

Chapter 2

Introduction to the Global Positioning System

In Chap. 1 we introduced some of the working principles of the GPS by analogy with a simple linear model. Now its time to move into more detailed discussion of the GPS satellite system and how it works. In this chapter we will cover the basics of how the satellites are configured, making the path delay measurements, a quick look at the user position solution equations, and a high-level block diagram of GPS receiver will be presented.

2.1 The Satellite System

GPS is comprised of 24 satellites orbiting the earth at a distance of 20,000 km as measured from mean sea level. See Fig. 2.1. Figure 2.2 shows a more detailed picture of a user getting position information from four satellites. Each satellite contains a very precise clock. All the clocks and hence all the timing signals associated with each satellite are in near-perfect synchronism. The basic principle of operation is similar to the Roadway Distance Measurement example of Chap. 1. A rough model is to imagine a system of satellites emitting a “pulse” of light at exactly the same instant as all the others (Important: GPS is NOT a pulsed system). If you were at the exact center of the earth and the satellites were in perfect circular orbits (GPS orbits are nearly circular) these pulses would all arrive at your receiver at the exactly same instant. For all other points the pulses would (typically) arrive at different times. If one has knowledge of the positions of the satellites and can measure the path delay (distance) from at least four satellites to the user receiver, the user position X, Y, Z in ECEF (Earth Centered Earth Fixed) coordinates and user clock bias can be solved for.

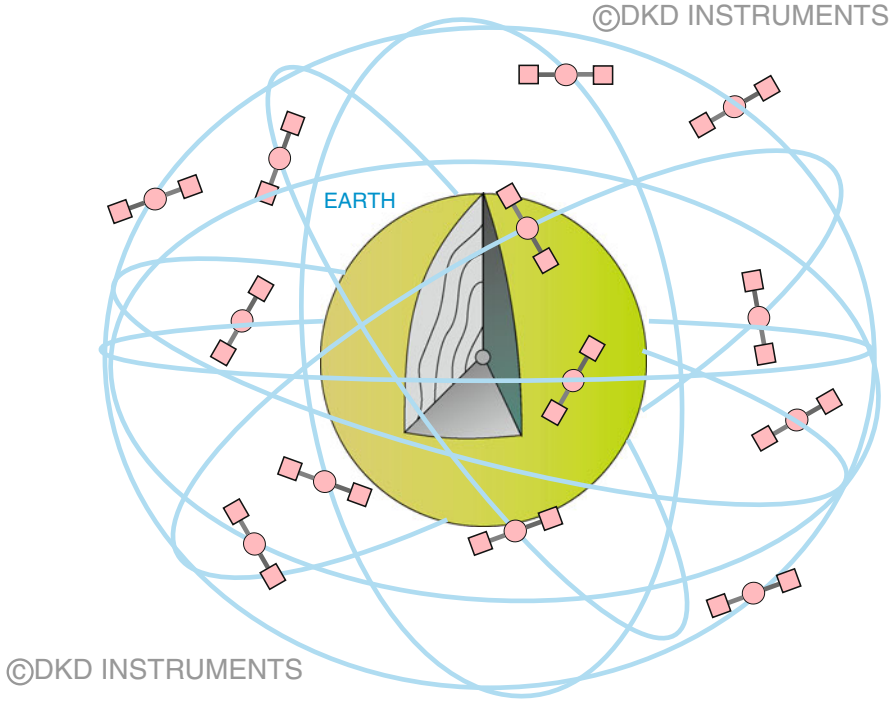


Fig. 2.1 The earth and its constellation of GPS satellites

2.2 Physical Constants of a GPS Satellite Orbit That Passes Directly Overhead

Figure 2.3 shows a simple model of one satellite circling earth directly overhead at the user zenith. We will use this simple model to compute some constants associated with a user at or near the earth's surface. We will assume that the user can be at a maximum altitude above sea level of 10 km. The minimum altitude will be assumed to be sea level. With this information and the known mean diameter of a GPS orbit we can calculate the minimum and maximum distance to the satellite as it passes overhead. Figure 2.3 shows the elements used in this calculation and other results. From this information we find that the maximum distance a GPS satellite is 25,593 km. The minimum distance is found to be 20,000 km. These distances correspond to path delays of 66 and 85 ms, respectively. As we have seen in Chap. 1, we can use this knowledge of maximum and minimum distance to set the receiver's reference clock to approximate GPS time.

