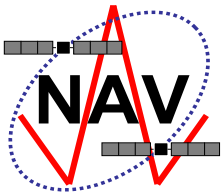


Technische Universität München
Lehrstuhl für Kommunikation und Navigation

Prof. Dr. Christoph Günther

Diploma Thesis

GNSS Carrier Phase Multipath Analysis

Supervisor: Dipl.-Ing. Sebastian Graf (NAV)
Author: Cand.-Ing. Markus Rippl
markus.rippl@mytum.de
Date of submission: 09.05.2007

Abstract

This Diploma Work proposes a method to estimate the presence of multipath in a GPS satellite signal. The method can provide a lower bound on the number of reflections, the attenuation factor of the individual paths with respect to the line-of-sight path. For the estimation, the satellite signal carrier is reconstructed by adding the phase feedback of the carrier tracking loop, and the I and Q outputs back together. The estimation technique is validated with measurements taken on the Spirent Constellation Simulator which is operated at DLR Oberpfaffenhofen. Measurements with the present GPS constellation which were done in an urban surrounding in Munich, suggest that the discovered effect is also present in reality.

Contents

Contents **5**

1 Introduction **7**

2 Direct signal characteristics **9**

2.1 Modulation 9

2.2 Signal Transmission: The Doppler Effect 12

2.3 Direct Path Signal 17

2.4 Theoretical Tracking Error Variance 18

2.4.1 Power Density Spectrum of the L1 C/A Signal 18

2.4.2 Autocorrelation Function 19

2.4.3 Code Tracking Performance 21

2.4.4 Carrier Tracking Performance 24

3 Multipath signal characteristics **27**

3.1 System Model of the Multipath Signal Reception Chain . . . 27

3.1.1 Direct Path Signal 30

3.1.2 Reflected Signal 31

3.1.3 Multiple Signals 34

3.2 Typical Properties of the Reflected Path Component 35

3.2.1 Reflection Geometry 35

3.2.2 Satellite Orbits and Angle of Arrival 46

3.2.3 Non-stationary Receivers 53

3.2.4 Summary of Multipath Signal Properties 53

4 Multipath Detection and Estimation **55**

4.1 Multipath Effects on Signal Processing 56

4.1.1 GNSS receiver 56

4.1.2 Correlator Function Distortion 57

4.1.3 Carrier Distortion 63

4.1.4 PLL Tracking Influence 74

4.1.5 Special Case: Sign Reversion 75

4.2 Multipath Mitigation by Carrier Phase Averaging 76

4.2.1 Carrier Phase Modulation 76

4.3 Multipath Estimation with Carrier Phase Analysis 79

4.3.1 Spectral Analysis of Distorted GNSS Carrier 79

4.3.2 Estimation of Reflected Path Parameters from Fourier Spectrum 93

5 Estimation with Signals **95**

5.1 System Configuration 95

5.1.1 Signal Generation 96

5.1.2 Signal Measurement 98

5.1.3 Signal Acquisition, Tracking and Post-Processing . . . 99

5.2 Estimation with Simulated Signals 104

5.2.1 Signals from MATLAB Signal Generator 104

5.2.2 SPIRENT Simulator 105

5.3 Estimation of Measured GPS Signals 109

5.3.1 Measurements Description 109

5.3.2 Measurement Results 109

6 Conclusion and Future Work **113**

7 Bibliography **117**

Chapter 1

Introduction

Global navigation is beginning to occupy a more and more important place in society - tomorrow, knowing about one's exact position will maybe have the same significance as knowing the time today. Instant positioning in the blink of an eye, with exact, reliable and available results, offers new opportunities in many applications.

Present Global Satellite Navigation Systems determine the user position by measuring the propagation time of the signals of at least four satellites to the user antenna. The quality of these position solutions depends on knowledge of the satellite position, the propagation speed and also the signal path - only if the signal travels on a direct line-of-sight path, the path length is equal to the distance between the satellite and the user. In practice, reflected signals alone are unusable to determine the distance between sender and receiver.

While in wireless communications engineers have taken advantage from multiple signal copies, each caused by a different reflection, the measurement of signal propagation delays leaves no option for such a method. It is therefore desireable to known whether the received signal is a pure line-of-sight signal, or whether it contains delayed signal components which would degrade the quality of the measurement.

Chapter 2

Direct signal characteristics

In this chapter, a mathematical model for the GPS C/A signal is introduced. The following description of the signal will cover its generation at the satellite, the transmission through space on a direct path, and the reception and down conversion of the signal in the user receiver. The signal model is used later in Section 3.1, where a system model for the received signal under multipath influence will be constructed.

2.1 Modulation

GPS satellites transmit a multitude of signals, both for military and civil use, and modulated onto different carrier frequencies. Recent professional GPS receivers in the civil field make use of multiple frequency bands; however, the scope of the present thesis only covers the analysis of the single frequency L1 C/A signal. This signal component is contained in the L1 carrier at $f_c = 1.57542\text{GHz}$, modulated onto the in-phase component. The quadrature component of L1 carries a P(Y) code, which is available to military users only [ns04].

The C/A signal used in GPS is a Direct Sequence Spread Spectrum (DSSS) signal, which uses a Gold code spreading sequence and BPSK data modulation [ME01]. The spreading sequence has a chipping rate of $f_C = 1.023\text{MHz}$ and a code period of $T_{CODE} = 1\text{ms}$. The length of this sequence, which is also called PRN (pseudo-random noise) sequence, is therefore 1023 chips per code sequence. The PRN sequences used in GPS are optimized with respect to minimum auto- and crosscorrelation, making it possible to concurrently use the same frequency for all satellites. The satellites transmit a navigation message at 50 bits/s. The data bits are modulated onto the spreading sequence with a bit length of $T_b = 20\text{ms}$, where bits may change synchronously with the start of a new chip sequence. Hence, one data bit

is constructed of 20 concluding chipping sequences. The relation between bits, PRN chips and the carrier cycles is shown in Figure 2.1.

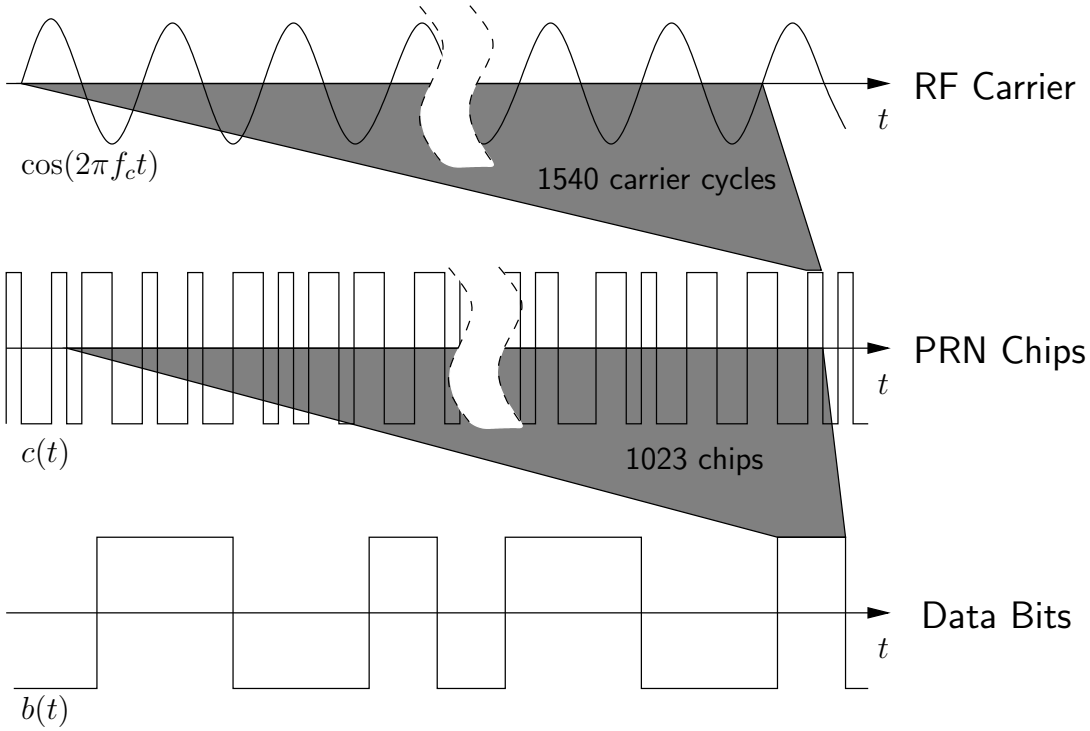


Figure 2.1: Timeline of the DSSS signal modulation used for the GPS C/A code

The transmitted signal is created by modulating the data bits and code chips onto the carrier. There are two ways to represent the modulation of chips with bits: Either, if possible bit values are zero and one, the modulation is represented by a modulus-2 addition. This operation is also equivalent to XOR and may be written as such.

If the bit representation is -1/+1, the modulation is a simple multiplication of carrier and bits, or chips, respectively. We will use this denotation in the following, because it is easier to handle and corresponds to actual signal amplitudes in a hardware implementation of the GNSS system.

Together, the signal modulation can be written as product of data bit, code chip, and carrier function:

$$s(t) = A \sum_{m=-\infty}^{+\infty} b_m \sum_{n=0}^{N-1} c_n p(t - mT_b - nT_c) \cos(2\pi f_c t + \varphi) \ , \qquad (2.1)$$

where the current data bit is stored in b_m, the chip value comes from c_n, and the cosine term applies the carrier oscillation to the spreading signal, making it the in-phase component RF signal at f_c which can be transmitted. The phase shift of the carrier frequency is given by the angle φ, and the signal has an amplitude of A. The unity impulse function $p(t)$ for a chip with duration T_c is

$$p(t) = \begin{cases} 1 & |t| \leq \frac{T_c}{2} \\ 0 & \text{else} \end{cases} . \tag{2.2}$$

In (2.1), the bit and chip sequences are written as a discrete series of possible values -1 or $+1$. Obviously, there is a relationship between the time t (that appears only in the carrier oscillation up to now) and the individual values of the b_m and c_n. The bit and chip sequence functions may also be written as a continuous functions of time, which results in the following representation of the modulated signal, and removes the need to use the impulse function:

$$s(t) = A \cdot b(t) \cdot c(t) \cdot \cos(2\pi f_c t + \varphi) \tag{2.3}$$

If $c_{PRN}[n] \in \{-1; 1\}$ is the discrete PRN chip sequence of the satellite, where $c_{PRN}[0]$ denotes the first chip value and $c_{PRN}[1022]$ stands for the last PRN chip, the time continuous chip function $c(t)$ can be written as

$$c(t) = c_{PRN}\left[\left\lfloor \frac{t}{T_{chips}} \pmod{T_{code}} \right\rfloor\right] . \tag{2.4}$$

The same method is used to determine the data bit value with respect to time:

$$b(t) = b_{NAV}\left[\left\lfloor \frac{t}{T_{bit}} \pmod{T_{NAV}} \right\rfloor\right] , \tag{2.5}$$

where the time interval $T_{bit} = 0.02$s is the duration of one bit, and T_{NAV} is the length of the navigation message, which is 12.5 minutes [ME01] for GPS. Figure 2.2 shows the modulation process in a block diagram.

We can see that the discussed part of the signal occupies the in-phase component of the RF signal only. The quadrature component of the L1 signal carries the P/Y code. This military P/Y signal uses a chipping rate ten times higher than that of C/A, and a much longer code duration of one week [ME01]. The correlation of P/Y code with the C/A code copy in the receiver, however, results in unbiased noise, so that the quadrature component has no effect on the in-phase component, even when the phase is tracked erroneously and some part of the quadrature signal is processed.

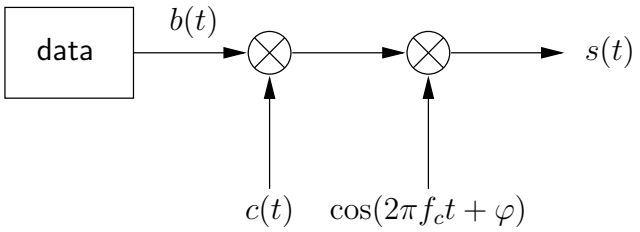


Figure 2.2: Modulation of the GPS C/A signal [Gü05]

Relativistic Correction

One important fact concerning the implementation of the space segment is the relativistic effect of time dilation, which makes the satellites' clock appear running slower from the user's point of view. This effect occurs because of the high velocity of the satellite, with respect to the user's coordinate reference frame (relative velocity). To compensate the time dilation, the on-board frequency normal is corrected accordingly to make the observed frequency appear correct. Instead of 10.23MHz, the board clock of the satellite is set to run at 10.22999999545 MHz prior to launch [Kap96].

In the scope of the following discussions, however, it is not necessary to consider this effect. The time dilation is fully compensated by the described frequency adaption, and thus we can just assume that the nominal frequency is transmitted, neglecting the time dilation effect together with its countermeasures.

2.2 Signal Transmission: The Doppler Effect

After the signal leaves the transmitting antenna, it travels a distance of about 20200km to 26000km before it reaches the user antenna. The signal's path leads mostly through vacuum, where the propagation speed is $c_0 = 299,792,458\frac{m}{s}$. The very last part of its journey, the signal enters atmospheric layers known as the ionosphere and the troposphere. Both contain particles, which slow down the signal. Compensation of this effect is a key feature of modern augmentation systems thats usually provide detailed data describing the distortion as specifically as possible.

Neglecting ionospheric and troposheric effects, the time between transmission and reception is denoted as

$$\tau = t_r - t_s \ , \tag{2.6}$$

where t_s is the time instant when the signal is transmitted at the satellite, and t_r describes the time when the signal reaches the user.

Annotation: Of course, both times are denoted in the same, virtual time system, which is the common GPS time,or GPST, in our case. The satellite clocks and the user clocks do not match exactly in their times in practice. However, the clock offsets with respect to the common base of all involved systems are estimated and considered in the measurement process. In the following, all times are defined to refer to that common time.

Apart from the transmission delay, the signal is also attenuated on the way to the user by path loss, antenna patterns, atmosphere and other factors. With the transmission delay given, the received signal can be written as a delayed and attenuated form of the transmitted signal:

$$r(t) = G \cdot s(t-\tau) = A' \cdot b(t-\tau) \cdot c(t-\tau) \cos(2\pi f_c(t-\tau) + \varphi) \; ; \qquad (2.7)$$

where the time t was substituted by $t - \tau$ in the formula to introduce the delay. The path loss and other factors between transmitter and receiver attenuate the signal amplitude, which is denoted as a gain factor $G < 1$. The corrected amplitude $A' = G \cdot A$ accounts for this path loss in the second part of the equation. A block diagram of the signal transmission can be seen in Figure 2.3. At this stage, the transmission delay τ is denoted as a time invariant value.

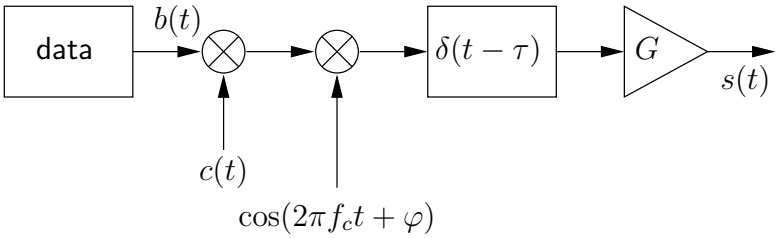


Figure 2.3: Signal transmission with time invariant transmission delay

In practice however, the propagation time τ is a value which depends directly on the distance ρ between the satellite and the user, that being a rapidly changing value. Assuming a fixed mean propagation speed along the signal path, τ may be considered directly proportional to the distance ρ between the satellite and the user:

$$\tau = \frac{\rho}{\bar{c}} \ , \tag{2.8}$$

where the propagation speed $\bar{c}$ was newly introduced. Because of atmospheric delays, it is not sufficiently exact to use the speed of light in vacuum in (2.8). The mean propagation speed, averaged over the full path from the satellite down to the user, is

$$\bar{c} = \frac{1}{\rho} \int_0^\rho c(\rho') d\rho' \ . \tag{2.9}$$

where the path dependent propagation speed function is $c(\rho)$, denoting the local propagation speed at the location which is at distance ρ from the satellite, and lies on the signal path from the satellite to the user.

For a single satellite, the mean propagation speed $\bar{c}$ may be assumed being constant within a reasonable time frame, because the propagation speed is locally correlated and the path track does not change very fast.

Because of satellite motion, earth rotation and also user motion, the distance ρ that separates user and satellite changes rapidly over time. It is hence necessary to use a time variant value for τ, which results in a modified Equation (2.7):

$$r(t) = A' \cdot b\big(t - \tau(t)\big) \cdot c\big(t - \tau(t)\big) \cos\Big(2\pi f_c\big(t - \tau(t)\big) + \varphi\Big) \ , \tag{2.10}$$

where the transmitted signal amplitude and the path attenuation are summarized to $A' = G \cdot A$. The time variant transmission time is accounted for in Figure 2.4:

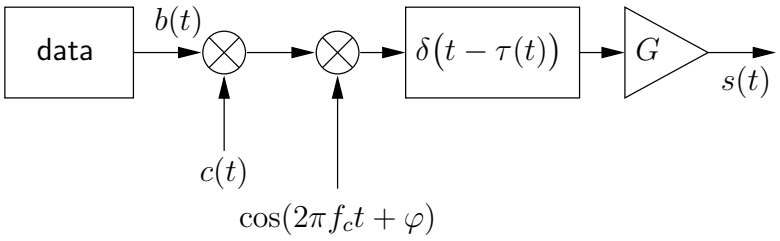


Figure 2.4: Signal transmission with time variant transmission delay

The path delay is a function of the distance, as was shown in (2.8). With v_{LOS} defined as the satellite's relative speed in the line-of-sight direction, seen from the user, ρ may be written as

$$\rho = \rho_0 + v_{\mathrm{LOS}}\ t \ , \tag{2.11}$$

and thus, $\tau(t)$ is

$$\tau(t) = \frac{\rho_0 + v_{\mathrm{LOS}} t}{\bar{c}} \ . \tag{2.12}$$

The delayed time $t - \tau$ can then be written as

$$t - \tau(t) = t - \big(\frac{\rho_0}{\bar{c}} + \frac{v_{\mathrm{LOS}} t}{\bar{c}}\big) = \big(1 - \frac{v_{\mathrm{LOS}}}{\bar{c}}\big)t - \frac{\rho_0}{\bar{c}} \ . \tag{2.13}$$

Plugging (2.13) into (2.10) returns then

$$\begin{aligned} r(t) \ &= \ A' \cdot b\big((1 - \frac{v_{\mathrm{LOS}}}{\bar{c}})t - \frac{\rho_0}{\bar{c}}\big) \cdot c\big((1 - \frac{v_{\mathrm{LOS}}}{\bar{c}})t - \frac{\rho_0}{\bar{c}}\big) \cdot \\ &\quad \cdot \cos\Big(2\pi f_c\big((1 - \frac{v_{\mathrm{LOS}}}{\bar{c}})t - \frac{\rho_0}{\bar{c}}\big) + \varphi\Big) \end{aligned} \tag{2.14}$$

The carrier component of the above equation can easily be rearranged to split the time dependent part from the constant part. Then, the equation of the carrier is:

$$r_C(t) = \cos\Big(2\pi f_c((1-\frac{v_{\mathrm{LOS}}}{\bar{c}})t-\frac{\rho_0}{\bar{c}})+\varphi\Big) = \cos\Big(2\pi(1-\frac{v_{\mathrm{LOS}}}{\bar{c}})f_c t+\big(2\pi f_c\frac{\rho_0}{\bar{c}}+\varphi\big)\Big) \ , \tag{2.15}$$

where we can now substitute

$$\widetilde{f}_c = (1 - \frac{v_{\mathrm{LOS}}}{\bar{c}}) f_c \ .$$

$\widetilde{f}_c$ is the received carrier frequency, resulting from the transmitted carrier frequency f_c with a correction factor depending on the change rate of the receiver distance $\tau(t)$. The user therefore receives a shifted carrier because of the relative motion between satellite and user. This phenomenon is known as Doppler effect after its discoverer, Christian Doppler; it is the key to the estimation technique that is proposed by this work.

The frequency term can also be split up into the original frequency f_c and a shifting component,

$$\widetilde{f}_c = f_c - \frac{v_{\mathrm{LOS}}}{\bar{c}} \cdot f_c = f_c + \Delta f \ , \tag{2.16}$$

resulting in the definition of the Doppler shift

$$\Delta f = -\frac{v_{\mathrm{LOS}}}{\bar{c}} \cdot f_c \ . \tag{2.17}$$

The velocity in LOS direction, v_{LOS}, has been defined such as to be positive if the the distance between satellite and user increases. The carrier signal (2.15) can now be written using the Doppler shift:

$$r_C(t) = \cos\Big(2\pi(f_c + \Delta f)t + \varphi'\Big) \ , \tag{2.18}$$

where the time invariant expression $\big(2\pi f_c w + \varphi\big)$ has been combined into φ' for a better readability.

The relative LOS speed v_{LOS} is the first derivative with respect to time of the path length ρ. From (2.12), it follows

$$\frac{d}{dt}\tau(t) \quad = \quad \frac{v_{\mathrm{LOS}}}{\bar{c}} \tag{2.19}$$

So, instead of the relative LOS speed v_{LOS}, the path delay change can also be described by the first derivative of the path delay $\tau(t)$:

$$\Delta f = -f_c \cdot \frac{d}{dt}\tau(t) \ , \tag{2.20}$$

or, with the distance ρ plugged in from (2.8), as

$$\Delta f = -f_c \cdot \bar{c} \cdot \frac{d}{dt}\rho(t) \ . \tag{2.21}$$

With the definition from (2.11), the distance between the satellite and the user increases when the relative speed is positive, resulting in a negative Doppler shift; the signal is "expanded" in space. Likewise, an approaching satellite produces a positive Doppler shift (the signal is "compressed" in space).

Because the GPS satellites fly in non-stationary orbits above the user, their relative speed towards the user usually changes constantly. Therefore, the received carrier signal (2.18) must again be written with a time variant Doppler shift:

$$r_C(t) = \cos\Big(2\pi\big(f_c + \Delta f(t)\big)t + \varphi'\Big) \tag{2.22}$$

If the user is assumed to be stationary and on the surface of the earth, the distance between satellite and user only depends on the satellite's orbit. For a satellite crossing the zenith of the user, their relative distance changes from about 26,000 km at its rise to about 20,200 km for the position in the zenith, and back to 26,000 km shortly before setting at the horizon. This means that a satellite that rises from the horizon is observed under a positive Doppler shift, because its distance decreases. A setting satellite is seen with a negative Doppler shift; its signals arrive slower than those of the first example, and are received at a lower carrier frequency.

With knowledge about the satellite's exact track, it is possible to find out its current position by measuring the Doppler shift of the carrier with a stationary receiver. This effect was taken advantage of in one of the early satellite based navigation systems called TRANSIT [Gü05].

2.3 Direct Path Signal

Section 2.1 highlighted the creation of the modulated signal, featuring an early frequency correction mechanism that compensates for relativistic effects. In Section 2.2, the transmission to a receiver on the earth surface was analyzed. The main effects are signal attenuation, signal delay and the Doppler frequency shift, which originates in the relative velocity between the satellite and the user.

The transformation from a time dependent transmission delay to a Doppler shift frequency was shown in the previous section. However, the Doppler shift in time occurs also with PRN chips and data bits.

Several representations of the Doppler shift exist: First, it can be given as an absolute value that is added to a signal's frequency; in the previous section, this value was called Δf. Second, the Doppler shift may be applied indirectly by using a changing delay in the time argument, as it was done with $\tau(t)$. Here, the Doppler shift is hidden in the slope of τ, as shown in (2.20). On the other hand, the downside of the first representation is that the Doppler shift for different signal components is not necessarily the same, because it depends on the component's frequency. The perturbation of the different components is not the same in absolute numbers, but only in their relative fraction with respect to the corresponding base frequency. For example, if the Doppler shift of the carrier is $\Delta f_c = +6kHz$, the much slower running chip rate is only increased by

$$\Delta f_{CH} = \Delta f_c \cdot \frac{f_{CH}}{f_c} = 6\text{kHz} \cdot \frac{1.023\text{MHz}}{1.57542\text{GHz}} \approx 3.9\text{Hz} \ . \tag{2.23}$$

A third possibility is therefore the declaration of a relative shifting factor, or perturbation factor [Psi01]. An advantage is that this factor is the same for all components. The downside is that we cannot read off the true absolute Doppler frequency shift of an individual component directly. The combinded signal using this perturbation factor is

$$r(t) = A' \ b\big((1+\eta)t\big) \ c\big((1+\eta)t\big) \cdot \cos\big(2\pi(1+\eta)f_c t + \varphi_{n,k}\big) \ . \tag{2.24}$$

Here, the time argument t is affected in the same way for all three components of the signal (bits, chips, and the carrier). For the carrier, we can read off its original frequency from the perturbation factorand the carrier frequency, because it is represented as a cosine oscillation. The bit and chip functions do not show their rates in this formula; however, if we know that the transmitted chip rate is f_{CH}, the received chip rate will be $(1+\eta)f_{CH}$. The perturbation factor η is derived from the absolute Doppler shift by

$$\eta = \frac{\Delta f}{f} \tag{2.25}$$

for a Doppler shift of Δf with respect to a base frequency of f.

