

**MODULE – 9 LECTURE NOTES – 3****GLOBAL POSITIONING SYSTEM (GPS)****1. Introduction**

Global positioning system (GPS) is also known as Navigation System with Time and Ranging Global Positioning System (NAVSTAR) GPS. Originally designed for military purposes, GPS is being increasingly used by civilians for various applications like marine navigation, surveying, car navigation. Signals made available for civilian use, known as the Standard Positioning Service (SPS) can be freely accessed by general public. On the other hand, the more accurate Precise Positioning Service (PPS) can only be used by authorized government agencies. The development of GPS system was mainly aimed for these aspects:

- a) To provide user's with locational coordinates
- b) To provide an accurate and continuous 3 dimensional positioning capability operating in all weather conditions over global extent
- c) To offer potential for various civilian applications

Some examples of space based GPS systems include GLONASS (an acronym for Globalnaya navigatsionnaya sputnikovaya sistema) which is operated by the Russian Aerospace Defence forces, BeiDou Navigation Satellite System (BDS) which is a Chinese satellite navigation system which has been operational since 2000, Galileo: global navigation satellite system currently built by European Union and European Space Agency, Indian Regional Navigational Satellite System (IRNSS) being developed by the Indian Space Research Organisation. The configuration of a GPS system is comprised of three distinct segments:

- i) Space segment: The space segment comprises of 24 satellites orbiting the earth at approximately 20200 km every 12 hours. There are 6 orbital planes with nominally four satellite vehicles in each orbit. The space segment is designed in such a way that there will always be a minimum of 4 satellites visible above  $15^\circ$  cut off/mask angle at any point of the earth's surface at any time. The reason will be clear by the end of this module. Each of the GPS satellites have highly precise atomic clocks on board which operate at a fundamental frequency of 10.23 MHz. These clocks are crucial to generate the signals which are

broadcasted from the satellite. Satellites generally broadcast two carrier waves which are in the L band. These carrier waves are usually derived using the fundamental frequency generated by the highly precise atomic clock onboard the satellite. The L band waves used are

a) L1 carrier wave at a frequency of 1575.42 MHz and L2 carrier broadcasted at a frequency of 1227.60 MHz.

These carrier waves have codes modulated upon it. The L1 carrier wave has two codes known as C/A code or coarse acquisition code and P code known as precision code. The C/A code is modulated at 1.023 MHz whereas the P code is modulated at 10.23 MHz frequency. The C/A code is based upon the time provided by a highly accurate atomic clock. The receiver also will contain a clock which is used to generate a matching C/A code. The incoming satellite code can be compared with the code generated by the receiver. The L2 carrier wave has just one code modulated upon it at 10.23 MHz. As the space segment consists of nearly 24 orbiting satellites, in order to distinguish between satellites, GPS receivers use different codes. The time for radio signal (L band) to travel from the satellite to any GPS receiver can be calculated using these codes.