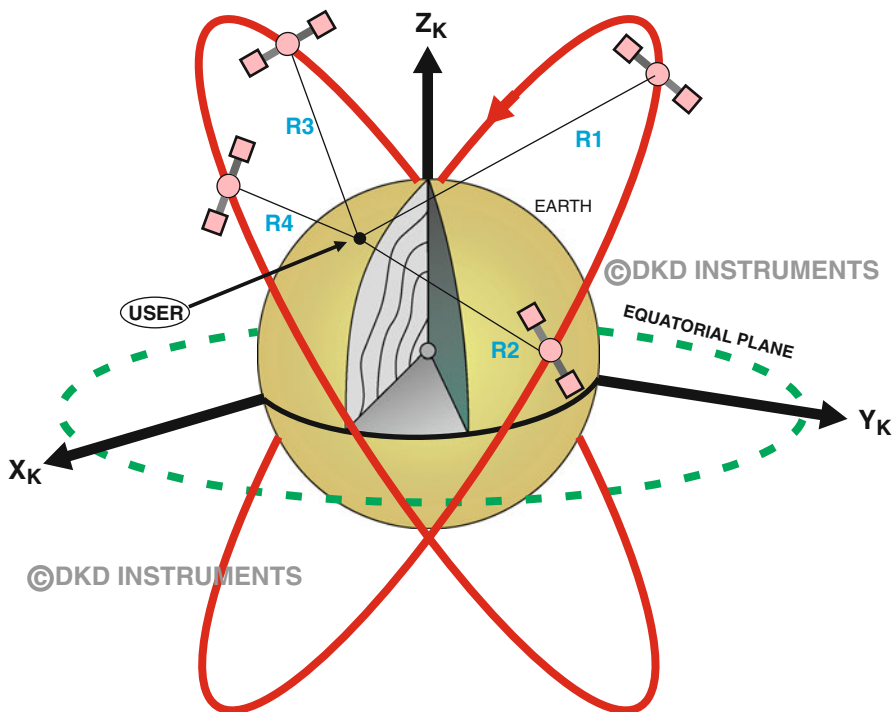


Fig. 2.2 User at earth's surface using four SV's for position determination

2.3 A Model for the GPS SV Clock System

Each GPS SV has its own clock that “free runs” with respect to the other clocks in the system. GPS uses a “master clock” method in which *all* the clocks are “referenced” to the master clock by the use of error terms for each SV clock. We discussed this already in Chap. 1. In this chapter we will use the clock model of Fig. 1.7 *minus* the outside second-counter dial. In addition to the omitted second-counter dial, we will not show the dial that indicates the week of the year, year dial, etc. We have not discussed these new dials but they are present in the time-keeping method employed by GPS. For the purpose of discussion and understanding GPS we will often use the clock model of Fig. 1.7 minus the dials above the 0–1 s time increment. If we show or discuss a SV or Receiver replica clock without *all* the dials, the reader will assume the missing dials are “present” but not shown for reasons of clarity.

The smallest time increment dial of our clock model will always be the 0–0.977 μs dial. As mentioned in Chap. 1 this dial has a very fine “effective” resolution as used in the SV. But this statement is an approximation. In reality the 0–0.977 μs dial is the one dial of all the clock dials in our model that does not have a direct physical counter part in the “true” SV clock. In other words, our clock model at the sub microsecond level is not a completely accurate model of the GPS clock.

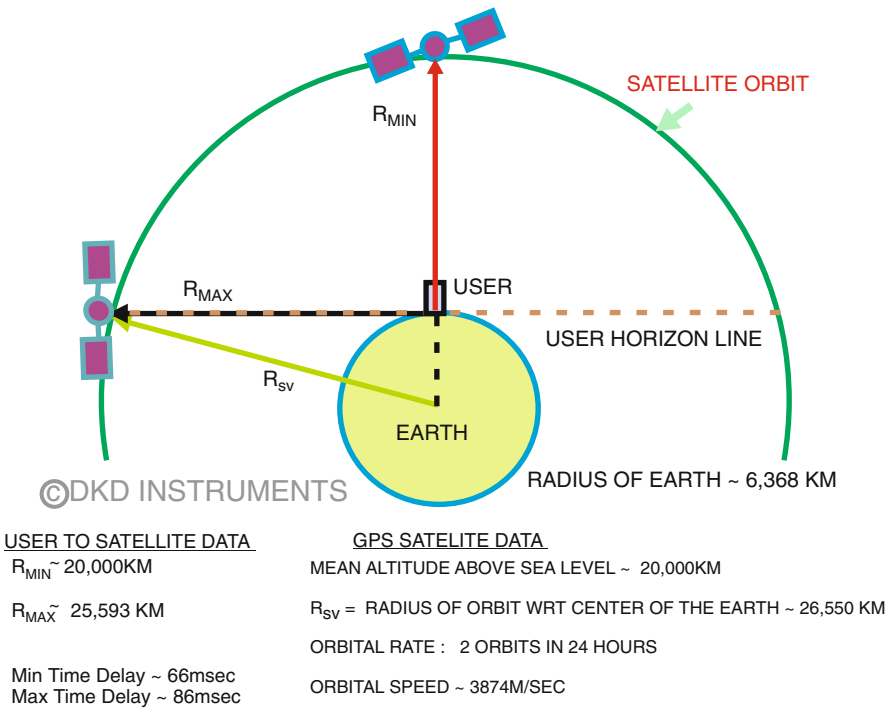


Fig. 2.3 The range from the receiver to a GPS satellite and physical constants (for a GPS satellite orbit that passes directly overhead)

When we use the clock model to describe the receiver’s reference clock (or SV replica clocks) we will see that the 0–0.977 μs dial does have a direct physical counter part in the receiver. The impact of having the smallest time increment dial not an exact model for the SV clock is not an important issue for this text as our goal is position accuracy of $\pm 100\text{ m}$.