Given the above formulas, we are now ready to examine the origination and effects of multipath. When analyzing the degradation of the multipath stressed signal, the interesting measure is the tracking accuracy. Since the error in tracking is a statistical process, it is best described with a mean value and the corresponding variance, or standard deviation. Before the error in the context of multipath is examined, it is interesting to know the theoretical tracking performance of this signal. This topic is covered in the following section.

2.4 Theoretical Tracking Error Variance

The signal form of the GPS C/A signal has been explained in the previous sections. For the analysis of multipath induced tracking errors, it is useful to know the tracking performance of a theoretically noise-free satellite signal which is exclusively received on the direct path. Given this value, the tracking performance with a certain multipath infliction will be a more concrete result.

2.4.1 Power Density Spectrum of the L1 C/A Signal

The GPS L1 C/A power density spectrum is a line spectrum, shaped with the envelope of a squared sinc function. The lines are equally spaced at the distance of 1kHz, which is caused by the repetition rate of the C/A code sequence of 1ms. The envelope shape follows a sinc^2 function, where nulls occur every 1.023 MHz. The main lobe's width is therefore 2.046 MHz, and all sidelobes are 1.023 MHz broad. This characteristic results from the chip length $T_c = 1/1023$ ms. Of course, the center of the signal spectrum is located at $f_c = 1.57542$ GHz, if the signal is not Doppler shifted.

The transmitted signal is specified with a L1 bandwidth of $24 \cdot 1.023$ MHz = 24.552 MHz [Gü05]. On the receiver side, a processing bandwidth of 20.46 MHz is usually chosen. After this bandwidth limitation, the spectrum at the receiver includes everything out to the 10th nulls, to the left and to the right.

The spectrum including the main sidelobe and nine sidelobes is drawn in Figure 2.5. The envelope following a sinc^2-function shape can be seen very clearly. In this overview, the line components of the spectrum cannot yet be distinguished in the chosen scale, because their spacing is only 1 kHz. One side lobe consists of 1023 such lines, each reaching to a different power level which is individual to the chosen code sequence. Only when averaged, the sinc^2 shape can be seen properly.

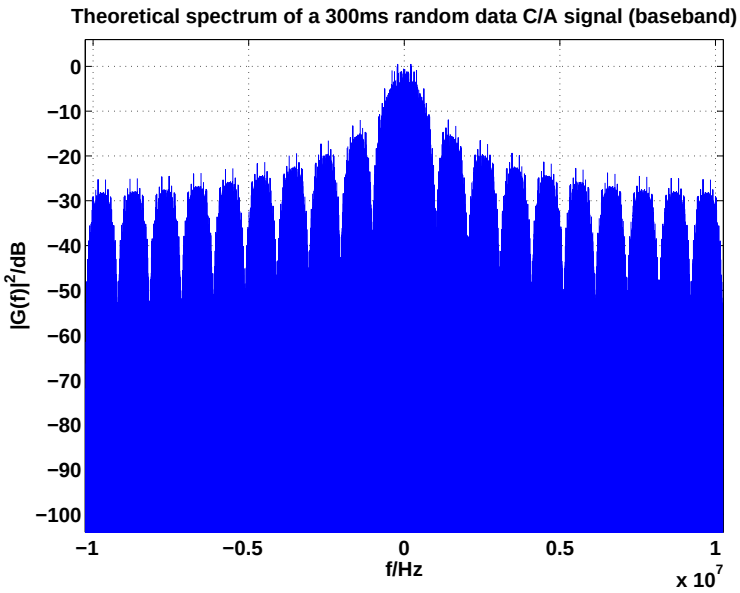


Figure 2.5: C/A spectrum at a 20 MHz bandwidth

Details of the spectrum are shown in Figure 2.6. The frequency range here spans from 0 to 12kHz, which is in the very first part of the main lobe. 12 peaks, one every Kilohertz and each at a different level, can be seen clearly. Ideally, the spectrum would appear as a set of completely distinct lines; but since this spectrum was derived using discrete signal processing techniques, the leakage effect which is implied by using discrete Fourier transforms distorts the shown spectrum [Lyo04].

2.4.2 Autocorrelation Function

The C/A signal's power density spectrum can be obtained by applying a Fourier transform to the autocorrelation function of the signal[Lan02]:

$$|S(f)|^2 = \mathcal{F}\{\varphi_{XX}(\tau)\} \tag{2.26}$$

Equivalently, the autocorrelation function can also be derived from the inverse Fourier transform of the power spectral density. To determine the bandwidth dependent tracking delay error variance, which is described in the following section, the autocorrelation function of the band limited GNSS signal was computed. This was done by constructing the exact frequency spectrum of a satellite's baseband signal by Fourier transforming the Gold code, made of rectangular pulses of duration T_c, into the frequency domain.

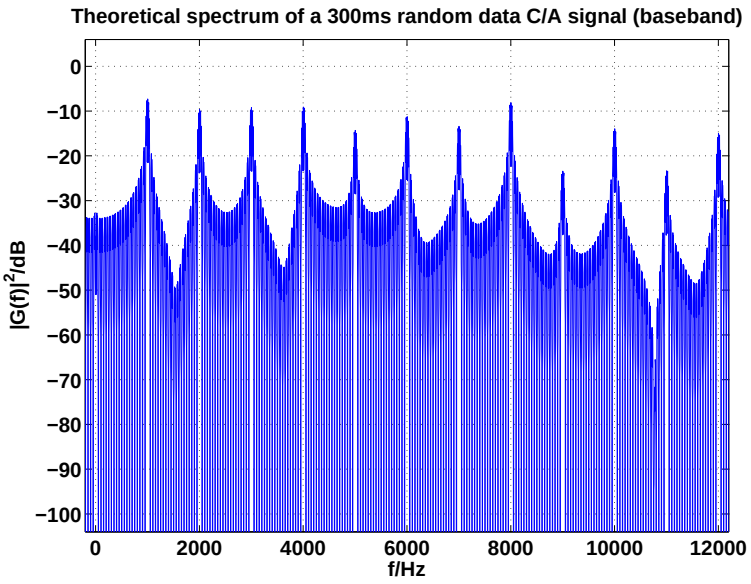


Figure 2.6: C/A spectrum detail showing the line components

The squared spectrum could then be cut off at the bandwidth limitation frequency, and transformed back into the time domain. The result is the autocorrelation function of the Gold code, where the processing bandwidth was limited. The ACF peak for different bandwidths is shown in Figure 2.7.

The result shows that, with decreasing available bandwidth, the correlation peak gets blunted, compared with the extremely sharp peak if the bandwidth is very high. If less than the main lobe is available as signal, the correlation peak degrades extremely. This degradation affects also the synchronization performance, which will be shown later in detail. An explanation for this behavior may be that bandwidth limitation of less than 1.023MHz (one-sided bandwidth) also limits the resolution in time to less than one chip length ($1/1023$ ms), and thus distorts the pulse shape signficantly, which leads to the observed effect. This massive degradation can be seen in the two lowest peaks of the figure, resulting from a (two-sided) bandwidth limitation of 1.023 MHz (half main lobe) and 511.5 kHz (quarter main lobe).

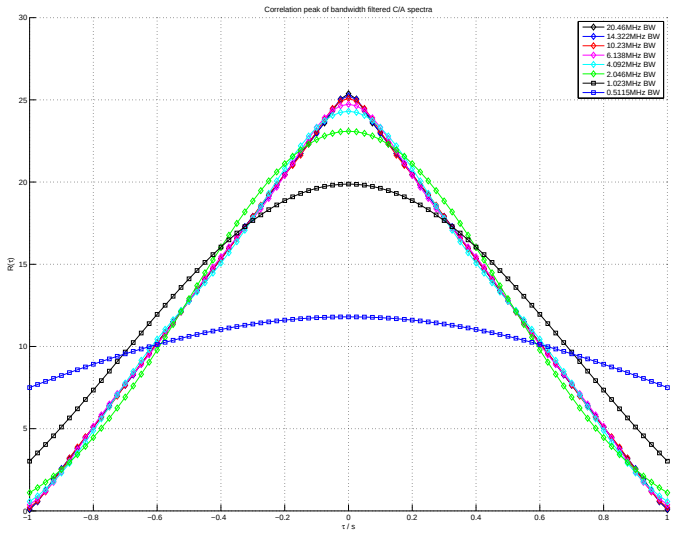


Figure 2.7: ACF of simulated C/A spectrum using different bandwidth filters

2.4.3 Code Tracking Performance

As described in [PS96], the tracking delay error of the quasi-optimal DLL can be approximated by

$$\sigma^2 \simeq \frac{1}{\mathrm{SNR}[-R_s''(0)]} \ , \tag{2.27}$$

where SNR describes the signal-to-noise ratio, and $R_s''(0)$ is the second derivative of the code's autocorrelation function with respect to time, at $\tau = 0$. It is therefore the curvature of the autocorrelation function at its peak. This becomes visually clear when the function of the quasi-optimal correlator is examined: It consists of a single correlation value directly at the value to be estimated, and therefore its performance depends on the available slope to the left and to the right of this peak. Because a bandwidth limited signal has a rounded peak (the pulse forms of the individual chips are rounded off, and therefore the ACF is rounded off as well), the correlator produces a smaller error signal for the same code offset. The noise that is applied to the signal is defined by the SNR, and can produce such an offset to the correlator output. The lesser the SNR or the curvature of the ACF at $\tau = 0$ is, the more tracking error is produced through the influence of noise.

The SNR used in (2.27) is the Signal-to-Noise Ratio after the correlation done by the DLL stage in the receiver. The following formulas describe the

transformation of the uncorrelated noise with a fixed power spectral density to the noise spectrum that is effective after modulation with the code.

The received signal consists of the delayed, attenuated signal from the satellite, plus noise. The noise is processed together with the wanted signal, and so it still exists in a modified form after demodulation and de-spreading.

Analyzing only the baseband process (the carrier is assumed to have been removed already), the modulation with the code sequence, denoted in the frequency domain, results for the received noise in

$$N_c(f) = N(f) * C(f) = \int_{-\infty}^{\infty} N(f')C(f - f')\mathrm{d}f' \ . \tag{2.28}$$

Because $N_c(f)$ is a band limited noise (the noise power above a certain frequency is cut off at the receiver input, as it is done with the signal power), the integral can be transformed to

$$N_c(f) = \frac{N_0}{2} \int_{-f_c}^{f_c} C(f - f')\mathrm{d}f' \ , \tag{2.29}$$

if the spectrum of the reference code is not band limited. N_0 is the power spectral density of the noise, which is a constant value for all frequencies within the filter range. The convolution of the rectangular noise spectrum and the squared sinc spectrum of the reference code can be approximated around $f = 0$ by $N_c(0) \approx N_0$, because the noise power density at low frequencies does not change much through the modulation of the code. This presumes, of course, that the code power is unity. The approximation is valid because most of the signal power lies in the main lobe of the signal spectrum, so a convolution at a low frequency includes almost all the signal power from the reference code.

After modulation and correlation, the receiver integrates the signal for at least one code periods. With knowledge about the synchronization between data bits and code chips, it is also possible to integrate for 20 code periods, which corresponds to an integrating time of $T_i = 20\text{ms}$. This integration, which can be described as a rectangular window function in the time domain, is a sinc-shaped low pass filter in the frequency domain, and it has its first null at $f_i = {}^1/_{T_i} = 50\text{Hz}$, for the denoted integration time. Approximated by an ideal low pass filter with the same cut-off frequency, the noise power that is mixed with the signal after integration can be obtained from

$$P_N = \int_{-f_i}^{f_i} N_c(f)\mathrm{d}f \approx \int_{-f_c}^{f_c} \frac{N_0}{2}\mathrm{d}f = 2f_i\frac{N_0}{2} \ . \tag{2.30}$$

For a noise power spectral density of $N_0 = -201$ dBW/Hz [ME01], and the mentioned filter bandwidth, the total noise power is

$$P_N \approx -184\text{dBW} \ . \tag{2.31}$$

The total C/A L1 signal power at the receiver input is specified as $P_S = -156$ dBW for a satellite in the zenith in [ME01]. Because the modulation with the reference code transports all signal power to the center of the spectrum, ideally all signal power is still available after the integration. The SNR can then be obtained from

$$\text{SNR} = \frac{P_S}{P_N} = 28 \text{ dB} \ . \tag{2.32}$$

The second derivative of the autocorrelation function can also be determined by making use of the relationships between the Fourier transform and the autocorrelation function. The autocorrelation function is the Fourier transform of the power density spectrum [PS96]:

$$\begin{aligned}
R(\tau) &= \int_{-\infty}^{\infty} G(f) e^{\mathrm{j}\omega\tau} \mathrm{d}f \quad , \qquad R(0) = \int_{-\infty}^{\infty} G(f) \mathrm{d}f = P_S = 1 \\
R'(\tau) &= \int_{-\infty}^{\infty} \mathrm{j}\omega G(f) e^{\mathrm{j}\omega\tau} \mathrm{d}f \quad , \qquad R'(0) = \int_{-\infty}^{\infty} \mathrm{j}\omega G(f) \mathrm{d}f \\
R''(\tau) &= \int_{-\infty}^{\infty} -\omega^2 G(f) e^{\mathrm{j}\omega\tau} \mathrm{d}f \quad , \qquad R''(0) = \int_{-\infty}^{\infty} -\omega^2 G(f) \mathrm{d}f \ .
\end{aligned} \tag{2.33}$$

The last term is also known as the Gabor bandwidth, and can be determined only numerically, because the function of the squared-sinc spectrum is not integrateable. If the spectrum is band limited, the integral limits are set accordingly, which lowers the resulting value. In Figure 2.8, the resulting delay error standard deviation in meters is given for a band limited receiver input.

As mentioned, the SNR after correlation with the reference code can be assumed to be constant, so this bandwidth dependency shown in the plot comes only from the different gabor bandwidths. The dashed line in the background signifies the one-sided PSD of the signal, and the shape of the spectrum can be guessed also in the devolution of the variance - in sections where the spectrum has low power, the variance slope is also low because an increase of the bandwidth does not result in much more information. As expected, the resulting delay error variance is highest if the processing

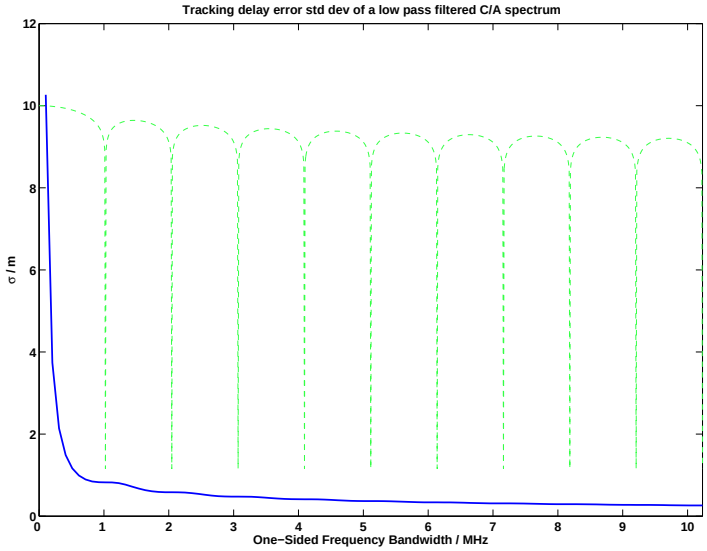


Figure 2.8: Code Tracking Delay Error Standard Deviation against Filtering Bandwidth

bandwidth is very low, and lowest for the maximum bandwidth analyzed, 20.46 MHz. For a full bandwidth of 24 MHz (not displayed), the standard deviation of the code tracking error is about 23 cm [HGI+].

The used spectra were generated using a Gold code and therefore consist of lines; their envelope follows a squared sinc-function. The sinc^2 function is therefore used frequently as an approximation of the spectrum. In fact, it corresponds to a random sequence of infinite length, where the individual chips have the same duration as those of the C/A PRN sequence.

An analysis comparing the delay error variance of the original spectrum with that of the approximated spectrum results in almost equal variances for both spectra. The ACF computed from the original spectrum, generated from C/A sequences, is compared to the ACF computed from the approximation envelope. Because the determined variances for a given bandwith limitation are equal for both spectra, it can be assumed that the approximation of the sinc^2 spectrum is valid in terms of computing the delay error variance.

2.4.4 Carrier Tracking Performance

For carrier tracking, the error variance can be computed in a similar way. Because the spectrum of the sinusoidal carrier is only a dirac, bandwidth

limitation does not affect the carrier phase error variance at all. The autocorrelation function of the cosine and its second derivative are

$$\begin{aligned} R_c(\tau) &= \int_0^{1/f_c} \cos(2\pi f_c t)\cos(2\pi f_c(\tau - t))\mathrm{d}t \\ R_c''(\tau) &= -4\pi f_c^2 \int_0^{1/f_c} \cos(2\pi f_c t)\cos(2\pi f_c(\tau - t))\mathrm{d}t. \end{aligned} \tag{2.34}$$

At $\tau = 0$, the second derivative yields

$$R_c''(\tau)|_{\tau=0} = -4\pi^2 f_c^2. \tag{2.35}$$

Converted in meters and for an SNR of 28 dB, corresponding to an integration time of 20 milliseconds as before, the standard deviation of the carrier phase tracking error is

$$\sigma_c = \sqrt{-\frac{1}{\mathrm{SNR}R_c''(0)}} \approx 1.2\mathrm{mm}\ . \tag{2.36}$$

Chapter 3

Multipath signal characteristics

In the previous chapter, the direct path signal has been analyzed. It was shown that the signal transmission from a moving satellite alters the observed signal frequency by applying a Doppler shift to the signal. The amount of frequency deviation depends on the relative motion velocity between satellite and user.

The following chapter explains how a scenario featuring multipath signal reception originates and which parameters define it; it then describes multipath effects on signal level.

First, a system model based on the signal description from the previous chapter will be constructed in Section 3.1. As it will be shown, multipath signal reception is basically the reception of multiple, time delayed and attenuated signal copies, and can be displayed mathematically as such.

In Section 3.2, the mentioned delay and attenuation parameters are examined with respect to a stationary receiver that is subject to reflections either on the ground, or through surrounding obstacles. Values for the path delay as well as its change rate due to the movement of the satellites are derived by analyzing orbits and the constellation geometry of the GPS satellites.

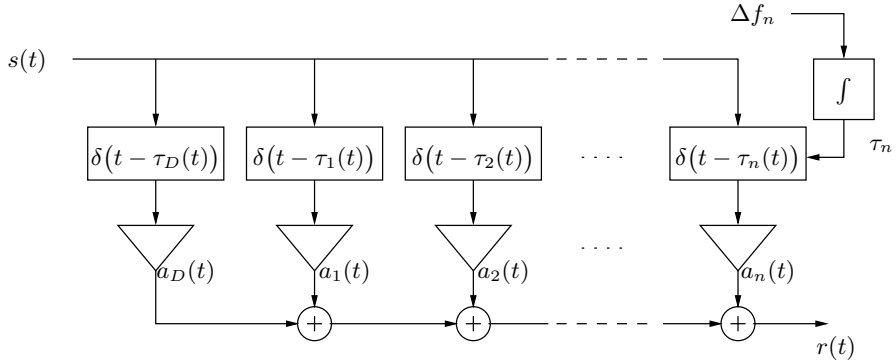
3.1 System Model of the Multipath Signal Reception Chain

Based on the signal description from Chapter 2, a system model is derived in this section. It covers the mathematical description of a single GPS satellite C/A signal on the L1 frequency. In contrast to the previous chapter, the

signal is now received concurrently on a direct path and multiple reflected paths.

Using this description, it is possible to find out how multipath delay afflicts the signal processing in the receiver, and what effects at the receiver output may be assumed in a scenario where multiple signal copies are received. Those effects will then be used as a basis for multipath estimation in Chapter 4.

The term *Multipath* refers to concurrent propagation of radio waves from the same transmitter along different paths, reaching the same receiver in the end. If the path lengths differ, the transmitted signal will arrive multiple times and add up at the receiver, resulting in a distorted waveform that can inhibit further signal processing. The individual path's components arrive at different times and usually with different amplitudes. Since radio propagation is comparable with transmission over a (very long) signal line, a frequently used model for the described multipath effect is the *Tapped Delay Line Model*, which is shown in Figure 3.1



The signal $s(t)$ is split up into several paths, displayed with delay elements and amplifiers. The delay parameters with parameters τ_i are time variant values, and they change according to the integral of the individual path's Doppler shift Δf_i. The parameters with the subscript D denote the direct path component of the signal, which is delayed only by the LOS path length.

Figure 3.1: Block diagram of the tapped delay line.

In a wireless data broadcast application, simultaneous signal reception of multiple, time delayed signal components is not necessarily a downside. If the broadcast channel is estimated correctly, the time-delayed copies can be coherently added together, making all the received components again usable for further processing. To make this possible, exact knowledge of the channel

in terms of the individual path delays is needed. The described principle is called a Rake receiver [Hag03] and is often used in CDMA systems.

However, in navigational applications, the broadcast data is of secondary importance. Here, the main focus is the estimation of the correct transmission delay, which is used to determine the pseudorange. If reflections occur, the estimated pseudorange may be distorted towards either a later or an earlier value, and this error propagates into the position solution. It is easy to see that the shortest path between satellite and user is always the line of sight, and every path that contains at least one reflection must therefore be longer. Still, as it will be shown later, it is possible that a signal copy that is received later than the original can influence the measurement result towards the early direction as well - then, the receiver measures a distance that is shorter than the LOS path length.

When speaking of multipath, two possible scenarios may be distinguished: If the direct path (also called Line-Of-Sight, or LOS path) is received along with multiple time delayed copies, we speak of a mixed LOS/NLOS scenario. Under certain circumstances, it may still be possible to isolate the direct component from the mixture and determine the LOS path length from the transmission delay of that component.

If the LOS is obstructed, only reflections reach the receiver antenna. In this case, there is no information about the path length of the LOS path in the signal, thus this condition may only be detected, but the signal is worthless. If the detection works reliably, this satellites' signal can be excluded from the position solution, inhibiting the propagation of this error into the derived position. Both mixed LOS/NLOS and NLOS scenarios are illustrated in Figure 3.2.

Depending on the attenuation caused by the reflection, the resulting tracking error may increase or decrease. In general, reflected components have smaller signal amplitudes, because any non-ideal reflector material absorbs some of the signal power. However, if multiple reflections are in-phase with each other and add up at the antenna, they may form a combined reflection that is received stronger than the LOS signal. This combination may be interpreted as one reflection that now appears as the strongest signal component at the receiver, and may result in a very strongly distorted measurement that basically tracks the reflected paths' common path length instead of the LOS length, even if the satellite may be clearly visible.

If the receiver faces such a mixed reception scenario, it is desirable to detect the existence of those delayed signal copies and estimate the impact on the measurement or its quality. Further counter-measures can be taken to ensure that the degraded measurement does not affect the quality of the position solution. Some of the effects introduced through multipath may also be mitigated, which will be shown later.

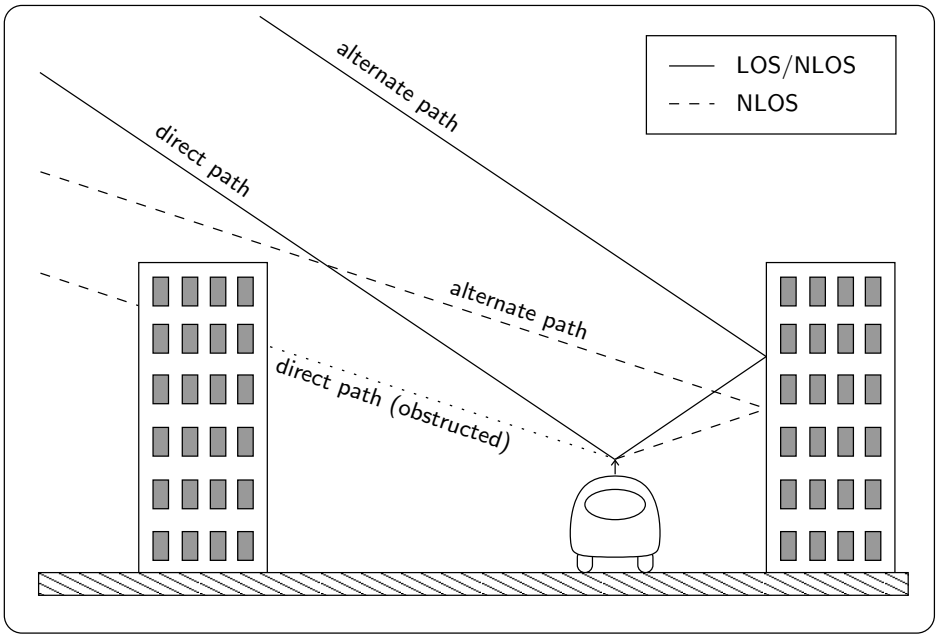


Figure 3.2: Mixed LOS/NLOS and pure NLOS receiver situation.