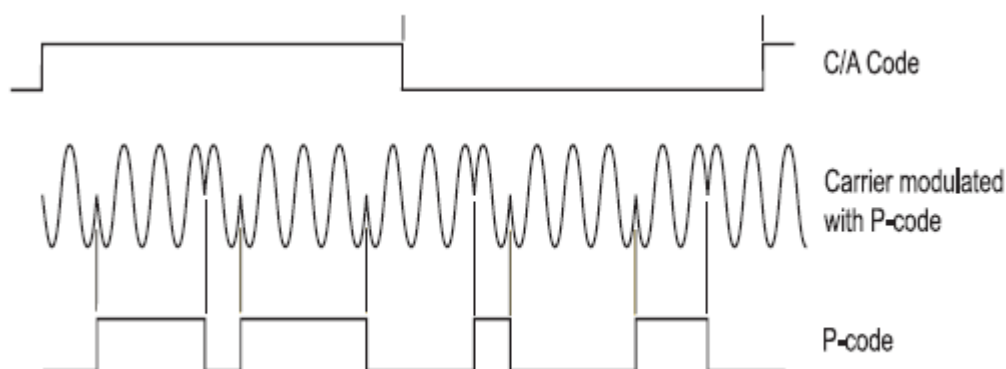


Figure 1: Schematic showing C/A and P code

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

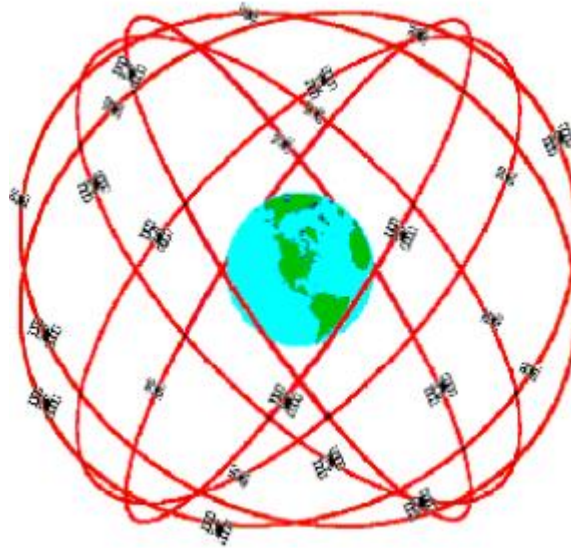


Figure 2: Space segment of GPS system

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

ii) Control segment: The control segment comprise of a master control station with 5 monitoring stations. These stations track/control the orbital's positions of satellites. The control stations are located at Hawaii, Colorado Springs, Ascension islands, Diego Garcia, Kwajalein. It is essential to estimate the orbit of each satellite in order to predict its path 24 x 7. This information is available in uploaded to each of these orbiting satellites which are subsequently broadcasted from them. The signals captured by user's with a GPS receiver enables to know the exact position of each of these satellites. The signals from satellites are read at the control stations which estimate the measurement errors. These errors are then transmitted to the master control station in Colorado Springs wherein processing takes place to determine any errors in each of these satellites. This information from the master control station is resent to the four monitoring stations which are then uploaded to these satellites.

iv) User Segment- GPS receivers are used to receive the GPS signals which can then be used for navigation and other purposes. Anyone who avails this facility comprises the user segment. The various applications for which GPS receivers can be used range from surveying, aerial/marine/land navigation, defense/defense machinery control etc.

## 2. Principle

The information broadcasted as a continuous stream of data by each satellite to the earth is termed as GPS navigation message. In order to calculate the current position of the satellites as well as to determine the signal transmit times, it is highly essential that we know this navigation message. The data stream is transmitted at 50 bits per second. Ephemeris and almanac data, the satellite orbits and the relevant coordinates of a specific satellite can be determined at a defined point in time.

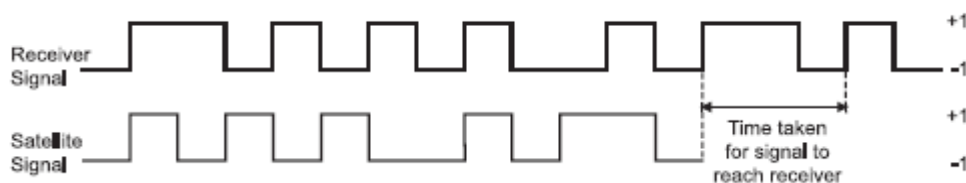


Figure 3: Schematic showing receiver and satellite signal time difference

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

GPS relies on different methods for estimating location coordinates which are dependent on the accuracy required by the user and the type of GPS used. GPS determines the pseudoranges and the time of arrival of signal. The simplest technique used by GPS receivers for instantaneous estimation of locational coordinates is analogous to the two point problem in plane table surveying. The basic concept is that if we know the distance of three points relative to our own position, we can determine our position relative to those three points. The navigation technique is known as “Trilateration”, which is based on the measurements of difference in distance to two or more stations (located at known coordinates) that transmit signals (at known time). This will result in an infinite number of locations, which when plotted form a hyperbolic curve. In order to narrow down on the exact user location, a second measurement needs to be taken to a different pair of stations which will produce a second curve and so on.

