BRIEF INTRODUCTION TO GPS

GPS is a satellite navigation system conceived, designed and operated by the US DoD. Originally intended to be used for precise positioning through the determination of pseudoranges from the satellites (of which there are ~28 in low earth orbit) to the (normally ground based) receiver. The key idea is that by measuring the time of flight of a radio signal from 4 or more satellites to the receiver, the position of the receiver may be accurately determined. In addition the time offset of the receiver (from composite clock GPS time) may be calculated from information within the orbit data (modulated onto carrier). By taking the time differential of these two quantities, the velocity of the receiver and the frequency offset of the receiver may be ascertained.

SATELLITE SIGNALS

The satellites transmit two L-Band (390-1600 MHz) carrier signals, L1 and L2. The carrier frequencies of L1 and L2 are 1575.42 and 1227.6 MHz respectively. Each carrier is turn modulated (phase shifted by a wave of lower freq. to convey signal) with one or more binary codes.

L1 is modulated with first the C/A (Coarse/Acquisition) code, which is the basis of the standard positioning service (civilian GPS provision). This is a pseudo-random (i.e. random like but actually not) but regularly repeating noise-like code. It has a chipping rate (rate at which binary digits are produced) of 1.023 MHz. The code modulation effectively spreads the spectrum of the carrier signal (i.e. over a far a wider frequency band than is actually required by the quantity of information sent). This gives it high resistance to interference and non- authorised jamming. The code length is limited to 1023 bits, giving a refresh rate (or duration of the code) of 1ms. The C/A code has a fast acquisition time and is easy for users to lock onto. Each of the ~26 active satellites modulates their L1 carrier with a satellite characteristic C/A code, enabling easy satellite identification through C/A code demodulation.

L1 is also modulated with a 50Hz navigation message, which provides GPS satellite orbits, clock corrections etc.

The Precise (P) code modulates both the L1 and L2 carriers, and has a far longer (7-day) duration than the C/A code. It has a chipping rate of 10.23 MHz. C/A code was designed partly to help users acquire the P code. Through a method called anti-spoofing (AS) the P-code is encrypted to form the user-restricted P(Y) code, available only to US military authorised users, through the use decryption keys.

The normal civilian users can all but forget about the P-code due to its encryption. Unfortunately the situation was made worse still in the 1990's through the introduction of Selective Availability (SA); a deliberate distortion of the satellite signals, preventing civilian users from fully utilising the full capabilities of even C/A code. SA is a time varying bias involving either manipulation of the data message (epsilon) and/or clock frequency, with the SA bias being different for each satellite. Just as the pseudoranges are combined, so must the SA biases from each satellite being tracked at a particular time be combined to form the navigational solution. The real problem for SA users is that SA is a time varying bias with low freq. terms in excess of a few hours. This makes averaging of individual pseudoranges (to effectively average away the SA effects) impossible for times less than a few hours. Fortunately for many time and frequency applications, a technique known as static positioning may be used. This allows for position determination using a stationary receiver, allowing implementation of averaging techniques, which greatly improved accuracy.

One of the advantages of the GPS system and indeed an essential feature of operation is that despite deliberate degradation of, and partial restriction to, the carrier and codes, the carrier and data modulating frequencies are held to very precise tolerances.

PURE C/A CODE RECEIVERS

Many low cost receivers track the low frequency (wrt. the carrier frequency) 1 MHz code phase. Internal synthesisers (to the receiver) produce SV specific PRN codes, which are then correlated, with the C/A code as received from each SV (with unique PRN) at the antenna. This method enables this arriving code phase to be evaluated (to within a 1ms ambiguity) using the auto correlation to within 10ns within an observation time of about 1s. By using the time tagged data within the navigation message it is possible to remove the final 1ms ambiguity.

The auto-correlation method (auto correlation is the method by which a signal is compared with itself to find the extent of correspondence between the signals) measures the difference between the propagation time as expected according to the orbital data and the propagation delay as actually measured at the receiver. This gives the time offset of the internal receiver clock relative to the (apparent) GPS time as realised using the satellites in view. It can be calculated that the 10ns code phase evaluation in 1s translates to a frequency determination capability between two successive code phase measurements 1 s apart of 10^{-8}. A pure code phase receiver is therefore only able to discipline an oscillator (say an OCXO) to within 10^{-8} of its nameplate frequency. This is simply not good enough for modern daytime and frequency applications.

Receiver noise limits this accuracy and this may be partially overcome by using averaging. The problem with averaging is that, as always, short-term frequency fluctuation detection is delayed according to the averaging time (similar to τ in the Allan deviation) used. This will mean that the receiver will have a slow response time to any frequency errors in the oscillator that it is disciplining. Therefore using only a pure C/A code receiver, only oscillators, which have a good inherent stability, are capable of being disciplined.

