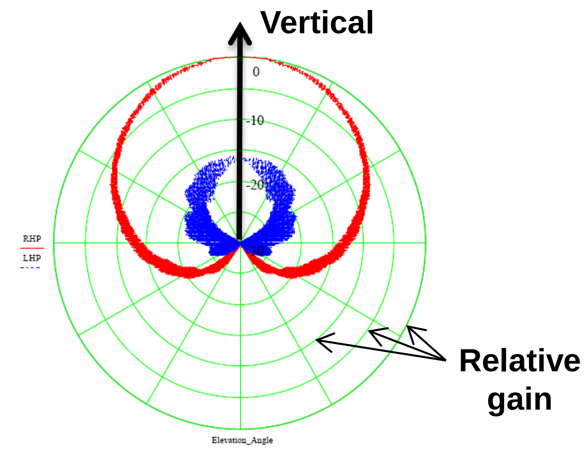
- Physical size and shape
- Ruggedness and resistance to weather
- Frequency sensitivity
  - L1 only vs. L1+L2 vs. multiple frequencies
- Radiation pattern (gain pattern)
  - Defines an antenna's sensitivity over some range of elevation and azimuth
- RHCP vs LHCP
- Phase centre stability
  - The phase centre is the point to which estimated position refers, and this point differs for satellite signals arriving from different directions
- Variations between units of the same design and/or manufacturer
- Multipath mitigation capability
  - Ability of the antenna to discriminate the direct signal from reflected signals

Antenna Gain Pattern

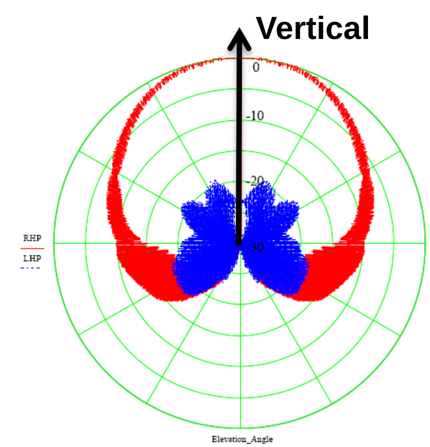
- The gain pattern defines how sensitive an antenna is in particular direction

- The examples to the right are from a typical high-quality GNSS antenna.
  - As you can see the antenna sensitivity is directed vertically
  - What benefits can be realized from this? Are there instances were this might not be desirable?

Example L1 Gain Pattern



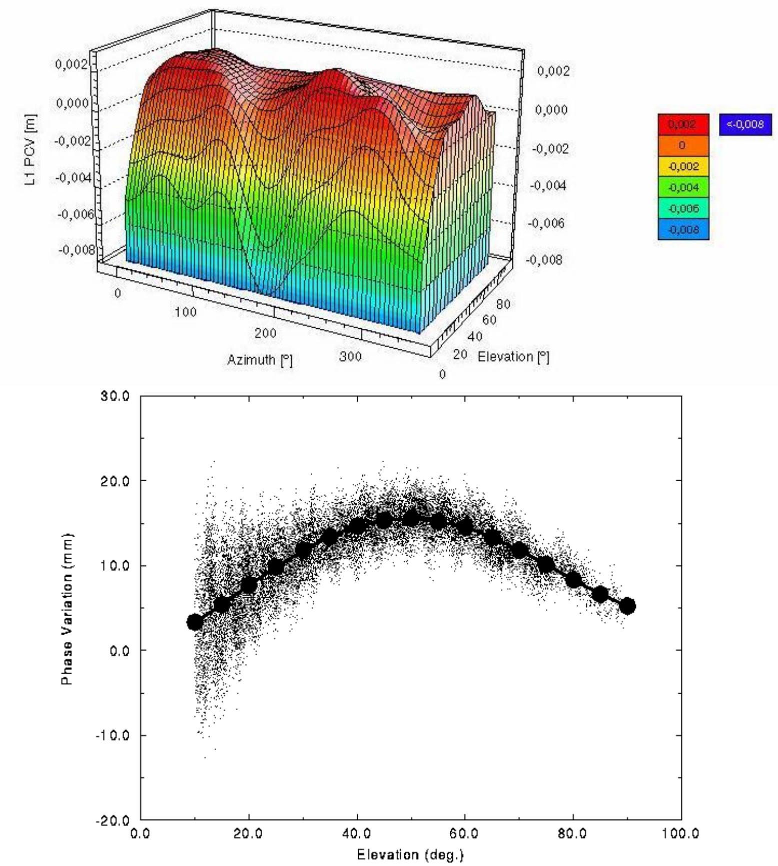
Example L2 Gain Pattern



Red is for RHCP signals
Blue is for LHCP signals

Antenna Phase Centre

- The antenna phase centre is not a physical point, but a virtual point that usually cannot be accessed with a tape measure. The exact location of the phase centre needs to be known to relate the estimated position to the point of interest. For dual frequency data, the phase centre may be different for L1 and L2.
- Ideally an antenna's phase centre is independent of the satellite signal's direction of arrival but in reality, the phase centre will vary depending on satellite elevation and azimuth. This can be particularly important when using different antenna types.



