Global Positioning System: Basic Taxonomy and Modeling

Luca Carlone

July 25, 2014

The Global Positioning System (GPS) is a cheap, widely available, and accurate localization technology provided by the U.S. government. This documents recalls basic concepts related to GPS taxonomy and modeling. The document is largely based on [2], where the reader can find an exhaustive treatment of the topic.

1 Taxonomy

A GPS system includes 3 *segments*:

- *space segment*: satellites vehicles (SV) orbiting in six 12-hour orbiting planes. SV are not in geosynchronous orbits. The satellites transmit using 2 carrier frequencies, called L1 and L2.
- *control segment*: 6 control stations + 1 master station. They compute orbital models and clock correction parameters, that are then broadcasted to the satellites.
- *user segment*: antenna and receiver. Receivers include 3 stages: radio-frequency front-end (amplification, filtering and shifting), baseband, and navigation portion. The output of this segment is an estimate of position, velocity, and time of the receiver.

GPS provides 2 levels of service:

- *standard positioning system* (SPS): since the 1993 this provides positioning service to all users, using the L1 frequency through the course acquisition code and a navigation data message. Navigation message includes satellite orbit and clock correction.
- *precise positioning system* (PPS): only for personnel authorized by U.S. government, and controlled using a technique called *anti-spoofing*.

There exists two types of GPS, with different levels of accuracy (Figure 1):

- *standard GPS*, which includes the majority of commercial GPS receives. Typical positioning errors are in the order of few meters;
- *differential GPS* (DGPS), which uses a base station to correct the common-mode errors (sometimes private or public sources make available the corrections, making unnecessary to setup a local station). DGPS can reach errors of $0.1 - 3.0$m for differential pseudorange, or few centimeters for carrier phase differential GPS.

GPS is a line-of-sight system. It cannot operate indoors, underwater, or under significant foliage.

GPS Mode	Typ. Range STD. (m.)	Min. Range STD. (m.)
C/A Code, SA off	9.0	8.0
C/A Code DGPS	3.0	0.1
L1 Phase DGPS	0.10	0.01

Figure 1: Range accuracy, comparing standard GPS (first row, with further details reported in the left column of Fig. 2), differential (pseudorange) GPS (second row), and carrier phase differential GPS (last row). Figure is from [2], pag. 333. From the measurements errors one can infer the positioning error using the *dilution of precision* (Section 5).

Pseudorange Sub-meter accuracy requires to measure time with error in the order of few nanoseconds. This would be too expensive to have on commercial receivers, so GPS time measurements are affected by an unknown time offset. For this reason, a GPS measurement from a satellite is called *pseudorange* (range + biases in user and satellite clocks). To compensate this issue, the GPS system estimates the time offset together with the position.

2 Measurement model and position estimation

If p is the position of the receiver and p^s is the position of satellite s, the *true* pseudorange is:

$$\rho_r^s = \|p - p^s\| + c\Delta t_r \tag{1}$$

where c is the speed of light and Δt_r is the time offset of the receiver.

The *measured* pseudorange is:

$$\bar{\rho}_r^s = \rho_r^s + \chi^s \tag{2}$$

where χ^s is the measurement error, called the *User Equivalent Range Error* (UERE) or *User Range Error* (URE). The URE is modeled as:

$$\chi^s = e_\rho^s + \eta_\rho^s \tag{3}$$

where

- e_ρ^s is the **common mode error**:

$$e_\rho^s \doteq E_s + c\delta t^s + \frac{f_2}{f_1} I_r^s + T_r^s, \qquad \text{and} \tag{4}$$

  - E_s is the *ephemeris* error induced by the inexact knowledge of satellite position (due to the space and control segment)
  - δt^s is the residual satellite clock error (due to the space and control segment)
  - f_1 (f_2) are the L1 (L2) carrier frequency
  - I_r^s is the error due to the *dispersive* atmosphere effects (ionosphere)
  - T_r^s is the error due to the *non dispersive* atmosphere effects (troposphere)

- η_ρ^s is the **noncommon mode error**:

$$\eta_\rho^s \doteq M_\rho^s + \nu_\rho^s, \qquad \text{and} \tag{5}$$

- M^s_ρ models multipath errors (this may change between L1 and L2, and it is connected to the siting of the antenna)
- ν^s_ρ models random measurement noise (this may change between L1 and L2).