2.4 Calculating Tbias Using One SV, User Position Known

Perhaps the simplest application of GPS is synchronizing a receiver’s clock when the receiver’s position is known. In this example we will form an first-order estimate of the Tbias term associated with receiver clock. Figure 2.4 shows a model of our example system. The SV has its own clock and sends its clock timing information to the earth-based GPS receiver using an encoded radio wave. The receiver has two clocks, a reference clock and a replica of the SV clock reconstructed from the received radio wave. Figure 2.4 shows the receiver clock corrected with the calculated Tbias term, i.e., it is in synchronism with the SV clock.

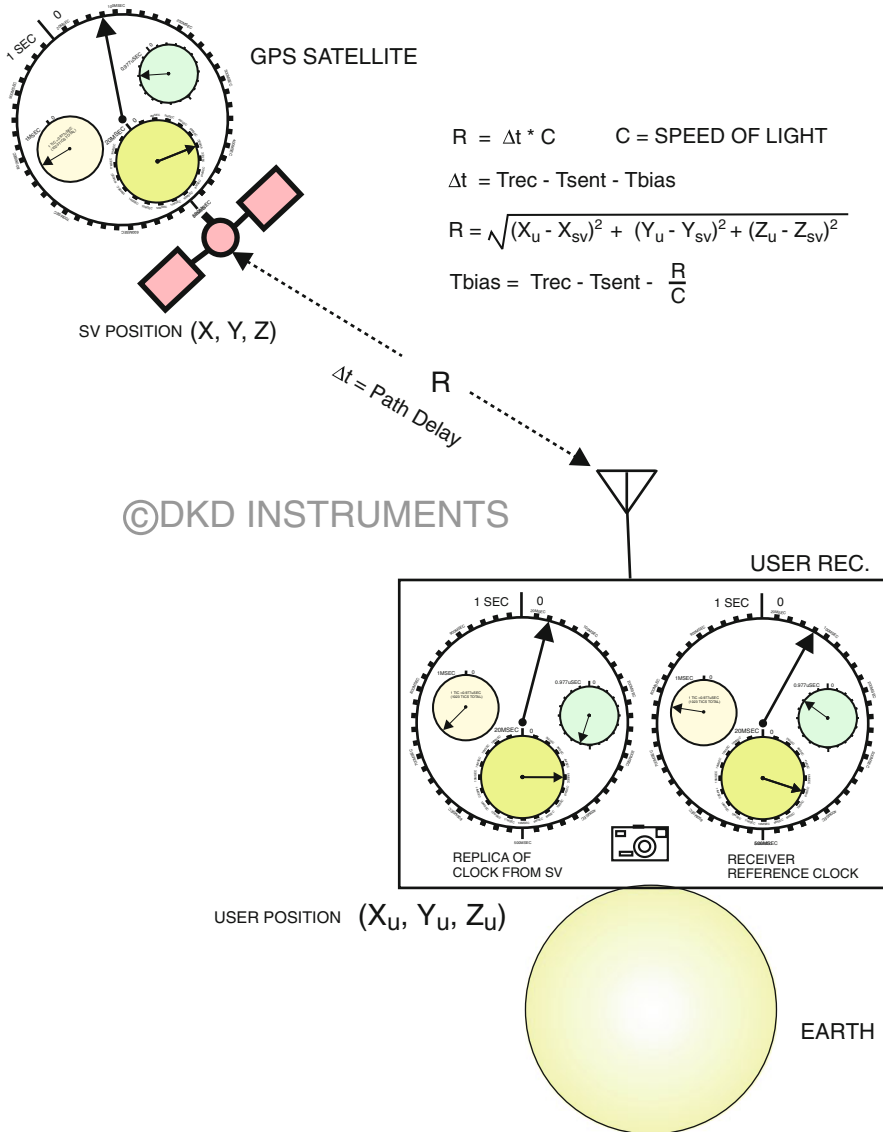


Fig. 2.4 Calculation of Tbias using a single SV, user and SV position known

In order to solve for Tbias we need to calculate the distance R from the SV to the User receiver and measure T_{rec} and T_{sent} as shown in Fig. 2.4. The calculation of R uses the distance equation which requires the user and SV positions in X, Y, Z . We have assumed the user position is known. In addition to the SV clock information the SV

position information is encoded onto the radio wave that is transmitted to the receiver. We will assume for now that this information is provided in X, Y, Z coordinates. We will see later that getting SV position information in X, Y, Z format is nontrivial.

The T_{sent} and T_{rec} information are obtained as before by just taking a “snapshot” of the receiver reference clock and the receiver’s replica of the SV clock. If we record the SV position at the same instant that we record T_{rec} and T_{sent} , we will have all the information needed to solve for T_{bias} . It is important to realize that due to SV motion we must “capture” the SV position data at the same moment we capture the state of the receiver’s clocks. If we do not properly capture SV position, T_{sent} , and T_{rec} , then the computed distance, R , would be incorrect for the measured path delay.