3.1.1 Direct Path Signal

As discussed earlier in Chapter 2, the modulated C/A signal gets delayed on its way to the receiver, depending on the propagation speed and the distance between satellite and user. For our analysis, it is sufficient to assume that the mean propagation speed $\bar{c}$ is a constant value, because we analyze only one satellite at a time, and the satellite position with respect to the atmosphere changes slowly. For a very long observation where the satellite's observed sky track spans a large distance, however, corrections to the propagation speed may be applicable and can be done with means of ionospheric and tropospheric corrections [Gü05].

The distance ρ, however, is time dependent for the non-geostationary orbits of GPS satellites. In (2.24), the direct signal was described with a consistent perturbation factor η. Looking at multipath, however, it is often beneficial to have a representation that directly displays the length of the individual paths with respect to each other. Given a constant propagation speed, this can be done choosing the absolute delay representation from (2.10).

The formula describing the received LOS signal therefore contains a time variant path delay and defines the received signal as

$$r_{LOS}(t) = A' \cdot b(t - \tau(t)) \cdot c(t - \tau(t)) \cos\Big(2\pi f_c\big(t - \tau(t)\big) + \varphi\Big) \ . \qquad (3.1)$$

The previously discussed Doppler shift Δf, perturbing the carrier frequency f_c, the chip sequence and the bit stream is now contained in the derivative of the time variant path delay $\tau(t)$. In other words: The time argument $t - \tau(t)$ runs at a certain different speed compared to that of t. This speed depends on the change rate of τ, which is its first derivative. The new subscript LOS points out the reference to the direct path signal.

3.1.2 Reflected Signal

Radio signal propagation is modeled as the propagation of electro-magnetic waves. Propagation always follows a line-of-sight path in free space, given that no refraction occurs due to the change of medium density along the path. If the propagation medium is not vacuum, but a gas with a density gradient along the path, the effect of refraction allows the signal path to be bent.

Radio waves are reflected from conductive surfaces like light in a mirror; this means that the Angle of Reflectance equals to the Angle of Incidence (see Figure 3.3). There are several factors that influence the way a wave is reflected in detail, but for the present work, the simplification of equal angles meets the required degree of modelling accuracy.

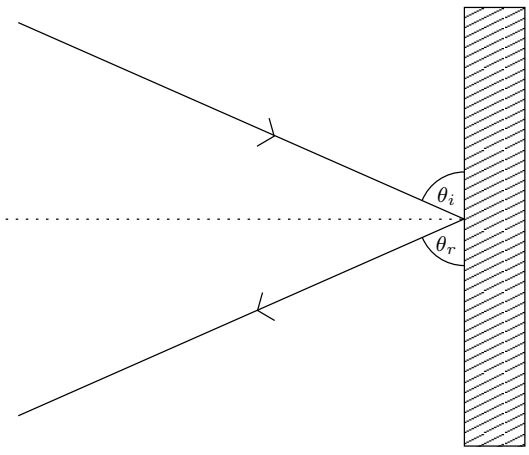


Figure 3.3: Optical reflection, $\theta_r = \theta_i$

With signal reflection defined so far, some simplifications can be made:

- Signal reflections occur only in close proximity to the user: It may be assumed that under normal circumstances, the free space between user and satellites is clear of significant obstacles that may produce a strong reflection. All important reflection are caused by obstacles within the range of several hundred meters from the user, which is a very small distance compared to the several thousand kilometers that separate the satellite from the user.

- Because the reflector distance is so small with respect to the path length, the LOS and NLOS paths may be assumed to arrive in parallel.

- The effect of atmospheric refraction may be assumed to be equal for all paths and is therefore not accounted for. The effect changes the angle of arrival at the point of reflection and the user, but because of the assumption of parallel signal paths from the satellite, the change is equal for all paths.

A more detailed analysis of the dependency between reflector distance, angle of arrival and the resulting path delay will be given in Section 3.2.

If all signals travel along straight paths, the shortest path must be the direct path, and all paths containing a reflection are longer than it. Thus, the direct path signal component always arrives first at the receiver, followed by one or more signal copies, each delayed individually according to the path length.

It is straightforward to see that the path delay, which is the difference between the direct path's and the reflected path's arriving time, is fully determined by the mutual position of receiver antenna, reflector and satellite. Since all paths are assumed to arrive in parallel from the satellite, the constellation can be described completely by two parameters: The reflector's distance from the user, and a difference in the LOS and NLOS paths' angle of arrival. In Figure 3.4, three points are used as helpers to define the scenario: (S, U, R), which are satellite, user, and reflector, define a plane in three dimensions. Signal paths are connections between those points and thus must lie exactly in that plane.

The drawing shows the direct signal arriving from S, and the signal to be reflected arriving from S', instead. The same satellite appears twice to depict the parallel signal propagation which was assumed. The path delay $\Delta\rho$ scales linearly with the reflector distance d. The angle γ, which is the angular difference between the signal paths' arriving directions, makes the path length difference equal to some value between zero (for $\gamma = 0$) and the double reflector distance $2d$ ($\gamma = \pi$, for a reflection opposite the satellite). Of course, this assumes that the reflector's orientation is always such as to direct the signal towards the receiver antenna at U.

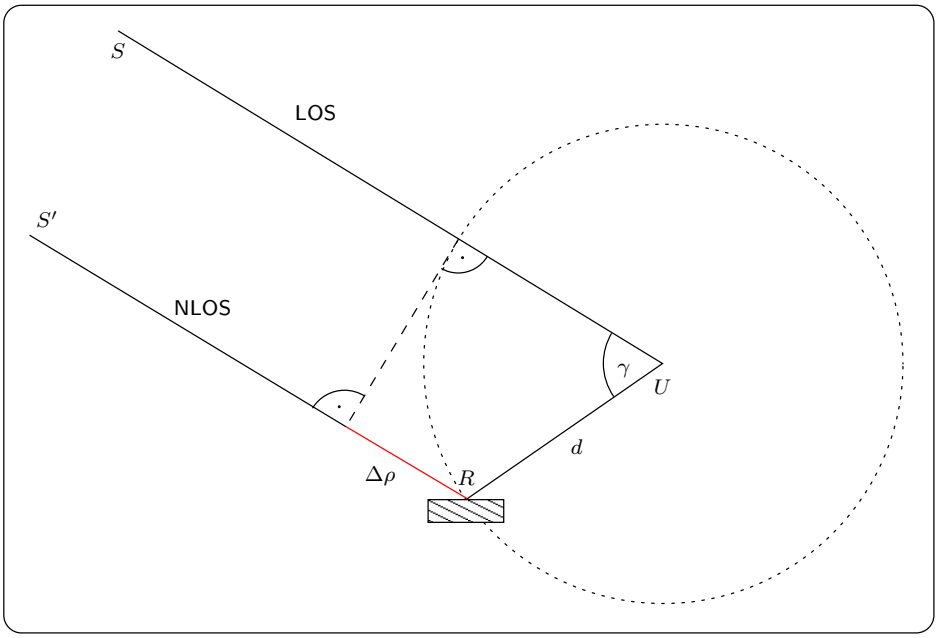


Figure 3.4: Reflector constellation defined by d and γ

So the additional path length introduced through a reflection depends only on the position of the reflector, and on the angle between satellite and reflector. If one or both of those parameters change, the resulting path length for the NLOS path will change according to it.

This change happens in addition to the path length change resulting from satellite movement, which was discussed in Section 2.1. However, the effect that is observable in the receiver is the same: The carrier frequency is Doppler shifted according to the changing rate of the path delay. For the mathematical representation, we therefore chose to include the additional, time variant and path specific path elongation $\Delta\rho_k$ corresponding to path k. It represents the extra path length introduced through the reflection, measured with respect to the length of the direct path. The path elongation $\Delta\rho_k$ is applied to the single, direct path signal from (3.1) as its temporal representation $\Delta\tau_k$, defined by

$$\Delta\tau_k = \frac{\Delta\rho_k}{c_1} \text{ , } c_1 \text{ is local propagation speed,} \tag{3.2}$$

resulting in the single reflected satellite signal:

$$\begin{aligned} r_k(t) &= A'_k \cdot b\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) \cdot c\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) \cdot \\ &\quad \cos\Big(2\pi f_c\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) + \varphi\Big) \end{aligned} \tag{3.3}$$

The overall path length of each path, and its corresponding temporal path delay are defined as

$$\begin{aligned} \rho_k &= \rho_{LOS} + \Delta\rho_k \text{ , and} \\ \tau_k &= \tau_{LOS} + \Delta\tau_k \text{ ,} \end{aligned} \tag{3.4}$$

where the LOS parameters are common to all of one satellite's paths.

The amplitude A'_k takes into account that a reflection usually attenuates the signal. Thus, all reflected copies of the signal will normally arrive at a significantly lower signal power than the direct signal. However, a NLOS scenario or interference of multiple reflection components with the same phase components may result in a higher combined path amplitude than the direct path.

3.1.3 Multiple Signals

The reflected signal was described as a time delayed, attenuated copy of the direct signal in (3.3). The received signal mixture at the user antenna is a combination of the direct signal with one or more reflections. Depending on the scenario, the number of distinguishable reflections may reach a large quantity - 50 and more concurrent echoes at a time in urban surroundings are possible according to [LS05]. Since radio waves are basically oscillating electromagnetic fields, the reflected components sum up as electrical voltages at the user antenna. Therefore, the received signal of a single satellite containing the direct signal and multiple reflections can be written as a sum of the individual signal components:

$$\begin{aligned} r_{MP}(t) &= \sum_{k=1}^{K} A'_k \cdot b\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) \cdot c\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) \cdot \\ &\quad \cos\Big(2\pi f_c\big(t - \tau_{LOS}(t) - \Delta\tau_k(t)\big) + \varphi\Big) \end{aligned} \tag{3.5}$$

where the index k denotes the number of the path, $k = 1$ referring to the direct path with

$$\begin{aligned} A'_1 &= A'; \\ \Delta\tau_1 &= 0; \end{aligned} \tag{3.6}$$

The formula becomes more concise when the LOS component and the path component of the path delay are combined according to (3.4):

$$r_{MP}(t) = \sum_{k=1}^{K} A'_k \cdot b\big(t - \tau_k(t)\big) \cdot c\big(t - \tau_k(t)\big) \cdot \cos\Big(2\pi f_c\big(t - \tau_k(t)\big) + \varphi\Big) \quad (3.7)$$

3.2 Typical Properties of the Reflected Path Component

In the previous section, a mathematical model describing the combination of direct and reflected signals at the receiver input has been derived. The reflections apply attenuation and temporal delay onto the signal, where both of these parameters may vary. It was demonstrated that the path delay between the LOS path and the reflected NLOS path are completely defined by the constellation of satellite, user and reflector. This constellation can be described by two parameters: Distance between user and reflector, and the angle between satellite and reflector, as seen from the user.

In the present section, these parameters will be analyzed with respect to the GPS satellite constellation, and possible reflector locations. More importantly, the changes of those parameters will be illustrated. A change in reflector distance or the angle of arrival means that the reflection geometry changes, which in turn alters the path delay between the two paths. The changing rate of the path delay is again the key value to make the proposed estimation algorithm usable, because it changes the relative phase between them.

3.2.1 Reflection Geometry

It was mentioned before that the appearance of a reflection at the receiver is basically a delayed and attenuated copy of the direct signal. In geodesy and for reference stations, the receiver is preferrably operated at a position where reflections from surrounding obstacles are unlikely. Still, ground reflection is an important factor, which can only partly be mitigated using counter-measures, for example, a multipath attenuating antenna (choke ring antenna).

In the following, the ground reflection will be analyzed with respect to the corresponding path delay, and its change rate. The main interest is to give a minimum path delay change rate, which defines how long it takes until the change in path delay exceeds the equivalent to one carrier phase cycle, which is $\approx$ 20cm for L1. At this point, a spectral analysis of the combined signal, as it will be proposed in the next chapter, is possible. Apart from

that, averaging over a time interval spanning at least the path delay change equivalent to one carrier cycle can be used to overcome the distortional effect of multipath.

Ground Reflection

Ground reflection means that the reflector is oriented horizontally below the receiver, and, within the scope of this inspection, extends indefinitely into every direction. The user antenna is usually mounted on a tripod or a mast at height h with respect to the reflecting ground. The influence of the antenna support considering signal reflection is neglected here. Satellites may be positioned anywhere on the top semi-sphere, so they are observed at an arbitrary azimuth angle and an elevation between zero and ninety degrees. Reflections on the ground occur provided that the angle of reflection equals the angle of incidence. Figure 3.5 shows the setup from aside.

The user receives both signals at an antenna (U) that is elevated from the ground by h. According to the simplifications defined in Section 3.1.2, both signal paths are assumed to arrive in parallel due to the great distance between user and satellite, as compared to the distance between the paths. The path length difference between $\overline{SU}$ and $\overline{S'RU}$ can be read off directly using p and q,

$$\Delta\rho = (\rho + p) - (\rho + q) = p - q \ , \tag{3.8}$$

which is made visible directly below the ground, in the reflection of the user antenna. The path elongation can be determined from the sine of the elevation angle and the antenna height by

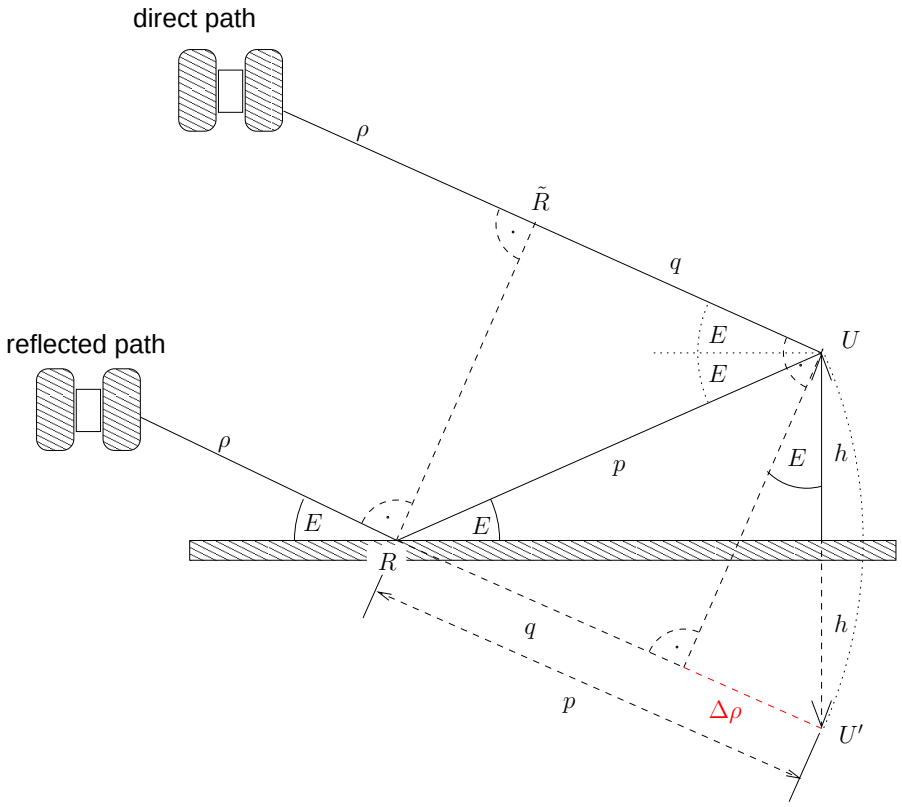
$$\Delta\rho = p - q = 2h \cdot \sin E \ , \tag{3.9}$$

and the corresponding path delay time is therefore

$$\Delta\tau(E, h) = \frac{2h}{c_1} \sin E \ , \tag{3.10}$$

where $c_1 < 299,792,458 \frac{m}{s}$ denotes the local speed of light near the receiver.

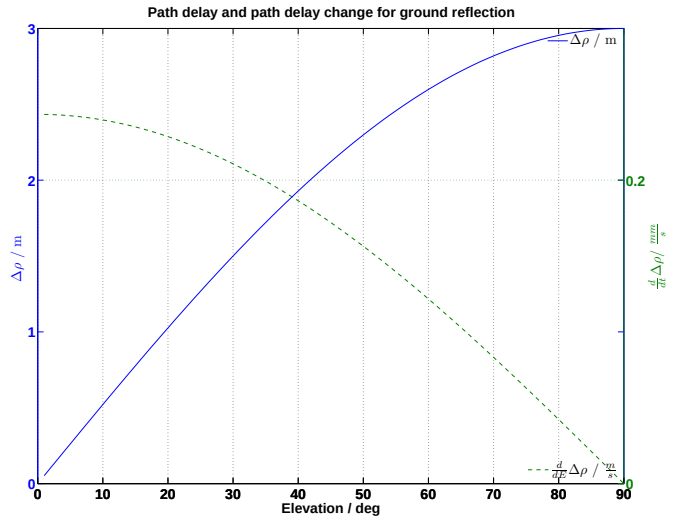
Some important results can be derived instantly from (3.9): First of all, the path delay is proportional to the antenna height. Secondly, the minimum path delay with respect to elevation is obtained for zero elevation. This is well understandable, because the paths do not differ much if the elevation is very low. In contrast, the highest path delay, corresponding to a NLOS path that is $2h$ longer than LOS, is obtained for a satellite right in the zenith, when the signal path actually passes through the antenna, gets reflected on the ground and travels back to the antenna.



Direct and reflected signal received at a stationary ground station. Note that both satellite signals are assumed to arrive in parallel prior to reflection. The signal paths have the same length at R and $\tilde{R}$; the total path length difference can be read off $p - q$.

Figure 3.5: Single ground reflection

Path Delay and Delay Rate. Figure 3.6 shows the path length difference between the direct and the reflected path with respect to the elevation angle. The antenna height is $h = 1.5$ meters above ground. The sine shape of the path delay slope is clearly visible, and it proceeds from zero to $2h$ (left axis, blue plot).



The left axis (solid blue) shows path elongation in meters. The right axis (green, dashed) shows the corresponding change rates for a fixed elevation rate of 0.081 millirad per second. For zero elevation, the elongation is zero, but changes fastest. A satellite at the zenith produces the most delayed reflection; but the path delay changes slowest.

Figure 3.6: Path Delay and Delay Rate for Ground Reflection

More interesting than the path delay itself is the change rate of the path elongation, however. As it will be shown more in detail in the next chapter, the change of a reflection's path length, with respect to the direct path, causes a change in the phase relation between the two path carriers as well. This phase change, which occurs as a sort of a phase modulation in the receiver, can be detected if the path phase relation between two paths has change for at least one carrier cycle while tracking the signal. This distance equals to the wavelength of the carrier, and is around 20 cm for the GPS L1 signal. So with the path elongation defined for every elevation angle in (3.9), the next step is to find out how much change in elevation is necessary to cause the phase relation to cycle around once.

The change rate of the path delay for a single ground reflection can be obtained by differentiating (3.9) with respect to elevation:

$$\frac{\mathrm{d}}{\mathrm{d}E}\Delta\rho(E,h) = 2h\cos E \tag{3.11}$$

If the elevation angle is given as a function of time $E(t)$, (3.9) can be written as a function of the time as well:

$$\Delta\rho(t,h) = 2h\cdot\sin E(t)\ , \tag{3.12}$$

and its derivation with respect to time is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Delta\rho\big(t,h\big) &= \frac{\mathrm{d}}{\mathrm{d}t}2h\sin\big(E(t)\big) = \\ &= 2h\cos\big(E(t)\big)\cdot\dot{E}(t)\ .\end{aligned} \tag{3.13}$$

The angular rate, or angular speed of the elevation has been defined here as $\dot{E}(t)$, and must be used in the unit Radian in this context. If degrees are used, the scaling factor has to be considered according to the chain rule of differential calulus. As an example, we define the antenna height as $h = 1.5\mathrm{m}$, the current elevation at $t = t_0$ as $E(t_0) = 30^\circ$ and its rate at $\dot{E} = \frac{\mathrm{d}}{\mathrm{d}t}E(t) = 0.081^{\mathrm{mrad}}/_{\mathrm{s}} \simeq 0.0046^\circ/_{\mathrm{s}}$. The path elongation change rate $\frac{\mathrm{d}}{\mathrm{d}t}\Delta\rho[^{\mathrm{m}}/_{\mathrm{s}}]$ is then

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\rho|_{t=t_0} = 2\cdot 2\mathrm{m}\cdot\cos(30^\circ)\cdot 0.081\frac{1}{s} \approx 0.281\frac{mm}{s}\ . \tag{3.14}$$

This value denotes the amount of path length the reflected path is changing with respect to the direct path length, in meters per second. It means that the delayed path gets longer by 0.281 millimeters per second, if the elevation of the satellite is 30 degrees above the horizon, and the satellite elevation changes by 0.081 milliradians per second. A short note on the elevation rate, which has not been commented yet: In the consecutive section, GPS orbits are analyzed with respect of the visible satellite movement from a randomly chosen location on the earth surface. The mean change of a satellite's elevation is exactly the used value.

This path delay rate is also plotted in Figure 3.6, where it is referred to the very same fixed rate of the elevation angle, but plotted for all valid elevation angles. The path delay resulting from a ground reflection changes fastest when the elevation angle is very low. In this case, the corresponding path delay is very low. The mixed LOS/NLOS signal is therefore subject to a *fast changing, low delay* multipath distortion.

When the elevation is high, the path elongation increases up to $2h$. In this situation, the change in path delay is only very slow. *For high elevation, the path delay is large, and changes slowly.*

The time required for the path delay to change one carrier wavelength's worth of distance can be determined from (3.13) by

$$T_{\mathrm{Cyc}} = \frac{\lambda}{2h\cos\big(E(t)\big)|\dot{E}(t)|} \ . \tag{3.15}$$

Within this time, the carrier phase relation between the two paths performs a full cycle, which means that the alternate signal processes through both constructive and destructive interference, and this change can be detected when the signal has been tracked along that interval. For the previously used example values at 30 degrees elevation, the duration $T_{Cyc}|_{E=30°,\dot{E}=0.081,h=2\mathrm{m}} \simeq$ 712s. The shortest time interval of 617s is achieved for zero elevation, where the path delay changes fastest. For very high elevations, the time can become indefintely long, because the cosine turns zero for 90°. This is an important key value to both estimation and mitigation of multipath, as it will be shown in the next chapter. As it turns out, the required measuring times are very long, and they only refer to the mean elevation movement. For satellites with much less change in their elevations, e.g. while at the elevation maximum, the path delay changes so slow that any attempt to detect them is hopeless. Only antennas much higher above the ground could deliver path delays changing fast enough for this technique. For mitigation of ground reflections, in any case, a special antenna form that attenuates signals from below the antenna provides a more efficient way to mitigate those reflections. The next analysis will cover reflection from obstacles around the scenery, and those path delays change much faster because the reflectors can usually occur at a much larger distance to the receiver than the ground does for stationary receivers.

Rotation symmetry of the ground reflection. Up to now, only a two-dimensional scenario has been analyzed. GPS satellites, however, are known not only to change their elevation, but also the azimuth under which they can be observed while passing the user's scope of view. But having a closer look, this azimuthal movement may be neglected without loss of generality:

In our previous model, the ground was defined to be a flat plane below the user antenna with sufficient extent in all directions. A coordinate system is now applied to the drawing from Figure 3.5, where its origin is defined to be lying in the reflecting plane, right beneath the antenna. It is therefore the projection of U onto the reflector plane. The z axis is defined to point upwards through the user antenna, and is therefore the zenith direction. The x axis is then chosen to lead through the point of reflection R, and the

y axis points orthogonally to x and z. The definition of azimuth is then defined as $A = 0$ for a point with positive x and zero y values.

Because the z axis is now normal to the reflecting plane, we can see that the whole scenario is rotation-symmetric around the z axis. Any three-dimensional scenario may be transformed into the case where all defined points lie in the (x, z) plane, like the basic scenario. This is done by performing a rotation about the azimuth angle, around the z axis.

This fact is useful, since it means that for a ground reflection, only the elevation of the satellite determines the resulting path delay. This simplifies upcoming analyses a lot; however, reflections at obstacles are much more complex and cannot be simplified that easy, as it will become clearer in the following section.

Other Reflections

Apart from ground reflection, reflections caused by house fronts or similar structures occur frequently, especially in urban locations. Concerning parameterization of those reflections, house front reflections are pretty similar to the previously discussed ground reflection, except that everything is rotated by 90 degrees. However, some important differences must be considered: The vertically positioned reflecting area can usually not be assumed to extend indefinitely, but only up to the height of the building. Apart from that, the symmetrical properties are more sophisticated than in the ground reflection model - there is no general independence of the satellite's azimuth angle. Rather, the reflector's orientation with respect to the user becomes important.