Usually one assumes that the measurements from different satellites are independent and that:

$$cov(\chi^s) = \sigma^2 I \tag{6}$$

Position Estimation Given 4 pseudorange measurements (2), receiver position and time offset $x = [p^\top c\Delta t_r]$ is computed via standard nonlinear least squares. Usually, the initial guess is the last position estimate (when available), or all variables are initialized to zero ($x^{(0)} = \mathbf{0}_4$).

Observability To observe a 3D position it would be enough to have 3 range measurements (3 spheres intersect at 2 points, one of which is not admissible, being far from Earth). However, since we have pseudoranges, we have to estimate one more quantity (time offset) and we need 4 satellites measurements.

Position Estimation Error Position estimation error has significant time correlation. This depends on the fact that some error sources (SV clock, ephemeris error, ionospheric and tropospheric errors) are slowly time varying. Details are given in Section 4 and 5.

Reference Frames Satellite positions is expressed in ECEF. Since ECEF is a rotating frame, then satellite positions at different time steps are in different reference frames, and need to be converted. Conversion is straightforward and it is discussed in [2], pag. 274.

3 How is the pseudorange measured in practice?

Each satellite broadcasts a sinusoidal signal, with amplitude depending on the PRN code assigned to the satellite, and with frequency depending on the carrier frequency. The receiver generates a local replica of this signal and searches for a phase offset and a Doppler frequency to align the received signal and the local replica. During tracking, the alignment error is small and enables a positioning error < 3m.

Pseudorange observable The pseudo range is computed as:

$$\Delta\tilde{t}^t_r = \tau_r - \tau_s, \qquad \bar{\rho}^s_r = c\Delta\tilde{t}^t_r \tag{7}$$

where $\Delta\tilde{t}^t_r$ is the interval between the instant in which the message was sent by the satellite (τ_s), and the instant in which it was received (τ_r). The pseudorange $\bar{\rho}^s_r$ is then computed by multiplying $\Delta\tilde{t}^t_r$ by the speed of light.

Carrier phase observable As mentioned before, the receiver also measures the phase shift ϕ^s_r in terms of number of cycles of offset. This is usually called *phase observable*. From the phase observable one can also compute a range measurement by multiplying the phase observable by the wavelength λ of the carrier:

$$\lambda(\phi^s_r - N_i) = \rho^s_r + e^s_\phi + \eta^s_\phi \tag{8}$$

where the common-mode errors are $e^s_\phi \doteq E_s + c\delta t^s - \frac{f_2}{f_1} I^s_r + T^s_r$ (same as e^s_ρ but with a minus sign); N_i is an unknown *integer phase ambiguity*, that has to be solved to compute the range estimate (e.g., using the LAMBDA method, see [1] and the references therein). The advantage of the phase observable is that the noncommon-mode error η^s_ϕ is usually smaller than 1% of the corresponding pseudorange error η^s_ρ. Therefore, after compensating the common mode errors e^s_ϕ, the phase observable enables positioning with centimeter accuracy.

Delta pseudorange observable (Doppler observable) The (satellite) carrier frequency f_s and the measured frequency at the receiver f_r are related via the Doppler frequency:

$$f_r = f_s \left(1 - \frac{\dot{r}}{c}\right) \tag{9}$$

where $\dot{r}$ is the rate of change of the range between the receiver and the satellite. In practice, the average rate of change of the pseudorange over the time interval T, called the *delta pseudorange observable*, is computed from differences of the carrier phase observables

$$\Delta\rho^s_r(\tau_r) = \lambda\left(\phi^s_r(\tau_r) - \phi^s_r(\tau_r - T)\right) \tag{10}$$

Further details are reported in in [2], pag. 280, while, for our discussion, we only mention that the Doppler observable allows observing the rate of change of the state vector x.

4 Measurement error model

Here we are interested to describe the URE (User Range Error). We consider each source of error in the following sections.