Now that we have all the information needed we can calculate R , Δt , and T_{bias} . If we continually update the measurements and calculations, the receiver reference clock will “track” the SV clock. This allows the GPS time receiver to replicate the stability of the SV clock. SV clocks are atomic based and so the stability is very high. If we modify the receiver’s reference clock to output a “pulse” every time the 1 s dial passes the 0 tic mark and a 1 pps signal will be generated. This is a common signal many GPS receivers provide.

2.5 GPS Time Receiver Using Master Clock and the Delay term T_{atm}

In the previous example we computed T_{bias} when the user position was known. In this example we will include the effects of SV clock error with respect to the GPS master clock and the additional delay caused by diffraction of the radio beam as it passes through the earth’s atmosphere.

Figure 2.5 shows the details of our new model. The receiver’s reference clock is shown corrected to the master clock time. The SV clock has a small error with respect to master clock. The error is less than a millisecond. As before we capture the state of the receiver’s reference clock and the replica clock. This information is used in conjunction with the computed path length R to form our estimate of T_{bias} . There is a difference from our first example and that is in the path delay. The expression for the path delay now has two additional terms. One, of course, is the SV clock error with respect to the master clock, $T_{\text{err_sv}}$. The other is the term T_{atm} .

The term T_{atm} is the extra delay experienced by the radio beam as it travels through the earth’s atmosphere. To a first approximation the atmosphere acts like a lens and “bends” the radio beam from the SV to receiver. This bending of the radio beam as it passes through the atmosphere causes the extra delay. Normally the delay T_{atm} is broken into two parts, one for the Ionosphere and one for the Troposphere layers of the earth’s atmosphere. Here we have combined them into one term with the sign convention following most of all GPS literature. How big is the added

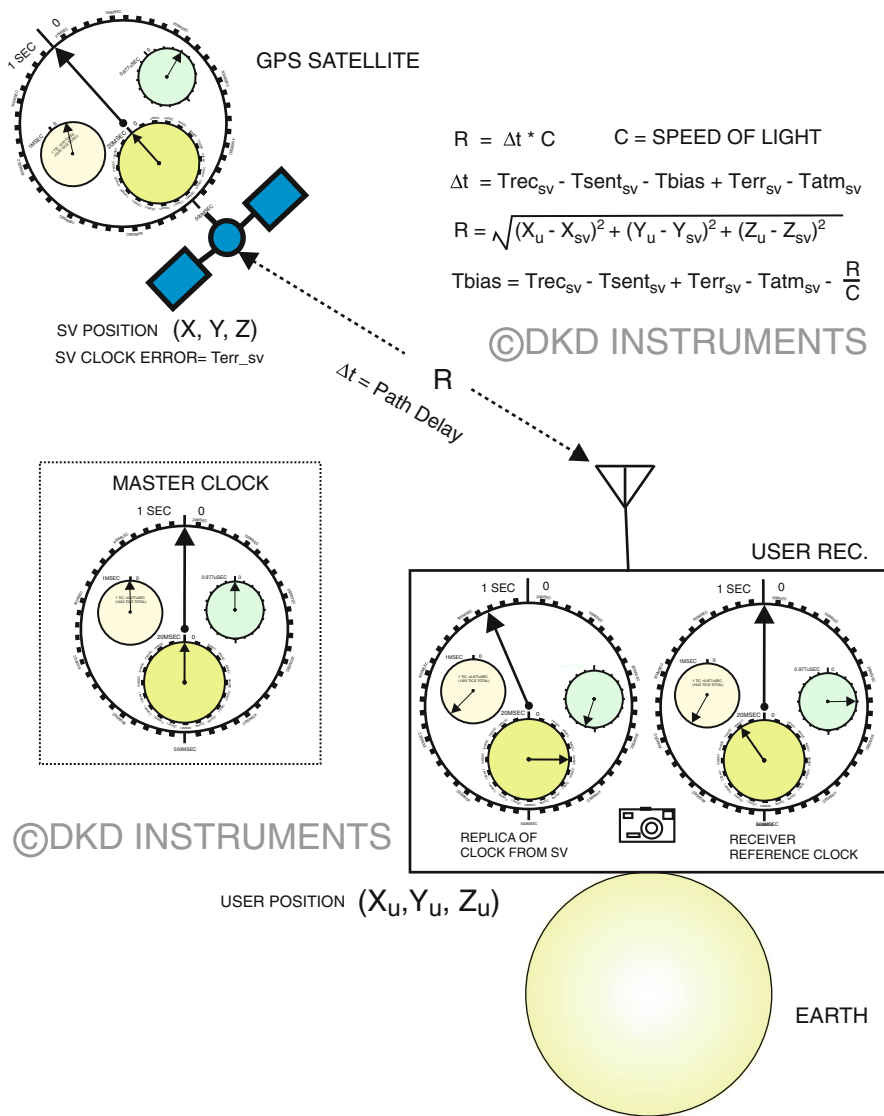


Fig. 2.5 Calculation of $Tbias$ using a single SV, user and SV position known with master clock and atmospheric delay term