For example, imagine that a user wandering across a desert has lost his way and would like to know his exact location using a hand held GPS receiver. The satellites orbiting far above the surface of earth passing over the aforementioned desert are continuously transmitting their own positions and clock times. Using the signal transit time to both the satellites, two circles can be drawn with radii as  $S_1$  and  $S_2$  around the satellites. Each radius represents the distance calculated to the satellite wherein all possible distances to the satellites will be located on the circumference of this circle. The location of the user's GPS receiver will be at the exact point where the two circles intersect beneath the satellites. But this positioning is only useful in a 2D plane where the aim is to find just the X and Y coordinates. Practically, for locating coordinates in a 3D space, an additional third satellite must also be available. This will easily give the sought after user location based on the intersection of all the three spheres, as shown in the diagram below:

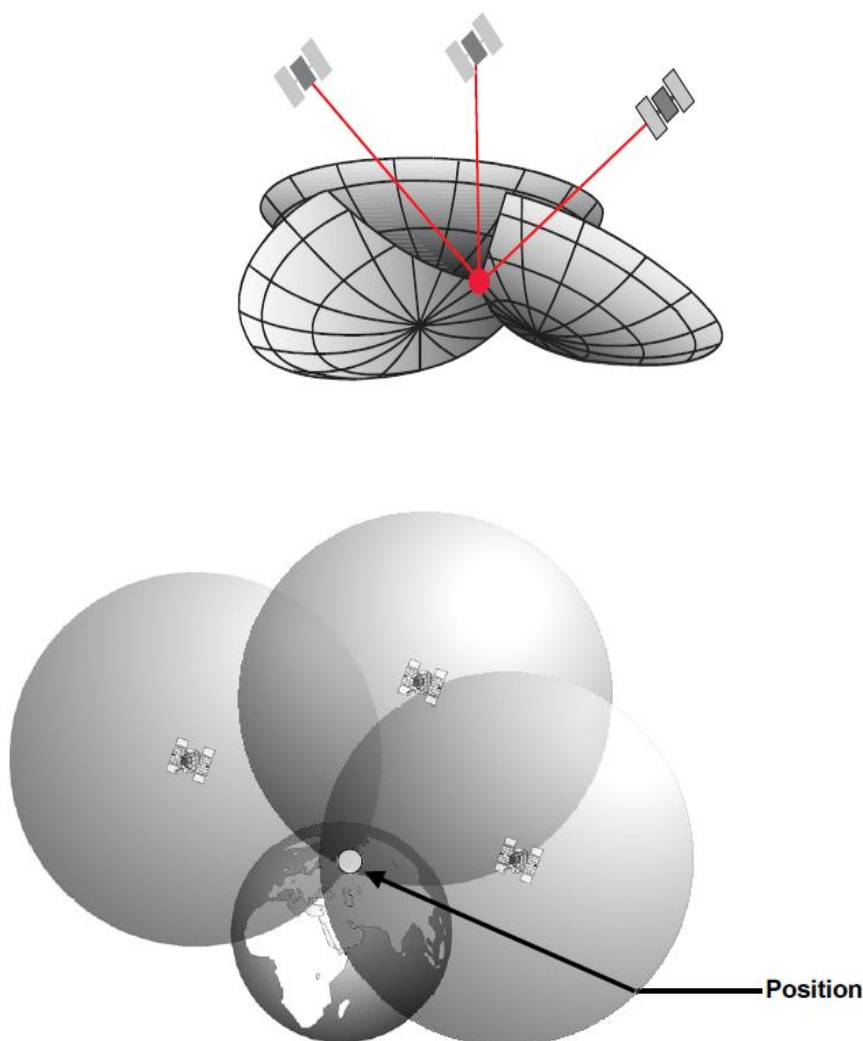


Figure 4: Principle of locating positions using GPS (Trilateration)