4.1 Common mode errors (δt^s, I^s_r, T^s_r, E_s)

Common mode errors are error sources that are common to all receivers operating within a limited ($\approx$ 15km) geographic region. These are slowly time-varying error sources with significant short-term correlation. They can be corrected when using differential GPS.

Clock errors (δt^s) In eq. (7) we used the measured time at the sender τ_s and at the receiver τ_r. In general these two measured times relate to the *true* time t (at which the signal is received) via

$$\tau_r = t + \Delta t_r, \qquad \tau_s = (t - \Delta t^s_r) + \Delta t_s \tag{11}$$

where τ_s is a noisy measurement of the true time $(t - \Delta t^s_r)$ at which the satellite transmitted the signal, and τ_r is a noisy measurement of the true time t at which the signal arrived at the received (clearly, Δt^s_r is the true, unknown signal propagation time); the errors Δt_s and Δt_r represent the sum of an initial offset and a clock drift. This error depends on the quality of the clocks.

The satellite clock error Δt_s is usually corrected by the control segment. The remaining error after correction is δt^s, and affects the pseudorange measurements as per eq. (4) (term $c\delta t^s$).

The receiver clock error Δt_r is estimated together with the position (it is included in the state vector x) and for this reason it is not included among the error terms. In general, there are three ways to compensate Δt_r:

- *single difference* across satellites: by computing the difference between 2 pseudoranges from 2 different satellites, the term Δt_r disappears and one can write a new measurement model using those differences as a function of the position of the receiver only.
- *double differences* across satellites: details are reported in [2], pag. 323. The use of double differences further reduces the effect of residual clock errors, but it creates correlation among pseudorange measurements.
- *clock bias dynamic modeling*: this includes Δt_r in the state vector and assumes that it evolves as
$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + noise, \qquad \Delta t_r = v_1 \tag{12}$$
which means that Δt_r is the integral of a random walk (called *frequency error*) plus a random walk. In practice the shape of this signal is similar to a ramp. Details on how to describe the noise in (12) is given in [2], pag. 285.

Atmospheric errors (I_r^s, T_r^s) Ideally, the signal sent by the satellite to the user travels in the vacuum at the speed of light c. In practice, the signal crosses the atmosphere (ionosphere and troposphere), which alters (slows down) the speed of propagation. The speed reduction is independent on the frequency of the signal for the troposphere (non-dispersive mean), while it depends on the frequency for the ionosphere (dispersive mean). This atmospheric effect is minimized when the satellite is at the zenith (max. when it is near the horizon).

The *ionosphere* (50-1000km of altitude) contains free electrons and charged molecules. The level of ionization depends on the solar activity, season, and time-of-day, and this in turn affects the signal speed. This induces an error in the pseudorange. This error can be partially compensated (e.g., reduced by 50%) with specific models or by using differential GPS (details in [2], pag. 290). In particular, since it is frequency dependent, it can be alleviated when using multiple frequencies (L1 and L2); see [2], pagg. 306-312. The remaining error after correction is denoted as I_r^s in (4).

The *troposphere* ($<$ 50km of altitude) contains electrically neutral particles. Temperature, pressure, and humidity in this region change, depending on the weather. Tropospheric error can be large ($\approx$ 30m) for satellites with low elevation and depends on local variables (receiver altitude, satellite elevation, etc.). This error can be alleviated with specific models or by using differential GPS (details in [2], pag. 291). The remaining error after correction is denoted as T_r^s in (4).

Ephemeris error (E_s) This error is due to the use of the estimated satellite position $\hat{p}^s$ in place of the true one p^s in (1)-(2). This error is typically $<$ 2m.

4.2 Noncommon mode errors (M_ρ^s, ν_ρ^s, N_i)

Noncommon mode errors are specific of each receiver.

Multipath (M_ρ^s, or M_ϕ^i when using carrier phase observable) These errors are due to the fact that the signal may reach the transmitter via multiple paths, causing constructive of destructive interference and altering the shape of the signal. Nominally, multipath errors are in the range $0.1 - 3.0$m (exceptional cases $\approx$ 100m) depending on antenna design and siting, and satellite elevation (the higher the better). For stationary receivers, these are systematic errors. Typically, for L1 carrier they are $<$ 5cm.