delay? Expressed as an increase to the R term this “extra” distance is in 50 m range. The maximum value occurs when the SV is low in sky and decreases as it rises. This delay is relatively large and consequently places the error introduced by $Tatm$ at the top of a long list of error sources. Estimating the value of $Tatm$ requires knowledge of the SV position in relation to the receiver, exact time of day, sun spot activity,

day/night, etc. Needless to say, the effects of the Troposphere and Ionosphere on a radio beam are extremely complex and beyond the scope of this text. The reader is directed to the bibliography for suitable references.

By including the master clock correction and a term for atmospheric delay we form a model that achieves greater accuracy than our previous model. We also put into place the primary components needed to do the full solution to the user position problem. If we were to look at a 1 pps signal generated from the receiver's reference clock it "track" the GPS master clock with greater precision.

This example also illustrates that by refining the path delay estimate we can achieve greater accuracy in the computed value of T_{bias} . This will also be true for computed user position. It is hoped the reader sees the pattern emerging of successive refinements to the path delay, which results in greater accuracy for both position and time measurements.

2.6 Solving For User Position Using Four Satellites

With the previous examples and models we have laid the groundwork for the task of determining user position, X_u , Y_u , Z_u in ECEF coordinates and the user clock error, T_{bias} . Figure 2.6 shows the system using four SV's and four SV clock replicas in the user receiver. We again assume the user receiver is at or near the surface of the earth (i.e., each path delay is 66–85 ms). The four SV clocks are referenced to a master clock. The receiver's reference clock is shown with its error removed. In other words the receiver reference clock is displaying master clock time. For each SV to user distance a path delay is measured/computed. Additionally a distance R_i is assigned to the equivalent distance corresponding to path delay multiplied by the speed of light.

Figure 2.7 shows the set of equations. Each distance, R_i , can be computed from two different methods. The first method uses the path delay, which contains measured, computed, and SV-supplied terms. Each path delay has its own unique $T_{\text{atm}_{\text{sv}_i}}$ term derived from estimates of user position. As user position estimates are refined, this term will also be refined in accuracy. The clock error for each SV is sent to the receiver and is identified by $T_{\text{err}_{\text{sv}_i}}$.

Normally the T_{rec} time would be the same for all the path delays. It is possible to have separate T_{rec} times if the user receiver does not move appreciably from one T_{rec} time to the next. As before we take a "snapshot" of the receiver's reference clock and the replica clock or clocks to record T_{rec} and $T_{\text{sent}_{\text{sv}_i}}$. If we choose to include all the clocks in our picture, then the value of T_{rec} could be the same for all SV path delay equations. This last point needs a bit more clarification. If the user receiver is stationary (or moving slow compared to the time to measure each path delay) then we can measure the path delay to each SV separately. That is, we could first do SV_1 , then SV_2 , etc. In other words, a sequential measurement using a single channel receiver. As long as the drift rate of the receiver's reference clock is low enough over the time all four measurements are made, this method will work fine. Many early GPS receivers used this method as they were single channel receivers.