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

The GPS measurement principle has been subtly explained assuming that one can estimate the signal transit time to a high degree of precision. Practically, for a hand held GPS receiver to give precise time, a highly synchronized clock is mandatory as the clocks onboard each of the orbiting satellites are all being synchronized (controlled by the control segment). Mathematically speaking, to estimate the values of  $N$  variables, an equivalent number of  $N$  independent equations are essential. Here, assuming that the GPS receiver time measurement has an unknown error term associated with it, the overall number of unknown variables will be four namely, longitude ( $X$ ), latitude ( $Y$ ), altitude/height ( $Z$ ) and time error ( $\Delta t$ ). In order to estimate these four unknown variables within a 3D space, four independent equations are needed. It should be noted that the 28 GPS satellite constellation are distributed around the globe in such a manner that at least 4 of them will always be “visible” from any point on the Earth’s surface.

For example, let  $S_1$ ,  $S_2$ ,  $S_3$  and  $S_4$  be the four satellites observed by a user carrying a GPS receiver. Let the positional coordinates of the user be  $x, y, z$  and let the coordinates of the four satellites be  $(X_1, Y_1, Z_1)$ ,  $(X_2, Y_2, Z_2)$ ,  $(X_3, Y_3, Z_3)$  and  $(X_4, Y_4, Z_4)$ . Let  $R_1$ ,  $R_2$ ,  $R_3$  and  $R_4$  denote the distance between the four satellites  $S_1$ ,  $S_2$ ,  $S_3$  and  $S_4$  with the user’s position. Also let the error in time measurement be represented as  $\Delta t_1, \Delta t_2, \Delta t_3$  and  $\Delta t_4$ .

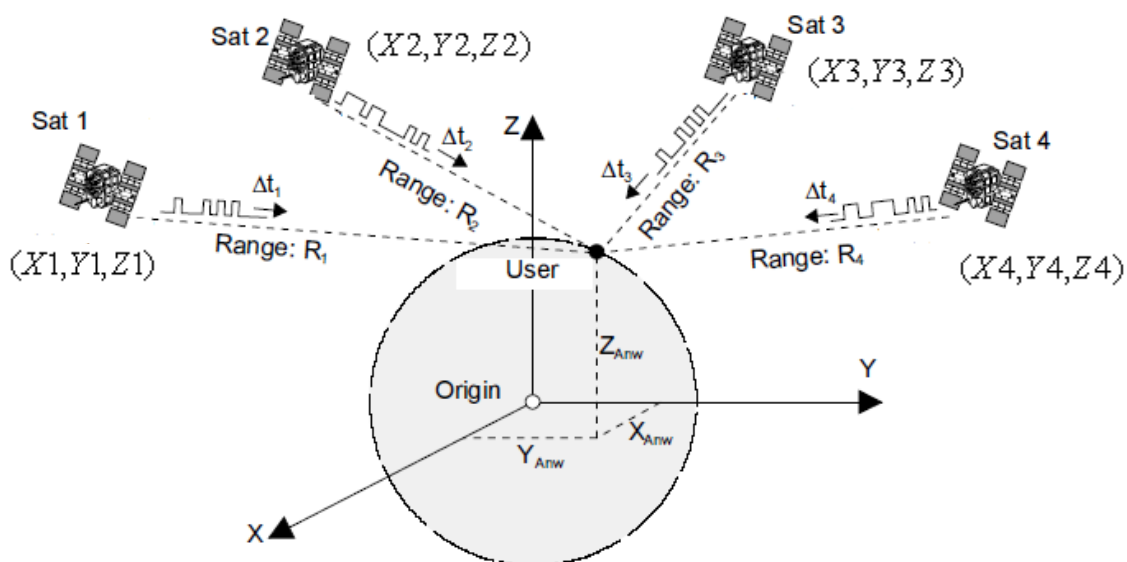


Figure 5: Estimation of location coordinates using GPS

[Source: <http://www.u-blox.com/>]