Example of Antenna Phase Center Variation

(Trimble 4800: Solid line is polynomial fit)

See **http://www.ngs.noaa.gov/ANTCAL** for the NOAA antenna calibration website

GNSS Antenna Quality

- The quality of an antenna can impact the performance of your system
- High quality antennas
  - Can help to mitigate multipath but usually at the expense of a relatively large size (typically 15 cm diameter or larger)
  - Sharp gain pattern roll-off near horizon
  - Can be quite expensive (>$1,000)
  - Phase centres for all frequencies are usually close to each other (and close to the physical centre of the antenna)
- Low quality antennas
  - Typically very small
  - Typically large gain patterns (i.e., not directed)
    - As a result, signal tracking may be a bit more difficult

Time Scales

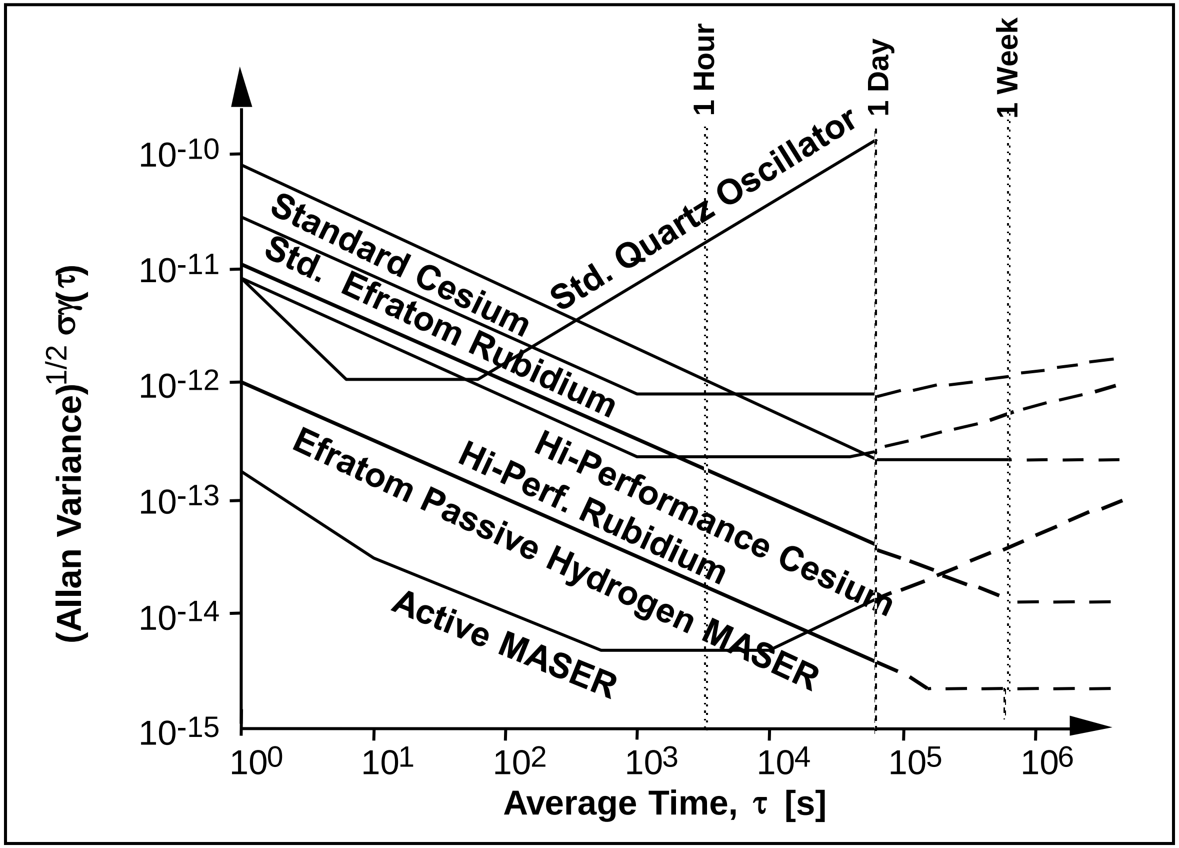
Atomic Time Scales

- GNSS systems are based on atomic time scales because of their extremely high stability
- UTC (Coordinated University Time) is the atomic approximation to the Earth's actual rotation (after accounting for polar motion)
- GPS itself also provides an atomic time scale determined by the satellite clocks plus five ground based atomic clocks. It is inherently transmit by every GPS satellite.
  - GPS time was coincident with UTC at the GPS standard epoch of January 6, 1980 (0 hours)
  - No integer leap seconds are introduced into GPS time scale, but leap seconds are introduced in UTC
    - GPS time = UTC + Correction
    - As of Jan 2014, the correction is 16 s but current information is available from http://leapsecond.com/java/gpsclock.htm

Reference: Beard, R. (2009) **The Future of the UTC Time Scale**, NAVIGATION Journal of the Institute of Navigation, Vol 56, No 1, pp. 1-8.

Frequency and Time Standards

- A frequency and time standard (FTS) is characterized by its ***Allan variance (AVAR)*** or ***Allan standard deviation***, which is a statistical measure to characterize the stability of FTS over an interval of time



This graph is intended to give a <u>*general*</u> idea of the relative performance of different types of oscillators. Large variations are possible even for a given type of oscillator.

AVAR and Clock Bias Prediction

- Allan Variance is not always intuitive, so consider a scenario where you want to predict a clock offset forward in time. The accuracy of the prediction is a function of the interval over which you are predicting (i.e. $t - t_0$) and the Allan variance of the oscillator.

$$\underbrace{\sigma^2_{Clock}(t)}_{\text{Clock uncertainty at time } t} = \underbrace{\sigma^2_{Clock}(t_0)}_{\text{Clock uncertainty at time } t_0} + (t - t_0)^2 \cdot \underbrace{\sigma^2_y(t - t_0)}_{\text{Allan Variance for time interval } \tau = t - t_0}$$

AVAR and Clock Bias Prediction – Example

- Assume you have a quartz clock with an Allan standard deviation of $\sigma_y(1\ \text{day}) = \sim 10^{-10}$ (see graph a couple slides back). Further assume you have perfect initial synchronization
  - $\sigma_{Clock}(t_0) = 0$

- What is the timing uncertainty after 10 seconds, 10 minutes and 1 hour?

- Why might you want to predict a clock offset forward in time?