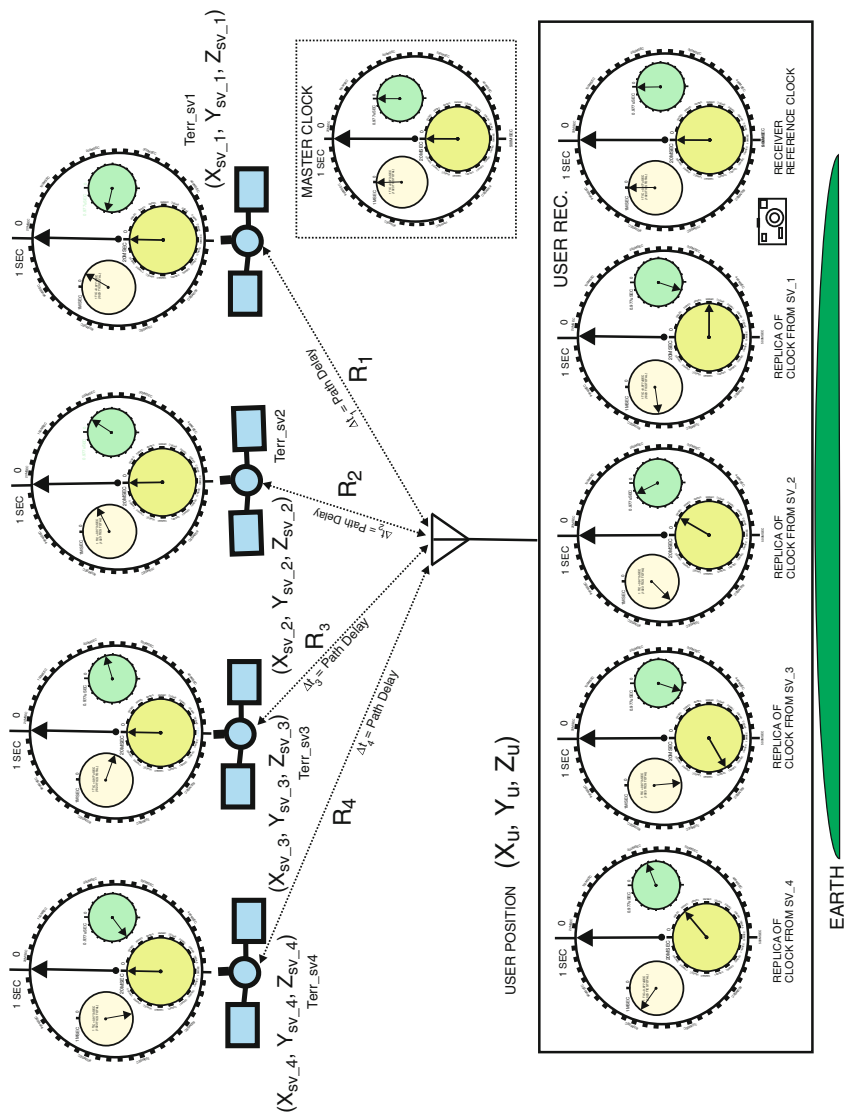


Fig. 2.6 Calculation of user position and Tbias using four SV's

Fig. 2.7 Equation set for solving for user position and Tbias using four SV's

$$R_1 = \Delta t_1 * C$$

$$R_2 = \Delta t_2 * C$$

C = SPEED OF LIGHT

$$R_3 = \Delta t_3 * C$$

$$R_4 = \Delta t_4 * C$$

$$\Delta t_1 = T_{rec} - T_{sent_{sv_1}} + T_{bias} - T_{err_{sv_1}} + T_{atm_{sv_1}}$$

$$\Delta t_2 = T_{rec} - T_{sent_{sv_2}} + T_{bias} - T_{err_{sv_2}} + T_{atm_{sv_2}}$$

$$\Delta t_3 = T_{rec} - T_{sent_{sv_3}} + T_{bias} - T_{err_{sv_3}} + T_{atm_{sv_3}}$$

$$\Delta t_4 = T_{rec} - T_{sent_{sv_4}} + T_{bias} - T_{err_{sv_4}} + T_{atm_{sv_4}}$$

$$R_1 = \sqrt{(X_u - X_{sv_1})^2 + (Y_u - Y_{sv_1})^2 + (Z_u - Z_{sv_1})^2}$$

$$R_2 = \sqrt{(X_u - X_{sv_2})^2 + (Y_u - Y_{sv_2})^2 + (Z_u - Z_{sv_2})^2}$$

$$R_3 = \sqrt{(X_u - X_{sv_3})^2 + (Y_u - Y_{sv_3})^2 + (Z_u - Z_{sv_3})^2}$$

$$R_4 = \sqrt{(X_u - X_{sv_4})^2 + (Y_u - Y_{sv_4})^2 + (Z_u - Z_{sv_4})^2}$$

The drawback is a better reference clock is needed as the method is relying on Tbias being approximately constant over the entire measurement sequence of the four SV's path delays. Being able to "measure" four Trec and four Tsent times *simultaneously* requires a four-channel receiver. This is the assumption made when writing the equations of Fig. 2.7.

The distance R_i can also be computed by using the distance formula as we did in the single SV example for obtaining Tbias. Each distance is computed from the user's position X_u, Y_u, Z_u and the position of the SV, which is assumed to be sent from SV as $X_{sv_i}, Y_{sv_i}, Z_{sv_i}$ coordinates. By using the distance equations combined with the path delay equations we have enough information to determine the receiver's position and the receiver clock error. Unlike our simple solution for Tbias using one SV, the equation set of Fig. 2.7 *cannot* be solved in closed form for Tbias and X_u, Y_u, Z_u . The reason is that the equation for distance contains the square root function. This introduces a nonlinearity which precludes a "closed form" solution. To solve the equations a iterative method is employed. An iterative solution for user position and Tbias is presented in Chap. 5.

2.7 The Pseudo-Range

When we subtract the Tsent reading of SV from the Trec reading we are computing a *measured* path delay. But our measured path delay has errors. Tbias is rarely, if ever, exactly zero. Because of this error the distance obtained by computing the range from known user position and known SV position *plus* the distance equivalent of the user clock error, Tbias, is called the “Pseudo Range.” This statement reflects the fact that this range is not exact because it has the error of receiver clock bias error expressed in it. In the terms used in this text and following conventional polarities of error terms, the Pseudo-Range for any given SV_i would be:

$$\text{Pseudo - Range}_i = [(x_u - x_{sv,i})^2 + (y_u - y_{sv,i})^2 + (z_u - z_{sv,i})^2]^{1/2} + \text{Tbias} \times C \quad (2.1)$$

A related measured path delay version/estimation of the Pseudo-Range is:

$$C \times \Delta t_i \geq C \times [\text{Trec}_{sv,i} - \text{Tsent}_{sv,i} + \text{Terr}_{sv,i} - \text{Tatm}_{sv,i}] \quad (2.2)$$

2.8 A Simplified Model of the GPS Receiver

Figure 2.8 shows a high-level block diagram of a GPS receiver. The receiver shown has four channels. Only one of the four channels is fully detailed. The other three are shown as repeated blocks. Each channel has an identical SV clock replication block. This block replicates the SV clock from the received SV RF signal. Only one receiver reference clock is needed. The RF carrier is typically converted to a lower frequency and then sent to each channel.