Of course, this class of reflectors is not limited to house fronts alone, but may also contain any other obstacles in the vicinity of the receiver which are large enough to be modeled using the principle of optical reflections. Small structures like lampposts, however, also introduce effects like refraction and are therefore not covered in this work. [LS05] describes a method to model reflections caused by lampposts and trees.

Geometrical Set-Up. As a start, lets look at Figure 3.7, where the reflection at a house front has been illustrated. In contrast to Figure 3.5, the reflector is now positioned vertically, and the horizontal distance of reflector and user antenna is fixed (g). The height h of the point of reflection varies with the elevation angle, where we assume for simplicity that the building is high enough to reflect satellites from all elevation angles under discussion.

One important assertion concerning the arrangement of the satellite, the reflector and the user must be explained here: In order to provide an easy-to-analyze reflection scenario, it is assumed that the plane defined by those

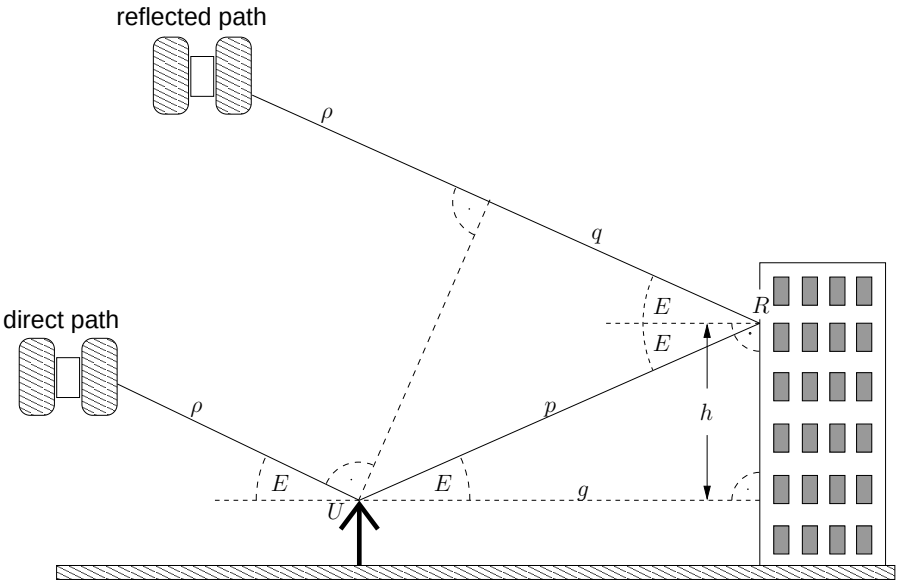


Figure 3.7: Single reflection at house front

three points is perpendicular to the ground; i.e. if the satellite's position is exactly in the north of the user, the reflector is defined to be exactly south. The orientation of the reflecting plane is then in east-west direction, so that the satellite signal is reflected exactly backwards. Furthermore, we assume that the satellite moves only within this plane, meaning that the azimuth does not change (A generalization of this scenario is provided later). The set-up is illustrated in a top view in Figure 3.8.

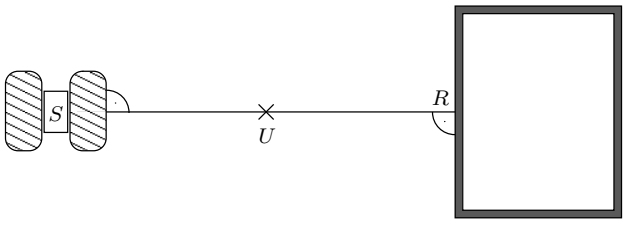


Figure 3.8: Top view of the house front reflection

As we can see from Figure 3.7, the path length difference results from p and q for the house front reflection:

$$\Delta\rho = p + q \tag{3.16}$$

With the predefined parameters g and E,

$$p = \frac{g}{\cos E} \tag{3.17}$$

$$q = p\cos 2E = \frac{g}{\cos E}\cos 2E \ ; \tag{3.18}$$

and therefore, the path length difference is

$$\Delta\rho = g\frac{1+\cos 2E}{\cos E} \ . \tag{3.19}$$

With the trigonometric identity [RW00]

$$\cos^2\frac{\alpha}{2} = \frac{1+\cos\alpha}{2} \tag{3.20}$$

(3.19) simplifies to

$$\Delta\rho = g\frac{2\cos^2 E}{\cos E} = 2g\cos E \ . \tag{3.21}$$

The corresponding path delay is then

$$\Delta\tau_k(E,g) = 2\frac{g}{c_1}\cos E \ , \tag{3.22}$$

where the local propagation speed c_1 is already known from (3.10).

Using the generic time-dependent elevation function $E = E(t)$, the derivative of (3.23) with respect to time leads to the path delay rate

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\tau_k(E,h) = -\frac{2g}{c_1}\sin\big(E(t)\big)\cdot\dot{E}(t) \ , \tag{3.23}$$

and we can see that both the path delay and the delay rate behave exactly opposite with respect to the ground reflection: For high elevation angles, the path delay is small, but fast changing. For low elevations, the path delay is high, but changes very slowly. Of course, the path delay becomes larger now if the elevation is decreasing, and the highest path delay is achieved for a signal that arrives in parallel to the horizon, at zero elevation.

Generalized Reflection An important assumption to the setup has been made for the preceding analysis: Since the scenario has basically been analyzed in the two-dimensional paper plane, it was presumed that the reflector plane is oriented perpendicularly to the user-satellite plane. It was also assumed that the satellite moves only in the paper plane, meaning that the angular velocity of its azimuth is always zero. Of course, this scenario occurs

rarely, and does not represent the real situation very exactly. In the following, the previously described setup is extended to fit the generic reflection, where location and orientation of the reflector are arbitrary.

Basically, the satellite movement has two degrees of freedom in the view of the user. We can think of the satellite being projected on a hemisphere atop of the user, where the satellite position is defined by elevation and azimuth - the distance is neglected in this case. Using this model, both azimuth and elevation may change, which is the normal case for the GPS constellation. Thus, every satellite movement can be split up into two orthogonal components. Elevation and azimuth are one possibility to split this up, and this separation has been used in the previous scenarios. In the following, the angular change of the observed satellite position will be split up into one part that changes the path delay, and an orthogonal part that doesn't change it.

Figure 3.9 shows again satellite S/S', receiver antenna U and a reflector plane. The current point of reflection, R_1, is on that plane and has distance d to the user. The circle C_R on that plane shows all other points R_i with the same distance to the user.

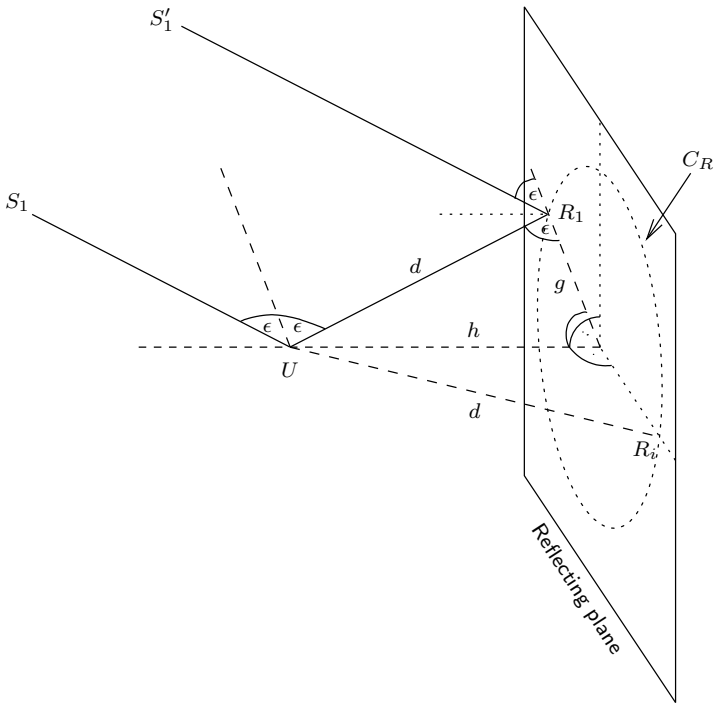


Figure 3.9: Generic reflection: Equal reflector distances

The three points S, U and R define a plane $P(U; R; S)$ for the reflection. The angle of incidence equals the angle of reflectance and is called ϵ. Defining h as the normal vector of the reflector plane which pierces U, the point C is the normal projection of U on the reflector plane, and simultaneously the center of C_R. With g linking C and R_1, the scenario resembles closely the ground reflection, where the equivalent to the elevation angle in the ground reflection is here ϵ; the length of h would then denote the antenna height. The rotational symmetry around h is now obvious - as long as the point of reflection lies on C_R, yielding the same reflector distance, the path delay stays the same.

This fact defines the direction of satellite movement that doesn't change the path delay already. Looking at the same scenario from a different viewing angle, it is easy to see that the angular satellite movement can be easily derived from the reflector movement, if we choose the right reference frame. Figure 3.10 shows the reflecting plane in the top view, meaning that the observe is situated somewhere on h.

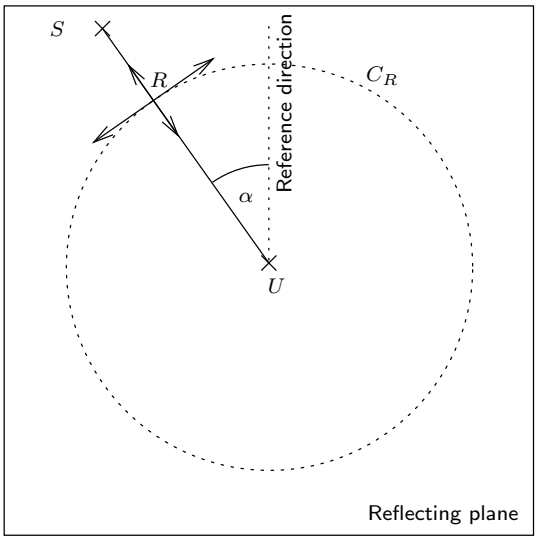


Figure 3.10: Generic reflection: Top View

The plane $P(U; R; S)$ is plotted as a dotted line here, and the signal comes from S directed into R, being reflected back to U. The angle of the satellite direction in the reflector plane, equivalent to the azimuth in the ground reflection case, is denoted as α here. If it changes, both S and R change the direction from U in the same way, and the path delay remains unchanged. From Figure 3.9, it can be seen that α is both the direction to R and to S in the reflector plane.

The orthogonal angular movement is equivalent to the elevation angle in the ground reflection. In the present model, the corresponding angle was called ϵ. If this angle changes, the radius of the new circle C_R will increase or decrease, and the path delay changes.

The described similarities with elevation and azimuth of the ground reflection can be used to construct a reference frame for the satellite position, where both α and ϵ appear. For the ground reflection, the satellite position is described using azimuth and elevation, and the path delay is defined by the antenna height h and the elevation angle E. For the generic reflection, the satellite position is described using α and ϵ, where ϵ is defined as $\angle(C;R;U)$ and defines the path delay, together with the distance h between the reflector plane, and U.

If the impact of a satellite position change on the path delay is to be determined, the position change has to be expressed in the polar $(\alpha;\epsilon)$ reference frame. This can be done by performing a coordinate transform on the angular (E, A) coordinates of the satellite.

With the satellite position expressed like this, the path delay can be determined by

$$\Delta\tau(\epsilon) = \frac{2h}{c_1}\sin\epsilon \ , \tag{3.24}$$

and the path delay change with respect to the change in ϵ is

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Delta\tau(\epsilon) = \frac{2h}{c_1}\cos\epsilon \ . \tag{3.25}$$

The minimum observation time that lets the phase between the paths rotate for exactly one cycle is determined similarly to (3.15):

$$T_{\mathrm{Cyc}} = \frac{\lambda}{2h\cos\big(\epsilon(t)\big)|\dot{\epsilon}(t)|} \ . \tag{3.26}$$

3.2.2 Satellite Orbits and Angle of Arrival

In order to have reliable values for the angular movement of ground observed GPS satellites, an analysis of the satellite orbits has been done. Results cover probability distributions of satellite positions defined by elevation and azimuth, and distributions of the angular speed of those satellite positions. The analysis was done as a simulation in MATLAB, which is implemented in detail as follows:

First, an exemplary GPS satellite orbit is determined in the ECEF coordinate frame. Therefore, a circular orbit in the ECI frame is assumed at

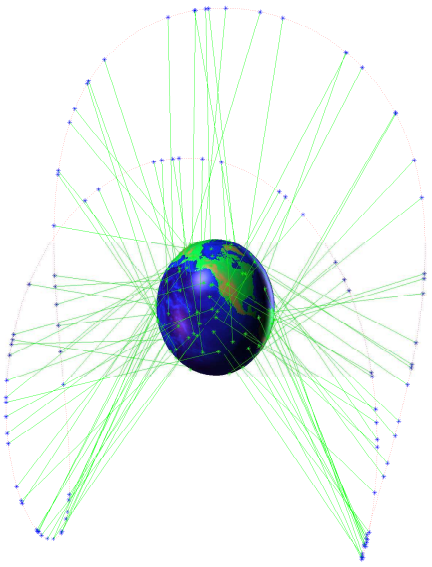


Figure 3.11: Random user locations and satellite orbit (100 simulated samples)

55° inclination, which is then transformed to ECEF. A satellite is placed randomly upon this orbit, where the equally distributed random parameter is the position on the orbit, called the mean anomaly. Having fixed the satellite position, its velocity is determined by transforming the velocity vector of the circular satellite orbit in ECI coordinate frame to ECEF. An exemplary satellite orbit with a number of random satellite locations is shown in Figure 3.11. Note that the orbit only represents a single orbital plane at an arbitrary longitude (LAAN, longitude of the ascending node, is used to parameterize the rotation of the orbit about the polar axis with respect to the Greenwich meridian [Gü05]). Handling the actual orbital planes and distributing satellites between them is unnecessary, because the user position is randomly chosen. Therefore, the satellite position in ECEF is defined to be located on one orbit, and the diversity of the user observed satellite position is reached by choosing random user positions. It is assured that the user can see the satellite; if not, the previous steps are repeated with a new random parameter, until this condition is met.

From there, the LOS vector can be determined. Given the velocity of the satellite in ECEF, the observed velocity can be determined by transforming the satellite position into the user local ENU coordinate frame. This returns the observed satellite elevation and azimuth angle, as well as its current angular speed, both in elevation and azimuth, and also as an absolute value.

Elevation Profile. The simulation yielded some interesting results. Firstly, the elevation probability shows a linearly descending slope for increasing elevation angles; this means that low elevation angles are more probable than high ones. Figure 3.12 shows the absolute probability for an elevation angle to occur, determined from 10^6 random samples of user and satellite positions.

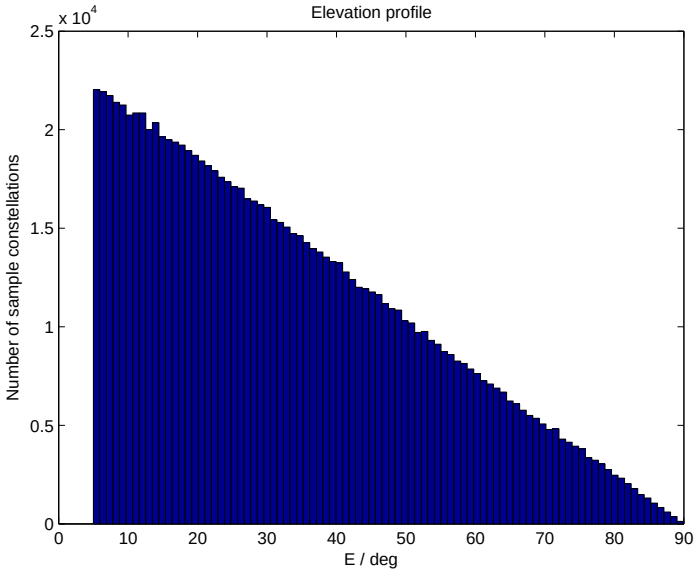


Figure 3.12: Elevation distribution profile (10^6 samples)

That is no surprise and can be explained by having a closer look at the user's point of view on the satellite orbit. In Figure 3.13, a simplified, two dimensional example of the simulated scenery is illustrated. The area of possible locations, which is a sphere around the center of earth in reality, is depicted as a dotted line. Satellites are defined to be equally spaced on the sphere, or that line respectively.

The two solid sections on the dotted line have the same size, so if a satellite position is chosen to lie randomly on the line, the two line sections have the same probability to contain a satellite. In three dimensions, the solid sections represent area elements of the same size. However, as it is easy to see from the picture, the corresponding angle α_1 is greater than α_2 - the section in the zenith appears much larger. This results from the shorter path distance ρ_1 (The satellite distance for a position at the zenith is around 22.000km) on the one hand, and on the other hand from the user seeing the portion of the line directly from aside. The line section is normal to the LOS vector between the user and this section. In contrast, the right line

section is seen under a slant angle, and the path delay ρ_2 is much higher - it may reach up to 26.000km if the satellite is close to the horizon. Thus, the resulting angle α_2 is smaller, and the area element, or line segment, looks smaller from the user's point of view.

With an equally distributed probability to chose an arbitrary location on the sphere for the satellite location, the probability of a satellite observation at a given elevation results therefore in the drawn distribution. This is of course transferable to satellite positions limited to an ECEF orbit, as it is the case in the simulation.

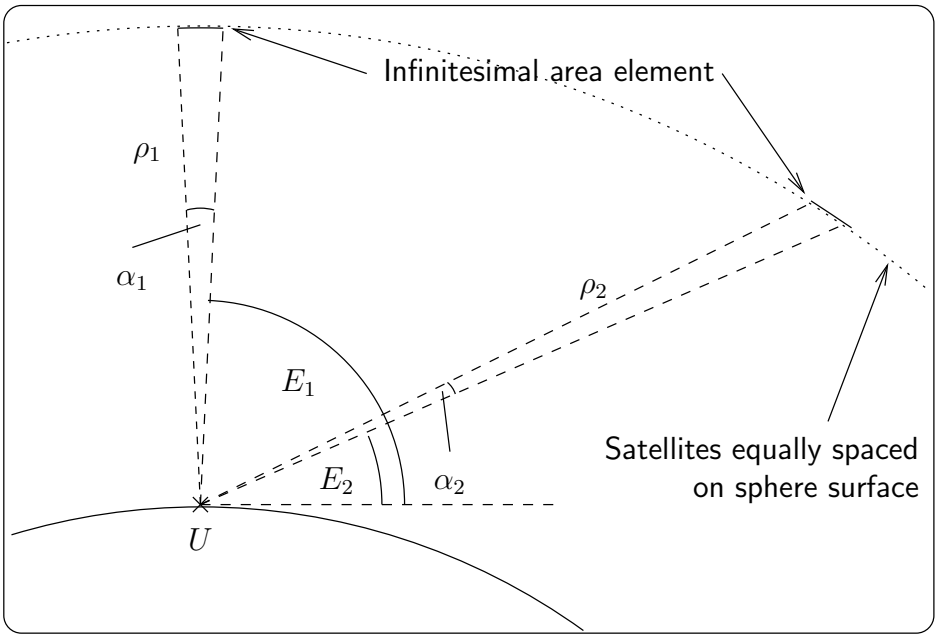


Figure 3.13: Explanation to the elevation probability density

Azimuth Profile. For the azimuth angles of the randomly sampled pairs of user and satellite position, the resulting distribution is shown in Figure 3.14.

Surprisingly, the probability is not equally distributed, but rather shows strong peaks at the directions north-east, north-west, south-east and south-west. For the analysis of reflections, however, this distribution plays only a minor role; we assume reflectors to be distributed randomly around the user, and thus the azimuth angle has no importance at all.

Elevation Rate. Figure 3.15 shows the obtained elevation rate for the total of all 10^6 simulated positions. A strong peak may be recognized at

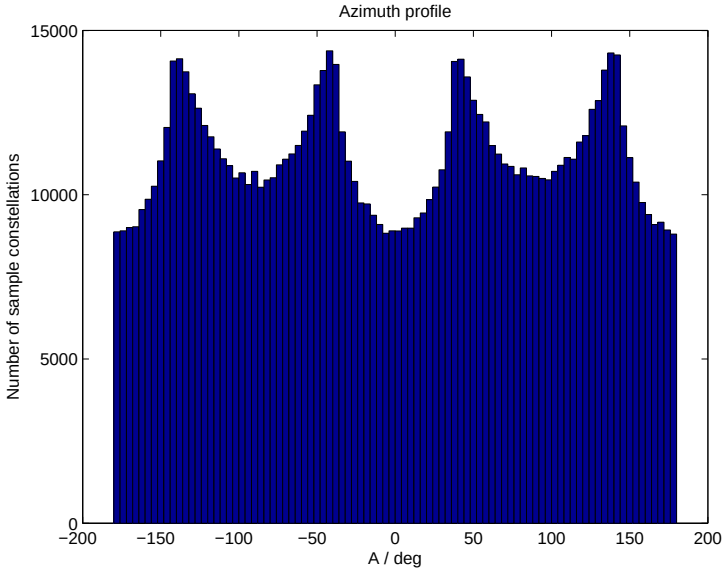


Figure 3.14: Azimuth distribution profile (10^6 samples)

$0.12\frac{mrad}{s}$, and the upper tail of the distribution is relatively steep. This shows that there is obviously an upper limit to the observable elevation rate, which is plausible because the satellite orbits at a given speed, and also the earth rotates at a constant rate. On the other hand, no minimum elevation rate can be seen from the profile. This result obviously issues from satellite observations which change mostly their azimuth while they reside at a specific elevation for a short time. The best imaginable example for this is a satellite which is about to reach its highest point of elevation, while the azimuth usually changes rapidly at that point.

The present distribution has its mean value at $0.081^{mrad}/_{s}$, and this value has already been used in the plot shown in Figure 3.6. Although the distribution of elevation rates makes the worst case for the changing rate of the path delay indefinitely slow, this value gives at least an estimate for average reflections.

Azimuth Rate. The shape of the azimuth rate plot shown in Figure 3.16 is quite similar to that of the elevation rate. The main difference is the flatter right tail, which means in rare cases, much higher azimuth angle rates may be observed. These high values result from satellites at high elevations, passing very closely to the zenith. In this situation, the regular velocity of the satellite in orbit causes its azimuth to change very fast towards the opposite direction.

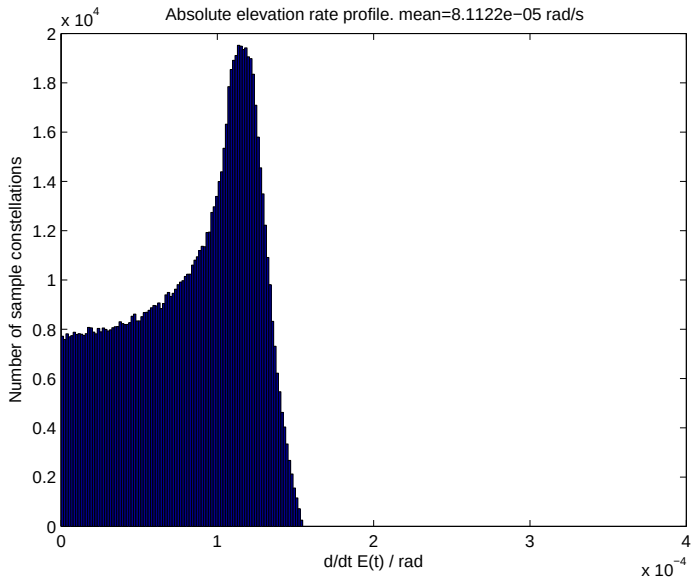


Figure 3.15: Absolute elevation rate profile (10^6 samples)

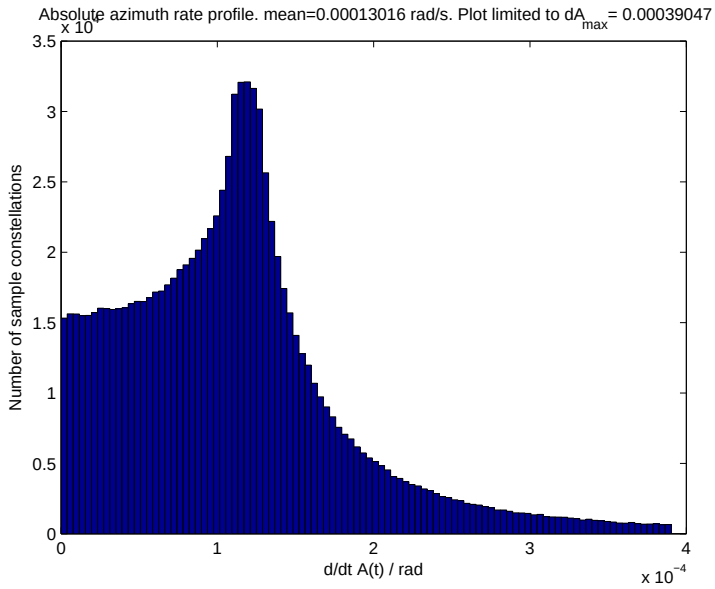


Figure 3.16: Absolute azimuth rate profile (10^6 samples)

Absolute Angular Rate In case of the ground reflection, it was shown that only the elevation rate is responsible for the path delay change between a direct and an indirect path. Having a reflector arbitrarily positioned, however, the fastest change is caused by an angular movement in one specific direction, which lies within the plane defined by the satellite, the user and the reflector. On the other hand, angular movement in the direction normal to that plane causes no change in the path delay.