The resultant error in time measurement causes inaccurate measurement of signal transit time as well as the distance (R). An incorrect distance which is measure is known as pseudo distance or pseudo range. The pseudo ranges are given by the relation:

$$PSR = R + (\Delta t_0 * c)$$

Where c is the speed of light and  $\Delta t_0$  is the difference between the satellite clock and user clock. From the above figure, the distance from satellite to user can be related using Cartesian system as follows:

$$R = \sqrt{(X_{SAT} - X_{USER})^2 + (Y_{SAT} - Y_{USER})^2 + (Z_{SAT} - Z_{USER})^2}$$

And pseudorange can be written as

$$PSR = \sqrt{(X_{SAT} - X_{USER})^2 + (Y_{SAT} - Y_{USER})^2 + (Z_{SAT} - Z_{USER})^2} + (c * \Delta t_0)$$

Here,  $X_{SAT} = X1, X2, X3, X4$  and  $Y_{SAT} = Y1, Y2, Y3, Y4$  and  $Z_{SAT} = Z1, Z2, Z3$ . The four equations obtained by observing 4 satellites simultaneously produces a non-linear set of equations which needs to be solved. The solutions will not be discussed in this module.

### 3. Errors in GPS measurements

Various sources of errors tend to degrade the quality of GPS positioning. These will be discussed here.

#### a. Ionospheric and Atmospheric errors

The atmosphere refraction remains the main limitation to the precision of GPS in navigation (meter precision) , Surveying Engineering (centimeter precision) , Geodesy and Geophysics (

millimeter precision). The presence of atmosphere introduces a propagation delay which depends on the ionospheric vertical total electron content (TEC) and on satellite elevation angle above the horizon. Due to it, the transmission time of signal varies introducing position errors. In Global Positioning System (GPS), the 8 parameter Klobuchar type ionospheric delay model is broadcast to predict the global vertical TEC distribution at a given time period for real-time correcting the ionospheric effect for single frequency GPS measurements. Signals while passing through ionosphere undergoes an effect similar to a light refracted through a glass block. Due to the slowing down of the signal, an error gets introduced into the calculation of pseudoranges. The delay caused does not remain constant but varies according to several factors.

-The ionosphere is the upper part of Earth's atmosphere between approximately 70 and 1000km wherein sufficient electrons and ions are present to affect the propagation of radio waves. The ionosphere is the first layer the signal encounters. It is a part of high atmosphere where sufficient electrons and ions are present to affect the propagation of radio waves. The generation of ions and electrons is proportional to the radiation intensity of the sun, and to the gas density. The ionosphere is not homogeneous. It changes from layer to layer within a particular area. Its behaviour in one region of the earth is liable to be unlike its behavior in another region. The impact of the state of the ionosphere on the propagation of waves is characterized by the Total Electron Content (TEC). The ionospheric effects contribute the second largest user equivalent range error (UERE) since Selective Availability (SA) was deactivated on 1 May 2000, the largest error source in precise positioning using GPS is the signal delay due to the ionosphere. During a low solar activity, the ionospheric delay ranges from a few centimeters to several tens of meters. With a high solar activity, the delay can reach up to 150 m, when the satellite is near the observer's horizon, the vernal equinox is near, and sunspot activity is at its maximum.

It is difficult to find a satisfying model for the TEC because of the various time dependent influences. The most efficient method is to eliminate the ionospheric refraction by using two signals with different frequencies. This dual frequency method is the reason why the GPS signal has two carrier waves L1 and L2. Several models like code pseudorange, carrier phase pseudorange, Klobuchar model etc are used to predict the vertical TEC at a given time so that ionospheric effect can be corrected.