Dividing down Receiver Reference Oscillator (RRO) forms the replica clock and the receiver reference clock. The receiver reference clock as shown can be identical in construction to the receiver’s replica clocks or it could be of different construction. Later in Chap. 6 we will introduce an alternate reference clock based on the tagged SNAP_SHOT method. This approach can result in a simpler reference clock implementation. The reference clock reset function shown is used to start the clock at predetermined time or to “subtract out” Tbias from the reference clock so that it displays Master Clock time (within the errors of receiver). The receiver’s computer can force Tbias to minimum by dynamically loading N₀, N₁, N₂ and N₃ values into the receiver reference clock registers. In some receivers this is optional and the error is held in software.

It is important to understand that the SV Replica Clock is never truly recovered. We can only form better and better estimates of the true SV clock as the receiver design is advanced. This same comment holds for the receiver’s reference clock. We can never get to the point when Tbias is exactly zero, the goal is to drive it to a minimum.

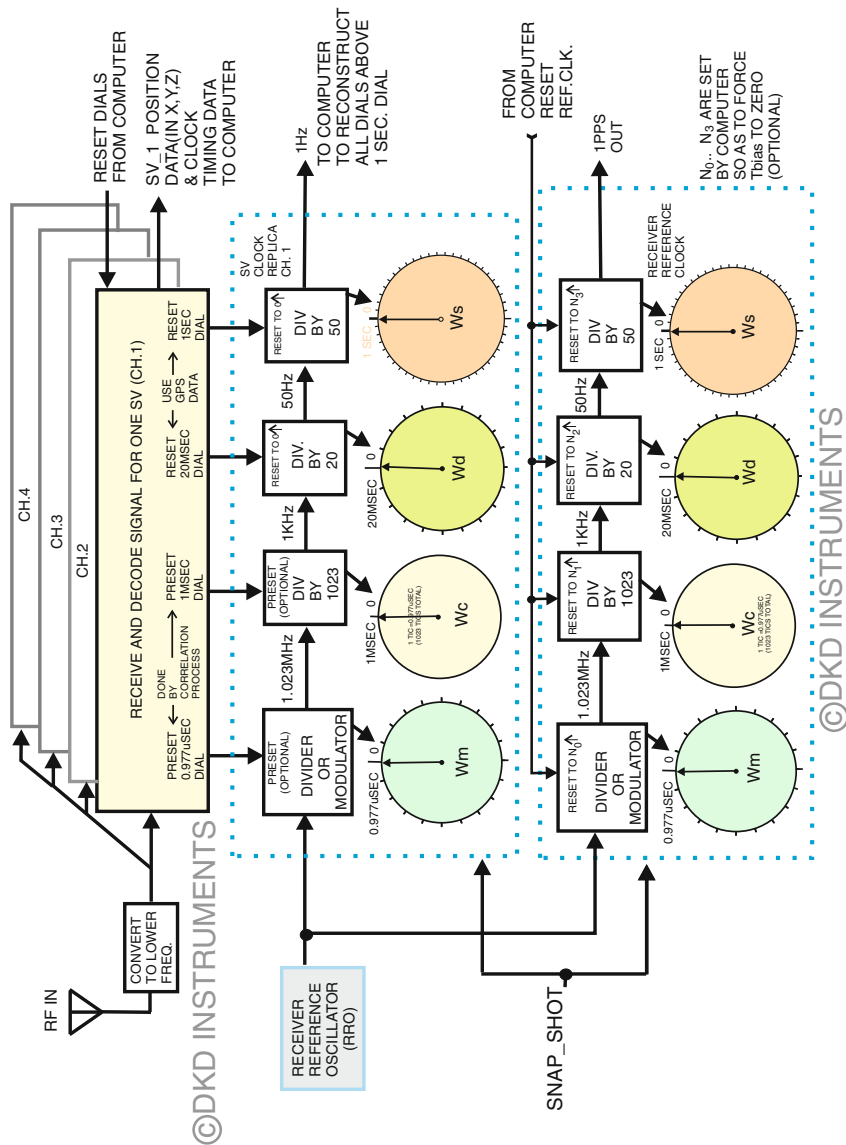


Fig. 2.8 Block diagram of four-channel GPS receiver

The SV replica clock must be properly reset in order to be a “faithful” delayed reproduction of the received SV clock. In the reception process the first two dials are correctly set by the *correlation* process used on the SV signal. In practice, the 1 ms dial is set by a quasi-dynamic process while the 0.977 μ s dial is dynamically controlled to maintain “lock” on the received SV clock. The Preset inputs on these two dials is optional as some receivers do not “preload” values here. Since the output of the 0.977 dial drives all the other dials locking this dial to the received SV clock “locks” all the subsequent dials. The 0.977 dial may be a fixed divider in combination with a modulator or a pure analog method, which would use a separate reference oscillator from the receiver’s reference clock (not shown).