To gain knowledge about angular changes in any direction, the simulated samples of elevation rate and azimuth range were recombined, resulting in an absolute value for the angular change. The distribution of these values is given in Figure 3.17, where one striking difference can be seen: Obviously, there is a lower bound on the absolute angular velocity. Since the examined GNSS satellite flies on a non-geostationary orbit, any user on the ground inevitably observes a movement of the satellite in the sky in one direction. The minimum angular speed that was observed in the simulation is $1.06 \cdot 10^{-4}$ rad/s. For an arbitrary reflector, however, there is no guarantee that this minimum satellite movement has any effect on the path delay, because the direction of the satellite movement is not necessarily in the plane defined by the satellite, the user and the reflector.

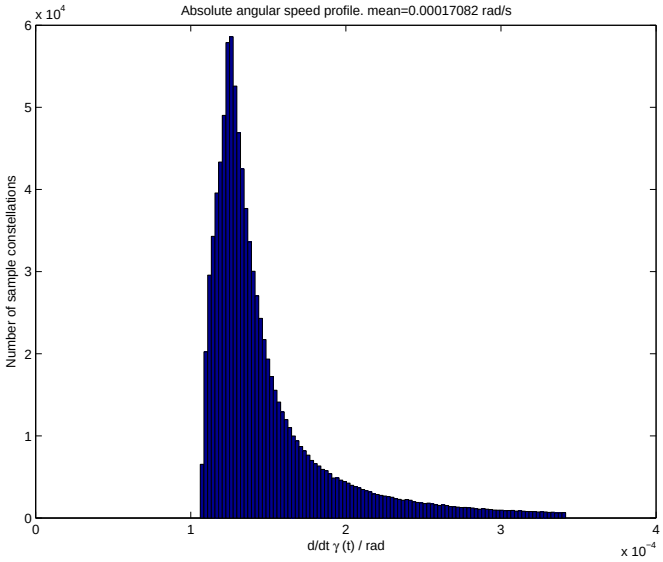


Figure 3.17: Absolute angular rate profile (10^6 samples)

3.2.3 Non-stationary Receivers

If the receiver antenna is able to move during the measurement, e.g. in a vehicle or a handheld receiver, multiple additional factors influence the path delays of individual reflected and direct paths. Although this brings additional complexity into estimation of multipath parameters, the movement of the receiver introduces an appreciable advantage: Since the direct signal and the reflected copy are extremely unlikely to arrive from the same direction, any movement of the receiver almost certainly alters the path delay between LOS path and any NLOS paths. This happens because while the distance to the satellite increases for instance, the reflector gets more distant, and thus, the path delays of the two paths change into opposite directions.

This leads to a rapid, but rather unpredictable modulation of both amplitude and carrier phase, especially with unsteady movement of the receiver [LS05]. It is unlikely that it is possible to correctly handle all that noise using the proposed estimation algorithm, but fast moving receivers benefit of the fact that the multipath applies an unbiased error to the carrier phase estimate, if phase relation is randomly distributed. However, this topic is not covered in the present work.

3.2.4 Summary of Multipath Signal Properties

In the previous sections, some basic types of reflections have been analyzed. The description of a ground reflection (Section 3.2.1) is complete using the elevation angle of the satellite, and the antenna height referred to the reflecting, flat ground. A reflection at a house front can be described using the ground distance between user and the reflecting structure, and an elevation angle, given that the reflecting surface is positioned upright and also normally to the plane defined by the satellite, the user and the reflector. This scenario was then generalized towards an arbitrary constellation of reflector and satellite, where the angular change of the satellite position was split up into a part that is essential to path delay, and one that is not. Any constellation can in fact be transformed into a ground reflection scenario, and using a satellite position description with two orthogonal angles α and ϵ, the path delay can be determined by the change of ϵ, which is the angle that describes the angular movement of the satellite position in the plane defined through user, satellite and reflection point.

In all reflection cases, the angular speed in a certain direction is vital for the change in path delay. The amount of change in the observed elevation angle or azimuth angle per time has therefore been analyzed in Section 3.2.2. Both azimuth rate and elevation rate have no lower bound, because with the MEO orbit used in GPS satellites, movements in the sky may be either observed in both directions, only in elevation or only in azimuth from certain

user positions. For the ground reflection, this means that in the worst case, the determining value of the elevation angle may not change at all for quite a long time, thus leaving the path delay at a constant value.

Although the absolute angular rate (that is, the absolute angular movement in the sky) of a satellite does have such a lower bound, the positive effect of this fact depends on the orientation of a reflector. If the angular change is directed such as to change the resulting path delay, the multipath effect may be mitigated or estimated; however, since this cannot be guaranteed, the existence of a lower bound in the absolute angular change rate is no general benefit.

Table 3.1: Path Doppler values for Ground reflection and Generic Reflection

Ground Reflection				
E / deg	$\dot{E}(t)/\dot{E}_{\text{mean}}$	h/m	Δf/mHz	T_{min}/s
10°	1.0	2 m	1.6	626
30°	1.0	2 m	1.4	713
45°	1.0	2 m	1.1	874
60°	1.0	2 m	0.81	1235
85°	1.0	2 m	0.14	7083
House Front Reflection				
E / deg	$\dot{E}(t)/\dot{E}_{\text{mean}}$	g/m	Δf/mHz	T_{min}/s
10°	1.0	15 m	2.1	47.4
30°	1.0	15 m	6.6	16.4
45°	1.0	15 m	8.5	11.6
60°	1.0	15 m	10.5	9.5
85°	1.0	15 m	12.1	8.3

Concluding, Table 3.1 provides a concluding overview of path Doppler differences with respect to reflection type and reflection parameters. For the elevation rate, the mean value that has been determined in the simulation before has been used, and is described by the relative factor $\dot{E}(t)/\dot{E}_{\text{mean}} = 1.0$. The antenna height for the ground reflection was defined as 2 meters, whereas the baseline to the house front in the second case was defined as 15 meters. The Dopplers behave linearly to those parameters, i.e. with a double reflector distance or antenna height, the Doppler would double as well, and the needed measuring time T_min becomes half the value.

Chapter 4

Multipath Detection and Estimation

In the previous chapters, the signal resulting from the reception of multiple paths has been illustrated intensely. Changes in satellite geometry may alter the path delay between the individual paths, and this changes the phase relation of their carriers. If that change happens at a specific rate, i.e. the path delay changes at a certain speed, it leads to a modulation of the carrier signal which can be made visible using spectral signal analysis. The present chapter discusses this effect and proposes an estimation technique that is able to isolate individual paths in the received signal mixture. If the path delays change at significantly different rates, the technique can provide an estimate on the total number of received signal copies, which can then be used to enhance knowledge about the signal quality, or the integrity of the navigation solution.

The first Section 4.1 illustrates the signal effects that are caused by the interaction of individual signal copies in the receiver loop. Both distortion of the recovered spreading code and distortion of the carrier phase are analyzed.

Section 4.2 shows that complete mitigation of the phase distortion is achieved if long the path length differences change steadily, and the observation time is sufficiently large.

In Section 4.3, the core topic of this Diploma thesis is introduced: Using spectral analysis of a carefully chosen combination of receiver outputs, it is possible to detect if a signal is received on multiple concurrent paths. Both the number of those paths and their signal amplitude, or relative power, can be detected.

4.1 Multipath Effects on Signal Processing

4.1.1 GNSS receiver

As a preliminary repetition, this subsection explains the GNSS receiver model that is used in the consecutive analyses. The GNSS receiver consists of a PLL synchronized demodulation stage removing the carrier, and a DLL driven de-spreading stage, which recovers the data bits from the baseband chip sequence.

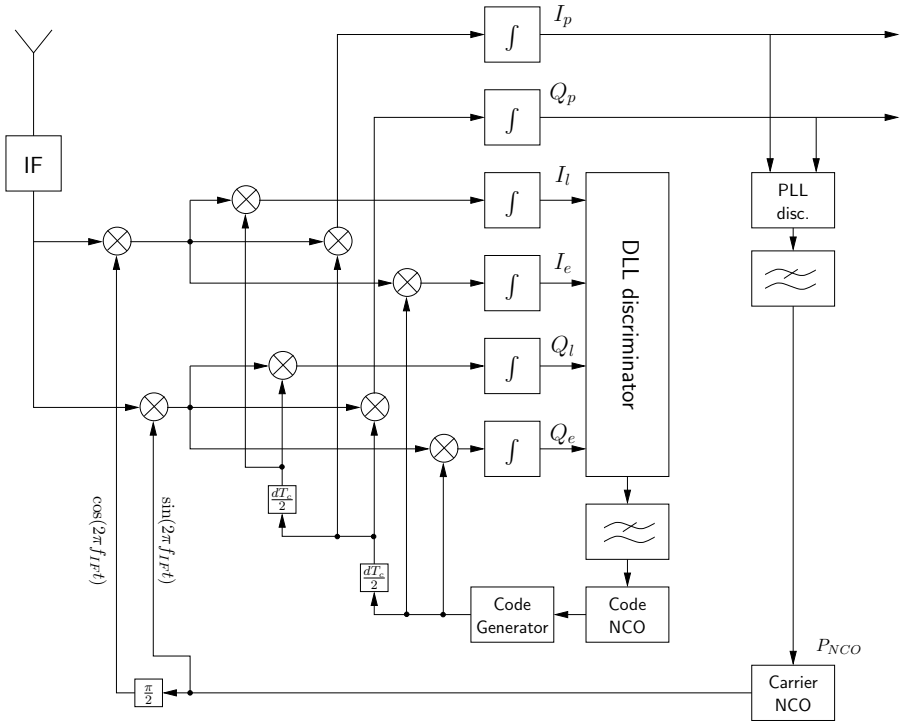


Figure 4.1: Schematic of the GNSS receiver [Gü05]

Figure 4.1 shows the schematic of a typical GNSS receiver. The antenna signal is mixed down to an intermediate frequency f_{IF}, and is then fed into the demodulation stage. Its objective is to remove the carrier coherently, and provide a baseband signal to the consecutive de-spreading stage. The signal is assumed to oscillate at a frequency close to f_{IF}, and thus an oscillation at the very same frequency must be multiplied on the signal. In practice, the exact frequency must be known from a previously performed acquisition of the signal, because the Doppler effect shifts the received carrier frequency in the range of ±6kHz. Because this Doppler shift may change rapidly due

to dynamics of the receiver or the satellite, it must be steered constantly by the PLL.

The demodulation is done for the in-phase component with a cosine, and for the quadrature component with the orthogonal sine signal, where both demodulated components are used to determine the phase estimate for the demodulation. This principle is called the costas loop [Cos56]. If the demodulation carrier is in phase with the received signal carrier, the quadrature output of the C/A code de-spread input signal is zero; otherwise, the phase difference between the demodulation signal and the received signal can be determined from the I and Q values. A discriminator function that can be used for the purpose of phase control is the inverse tangens [ME01], because it automatically eliminates the sign of the signal (tangens is periodic with π, instead of 2π). $\tan^{-1}(Q/I)$ returns the phase angle between the received signal carrier and the in-phase axis. In practice, however, this function is not used frequently because its computational burden is very high.

The chip sequence is then removed in a similar way in the de-spreading stage, where the reference code sequence is multiplied onto the signal. The code phase of the arriving signal is tracked with a delay-locked loop, which uses an early-late correlator as basis for the code phase control. This correlator type compares the output of a correlation with an early reference code with the output of a late correlation, and determines an offset that is forwarded to the code generator for the reference code. The detailed function of an early-late-correlator is explained in Section 5.1.3, and in a more complete and general way in [PS96].

4.1.2 Correlator Function Distortion

The satellite signal which is received under the noise floor has to be correlated with its own code sequence in order to make the data bits detectable again. A much more important "side-effect" of this autocorrelation is that, because the code phase has to be known exactly, the receiver already knows quite precisely about the time of reception for the corresponding signal when it has successfully demodulated the code sequence. Only synchronizing with bits would provide a time measurement at the accuracy of about 6000 km - this is clearly to large for reasonable positioning. With synchronization within the range of one chip duration, the accuracy becomes 300 meters. But because the detected code phase is used to determine the propagation delay, any distortion to this estimate will lead to a positioning error. Multipath can impose such a distortion on the shape of the autocorrelation function, and therefore lead to a misleading estimate. How this happens is described in the following.

Signal Correlation

The code phase of the received signal is found out by correlating it with a copy of the used pseudo-random noise sequence. The correlation returns a measure for the similarity of two signals, and since the sequence is almost random, similarity is expected to be almost zero for everything else than the same code sequence, with the same delay for both received signal and signal copy.

The received signal of a single path is assumed to arrive with a unknown delay τ, so it can be written as $r(t) = s(t - \tau)$. The parameter τ is to be estimated, since it defines the exact delay of the signal with respect to a fixed time frame. The correlator needs a copy of the signal, delayed by the estimate $\hat{\tau}$, so the signal copy is $r_c(t) = s(t - \hat{\tau})$. The difference between the estimated delay and the true delay is defined as $\Delta\tau = \tau - \hat{\tau}$. The correlation function $\varphi_{SS}(\Delta\tau)$ of a function $s(t)$ is defined as the multiplication of the signal with its delayed copy $s(t - \tau)$,

$$\varphi_{SS}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} s(t) \cdot s(t - \tau) \mathrm{d}t \ . \tag{4.1}$$

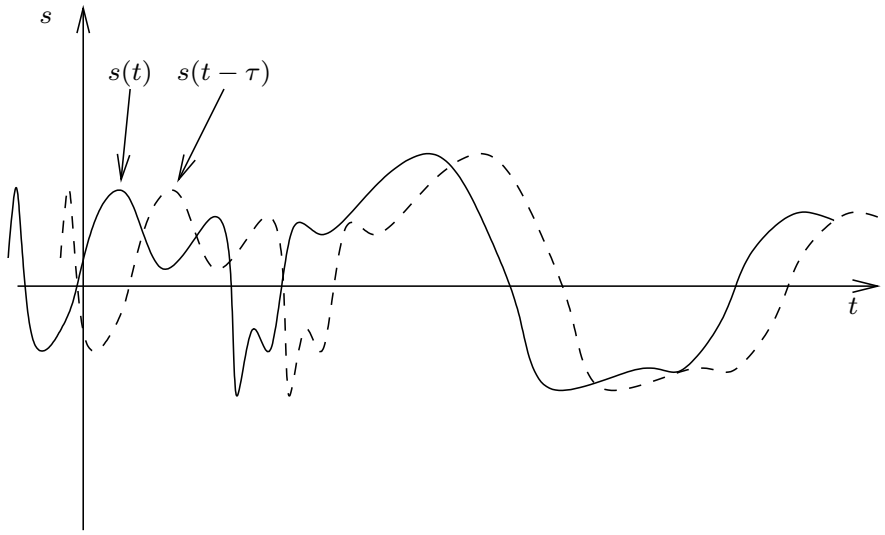


Figure 4.2: Signal and delayed signal

In Figure 4.2, a random signal of infinite length with zero mean is displayed. Zero mean is defined by the expectation of this signal, which is $\mathcal{E}\{s\} = \int_{-\infty}^{\infty} s(t)\mathrm{d}t = 0$. Furthermore, the signal is defined to be completely uncorrelated, meaning that there is no statistical interdependency between

the occurence probability of any two samples s_i and s_j - if the first pick was s_i, this does not state anything about whether the second pick will be s_j.

If the signal is multiplied with an exact, undelayed copy of itself, $\varphi_{SS}(0) = \int s^2(t)\mathrm{d}t$. If the signal copy is delayed, $\varphi_{SS}(\tau \stackrel{!}{=} 0) \equiv 0$ because the signals $s(t)$ and $s(t-\tau)$ consist of different sample sets, which are uncorrelated:

$$\mathcal{E}\{s(t) \cdot s(t-\tau)\} = \mathcal{E}\{s_i \cdot s_j\} = \mathcal{E}\{s_i\} \cdot \mathcal{E}\{s_j\} = 0 \ . \tag{4.2}$$

The autocorrelation is then zero:

$$\varphi_{SS}(\tau) = \int s(t) \cdot s(t-\tau)\mathrm{d}t = \int \mathcal{E}\{s(t) \cdot s(t-\tau)\}\mathrm{d}t = 0 \ . \tag{4.3}$$

This means that maximizing the autocorrelation function of a signal with unknown offset yields the offset as argument to the correlation function:

$$\hat{\tau} = \arg\max_{\hat{\tau}'} \varphi_{SS}(\hat{\tau}') = \arg\max_{\hat{\tau}'} \left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} s(t-\tau)s(t-\hat{\tau}')\mathrm{d}t \right] \tag{4.4}$$

where the parameter to be estimated is $\hat{\tau}$, and the signal's true delay is τ. Chosing $\hat{\tau} = \tau$ results in a maximum autocorrelation of the two signals, and this is how the GPS receiver finds out the code phase of the signal. In the acquisition stage, this has to be done by means of a full search for all possible delays - fortunately, the code sequence duration is only one millisecond, and thus all higher delays return the same correlation; it is therefore a cyclic autocorrelation function. When the code phase has been acquired initially, it can be tracked by carefully following the code phase such that the correlation stays at its maximum.

The autocorrelation function of the *pseudo*-random code sequences used in GPS satellites, however, does not have perfect autocorrelation; it only resembles this property very closely. The autocorrelation function of the Gold code is shown in Figure 4.3. This shape results from the shape of the code sequence, which is composed from rectangular pulses with amplitudes $\{-1; +1\}$. Thus, a delay of the signal for only a fraction of the pulse length does not immediately remove any correlation, but only degrades it slowly. In addition, the autocorrelation for offsets greater than the length of one chip is not perfectly zero, but may take one of three distinct values near zero [ME01]. In contrast to that, a completely random signal with zero mean has an autocorrelation function that is basically a dirac at $\tau = 0$, and zero otherwise.

Tracking the code phase is usually done with an *Early-Late Correlator* [Gü05]. Here, two correlators are tracking the signal continuously: One is

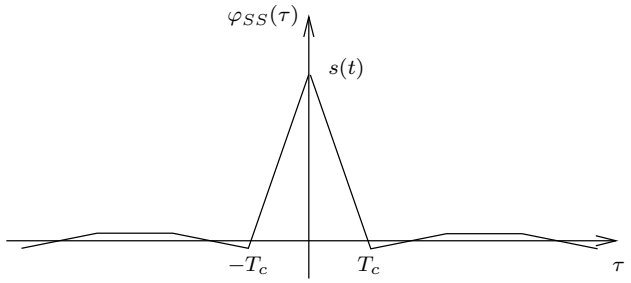


Figure 4.3: Autocorrelation function (ACF) of the Gold code

using a signal copy that is a little early, and one is using a late signal copy. Both signals are shifted equally far from the prompt signal, positioned at a distance such that the correlatos lie on the two edges of this characteristic correlation peak, if the estimated delay is correct. If the code is received at a different delay than the estimate, the correlators return different values, because one is closer to the peak and one is more distant. The spacing between the correlators equals either one chip length, or may else be decreased to enhance multipath mitigation properties [vDFF92]. While the standard spacing of one chip with is called "wide correlator spacing", a smaller value makes the correlator a "narrow correlator".

Figure 4.4 shows the correlator pair for a spacing of one chip length ($d = 1$; $\tau_L - \tau_E = d \cdot T_c$) in two situations: On the left, the estimate of the delay is correct: $\hat{\tau} = \tau$. Both correlators return therefore the same value, and their difference is zero. The right side shows a signal that was shifted to the right, meaning that the signal delay is larger than its estimate by the estimation error ε. The correlators return different values, and their difference can be processed to steer the estimate back to the true code delay.

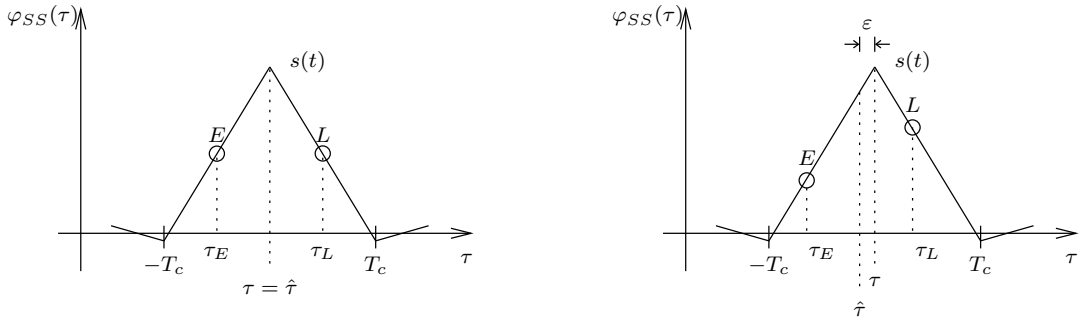


Figure 4.4: Early-Late correlator

Superposition of Correlation Functions

If multiple signal copies are received simultaneously, their signal sum that is processed by the receiver can be written as

$$s_{\sum}(t) = \sum_{k=1}^{K} A_k s(t - \tau_k) \ , \tag{4.5}$$

where the individual signal of path k has a path specific amplitude of A_k and a delay τ_k. The correlation function of the sum signal with the code copy is then

$$\begin{aligned} \varphi_{SS}(\hat{\tau}) &= \int_{-\infty}^{+\infty} s_{\sum}(t) s(t - \hat{\tau}) \mathrm{d}t \\ &= \int_{-\infty}^{+\infty} \sum_{k=1}^{K} \left(A_k s(t - \tau_k) \right) \ s(t - \hat{\tau}) \mathrm{d}t \\ &= \sum_{k=1}^{K} A_k \int_{-\infty}^{+\infty} s(t - \tau_k) s(t - \hat{\tau}) \mathrm{d}t \\ &= \sum_{k=1}^{K} A_k \varphi_{S_k S}(\hat{\tau}) \ . \end{aligned} \tag{4.6}$$

The correlation function $\varphi_{S_k S}(\hat{\tau})$ denotes the correlation function of the delayed and attenuated path copy k with the original receiver-internal code copy. The identity above shows that if multiple signals of the same origin, but with different delays are added, the correlation function of the combined signal results from the sum of their individual correlation functions. Therefore, the correlation slope that is effective when multipath is received can be constructed by addition of copies of the delayed and scaled peak. The example in Figure 4.5 shows the ACF of one direct signal and of a component delayed by half a chip length, corresponding to a 150m path delay. The code phases of the two signals are τ_1 and τ_2. The solid line shows the addition of the two curves, and the initial positions of the early and late correlators are depicted on that line.

As it can be seen, the influence of the second path signal on the correlation function causes a difference in the Early-Late correlator output. In this special situation, the second signal arrives just with enough delay as not to influence the early correlator any more; however the late correlator value is increased significantly by the reflection. When the delay locked loop tries to

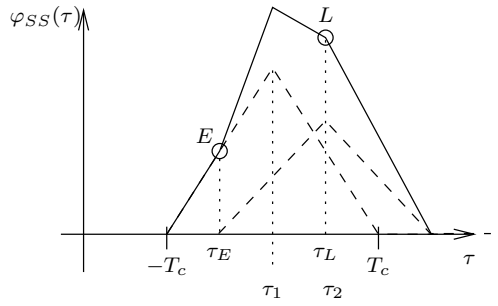


Figure 4.5: Distorted ACF with Early-Late response

equalize the correlator difference, it will do so by increasing the delay of the reference code until the difference is zero again. This will result in a code phase estimate that is too high with respect to the code phase of the first signal, which is the LOS signal to be measured.

Multipath Carrier Phase Influence on ACF

Until now, the analysis has been done using plain signals without modulation. As it was shown before, the correlation is done with the demodulated signal, i.e. the carrier has been removed before. Therefore, the receiver tracks the carrier phase in order to coherently demodulate the signal, which then consists of chips and bit modulation.

Assuming that the tracking loop of the carrier demodulation stage manages to keep track of the first, line-of-sight path's carrier phase; then the first path signal is demodulated by

$$\begin{aligned} r_{d,LOS}(t) &= A(t)\ b(t)\ c(t)\ \cos(2\pi f_c t + \varphi) \underbrace{2\cos(2\pi f_c t + \varphi)}_{\text{Demodulation}} = \\ &= A(t)\ b(t)\ c(t)\ (1 + \mathcal{R})\ , \end{aligned} \tag{4.7}$$

where $\mathcal{R}$ denotes the double-frequency residual, which is removed by a low pass filter consequently.