CARRIER PHASE

The C/A code correlation length of 1µs limits dramatically the resolution of the C/A measurement. The substantially higher frequency of the L1 carrier (as compared to the C/A code), and the resulting shorter cycle of 635 ps, will reduce its sensitivity to jamming and also improve the resolution 10000 fold over C/A code measurement. A 1-% noise induced change in the carrier and code signal amplitude results in a phase shift of 10ns and 1ps in the code and carrier respectively.

The advantage of carrier phase tracking is that frequency measurements are achievable with almost no receiver noise contribution. This enables relative frequency determination with uncertainties of a few 10^{-11} within fractions of a second. The short dwell times (on each satellite signal) enable a single time multiplexing channel (tracking of multiple satellite signals by using a rapid sequencing process) instead of the costly multichannel method, with better results

CARRIER AND CODE PHASE

The problem of the carrier phase evaluation method is that different cycles are incapable of being distinguished from each other. This makes it impossible to determine the propagation time of the signal. In a normal time and frequency oriented (i.e. not a costly geodetic receiver where different techniques are often used) receiver the modulated coded sequence must be utilised to determine the propagation time (from which all other properties are derived). The advantage with measuring the carrier phase is that it yields a very precise calculation of the rate of change of the time of flight (i.e. the time differential of the propagation time). Integration of carrier phase gives you a very accurate propagation time.

Therefore the ideal solution is to somehow combine the code and carrier phase measurements so that you get the absolute but noisy information from the C/A code and the extreme (relative) precision from the carrier phase. This will give you the smoothed propagation time without time delay (which results from averaging). This method reduces the receiver noise to nearly zero, making the accuracy of the evaluation not receiver dependent but signal dependant.

The receiver actually performs several independent carrier phase measurements once a second dwelling on each satellite for approximately 80 to 640 ms (quasi simultaneous satellite tracking), the results from which are averaged. By performing an Allan deviation on this measurement method the limiting effect seems to be white frequency modulated noise and not some systematic error. As stated earlier this method enables you to get away with one time multiplexed channel with parallel evaluation.

DIFFERENTIAL GPS

However good you make your receiver, if you operate it in stand-alone mode (i.e. as a single receiver) the accuracy available to you, as a user will always be limited by certain systematic factors, such as SA and ionosphere delays. The effects of SA can be partially or almost totally removed through static positioning and averaging techniques. Whilst this will improve the long-term performance, the short-term stability will still be affected (on the most basic of levels, without correction factors). The effects of the delay due to the ionosphere may be partially eliminated by modelling the local conditions, but in stand-alone receiver this will never be completely removed. Therefore the user interested in top end time and frequency GPS usage must resort to differential GPS, the referencing of the users GPS to a local atomic clock synchronised GPS receiver. This GPS receiver will measure the clock offsets of all satellites in view (remember it's clock offset is zero due to its synchronisation to a local atomic clock, which is not subject to the delays like SA and the ionosphere). This useful data can then be made available to the user interested in quantifying his systematic delays. This can then be used to calibrate out the contribution of SA and the ionosphere (i.e. errors which are –roughly- the same magnitude at the reference and user positions). This necessitates the reference position being 'quite' nearby for this technique to be of any use.

One test of the accuracy of your receivers is not their absolute accuracy, rather the ability of two co-located (i.e. subject to the same systematic delays like SA and ionosphere) to agree. Each receiver is assumed to be independent (which in a sense it is not because each is subject to the same systematic errors) and tracks a satellite with SA activated. The resulting plots of the time development of the internal clock offset of each receiver clearly show the results of the SA perturbation of the signals. This is apparent for each receiver. Taking a closer look at the difference between the code phase measurements made by each receiver, reveals that whilst a certain common factor is removed (i.e. each receiver suffers from similar though not identical delays and perturbations) the remaining amount does not show noise-type characteristics. It is probably due to multipath reflections (i.e. the signal can be received at the antenna after reflecting off an object not by the direct route), which differ between receivers. This illustrates the importance of carefully selecting antenna positions for timing applications and the use of quad helix antennas.