The 20 ms and 1 s dials must be set by examination of the received data from the SV for that channel. Once the dials are initially set to the correct initial value they “free run” except for the 0.977 dial, which is under a servo loop control to keep lock on the received SV clock. Each channel must independently set its dials as per the method just described. After all channel clock dials are properly set then a SNAP_SHOT can be taken which records the state of all replica and reference dials (details of capture method not shown).

The data captured from replica and reference dials is now sent to the computer. In addition, each channel decodes the SV position information for the time the signal left the SV (i.e., T_{sent}). The computer now has SV position, T_{rec} , and T_{sent} for all four channels. Now the computer can solve the equations of Fig. 2.7 (by iteration methods) for user position and receiver clock bias.

In the hardware section of the text we will get into the details of exactly how all this is accomplished. In Chap. 3 we will see how the transmitter, or SV, clock is constructed.

2.9 The Receiver Reference Oscillator

The reference oscillator used in the GPS SV is an atomic-based system or “atomic clock.” Atomic clocks are extremely accurate and stable over long periods of time. Due to reasons of cost and size the RRO used in “lower end” GPS receivers cannot be an atomic clock. Typically, lower-cost GPS receivers use oscillators accurate to ± 0.5 ppm, while the SV atomic clocks will exceed ± 0.00001 ppm accuracy.

To investigate the effect of RRO error suppose we reset the receiver’s reference clock such that it reads exactly 0 at the same instant as the GPS Master clock. We then let the reference clock “free run.” After a period of time (an interval dependent on the PPM rating of the RRO) our reference clock would typically drift from the GPS Master clock resulting in a difference of displayed time on the dials of the two clocks. This occurs because our reference oscillator has a static frequency offset and drift rate. The result is the receiver reference clock will lose synchronism with respect to the Master clock as time advances. Once the error grows to approximately 1 μ s we

would see an equivalent distance error of 300 m in our path delay measurements. In general, higher-quality GPS receivers use higher-quality reference oscillators in order to mitigate the problems associated with RRO errors.

2.10 Satellite Position Information

In our model we assumed that the XYZ data was sent down to receiver on an encoded RF signal. The format was assumed to be the satellites position expressed in ECEF coordinates, X , Y , Z . Unfortunately it is not that simple. Rather than sending its position in X , Y , Z coordinates each SV sends a very complex set of *Orbital Parameters* for its particular orbit. This set of orbital data for GPS is called the *Ephemeris* data.

Orbital parameters can be used to describe the precise orbit characteristics of any orbiting body about the earth. The parameters contain the diameter of the orbit, its deviation from circular, its angle w.r.t. the polar axis, etc. If you know the orbital parameters of an orbiting body precisely, you can predict where it is and where its going, provided you know what *time* it is.

The SV orbits are fairly stable and repeatable. So once you have the orbital parameters for all the satellites at a particular day and time you can compute their positions at a later date (not too much later) with a fair degree of accuracy, using the same set of orbital data. In fact, as we will find out later each SV sends down the orbital parameters for not only itself but for all the other SV's of GPS but with reduced precision (Almanac Data). This allows the receiver that has a successful contact with just one SV to obtain position information for all the other satellites. Together with the initial estimate of present GPS time this information can assist the receiver in obtaining its first estimate of user position.

Once the receiver gets the orbital data it now has the task of converting this data to X,Y,Z in ECEF. In order to do this accurately it must know what "time" it is. We can get this information from the T_{sent} data present on the receiver SV replica clocks. Once we have the T_{sent} time we can use the orbital parameters to work out the X,Y,Z position of the SV at the T_{sent} instant of time.

2.11 Summary

Its time to take a step back and examine where we are in our quest to understand the GPS receiver and to get our first-order solution of the user's position. The first two chapters have laid the groundwork for understanding the basic principles of the using synchronized clocks to measure distance by measuring path delay. In this chapter we have discussed some of the details of the satellite system. It is hoped that the reader can start to see the real complexity involved and now has a grasp on the basic principles and working mechanisms behind GPS.

The first two chapters were written on a lay basis so as to insure that as many readers as possible can get the basics of GPS. From here on some readers may not have the background to proceed. Those readers may need to consult other texts to get the needed background knowledge to understand the discussions ahead. The GPS signal structure will be the next topic.



<http://www.springer.com/978-1-4614-0408-8>

Fundamentals of GPS Receivers

A Hardware Approach

Doberstein, D.

2012, XVI, 329 p., Hardcover

ISBN: 978-1-4614-0408-8