Because the demodulation of all additional path components is done with the very same demodulation carrier, the phase relation between demodulator signal and received signal has an influence on the output. The delayed signal $\cos(2\pi f_c(t-\tau)+\varphi)$ is demodulated:

$$r_{d,NLOS}(t-\tau) = A(t-\tau)\ b(t-\tau)\ c(t-\tau)\ (\cos(2\pi f_c \tau) + \mathcal{R}) \tag{4.8}$$

The low pass filtered result is now prepended with a cosine-shaped factor $\cos(2\pi f_c \tau) \in [-1; +1]$, and thus the phase difference of the alternate path makes the resulting demodulated signal cycle somehow between its maximum and its minimum. For the correlation function of this individual component, this means that the positive reference code sequence is eventually multiplied with the inverted received signal, which results in an inverted correlation curve as well. The resulting correlation function sum of direct path and the reflected path is then not distorted upwards, but bent downwards. Referring back to the example from Figure 4.5, this would lower the late correlator output and thus shift the estimate for the code delay left. So a delayed signal can also influence the correlator to estimate the code phase too early.

Conclusion to Code Phase Distortion

The estimated code phase is distorted massively if the delay between the paths is small enough. This degrades overall receiver performance and introduces measurement errors which propagate into the position solution. However, several counter-measures are known that can at least detect this kind of degradation: As already mentioned briefly, the correlator spacing can be lowered, making reflections from higher distances vanish in the noise. The criterion for this to happen is $\Delta\tau \geq \frac{(d+1)T_c}{2}$; for the wide correlator spacing of $d = 1$, the expression reduces to $\Delta\tau \geq T_c$. A second technique is to use multiple correlator pairs, and thus reconstruct the curve of the correlator function. If each side of the peak is at least sampled at two distinct points, the gradient of the rising and the falling edge can be compared, and differences can be taken as a hint for existing multipath.

4.1.3 Carrier Distortion

Apart from the shape of the correlation function, multipath reception also distorts the carrier phase of the effectively received signal. Again, the visible signal at the receiver originates from the sum of all signal paths. The path delays of the individual components can be translated to phase offsets, and the sum signal that is visible at the receiver has a distinct phase offset that depends directly from all the component phases. In the following, the origination of the resulting phase offset is illuminated by examination of the signal addition in complex numbers.

Path Carrier Phase and Phase Relation

The GPS L1 signal reaches the receiver antenna in its RF modulated form, i.e. multiplied onto a cosine carrier at $f_c + \Delta f$. Demodulation is done by

a coherent multiplication of the signal with the reference carrier oscillation, and it results in a baseband component and a component at the double carrier frequency. Since the signal was modulated with a cosine, we speak of the in-phase component and use again a cosine to remove the carrier.

Coherent Direct Path Demodulation. The received signal $r(t)$ was previously defined as

$$r(t) = A(t)\ b(t)\ c(t)\ \cos(2\pi f_c t)\ , \tag{4.9}$$

where Doppler shift Δf, the path delay τ and the phase offset φ have been removed for simplicity. After demodulation, the in-phase signal is

$$\begin{aligned} I(t) &= r(t) \cdot 2\cos(2\pi f_c t) \\ &= A(t)\ b(t)\ c(t)\ \big(\underbrace{\cos(2\pi f_c t - 2\pi f_c t)}_{=1} 1 + \underbrace{\cos(2\pi f_c t + 2\pi f_c t)}_{double frequency}\big)\ , \end{aligned} \tag{4.10}$$

consisting of the baseband portion and a term at the frequency $2f_c$. The high frequency part is removed by a low pass filter H_{LP}, so the remaining demodulated baseband signal is

$$I^*(t) = I(t) * H_{LP}(t) = A(t)\ b(t)\ c(t)\ . \tag{4.11}$$

If the receiver tracks the phase of the signal exactly, the remaining baseband signal is the recovered modulation payload, with the carrier completely removed. The coherent quadrature demodulation of a in-phase signal is zero after low pass filtering:

$$\begin{aligned} Q(t) &= A(t)\ b(t)\ c(t)\cos(2\pi f_c t) \cdot 2\sin(2\pi f_c t) = \\ &\quad A(t)\ b(t)\ c(t)\big(\underbrace{\sin(2\pi f_c t - 2\pi f_c t)}_{=0} + \underbrace{\sin(2\pi f_c t + 2\pi f_c t)}_{\text{filtered}}\big) \\ Q^*(t) &= Q(t) * H_{LP}(t) = 0 \end{aligned} \tag{4.12}$$

The carrier phase is usually tracked in a phase locked loop (PLL) using a Costas discriminator [PS96]. In the case of GPS, the quadrature component of the L1 signal contains the P/Y code, spread at a different rate and with a different code sequence. Therefore, after de-spreading with the C/A spreading sequence, the recovered quadrature signal is zero [ME01]. The discriminator function is the inverse tangens of quadrature component divided by in-phase component, so the discriminator output is:

$$d = \tan^{-1}\left(\frac{Q^*}{I^*}\right) \tag{4.13}$$

The numerically controlled oscillator is steered with the output d of the discriminator loop as its input phase correction. A single path is therefore tracked correctly, when the PLL loopback steers the demodulation carrier frequency so that the quadrature signal after de-spreading and demodulation is kept zero. The in-phase component contains only the data, because the chips are removed by despreading in the DLL.

This discriminator returns the angle of the complex signal with respect to the I axis. It does not depend on the sign of the in-phase component and is therefore usable without knowledge of the data. The downside is that if the phase runs away for more than $\pi/2$, the loop looses track and erratically locks onto the signal with a wrong phase which is shifted by half a cycle.

Demodulation of Delayed Signal Components. To understand the effect of delayed paths adding up onto the direct signal, demodulation of a single delayed path signal with the carrier referring to the direct path is analyzed first. Only the carrier of the signal is examined, and the effect of non-ideal de-spreading will be shown later.

The carrier of the delayed signal, $r_c(t-\tau) = \cos(2\pi f_c(t-\tau))$, is decorrelated with the undelayed carrier $\cos(2\pi f_c t)$, which refers to the undelayed line-of-sight signal. *(Note: In this section, the term "delayed" refers to the delay between the LOS and NLOS paths. The common path delay introduced through the transmission is neglected here. The signal is therefore assumed to arrive with no delay on the direct path, and with only the path specific delay τ_k on path k.)* The demodulation of the signal carrier with unity amplitude and without modulation yields

$$\begin{aligned} I_{MP}(t) &= r_c(t-\tau)\cdot 2\cos(2\pi f_c t) \\ &= \cos\left(2\pi f_c(t-\tau)\right)\cdot 2\cos\left(2\pi f_c t\right) \\ &= \cos\left(2\pi f_c \tau\right) + \underbrace{\cos\left(2\pi(2f_c)(t-\frac{\tau}{2})\right)}_{\text{filtered}} \\ I^*(t) &= \cos\left(2\pi f_c \tau\right)\ , \end{aligned} \tag{4.14}$$

and

$$\begin{aligned} Q_{MP}(t) &= r_c(t-\tau)\cdot 2\sin(2\pi f_c t) \\ &= \cos\left(2\pi f_c(t-\tau)\right)\cdot 2\sin\left(2\pi f_c t\right) \\ &= \sin\left(2\pi f_c\tau\right) + \underbrace{\sin\left(2\pi(2f_c)(t-\frac{\tau}{2})\right)}_{\text{filtered}} \\ Q^*(t) &= \sin\left(2\pi f_c\tau\right)\ . \end{aligned} \tag{4.15}$$

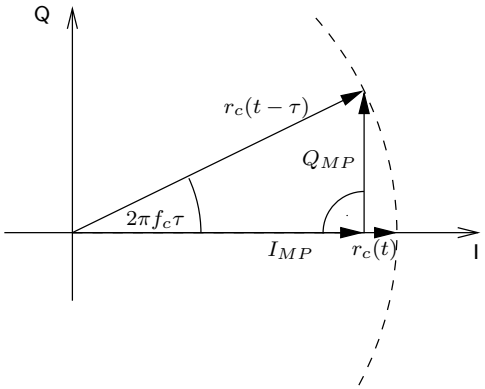


Figure 4.6: Carrier Phasor Rotation

Instead of the unity carrier amplitude, the in-phase component I_{MP} is reduced by a cosine factor depending on τ. The quadrature component Q_{MP} is no more zero, but increases with the sine of the delay in cycles. A descriptive explanation of this is the complex phasor diagram of the signal. Figure 4.6 shows the in-phase and quadrature outputs of the receiver in orthogonal axes, and it can be seen that the vector norm of the delayed signal vector $r_c(t-\tau)$ is still unity, whereas the measured I and Q values depend on the delay τ. The signal phase can be read off the diagram, and depends directly on the path delay:

$$\Delta\varphi = 2\pi f_c\tau = 2\pi\frac{\tau}{T_c}\ , \tag{4.16}$$

where special attention should be paid to the alternative use of T_c: Here, this parameter denotes the period of the carrier frequency f_c, not the duration of a code chip.

So the delay of the signal appears as a rotation of the demodulated signal in the I/Q plane. The rotation angle is periodic, so any two paths with a path delay difference of $k\cdot T_c$ result in the same demodulated carrier signal. Speaking in path lengths, this delay corresponds to $\lambda = T_c\cdot c_0 \approx 20$cm.

To analyze the addition of multiple signal paths, the introduced I/Q diagram can be displayed numerically by making use of complex numbers. The in-phase part of the signal is therefore defined as the real part, and the quadrature part delivers the imaginary value of a complex signal vector. Then, the demodulated signal is written as

$$\begin{aligned} \mathbf{r_d}(t) &= I_{MP} + j\ Q_{MP} \\ &= \cos(2\pi f_c \tau) + j\ \sin(2\pi f_c \tau)\ , \end{aligned} \tag{4.17}$$

or, using the complex notation of $\cos + j \sin$ [RW00],

$$\mathbf{r_d}(t) = e^{j2\pi f_c \tau}\ . \tag{4.18}$$

Note that the time parameter t does not appear on the right side of those expressions; the demodulated, low pass filtered signal is constant with time and only depends on the path delay, and the amplitude.

Demodulation of Signal Sum. Of course, if only the NLOS signal with the phase offset is received as described above, the Costas loop will adapt the carrier phase of the demodulator to match the signal phase, and the quadrature output will return to zero. The complex signal is then reduced to a real signal. If the path delay changes rapidly, the loop response might be too inert to compensate the phase change, and a signal component can be observed at the quadrature output.

However, when direct and indirect signals sum up at the receiver antenna, their corresponding complex demodulated signals may be added directly. The resulting I and Q values correspond to a virtual carrier signal, that is actually not received, but made up by the combination of the received components. In the following, the receiver carrier tracking is assumed not to react on the delayed signals, but to follow the direct path's signal phase without distortion, so that we can see the multipath's effect directly in the quadrature component. In practice, the receiver PLL always tracks the demodulation carrier phase, so this effect is only visible after reconstruction of the phase by subtraction of the phase feedback. This technique, however, will be explained in detail in a later section. For now, we just assume that the receiver has knowledge about the direct path's carrier phase, and uses it to demodulate all received signals.

The cumulative signal reaching the user antenna can be written as a sum of delayed cosines:

$$r_{\sum}(t) = \sum_{k=1}^{K} \cos\left(2\pi f_c (t - \tau_k)\right) \tag{4.19}$$

where again, only the carriers are considered. Demodulation with the undelayed carrier and low pass filtering results in the in-phase and quadrature components of the sum signal:

$$\begin{aligned} I_{\sum}(t) &= \sum_{k=1}^{K} \cos\left(2\pi f_c \tau_k\right) \\ Q_{\sum}(t) &= \sum_{k=1}^{K} \sin\left(2\pi f_c \tau_k\right) \ , \end{aligned} \tag{4.20}$$

or in complex notation,

$$\mathbf{r}_{\mathbf{d},\sum}(t) = \sum_{k=1}^{K} e^{j2\pi f_c \tau_k} \ , \tag{4.21}$$

where the path delays of the individual multipath components are τ_k. The LOS path is included as path $k = 1$, and has a path delay implicitly defined as $\tau_1 = 0$.

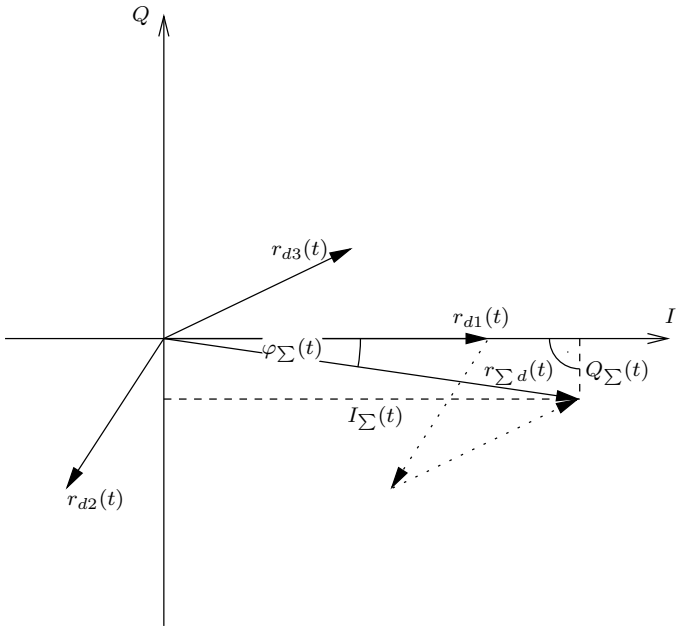


Figure 4.7: I/Q phasor diagram of multiple demodulated signal carriers

Figure 4.7 shows the superposition of three demodulated signals in the I/Q plane. The demodulation was done using the carrier phase of the LOS signal, so the LOS phasor is in parallel to the I axis. The addition of the

LOS vector and the two NLOS vectors, however, results in a virtual carrier signal with an imaginary component. It is important to recognize that the receiver does not see any of those individual path phasors, but only the resulting signal in the form of the receiver outputs $I_{\sum}$ and $Q_{\sum}$. The phase of this signal can again be derived using the inverse tangens:

$$\varphi_{\sum} = \tan^{-1}\left(\frac{I_{\sum}(t)}{Q_{\sum}(t)}\right) \tag{4.22}$$

Obviously, path components with a delay not matching a full or a half cycle duration distort the phase of the carrier tracking loop. As the receiver has no initial knowledge about the LOS signal's phase, it will therefore track the signal sum's phase. If any path's delay changes, the resulting phase will change as well. This is shown in the following subsection.

Changing Path Delays and Doppler Differences

In the previous analysis, the path delays of the indirectly received signal components were considered static, thus resulting in a static signal carrier phase at the receiver. As a next step, the effect of changing path delays will be examined. As was already shown, the path delay change rate can be expressed as a Doppler shift of the signal:

$$\frac{d}{dt}\tau(t) = \dot{\tau}(t) = \frac{\Delta f}{f_c} \tag{4.23}$$

In the I/Q plane from Figure 4.6, the delayed carrier signal was shown as a rotated phasor, and the phasor rotation was a function of the delay τ. The path delay is now assumed to change at a locally constant rate, and is then written as

$$\tau(t) \simeq \dot{\tau}(t) \cdot t + \frac{\varphi_0}{2\pi f_c} \ , \tag{4.24}$$

where φ_0 is chosen arbitrarily.

Plugging the path delay into (4.17), the signal phasor is

$$\begin{aligned} \mathbf{r_d}(t) &= \cos\left(2\pi f_c\left(\dot{\tau}(t)\cdot t + \frac{\varphi_0}{2\pi f_c}\right)\right) + j\sin\left(2\pi f_c\left(\dot{\tau}(t)\cdot t + \frac{\varphi_0}{2\pi f_c}\right)\right) = \\ &= \cos(2\pi\Delta f\ t + \varphi_0) + j\sin(2\pi\Delta f\ t + \varphi_0)\ . \end{aligned} \tag{4.25}$$

(4.25) constitutes a complex oscillation at the Doppler frequency Δf, and can be also written in complex exponential notation as

$$\mathbf{r_d}(t) = e^{j(2\pi\Delta f\ t+\varphi_0)} \quad . \tag{4.26}$$

With the complex signal phasor written as such, it can be seen that for changing path delays, the demodulator output gets modulated by the Doppler shift of the received signal, which is defined by the change rate of the path delay (or vice versa). Using the exponential denotation, the signal sum of multiple path delayed components can be written as

$$\mathbf{r}_{\sum \mathbf{d}}(t) = \sum_{k=1}^{K} e^{j(2\pi\Delta f_k\ t+\varphi_k)} \ , \tag{4.27}$$

and this reveals the main interest of the present work: Individual signal paths can be detected by analyzing the demodulated signal for distinguishable oscillations. This is easier done in the frequency domain, and will be explained in detail in Section 4.3.

Multipath Attenuation in the DLL

In the previous sections, the effect of multipath on the phase of the demodulated signal carrier has been analyzed. Therefore, we assumed that no PRN chip modulation is present, and the signal is received well above noise, so that the carrier is actually visible to the receiver at any single time instant. The GPS signal, however, is received below noise and must be recovered by de-spreading it with the correct code sequence, at the correct code phase. The code phase is tracked with a delay-locked loop [Gü05], which basically maximizes the returned signal power by shifting the reference code sequence which is multiplied with the received signal. The autocorrelation function of the used Gold code gives almost zero correlation for any code phase that is at least one chip length distant to the received code phase. In fact, there are 3 different distinct correlation levels apart from unity correlation (for zero phase shift), and this effect is caused by the mathematical properties of the used code family [ME01]. Figure 4.8 shows the autocorrelation function in a detailed view around $\tau = 0$ for the satellite PRN 23. The difference to the autocorrelation function of a completely random sequence is visible at delays $\tau > T_c$, where a small amount of correlation remains, whereas the random sequence has absolutely no correlation a these positions.

The figure shows that the correlation decreases continuously from its maximum when the code phase between the two compared sequences is increased up to one chip length. Any value higher than that results in a very low cross-correlation, which doesn't change significantly any more. This property of the autocorrelation function allows for the receiver to find

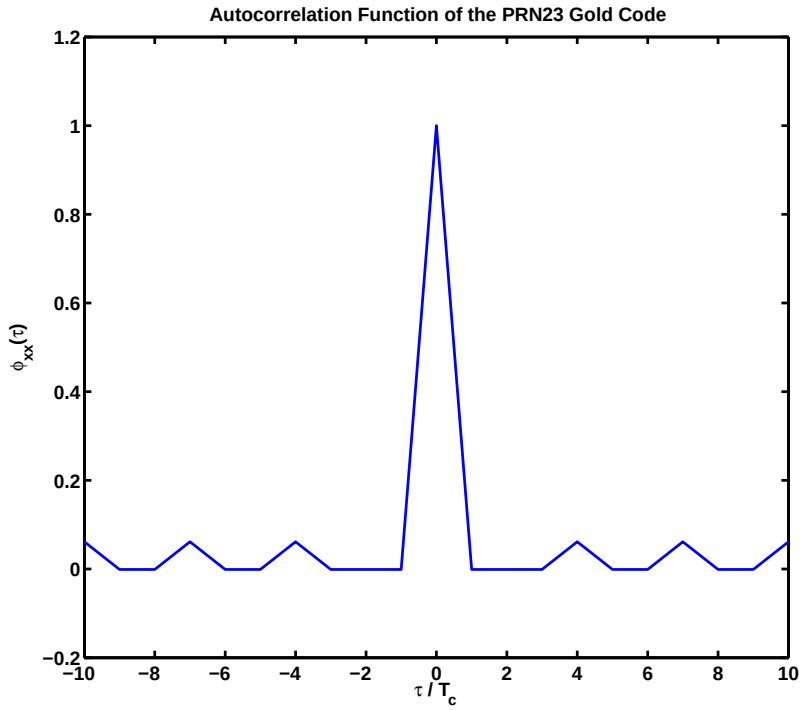


Figure 4.8: Autocorrelation function for the Gold code of PRN 23

out the true code phase very precisely, which is a prerequisite for the de-spreading mechanism which reconstructs the signal, originally received below the noise floor.

What is vital for the recovery of the very weak GPS signals, has another benefit concerning multipath mitigation: Any signal that arrives at least one code chip's duration $T_c = 1/1023$ ms later than the first signal is completely mitigated in the despreading stage of a receiver which tracks the LOS code phase correctly. This means that if the receiver tracks a signal at code phase τ_1, any signal copy at the delay $\tau_2 | \tau_2 - \tau_1 > T_c$ does not result in any correlated signal power, and is thus pretty much the same than any other uncorrelated noise. It therefore does not affect the tracking behavior of the loop directly, but only degrades signal-to-noise ratio.

Moreover, this effect not only begins suddenly at $\tau = T_c$, but any signal less distant from the LOS is already mitigated according to its path delay's fraction of T_c. To explain this, let's have a look at Figure 4.9. The upper part of it shows a detailed view of a fictive Gold code sequence $g(t)$, multiplied in-phase with the reference code $g_r(t)$. The result of the multiplication is plotted in the third timeseries, and since only values of the same sign are multiplied, the output is always positive. The DLL discriminator uses the integrated value of one or multiple whole code periods to determine the total correlation. It is easy to see that the illustrated multiplication of perfectly in-sync code sequences therefore results in the maximum integrator output, which is the objective of the DLL's controlling effort.

The lower half of the figure shows the de-spreading process of a delayed signal, e.g. of a reflected signal where the DLL is tracking the direct component, and thus the code phase is not correct for this replicated instance. If the integration is done continuously over the full code sequence duration, the fraction of the reference signal that gets gets de-spread correctly is only $\frac{\tau}{T_c}$. The portion of each chip where the reference sequence has already advanced to the next chip, but the received signal has not, results in unbiased noise. This is true because there is almost no correlation between the individual chip values in the correlator replica and the received signal.

The moments when the received signal is at the beginning of a new chip are depicted as dotted lines; the moments when the reference code sequence is at the beginning of a new chip are dashed lines. It can be seen that only if the new chip leads to a sign alteration, the interval between the two corresponding points in time leads to a negative correlator output; otherwise, the correlation is still positive because the neighboring chips are effectively indistinguishable.