Satellite signals need to travel an increasingly large distance to pass through the atmosphere. Due to this reason, signals from the low elevation satellites will be greatly affected by satellite's elevation that signals from the higher elevation satellites. The presence of atmospheric water vapour molecules also influences the GPS signal. These tend to cause position degradation which can be reduced by using atmospheric models.

### **b. Satellite and Receiver clock errors**

The clocks employed in satellites are high precision atomic clocks which are very accurate to about 3 nanoseconds. Even then, sometimes they might drift slightly, resulting in errors of small magnitudes. The United States Department of defense monitors these clocks using the control segment. Their duty is to correct any drift, if found.

### **c. Multipath Errors**

Multipath errors are caused when the signal tends to travel multiple paths before reaching the user's GPS receiver. Whenever the user GPS receiver is positioned closer to a number of reflecting surfaces like lakes/buildings in urban areas, the satellite signal instead of travelling directly to the GPS antenna, tends to hit these surfaces first and then get reflected into the antenna. This creates a false measurement. Multipath errors can be reduced by using a choke ring antenna which has 4 to 5 concentric rings around that enable trapping of any indirect signals. One thing to note is that this error tends to affect high accuracy, survey type instruments. Hand held GPS receivers do not employ such techniques.



Figure 6: Choke ring antenna to eliminate multipath errors

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

#### d. Dilution of Precision (DOP)

Satellite ranging errors can be magnified by the effect of Dilution of precision. DOP is essentially a measure which shows the strength of satellite geometry. This is related to the spacing and position of the satellites in the sky. Well spaced satellites will usually have lower uncertainty of position than poorly spaced satellites. There are different types of DOP like the vertical dilution of precision (VDOP), horizontal dilution of precision (HDOP). Both VDOP and HDOP gives accuracy degradation in vertical and horizontal direction respectively. Positional dilution of precision (PDOP) provides accuracy degradation in 3dimensional position while Geometric dilution of precision (GDOP) provides accuracy degradation in both position and time respectively. Of these, the most important is GDOP as this represents a combination of all the other factors. Observing as many satellites as possible will minimize the value of GDOP. However, it should be noted that the satellite signals broadcasted from low elevation satellites will generally be influenced by greater error sources. The thumb rule is to best observe satellites which are 15° above the horizon as positional coordinate's precision will be highest when GDOP is low.



Figure 7: Dilution of precision

[ Source: [http://www.leica-geosystems.com/en/page\\_catalog.htm?cid=227](http://www.leica-geosystems.com/en/page_catalog.htm?cid=227) ]

#### e. Selective Availability (SA)

In order to intentionally deny the civilians and other hostile foreign powers from fully availing precise accuracy measurements using GPS, the U.S Department of Defense have subjected the satellite clocks to “dithering”. This alters the broadcasted time as well as the satellite ephemeris (path) slightly. As a direct consequence, a degradation will be caused in

positional accuracy. Selective availability has been switched on May 1, 2000. To ensure confidence of general users, US government has decided in 2007 to launch GPS satellites (known as GPS III) without selective availability.

#### **f. Anti Spoofing**

This is similar to SA with the difference that anti spoofing tends to encrypt the P code into a signal called the Y code which can only be decrypted by using military GPS receivers. The users of military GPS receivers will usually be privileged with high positional accuracy of around 5m, whereas the civilian users will only get accuracies between 15-100 m.