This problem can be partially eliminated if carrier phase measurements are taken into account. The higher frequency of the carrier c.f. the code reduces the effect of reflections and improves the accuracy between two co-located receivers to ~5-10ns wrt apparent GPS time. This is an excellent demonstration of the ability of combined carrier and code phase evaluation to deliver high accuracy (agreement between two co-located receivers) in short observation times

THE QUARTZLOCK GPS-DO SERIES

The Quartzlock line of GPS disciplined oscillators is based upon the type of combined carrier and code evaluating receivers described above.

i) **Down-converter**:

This is not just a signal preamplifier; it has been designed as an integral part of the receiver. Its purpose is to reduce the frequency of the signals travelling down to the receiver. The arriving signal at the antenna is referenced to the receiver local oscillator. Travelling up the down-converter cable to the receiver will be a 92.07 MHz reference frequency from the local oscillator and DC power from the receiver. The 2nd IF signal (the first being at 102.3 MHz and is confined to the receiver) at 10.23 MHz travels down the cable to the receiver. This will reduce the cable loss cf. the 1.6 GHz carrier signal frequency and thereby allow the use of lighter and more flexible cable.

ii) **Antenna**:

This is a quad-helix antenna and is mounted to the down-converter by means of type N connectors (due to the frequency being transmitted between them). If cable must be fitted between the antenna and down-converter then the maximum (theoretical) loss when carrying the 1.6 GHz carrier signal should be not greater than 3 dB. Thus a short length of RG213 (<2.5m) would acceptable. In order to obtain the peak performance out of the unit, the antenna must have a good view of the sky. Ideally the antenna position should be known accurately (i.e. to within +/- 2m latitude/longitude and +/- 4m altitude) before operation, as this will reduce the time to first fix (TTFF). One advantage of using the quad helix antenna is that troublesome multipath effects are all but eliminated. Multipath is a signal arrival at a receiver's antenna by way of two or more different paths such as a direct, line-of-sight path and one that includes reflections off nearby objects. The difference in path lengths causes the signals to interfere at the antenna and can corrupt the receiver's pseudorange and carrier-phase measurements. Multipath error is the GPS positioning error caused by the interaction of the GPS satellite signal and its reflection. The positioning error is due to interference between the radio signals, which pass from the transmitter to the receiver by two paths of different electric length.

iii) Time constant

A short (loop) **time constant** will give a very fast response time, to time errors. The problem with this is that little or no averaging is done to eliminate SA, leading to a significant degradation of the short-term frequency stability. (c.f. free running oscillator stability). This is fine for timing applications. A long time constant will allow for slow response to time errors but will 'ride' over many of deleterious effects of SA and allow the short-medium term stability to be primarily determined by the local oscillator. However, the user must be careful that he does not select a large maximum time constant without knowledge of the (frequency) performance of the local oscillator. The loop time constant may approach the maximum time constant too quickly for the LO, and not correct for errors in the LO. This could cause LO time errors to exceed a certain threshold for a small period of time. However, if this is only for a short period of time, time coherence will not be lost (time accuracy will be restored without any loss or gain of cycles-cycle slips- at a frequency output wrt the 1pps output). This ensures coherence between the time and frequency outputs. A far worse situation will occur if the unit has undergone a power failure. In this case, the error that will have build up in the LO that only way to restore synchronisation will be to reset the time counters in the receiver. Such total loss of synchronisation would cause a red LED to flash, alerting the user to this problem. If the apparent time error δt relative to apparent GPS time has exceeded a predetermined threshold for more than a set time, then the oscillator control time constant is automatically reduced 1s/s until the error drops back below the threshold value.

iv) Positioning-

Ideally the antenna position should be known accurately (i.e. to within +/- 2m latitude/longitude and +/-4m altitude) *before* operation, as this will reduce the time to first fix (TTFF). If this is impossible (likely), then the receiver must attempt to estimate the position for itself. This necessitates at least 4 satellites being visible (3 for spatial dimensions and 1 time dimension). In order to be able to assist the receiver, the user has the option of entering an approximate position (and also approximate time) which will help the receiver search for satellites that *should* be visible according to the almanac stored in the receiver. This is a set of parameters similar to the more precise ephemeris data, which is used for approximating GPS satellite orbits. In order for the required 4 satellites to be visible for an acceptable period of time, the aforementioned antenna position must be good. An obstructed view of the sky will reduce the time when at least 4 satellites are visible will drop, the geometry of constellation will be degraded (complicated and difficult to model) and the overall time to determine an accurate position and time will increase.

Studies have shown that there appears to be a 'hole' in the GPS constellation-looking north. It is therefore doubly important that the antenna have a 'good look' south. In order to eliminate single position estimate errors an position averaging procedure is carried out in receiver, with up to a day's worth (86400) of 1-second estimates capable of being averaged. Due to the way memory is assigned in the unit, no further updates are made subsequent to this. One-way of improving the position determination is to turn the unit on for 2 days, note the ~24hr average and then repeat the process as many times as you see fit. Taking the standard deviation will give the precision of the position determination, and will allow manual entry at switch on. The receiver can be forced to then operate on this (entered) position. Determination of accurate altitude, whilst more difficult to do, is more important. This is because the satellites are always at a positive altitude.