The chip values of Gold codes are pseudo-randomly distributed and almost balanced [ME01], meaning that there are nearly as many -1 chips than there are $+1$ chips. Nearly, because with a total of 1023 chips, the

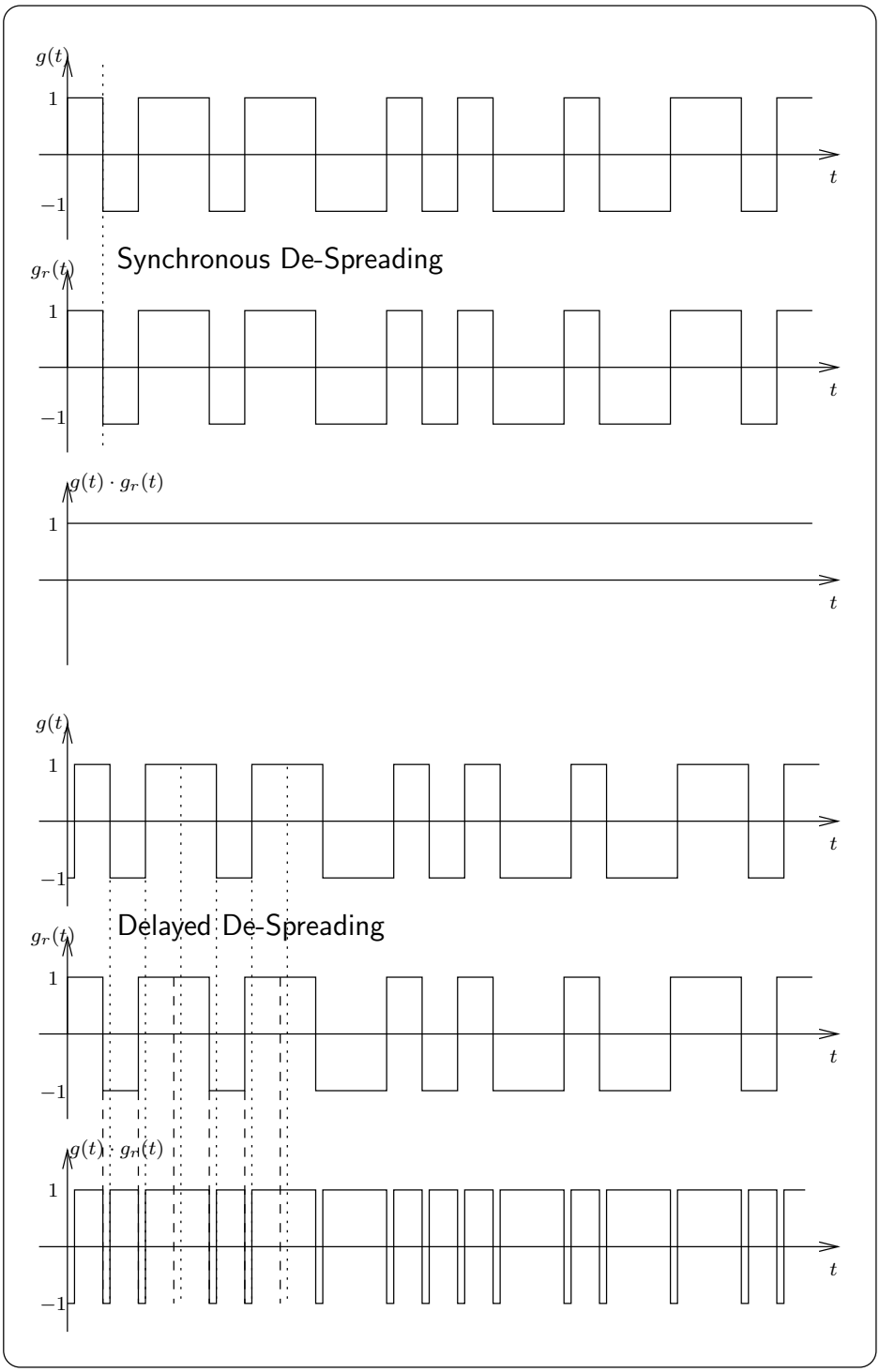


Figure 4.9: Despreading of LOS and NLOS signals

distribution is required to be unbalanced by one chip. More specifically, the Gold code has about as many sign changes as it has sign retentions during a whole code sequence. This means that the chance that the non-synchrouous chip change contributes positivly is about 50%, and about 50% that its contribution is negative. The total contribtion of the asynchronously de-spread components is therefore approximately zero, and this removes the influence of these parts to the correlator output completely. The conditional probability of a chip value can be written as

$$P\big(c[n] = +1 \;\big|\; c[n-1] = +1\big) = 0.5 \tag{4.28}$$

For a delayed signal with amplitude A_{NLOS} and path delay $\tau_{\mathrm{NLOS}} \leq T_{\mathrm{ch}}$, the amplitude after de-spreading must therefore be corrected:

$$A_{\mathrm{MP}} = A_{\mathrm{NLOS}} \cdot \left(1 - \frac{\tau_{\mathrm{NLOS}}}{T_{\mathrm{ch}}}\right) \;\; ; \quad \tau_{\mathrm{NLOS}} \leq T_{\mathrm{ch}} \; . \tag{4.29}$$

Reflections with a shorter path delay have therefore a much stronger impact on the total signal than reflections with a longer path delay. Not only does this afflict code correlation and therefore the DLL, but also the PLL and the tracked carrier phase: Since the phase tracking is done with the de-spread I/Q samples, the influence of delayed paths is only that of the correctly de-spread fractions of them. The difference between code and carrier in this context, however, is that for the carrier, the amount of path delay affects the amount of phase shift in a totally different way: The code phase is influenced stronger by a more delayed signal, while the carrier phase distortion depends only on the phase relation between the paths, which is periodic with 20cm. In particular, this means that concerning carrier phase distortion, the most "dangerous" echoes are those with a very small path delay, because they are almost fully reconstructed and thus have the greatest distorting impact. In contrast, for code phase measuremens, large delays are most harmful.

4.1.4 PLL Tracking Influence

In previous sections, the tracked carrier phase was assumed to stay at the LOS phase. In practice, this is not the case; the PLL rather tracks the phase of the signal combination. This means that the receiver always tries to adapt its demodulation carrier in a way that the combined, virtual signal component lies completely within the in-phase axis. This effect has only an impact on the carrier phase, since the code phase tracking discriminator function uses the envelope from $\sqrt{I^2 + Q^2}$. To correctly analyze the signal carrier's rotation components, however, the phase of the LOS signal must be reconstructed somehow. In detail, this can be done by removing the impact

of the PLL on the phase by subtracting the feedback phase. Another possibility is to remove the steady component of that feedback phase. Different techniques are analyzed and evaluated in a later section.

4.1.5 Special Case: Sign Reversion

A probably very rare, still serious case should be illuminated shortly: Previous sections made clear that the complex signal phasors of the demodulated multipath components superpose in any directions, according to their path delays. Therefore, the received signal carrier combination may differ both in phase and also in amplitude from the line-of-sight path. If enough reflected paths with a phase opposed to the LOS path add up, the resulting combination phasor will finally point into the left semi-plane, and the in-phase component of its value will be negative. In the receiver, this leads to the same effect as a bit change, and the carrier PLL will then steer the output signal of the demodulator towards the negative in-phase maximum $-I$ instead of $+I$, as is shown in Figure 4.10. It should be noted that the described effect is only possible if the path delay of the reflected components changes fast enough so that the PLL does not equalize the rotation of the resulting signal before it crosses the Q axis. Scenarios where the multipath components rotate fast enough, so that the resulting signal phasor lies alternatively in the positive and negative I semi-plane, is likely to produce errors in the navigation message.

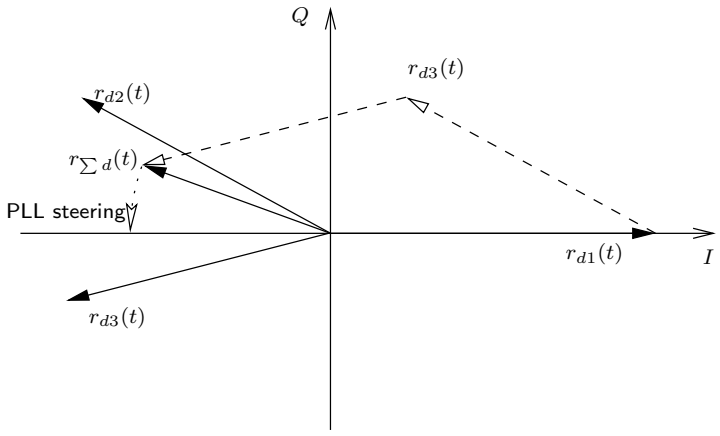


Figure 4.10: Sign Reversion by Addition of Multiple Components

4.2 Multipath Mitigation by Carrier Phase Averaging

The previous section showed the effects of multipath reception on the receiver-determined code phase and carrier phase. For both values, the carrier phase difference of the alternate paths is a key value, because it decides on the direction of both the distortion effects. This section illustrates that the distortion of the carrier phase can be mitigated by averaging it, if certain requirements are fulfilled.

4.2.1 Carrier Phase Modulation

As it was shown in (4.21), the combined reflection of multiple delayed signals can be described in the complex as a sum of rotating phasors. The rotation of a phasor around the tip of another results in distortion of the resulting measured carrier phase, because the sum phasor rotation either has a greater phase that the original signal, or less phase.

For a single reflection, and under the assumption that the line-of-sight carrier phase is tracked correctly, the complex signal can be written as

$$r_{\sum} = r_1 + r_2 = A_1 + A_2\big(\cos(\Delta\varphi) + \mathrm{j}\sin(\Delta\varphi)\big)\ , \tag{4.30}$$

where the phase difference between the line-of-sight phasor and the reflected phasor is denoted as $\Delta\varphi$, and the amplitudes of the two paths are A_1 for the LOS component r_1, and A_2 for the NLOS path r_2, respectively. Figure 4.11 shows the scenario; the LOS phasor lies in the in-phase axis, because it's carrier phase is tracked correctly. This is the case if $\Delta\varphi$ changes fast enough for the loop not to follow any more.

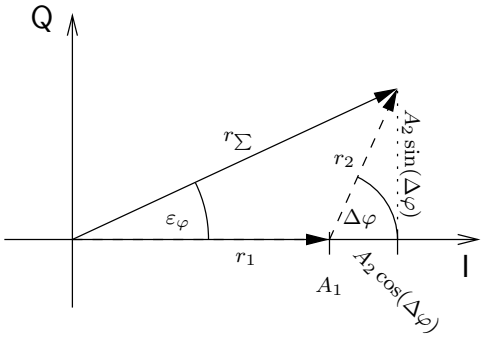


Figure 4.11: Carrier Phase Distortion by a Single Reflection

Given the phase difference $\Delta\varphi$, the in-phase and quadrature components of the resulting signal $r_\sum$ can be determined by

$$\begin{aligned} I_\sum &= A_1 + A_2\cos(\Delta\varphi) \quad \text{, and} \\ Q_\sum &= A_2\sin(\Delta\varphi)\ . \end{aligned} \tag{4.31}$$

The phase angle distortion of the resulting signal phasor is then

$$\varepsilon_\varphi = \tan^{-1}\left(\frac{Q_\sum}{I_\sum}\right) = \tan^{-1}\left(\frac{A_2\sin(\Delta\varphi)}{A_1 + A_2\cos(\Delta\varphi)}\right)\ . \tag{4.32}$$

The inverse tangens is valid in this formula if $I_\sum > 0$, which is assured by $A_1 > A_2$. This can be assumed, because the reflection attenuates the alternate signal path, and thus the direct path is always stronger.

With (4.32), the phase distortion for a given path phase difference has been described. As the phase difference changes, the phase distortion changes as well, and it is possible to determine a mean phase distortion that results from averaging over a set of measurements, where the probability density function of the phase difference is given. In the following, it will be showed that for equally distributed phase differences, the phase distortion is zero in average.

Because the phase difference $\Delta\varphi$ is periodic at 2π, the resulting phase distortion function is periodic as well:

$$\varepsilon_\varphi(k\cdot 2\pi + \varphi_0) = \varepsilon_\varphi(\varphi_0) \qquad k\in\mathbb{N} \tag{4.33}$$

Thus, a phase difference angle normalized to $|\Delta\varphi| < \pi$ is assumed in the following, making it easier to provide a probability since now the range of valid angles is limited in both directions.

It has been shown that the change rate of the path delay can be expressed as a difference in the path Doppler in (4.23). Assuming a constant Doppler difference, the path delay changes at a constant rate and can be expressed by a linear equation containing the change rate and an offset:

$$\Delta\tau(t) = \Delta\dot{\tau}t + \Delta\tau_0\ , \tag{4.34}$$

so that the path length difference between the two paths, direct and reflected path, changes at a constant rate. The resulting phase angle between the paths is then

$$\Delta\varphi(t) = \frac{2\pi c_0}{\lambda}\cdot\Delta\tau(t) = t\cdot\frac{2\pi c_0}{\lambda}\Delta\dot{\tau} + \varphi_0 \tag{4.35}$$

For this constant phase rate, the probability for the nomalized path phase angle is equally distributed. The signal phasor of the reflected path

rotates around the tip of the direct path phasor, and its rotation speed is constant. This equal distribution means that the probability of occurence for any given two path phase angles $\Delta\varphi_1$ and $\Delta\varphi_2$ is equal:

$$P(\Delta\hat{\varphi} = \Delta\varphi_1) = P(\Delta\hat{\varphi} = \Delta\varphi_2) \tag{4.36}$$

The resulting phase distortions of two path phase angles which are equal in absolute value, but have a different sign are described by the negative identity:

$$\begin{aligned}
\varepsilon_\varphi(-\Delta\varphi) &= \tan^{-1}\left(\frac{A_2 \underbrace{\sin(-\Delta\varphi)}_{\text{odd funct.}}}{A_1 + A_2 \underbrace{\cos(-\Delta\varphi)}_{\text{even funct.}}}\right) \\
&= \underbrace{\tan^{-1}\left(\frac{A_2 \cdot \big(-\sin(\Delta\varphi)\big)}{A_1 + A_2\cos(\Delta\varphi)}\right)}_{\text{odd function}} \\
&= -\tan^{-1}\left(\frac{A_2 \cdot \sin(\Delta\varphi)}{A_1 + A_2\cos(\Delta\varphi)}\right) = -\varepsilon_\varphi(\Delta\varphi)\ .
\end{aligned} \tag{4.37}$$

This transformation is valid because $\sin(-x) = -\sin(x)$, $\cos -x = \cos x$ and $\tan^{-1}(-x) = -\tan^{-1}(x)$. Using $\Delta\varphi_2 = -\Delta\varphi_1$ in the identity (4.36), we get

$$P(\Delta\hat{\varphi} = \Delta\varphi_1) = P(\Delta\hat{\varphi} = -\Delta\varphi_1)\ , \tag{4.38}$$

and thus,

$$\begin{aligned}
P(\hat{\varepsilon} = \varepsilon) = \sum_{\Delta\varphi|\varepsilon_\varphi(\Delta\varphi)=\varepsilon} P(\Delta\hat{\varphi} = \Delta\varphi) \qquad &= \\
= \sum_{\Delta\varphi|\varepsilon_\varphi(\Delta\varphi)=\varepsilon} P(\Delta\hat{\varphi} = -\Delta\varphi) \quad &= \quad P(\hat{\varepsilon} = -\varepsilon)\ .
\end{aligned} \tag{4.39}$$

The above formula states that the probability to obtain a certain phase distortion ε is the sum of the probabilities to obtain a phase difference $\Delta\varphi$ which causes this certain phase distortion ε. From 4.38 follows that the same probabilities exist for the occurence of phase differences $-\Delta\varphi$; their sum results in the probability to obtain a phase distortion $-\varepsilon$.

Because all phase difference probabilities are defined to be equal due to the constant Doppler, the probability for any phase distortion ε is equal to the probability for its complementary phase distortion $-\varepsilon$.

Of course, this rule is also valid for any other situations where the phase difference between two paths is equally distributed. For fast fading scenarios using a mobile GNSS receiver, this would be also the case; although the resulting phase distortion is only zero mean because of the statistic nature of the phase diffences.

4.3 Multipath Estimation with Carrier Phase Analysis

This section discusses the mathematical background of the proposed estimation technique, which is able to detect the number of reflections from tracking loop generated output signals while a GNSS receiver is tracking the satellite.

The general properties of the signal carrier have been illustrated already before. In this section, emphasis is placed on the modulating effects of constantly changing path delays, which can in fact be measured in the tracking loop. Those effects are best visible in the frequency domain, thus a Fourier transform is applied to the loop outputs.

4.3.1 Spectral Analysis of Distorted GNSS Carrier

As it was shown in previous sections, the reception of multipath signals appears as an addition of rotating carrier signal phasors in the complex signal plane. The rotation speed of each phasor is given by the path Doppler, which in turn is proportional to the change rate of the path length.

The following sections show the signal processing steps necessary to isolate components of the received signal which prove the existence of distinct reflection paths in the signal mixture. According to (4.21), the L1 signal carrier of a single satellite can be written as

$$\mathbf{r}_{\sum}(t) = \sum_{k=1}^{K} A_k e^{j(2\pi(f_c+\Delta f_k)t+\varphi_k)} \ , \tag{4.40}$$

where the path specific amplitude is assumed to be stationary for simplicity. The databits b_k and PRN chips c_k are again not considered here.

The path Doppler Δf_k of each path results from the path length change, and is individual to every reflection. In the complex signal plane, the result is a two-dimensional addition of the signal phasors of the individual paths , each of them with its vector norm A_k and a direction defined by the argument $2\pi(f_c + \Delta f_k)t + \varphi_k$. The first part of this term is the time-variant component of the direction angle, which is a rotation at the Doppler-shifted path carrier frequency. The second part φ_k is a phase offset, which is defined by the initial length of the path.

Modulation of Signal Amplitude and Phase

In the RF representation, all phasors rotate constantly at a very high rate, their path carrier frequency, but the individual paths rotate at slightly dif-

ferent speeds. The demodulation removes the common part of that rotation, leaving only the difference in rotation speed to the path phasors.

Demodulation is done using an estimated carrier frequency $\tilde{f}_c$, which is determined by maximizing the in-phase output of the demodulator, while minimizing the quadrature output. Particularly, the Costas loop maximizes the absolute in-phase signal, because it is unsusceptible to the sign reversion introduced by the data bit.

The demodulated signal carrier is again written as a sum of complex phasors:

$$\mathbf{r_{D\sum}}(t) = \sum_{k=1}^{K} A_k e^{j(2\pi(f_c+\Delta f_k-\tilde{f}_c)t+\tilde{\varphi}_k)} \ , \tag{4.41}$$

with the new path Doppler relating to the sum phasor's frequency, $\Delta\tilde{f}_k = \Delta f_k + (f_c - \tilde{f}_c)$.

Neglection of data and code modulation. At this point, it should be noted again that the modulation with data bits and pseudo-random chips has been completely disregarded up to now. The reason for this is that a sufficiently working de-spreading stage is assumed, which removes the chips from the signal. Only after this operation, the signal is usable; before, it is far below noise and cannot be detected. Thus, all the demodulator output can process is an integrated, and therefore averaged value, consisting of all samples during at least one whole code period, and with the chips already demodulated from the signal. Since correct removal of the chips is critical for the PLL demodulator to work, it is sufficient to neglect it in this context, and to assume an unspreaded signal above the noise instead.

The data bits change the sign of both in-phase and quadrature outputs, and thus rotate all complex signal phasors by π. This may happen every 20 milliseconds, and certainly impacts the phase of the carrier signal. But with the Costas loop used in common GNSS receivers, the output of the PLL, steering the demodulation phase, is not affected by this phase change. If signal copies arrive at different delays, the bit rotation occurs at different times for each path, which leads to a short distortion of the carrier phase. Here, the integrating property of the DLL is effective again: Since the delay between those paths is at most one chip length, it is much smaller than the integration interval. This effect is therefore mitigated mostly, and does not have to be considered.

Phase following PLL. An important fact to remember is that the receiver does not cope with the individual phasors, but tracks a combined phasor instead, which results from the addition of all received paths. The

PLL tries to track the rotating frequency of this combined phasor, and if the PLL filter bandwidth is set sufficiently high, it will succeed adjusting the demodulator phase just so as to keep the phasor in the in-phase axis. The result of this is that at the receiver output, no rotation or oscillation will be seen at all; instead, the phase feedback will contain this information completely. What can be seen at the receiver output is the amplitude modulation of the phasor, caused by the impact of the delayed signal copies.

Figure 4.12 shows the addition of three signal carriers in the complex signal plane, where the original I axis is that of the LOS path. The I' and Q' axes are those estimated by the receiver, because it tracks the phasor sum. Since the phase distortion ε_φ is followed all the time, the quadrature output will be zero at all times, and only in the in-phase output, a modulation can be observed. The phase feedback of the PLL has the phase distortion added to the phase it would track if only the LOS path was received.

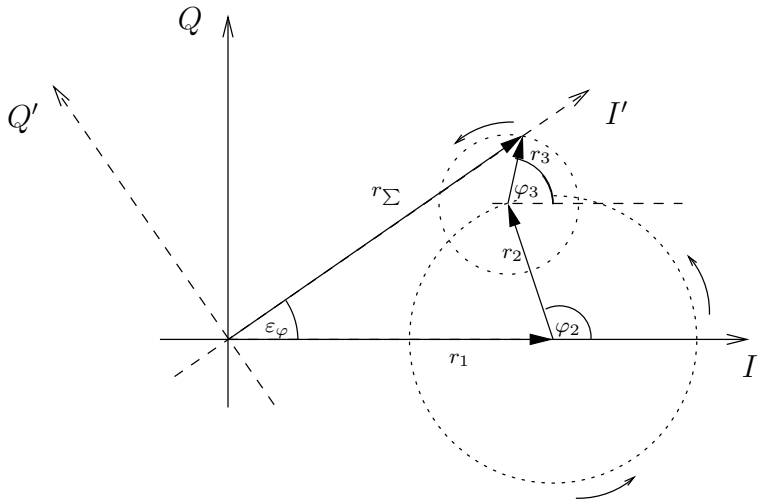


Figure 4.12: Addition of rotating carrier reflections in the I/Q plane

The in-phase and quadrature outputs of the receiver can be determined by adding the signal components, but neglecting the resulting phase. The path I_k and Q_k components of each signal path k are

$$\begin{aligned} I_k &= A_k \cos(\varphi_k) \\ Q_k &= A_k \sin(\varphi_k) \ . \end{aligned} \tag{4.42}$$

The amplitude of the signal sum can then be determined by the magnitude of the sum of all inphase and quadrature components, and since the phase is tracked, the in-phase output returns the exact amplitude:

$$I = A_{\sum} = \sqrt{\left(\sum_{k=1}^{K} I_k\right)^2 + \left(\sum_{k=1}^{K} Q_k\right)^2} \ . \tag{4.43}$$

With the path phase φ_k described by the path delay plugged in from (4.35), the amplitude can also be written as

$$A_{\sum} = \sqrt{\left(\sum_{k=1}^{K} A_k \cos(2\pi \Delta f_k t)\right)^2 + \left(\sum_{k=1}^{K} A_k \sin(2\pi \Delta f_k t)\right)^2} \ . \tag{4.44}$$

The phase distortion, which is contained completely in the feedback, is then determined by the inverse tangens of the inphase and quadrature sums:

$$\varepsilon_{\varphi} = \tan^{-} 1 \frac{\sum_{k=1}^{K} Q_k}{\sum_{k=1}^{K} I_k} \tag{4.45}$$

Utilizable Signals. Although the phase is followed by the PLL, this phase feedback is usually available, and can be used for further processing. In the case of multipath, it contains the phase distortion which the PLL has to mitigate. With only one rotating alternate path carrier phasor, the phase feedback contains the common phase trend of all paths, plus an oscillating component caused by multipath. This oscillation can be detected, and is a strong indicator for existing multipath distortion.

The amplitude of the signal carrier sum can be taken directly from the inphase output of the PLL, if the carrier phase is tracked completely. It is also subject to an oscillating distortion, if multipath exists.

Both amplitude and phase contain oscillating components at the Doppler frequency difference $\Delta\tilde{f}_k$ for each path k. Frequency components are best detected in the frequency domain, and therefore in the next step, a Fourier transform is applied to those signals.

Fourier Transform of Phase and Amplitude Output

The Fourier transform converts a timeseries signal into the frequency domain, and is therefore a powerful tool to analyze any signal with respect to periodic components. The Fourier transform of a function $f(t)$ into the frequency domain is defined as

$$\mathcal{F}\{f(t)\} = F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \mathrm{d}t \ . \tag{4.46}$$

The rotation of the multipath phasors can be seen in a Fourier spectrum at the indicated frequencies Δf_k; but since the rotation is not a pure sinusoidal oscillation, the spectrum of each path component consists of a peak at the path Doppler, plus several harmonics. This effect occurs both for the Fourier transform of the phase distortion, and the Fourier transform of the amplitude distortion. Figure 4.13 shows the case where only one reflected phasor rotates around the LOS phasor, with the phase positions where the maximum amplitude distortions and the maximum phase distortions occur marked with small circles. Given the assumed constant cycling speed of the phasor, it can be seen that both phase distortion and also amplitude distortion do not change completely sinusoidal:

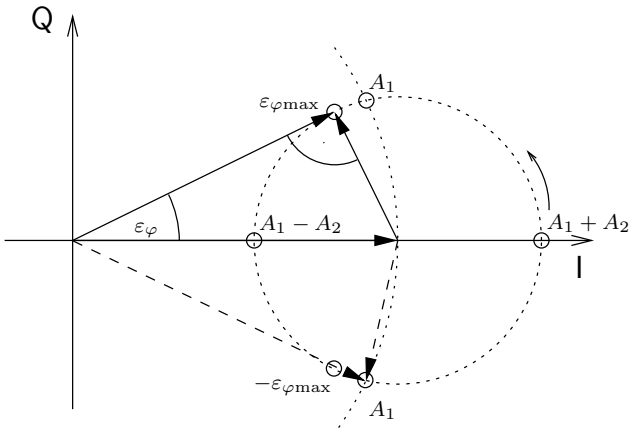


Figure 4.13: Positions of Maximum Phase Distortion and Maximum Amplitude Distortion

The **amplitude distortion** moves slower when the reflected path adds signal amplitude, because the angle, enclosed by the two points where the resulting amplitude equals the LOS amplitude A_1, is greater than 180°. This means that the zero crossing of $A_\sum - A_1$, which is the amplitude difference with respect to the LOS amplitude, is shifted towards the lower half-wave. For the average value of the distorted amplitude, this means that multipath results in a slightly higher mean amplitude than without multipath.

Secondly, the **phase distortion**, which cycles between $\varepsilon_{\varphi\max}$ and $-\varepsilon_{\varphi\max}$, has a faster falling edge than the rising edge under the indicated rotation direction of the phasor.

Both these signal shapes, which are depicted in Figure 4.14, can be approximated by a Taylor series at the base frequency Δf_k. Therefore, the theoretical spectrum contains many harmonics for every path, which can be seen in Figure 4.15. For the shown spectrum, a generated carrier signal consisting of two carriers has been processed. The path Doppler of the

carriers was 0.25Hz, and the amplitude ratio was set to 1 : 0.7, in order to get a good visualization of the effects. While for the amplitude spectrum, the DC component of the first path is clearly visible at $f = 0$, the phase spectrum has no such component.