### **4. Differential Global Positioning System (DGPS)**

Differential phase GPS enables highly precise relative positioning accuracies of 5-50 mm wherein, a minimum of two GPS receivers are always used simultaneously. DGPS is a differential measurement technique which allows the civilian user to increase the positional accuracy from around 100m to about 2-3 m or even less. This would aid in many civilian applications. DGPS will essentially consist of a reference receiver and a rover receiver. The reference receiver will be mounted on a previously known point (i.e., known coordinates). When switched on, the reference receiver will begin to track orbital movements of satellites precisely as it is located on a known coordinate. It can work out the difference between the computed and measured range values. The rover receiver is built to receive the range corrections broadcasted by the reference receiver. The rover also calculates ranges to satellites by applying the range corrections received from the reference receiver. There can be multiple rover receivers receiving corrections from one single reference receiver. DGPS is most commonly used for inshore marine navigation, precision farming etc.

a) Determining correction values

As the location coordinates of reference station are known precisely, the true distance to each of the GPS satellites can be estimated. The difference taken between the true value and the pseudo range will give the correction value. This value will usually be different for every GPS satellite.

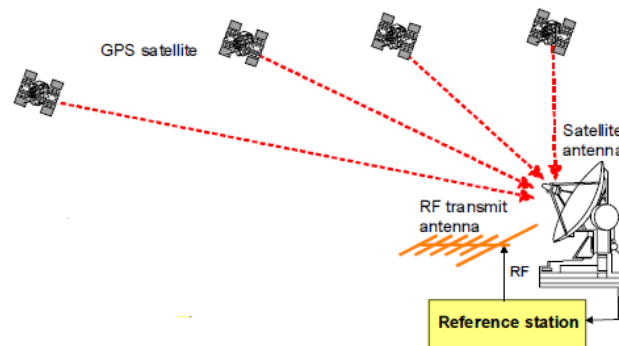


Figure 8: Determining the correction values [Source: <http://www.u-blox.com/>]

#### b) Transmitting of the correction values

The correction values estimated can be used to correct measured pseudo ranges. Hence, these are relayed using a suitable medium.

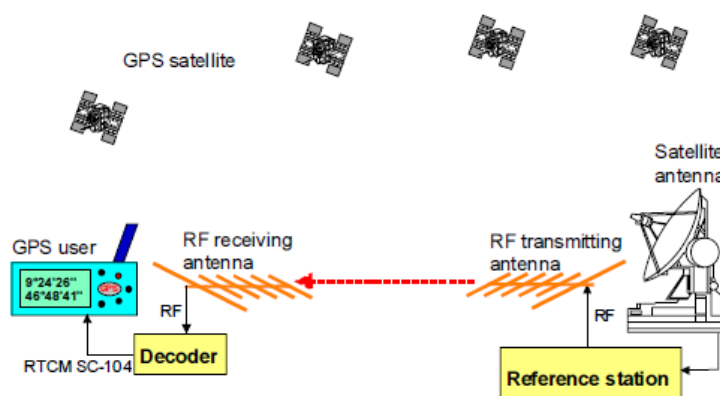


Figure 9: Relaying the correction values

[Source: <http://www.u-blox.com/>]

c) Once the correction values are known beforehand, a GPS user can easily determine the true distance using the measured pseudo range measured. This technique of differential GPS eliminates most of the errors except for those caused due to receiver noise and multipath.

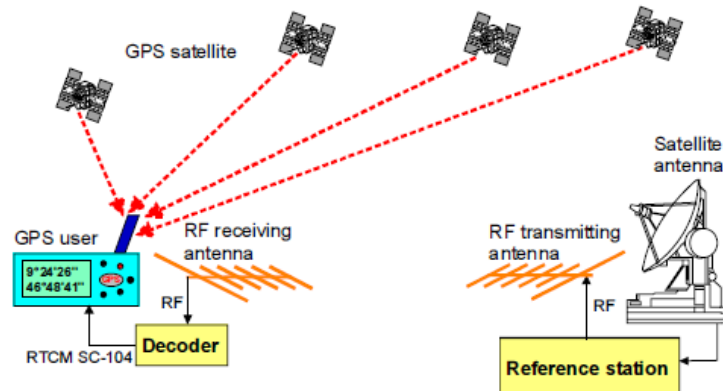


Figure 10: Correcting the measured pseudo range [Source: <http://www.u-blox.com/>]