Tests have shown that using this method the Quartzlock model can ascertain it's position to within +/- 2m latitude/longitude and +/-4m altitude to within a 95% (2σ) certainty. The last digit on the display in the position menu has a resolution of about 1.8m longitude and 1.8cosX, where X is the latitude of the receiver. At 55°N this gives about 1.2m. This means that if two co-located receivers agree down to this last digit, they agree to within to within 1.8m and 1.2m in longitude and latitude respectively. However, in the serial port data, which is viewable through specially designed monitoring software, an extra digit is supplied, allowing precision down to the ~10cm level. By performing the manual averaging procedure at this level of precision the user gets unrivalled antenna position determination, and the associated benefits this brings.

Like most GPS receivers of this type, this series references its position to the world geodetic system 1984 (WGS 1984) with the altitude being relative to the geoid. Whilst the geoid is much more complex than the simple ellipsoid, it is an approximation of the true shape of the earth and is therefore closer to mean sea level for most places on earth. The difference between the ellipsoid and geoid is stored within the almanac data contained in the navigation message. The receiver has a plausibility checking method to ensure that erroneous entered positions are not used, making continuous comparisons against it's own averaged position. This should prevent undue timing errors resulting. The user should also ensure that after 24 hours the averaged position agrees with the entered position to within +/- 3m latitude/longitude. If this is not the case, the entered position was wrong! It is also important to tell the unit what position to use. If the user has a very accurate position, determined either through a geodetic survey or repeated position averages (both with same unit and different co-located similar units), then the user must instruct the unit to use this (not the result from the last position average estimate)

v) **Warnings**

In order for the unit to operate properly the unit must be set up as detailed in **[5]**. If a fault exists at the power up of the unit, indication is likely to be given via a front panel warning display. An common example is if the entered position failed the plausibility check, i.e. the receiver has switched to using averaged position or if satellites have been found that *should* be below the horizon according the almanac data for that position). Many of these initial problems will go away as the averaged position is used or the almanac data is updated (after about 15-20 minutes after switch on). Other problems like a missing antenna; down-converter or cable (or indeed if any of these are faulty) will only be detected once the internal LO (normally OCXO or rubidium) has warmed up. These can take up to 10 minutes depending on the type of oscillator used. Such messages would need to be thoroughly investigated if normal operation is to be obtained from the unit. Indeed, the amount of noise detectable is a good indication of the health of the receiver. A partially or totally obscured antenna could cause insufficient or no satellites to be seen.

However, quantification of GPS-DO signal degradation due to different levels of antenna obstruction is difficult, but work is on hand to do this. It is important to remember that the RMS errors in apparent (i.e. as realised locally) GPS time will be greater than with a full constellation. The solution, as always, is to improve the antenna position, to reduce the instances when this might happen. Other errors may occur if the user was tracking a particular satellite and it went (temporarily) out of view. The receiver in this case would resort to an ‘all in view’ mode.

Another important message indicates whether the control voltage, which is applied to the local oscillator to correct for frequency excursions, is above a certain normal threshold. Abnormally high voltages being applied (c.f. what the receiver believes should be applied) indicate problems with the local oscillator associated primarily with drift/ aging of quartz crystals or possible failure of the rubidium physics package. Normal operation (i.e. output precision is not necessarily adversely affected) is continued whilst this message is displayed, but future investigation should take place. If it occurs during the first few minutes of operation in a rubidium oscillator, it may well simply be a result of the control DAC limiting. This is due to the very small adjustment range of the Rb oscillator (the software for the rubidium option is different for the rubidium LO than for the OCXO LO, to account for their different characteristics). This message will then go away after the rubidium has warmed and settled.

vi) Delay

The **delay** option enables the user to select the delay to be applied to the 1pps output with a maximum of ±500ms thus effectively providing any required time offset with 1ns resolution. This is designed to calibrate out ionosphere and troposphere delays, and antenna/down-converter/cable delays. This is important if the GPS-DO is to be used for time dissemination i.e. not as a frequency standard. Note that changing the delay once the "locked" condition has been achieved may result in loss of lock and will almost certainly cause transient timing and frequency errors. The 1pps delay should be corrected before lock has been achieved.

vii) Time

The time will be displayed including seconds as soon as the receiver has started tracking at least one satellite. Upon turn on the seconds are suppressed because of the uncertainty associated with only having the internal back-up clock as a reference. During the period between power up and satellite tracking commencement, the user is free to alter the time manually because the internal master clock is not ‘set’. Setting involves confirmation by a satellite. Manually altering the time to within ~30 minutes of GPS time reduces the TTFF by providing a time estimate for the receiver to ‘going-on’ with. Occasionally almanac data stored in the receiver will be too old for UTC to be calculated from GPS (i.e. once locked) until new almanac data is downloaded from the space vehicles. During the time taken to achieve lock (i.e. δt & $\delta f/f$ are within prescribed limits-different for each LO and are of opposite signs) the 1pps remains inactive.