Random receiver noise (ν_ρ^s, or ν_ϕ^i when using carrier phase observable) These errors model various factors inside antenna, cabling, and receiver, affecting the measurements. For instance, they mode thermal fluctuations, extraneous RF signals, quantization, and sampling. They are usually modeled as white noise (independent among satellites).

Integer ambiguity (N_i, only for carrier phase observable) This is connected to the cycle ambiguity in aligning the signal from the satellite to the local signal (roughly speaking, the signal is periodic and one can be wrong if does not count correctly the number of periods or cycles). During *phase lock* the cycle ambiguity can be solved accurately, while the capability of detecting the loss of phase loss depends on the signal to noise ratio (details in [2], pag. 300).

Selective Availability (SA, now turned off) These errors were purposely added by the U.S. government to degrade position accuracy. Now these errors are turned off.

A summary of the errors (and typical standard deviations) in pseudorange measurements is given in Figure 2. Note that these are measurement error, while the user is usually interested in the *positioning error*, which is the argument of the following section.

Common Mode Errors	L1 C/A, σ, m	L1/L2, σ, m
Ionosphere	7-10	–
Troposphere	1	1
SV Clock	2	2
SV Ephemeris	2	2
Non-common Mode Errors		
Receiver Noise	0.1 - 0.7	0.1 - 0.7
Multipath	0.1 - 3.0	0.1 - 3.0
URE	8-11	3-4

Figure 2: Typical error standard deviations for **pseuderange measurements**. For phase observables, common mode errors are the same, while noncommon mode errors are 2 orders of magnitude smaller. Figure is from [2], pag. 301. From the measurements errors one can infer the positioning error using the *dilution of precision* (Section 5).

5 Positioning error model

The concept of *dilution of precision* (DOP) relates the measurement error (URE) to the positioning error, which is the one the user is usually interested in.

As discussed in Section 2, the position estimate is obtained via nonlinear least squares. Accordingly, it is possible to show that the estimate is unbiased and has covariance matrix:

$$P = \sigma^2 \left(H^\top H\right)^{-1} \tag{13}$$

where H is the Jacobian of the measurement model and σ is the same defined in (6).

Warning The unbiasedness of the estimate means that if we average the estimate for a long enough time, the error should converge to zero. However, The short term average shows significant correlation (due to δt^s, I_r^s, T_r^s, E_s, M_ρ^s) implying a **bias in the estimate**.

Note that $\left(H^\top H\right)$ is assumed invertible; the case in which it is not invertible is called *GPS chimney*, and corresponds to cases in which the position becomes unobservable from the measurements. This phenomenon is not common nowadays as receiver usually tracks more than 4 satellites, and using all the measurements results in a well-behaved P.

The matrix $G = \left(H^\top H\right)^{-1}$ is determined by the geometry and number of available satellites (hence it is time varying). G defines the DOP, which can be defined in different ways, e.g.:

$$VDOP = \sqrt{G_{33}}, \quad HDOP = \sqrt{G_{11} + G_{22}}, \quad PDOP = \sqrt{G_{11} + G_{22} + G_{33}}, \quad GDOP = \sqrt{\mathrm{tr}(G)} \tag{14}$$

Essentially, the expressions of the DOP quantify how the URE is magnified in the vertical direction (VDOP), horizontal directions (HDOP), etc. Typical (standard) receivers enable positioning with a nominal error < 2.5m.

6 Additional information

Civilian users are able to track L1 signals; modern receivers are also able to use L2 signals, while the latter can be fully exploited only by authorized personnel. A GPS modernization phase is in progress, and this will give (before 2015) to civilians a full use of L2 signals plus 2 additional carriers. This will improve positioning availability, integrity (by using protected aviation frequencies to reduce interference), and accuracy (better compensation of ionospheric effect, and more effective solution to integer ambiguity).

The GPS is provided by the U.S. government, while other positioning systems do exist, including the European GALILEO and the Russian GLONASS.

References

[1] X.W. Chang, X. Yang, and T. Zhou. MLAMBDA: a modified lambda method for integer least-squares estimation. *Journal of Geodesy*, 79(9):552–565, 2005.

[2] J.A. Farrell. *Aided Navigation: GPS with High Rate Sensors.* McGraw-Hill, 2008.