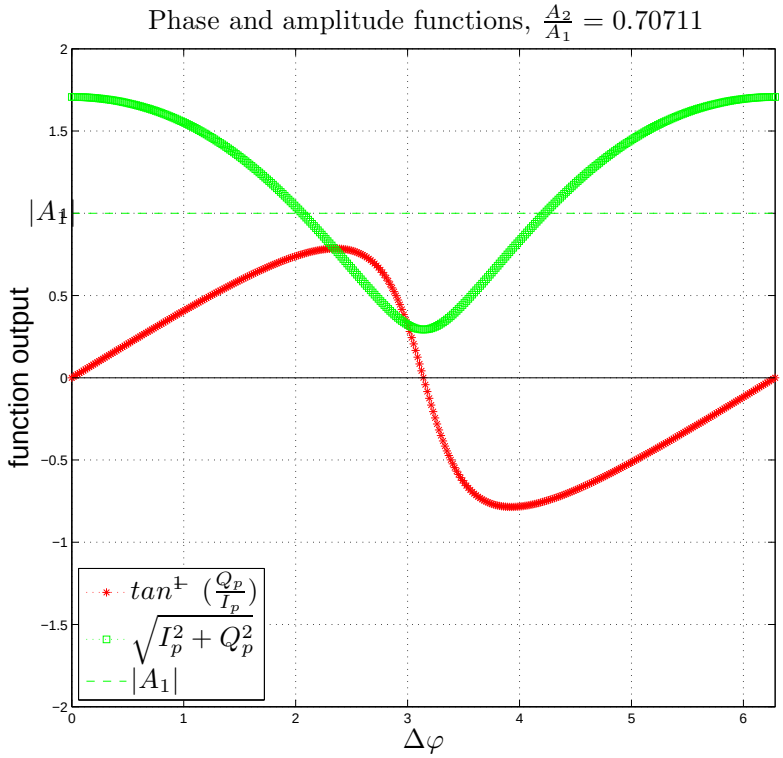


Figure 4.14: Phase Distortion Function and Amplitude Distortion Function

This is a serious drawback for detection of multipath reflections. Since almost every frequency bin corresponding to a multiple of the original detected path frequency is occupied by a harmonic, detection of additional paths on those frequency bins is not possible using the spectrum only. An approach which uses iterative decomposition of the spectrum has been analyzed for spectra derived from noise-free generated signals, where the detection success was only of medium quality. For noisy signals like the ones measured, the detection of paths using the amplitude distortion spectrum or the phase distortion spectrum was not possible. Apart from harmonics occupying lots of frequency bins, the available signal power in the main peak is decreased significantly.

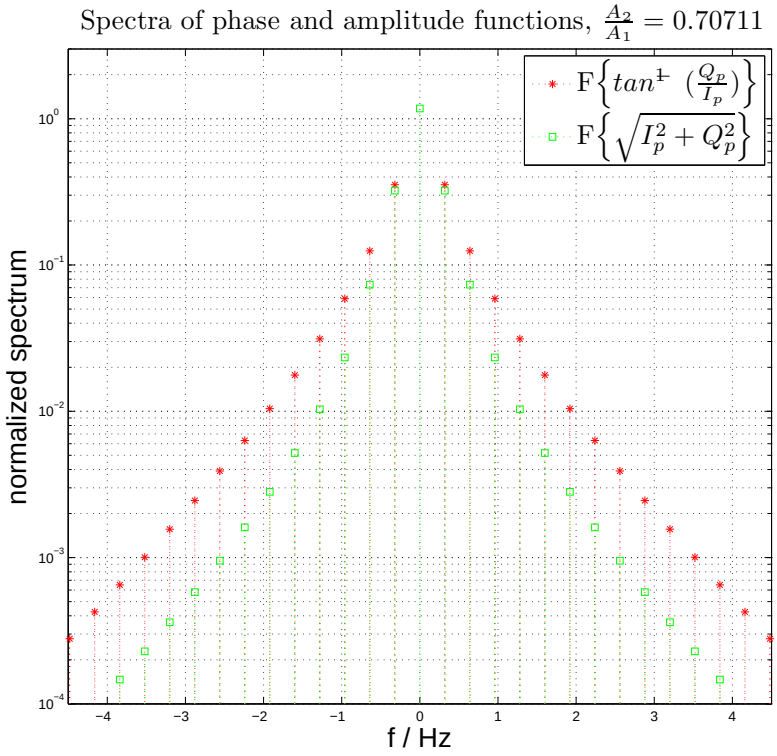


Figure 4.15: Theoretical Fourier Spectrum of Distortion Functions

The Complex Receiver Output

Because the previously analyzed distortion functions do not image the Doppler differences good enough, a signal representation that is closer to the sinusoidal shape has to be found. The complex signal plane, which has already been used to illustrate the combination of the individual path components, meets exactly this requirement. When the complex carrier signal is used, real part and imaginary part perform orthogonal, sinusoidal oscillations for every single path with a constant path Doppler, if the signal is noise-free. Moreover, using the complex value itself as an input function to the Fourier transform, we can even determine the sign of the Doppler frequency, which corresponds to the rotation direction of the phasor.

The complex carrier signal can not be taken directly from existing outputs of the tracking loop. With the in-phase output assigned to the real part of the complex signal, and the quadrature output providing the imaginary value, the phase-corrected complex signal is already available:

$$r_{\text{cplx}} = I + \mathrm{j}Q \tag{4.47}$$

Because the PLL tries to keep the signal phase zero, this signal will be mostly of real nature. The phase distortion, which rotates the signal away from the in-phase, or real axis, has been removed in parts by the phase feedback, which corrects the demodulation oscillator. In order to get an undistorted rotation of the phasors, the complex signal must refer to a fixed phase reference. The PLL basically tilts the I/Q coordinate frame every time the phase feedback changes. This rotation of the reference frame can be reversed by rotating the complex signal by the phase feedback:

$$r_{\text{cplx}} = \big(I + \mathrm{j}Q\big)e^{j\Phi_{\text{NCO}}} \ , \tag{4.48}$$

where the phase feedback Φ_{NCO} controls the oscillator and contains the phase distortion, assumed that the PLL loop filter bandwidth is high compared with all involved Doppler differences. If Doppler differences exist that can not be followed by the PLL, they cause signal on the Q output. Because Q is considered in the reconstruction of the phasor, this means that any Doppler difference cause by multipath will be included in the reconstruction, regardless of its frequency with respect to the filter bandwidth.

A typical phase plot of Φ_{NCO} is depicted in Figure 4.16. It shows the measured carrier phase, which was reconstructed from the phase feedback plus the residual I/Q phase. The straight line shows the theoretical LOS phase, of which the slope results from the inexact initial estimate of the carrier frequency, or Doppler shift to f_c, during acquisition. This means that using a PLL, the demodulator phase has to be re-adjusted constantly. In this

case, the initial frequency estimate was too small, so the rising demodulator phase effectively increases the resulting demodulation frequency.

The oscillation of the phase can be seen very clearly, and the shape shows the mentioned effect of different positive and negative edges. For this example, the alternate path amplitude has been set very high to visualize this effect, but common attenuation factors for a reflection are normally -5 to -15 dB with respect to the direct path, except some very rare, but still possible cases of low reflection energy. However, in this uncommonly strong reflection, we can see clearly that the oscillation is not perfectly sinusoidal, which decreases its qualification for being the sole input for detection of the frequency components.

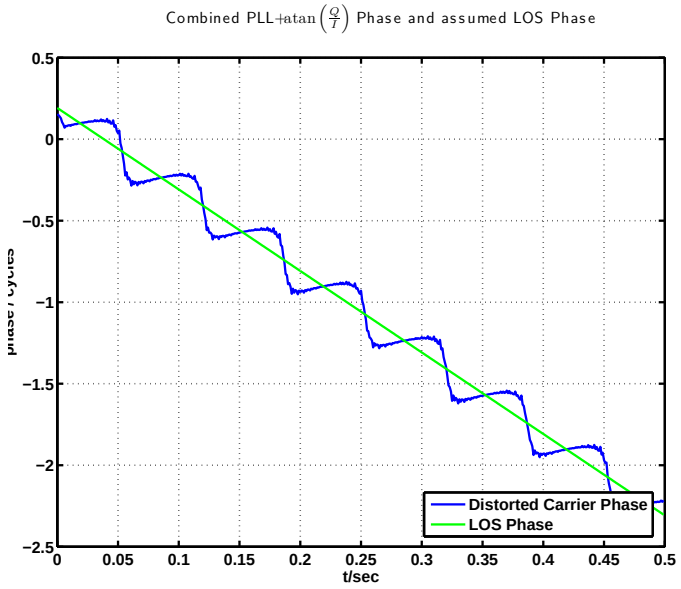


Figure 4.16: Typical Phase Feedback Distortion in case of Single Reflection

What happens if this phase correction is applied to the I/Q complex signal is that a slow modulation is re-applied to the phasor, making it effectively rotate. This behavior is undesireable, because it results in a frequency

shift of the Fourier spectrum. To recognize reflections in the spectrum, the LOS path should show up as the DC part of the complex signal, which corresponds to a phasor fixed to the real axis. In order to achieve this, the phase correction Φ_{NCO} is pre-processed before it is applied to the complex signal. In the scope of this Diploma work, two methods have been analyzed: Subtraction of a polynomial approximation, and high pass filtering.

The first method removes the trend from the phase correction by approximating the signal with a polynomial with a low degree of typically 2 to 5. This approximation can then be subtracted from the original phase feedback, and the result contains only the fast and periodically changing components. For noise-free signals of medium length, this method returns good results; however, the choice of the ideal polynomial degree depends heavily on the length of the signal to be analyzed, and is generally more difficult for noisy signals.

A better approach seems to be using a high pass filter removing the DC part and very slowly changing parts of the phase correction. It is important to notice that the present signal represents the phase distortion, and not the Doppler shift of an individual path. That means that any component in this signal that does not oscillate results from a rotation of the whole sum phasor. This is caused exclusively by the estimation error of the carrier frequency, and not by a reflection. Therefore, the cut-off frequency of the high-pass filter should be chosen as low as possible, while retaining a very steep filtering edge. Because reflections at a stationary receiver produce only very small Doppler differences, the effects of the reflection would otherwise be wiped off the phase distortion.

The complex phasor signal, rotated back into a fixed-reference complex plane by the filtered phase feedback, is then written as

$$r_{\mathrm{cplx}} = \big(I + \mathrm{j}Q\big) e^{j\left(\Phi_{\mathrm{NCO}} * h(t)\right)} \ , \tag{4.49}$$

where $h(t)$ depicts the timeseries representation of the described high-pass filter.

The effect of the data bit modulation has been neglected in the analysis up to now. With the phasor reconstructed as described, a bit change would rotate the phasor by 180°. Although the PLL is insensitive to sign changes, the phasor spectrum is severely degraded because the modulation is equivalent to a periodic step function. The result in the frequency domain is a sinc shaped envelope the spectrum, making it basically unusable for further processing.

Under the assumption of no sign flips due to multipath ($\sum_{k=2}^{K} A_k < A_1$), the data bit can be read off the sign of the receiver I output. The modulation

is undone by rotating the complex signal by 180° if the sign is negative, and by not rotating it if the sign is positive. With sign(x) defined as

$$\text{sign}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases} , \tag{4.50}$$

the sign is removed from (4.49) with

$$r_{\text{cplx}} = \big(I + \text{j}Q\big) e^{j\left(\Phi_{\text{NCO}} * h(t)\right)} e^{j\left(\pi \frac{1-\text{sign}(I)}{2}\right)} . \tag{4.51}$$

Figure 4.17 finally shows a simulated trace of the reconstructed complex signal. This plot was generated from a noise-free generated GNSS signal consisting of one LOS component and one NLOS component. The signal was acquired and tracked, and the tracking loop outputs I_p, Q_p and Φ_{NCO} were used according to (4.51) to reconstruct the complex phasor. The trace shows a circle which results from the NLOS phasor rotating around the LOS phasor very precisely. From this noise-free example, the true phase and the amplitude of the LOS component can be estimated perfectly from the position of the circle center.

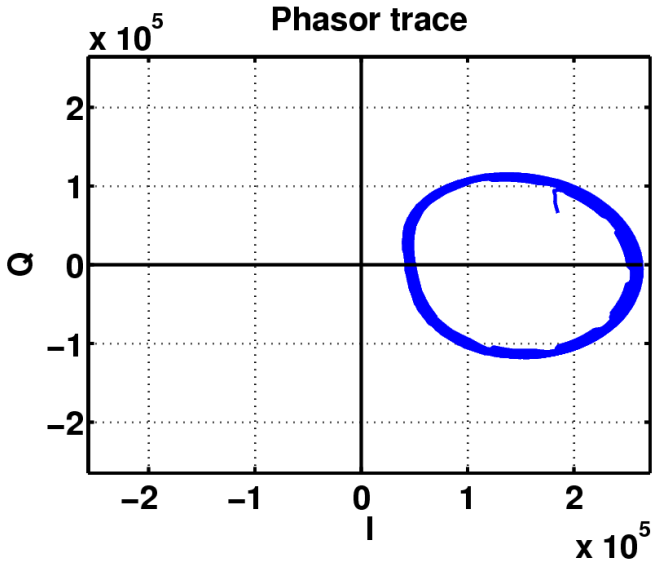


Figure 4.17: Phasor Trace of a Generated Signal

Fourier Transform of Complex Receiver Output

The Fourier spectrum of the reconstructed complex phasor outperforms the other previously analyzed spectra in several features: For noise-free signals, it has only one peak per reflection, and the available energy is fully

concentrated on that peak. Furthermore, because the Fourier transform is performed with a complex signal, the direction of the path Doppler can be determined from the spectrum.

Continuous Transform. For a complete reconstruction of the signal using the I, Q and Φ_{NCO} outputs of the receiver, where the Doppler of the first path has been removed sucessfully, the available signal is

$$r_{\sum}(t) = \sum_{k=1}^{K} A_k e^{\mathrm{j}(2\pi\Delta f_k t + \varphi_k)} \tag{4.52}$$

and represents the sum of all signal carrier phasors in the complex plane. It consists exclusively of sinusoidal oscillations, and the Fourier transform of that signal is

$$\begin{aligned} R_{\sum}(\omega) &= \int_{-\infty}^{+\infty} \sum_{k=1}^{K} A_k e^{\mathrm{j}(\Delta\omega_k t + \varphi_k)} e^{-\mathrm{j}\omega t} \mathrm{d}t \\ &= \sum_{k=1}^{K} \left(A_k e^{\mathrm{j}\varphi_k} \int_{-\infty}^{+\infty} e^{\mathrm{j}(\Delta\omega_k - \omega)t} \mathrm{d}t \right) \end{aligned} \tag{4.53}$$

where the angular frequency of the path k, $\Delta\omega_k = 2\pi\Delta f_k$ has been used. The value of the Fourier transform $R_{\sum}(\omega)$ is only non-zero if the argument ω matches one of the signal frequencies. This is shown in the following with the transform of a single oscillation:

$$\begin{aligned} f(t) &= e^{\mathrm{j}\tilde{\omega}t} \\ F(\omega \neq \tilde{\omega}) = \mathcal{F}\{f(t)\} &= \lim_{T\to\infty} \int_{-T}^{T} e^{\mathrm{j}\tilde{\omega}t} e^{-\mathrm{j}\omega t} \mathrm{d}t = \lim_{T\to\infty} \int_{-T}^{T} e^{\mathrm{j}(\tilde{\omega}-\omega)t} \mathrm{d}t = \\ &= \lim_{T\to\infty} \sum_{n=\frac{-T}{2\pi}}^{\frac{T}{2\pi}} \underbrace{\int_{n\frac{2\pi}{\tilde{\omega}-\omega}}^{(n+1)\frac{2\pi}{\tilde{\omega}-\omega}} e^{\mathrm{j}(\tilde{\omega}-\omega)t} \mathrm{d}t}_{=0} = 0 \\ F(\omega = \tilde{\omega}) &= \lim_{T\to\infty} \int_{-T}^{T} e^{0} \mathrm{d}t = \lim_{T\to\infty} \int_{-T}^{T} 1 \mathrm{d}t \ . \end{aligned} \tag{4.54}$$

With a normalzation, the spectrum results 1 for the exact frequency of the oscillation, and zero otherwise. If the oscillation is phase shifted, the

phase shift returns in the spectral component as a complex value of the form $e^{j\varphi}$ instead of 1.

The spectrum $R_{\sum}(f) = R_{\sum}(2\pi\omega)$ is then

$$R_{\sum}(f) = \begin{cases} A_k e^{j\varphi_k} & f = \Delta f_k, \quad k \in \{1 \dots K\} \\ 0 & \text{else.} \end{cases} \tag{4.55}$$

If all signal paths consinst of pure sinusoidal oscillations with arbitrary phase offset and amplitude, the spectrum of that signal consists of distinct lines at the frequencies of the individual oscillations, and is zero everywhere else. This gives a very good basis for estimation of those oscillations, even under noisy conditions, where also the Doppler spectrum is affected by the received noise.

Signal Spectra Figure 4.18 shows the Fourier spectrum derived from the complex signal of a two-path simulated GNSS L1 signal, which was tracked with a GNSS receiver implementation. The phase reconstruction has removed the common part of the phase change rate by high pass filtering the phase feedback Φ_{NCO}. Two distinct peaks can be seen from the spectrum: The component at $f = 0$ is the direct path with the stronger amplitude. The secondary peak at about 0.3Hz comes from the reflected path. It results from the constant change rate of the phase relation between LOS and NLOS path, which occurs at a frequency of 0.3 full cycles per second. This means that the path delay of the alternate path changes by 6 cm per second with respect to the length of the direct path. As a reference, the two previously analyzed spectra of amplitude and phase are plotted again.

Discrete Fourier Transform DFT Unfortunately, the above is only valid for contiuous signal processing. In order to perform the described analysis on a GNSS receiver, the signal has to be transformed as a discretely sampled timeseries. The discrete Fourier transform (DFT), however, has special properties that degrade the quality of the estimate.

A time-discrete N-point signal $x[n]$ is transformed into the frequency domain by a discrete Fourier transform [Lan02]

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn} \ , \tag{4.56}$$

where k denotes a discrete frequency bin corresponding to a frequency of $f(k) = k \cdot \frac{f_s}{N} = {}^{k}/_{T_s}$, T_s is the sampling frequency, and the exponential sequences W_N^{kn} are defined as

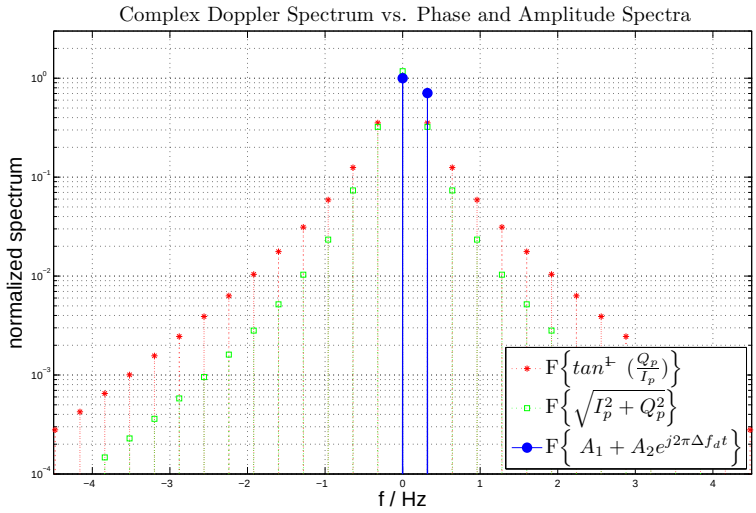


Figure 4.18: Complex Doppler Spectrum compared with spectra of phase and amplitude

$$W_N^{kn} = e^{-\mathrm{j}\frac{2\pi kn}{N}} \quad . \tag{4.57}$$

The discrete Fourier transform returns a periodic spectrum, which means that the upper half of the returned values can be seen as the lower half of a centered spectrum, or its first image spectrum shifted by $1/T_s$, alternatively.

An important difference to the continuous transform is the effect of "leakage". Since both time and frequency are condensed into bins, the transform can only return a distinct peak in the spectrum, if the input frequency of the oscillation lies exactly on the center frequency of one input bin. If not, the signal power leaks into neighboring frequency bins [Lyo04].

This leakage effect causes a degradation of the spectrum which inhibits proper detection of paths that are close together, i.e. with a very similar path Doppler. Windowing the timeseries signal can reduce the effect of leakage, however.

Another important fact that discriminates the discrete transform from its continuous equivalent is the symmetry of the spectra. The complex Doppler spectrum intruduced in Figure 4.18 shows the discovered path only at one distinct frequency in the spectrum. Due to sampling and discrete signal processing, however, the Fourier spectrum that results from the transformation is a periodic one and becomes symmetric, meaning that each path will result in two peaks at $\pm\Delta f$. Effectively, this removes the possibility to detect two independent paths with frequencies $\Delta f_1 = \Delta f; \Delta f_2 = -\Delta f$, resulting from

two reflections where one path delay is decreasing at the same rate as the other path is elongated.

The following section illustrates shortly what kinds of estimation are possible with this technique, and what kinds are not possible. After that, some results from simulations and measurements are presented along with a more detailed insight on the implementation in the next chapter.

4.3.2 Estimation of Reflected Path Parameters from Fourier Spectrum

The previous section derived the optimal signal form to be transformed into the frequency domain, and explained how this frequency representation of the receiver outputs can be used to detect active reflections in a satellite's received signal. This section provides a sum-up of the detecting possibilities of the proposed technique.

Multipath Existence Estimation

First of all, a strong peak in the complex signal spectrum is a strong hint that multipath exists, and the receiver is currently receiving a satellite at multiple times. If only one single path is received, the Fourier spectrum of the complex analysis signal has no strong peaks but rather a noisy shape, which depends on the used filter parameter in detail. There is a strong DC component in the spectrum, which represents the absolute value of the first path, i.e. the amplitude A_0. From there, the spectrum decreases with higher frequencies. High pass filtering can help getting rid of the steep slope in the spectrum around very low frequencies, and thus make the detection of peaks easier.

Path Quantity Estimation

If a multipath is sufficiently strong with respect to the direct path, it can usually be detected from the spectrum. An exception must be considered for multiple reflections with the same, or a very similar, Doppler shift: They apply the same kind of distortion to the carrier phasor, and are therefore also not visible independently in the spectrum. Depending on their mutual phase relation, they either amplify or cancel each other, known as constructive or destructive interference.

This limitation holds not only for exactly the same Doppler shift, but must be applied to similar path Dopplers as well: As already explained, the effect of leakage blurs each identifiable peak if the path Doppler is not

exactly the center frequency of a frequency bin, defined by the discretely sampled signal and its discrete Fourier transform.

Path Strength Estimation

The amplitude of the reflected signal's carrier, and therefore of the whole GNSS signal, is contained directly in the amplitude of the corresponding spectral peak. The ratio of the signal amplitudes of a detected NLOS path and the LOS path are the ratios of the peak values in the spectrum; and the reflected signal power to direct signal power ratio is therefore the squared ratio of the spectral peaks.

Mitigation of Carrier Phase Distortion

As has been shown in Section 4.2, in contrast to the code phase distortion, the distortion on the carrier phase through the described type of multipath is a non-biased one. The detection mechanism can therefore be used to activate an averaging process of the carrier phase, which effectively mitigates the phase distortion effect if the observation time is long enough. Since typical path Dopplers are very low frequencies at the fraction of a Hertz, it is yet questionable whether the phase of the direct path is stable enough during the required time interval, so that this mitigation can be performed properly.

Chapter 5

Estimation with Simulated and Measured Signals

In the previous chapter, the method of multipath estimation using Fourier spectra gained from various receiver outputs have been explained. The spectra visualize the oscillation of the carrier amplitude, and the carrier phase, which happens due to the changing inter-phase relationships between the individual direct or reflected paths.

In the frame of this work, different GPS L1 signals have been processed in order to validate the proposed estimation technique. These signals were either obtained from simulations, a signal generator, or from the present GPS satellites using field measurements. For all signal types, the resulting spectra could reveal multipath up to a specific level of path attenuation; however, for the measured signals, an airtight statement about contained multipath in the measured signals can hardly be given, because we don't know the corresponing channel impulse response for the real-life measurements.

In the following section, an overview and some detailed descriptions of the used components are given. Those components consist of the signal generator, which is provisionally described as a model; furthermore, the signal acquisition and tracking stage, which was implemented in MATLAB, is introduced in an extra subsection. Some implementation specific details about the estimation process and the necessary signal processing are described thereafter.

5.1 System Configuration

Initially, the proposed estimation algorithm has to be validated with a signal where the Channel Impulse Response (CIR) is known. For first tests

and optimization of the signal post-processing, a very simple signal generator has been implemented in MATLAB. To maintain compatibility between the different modules of signal generation, tracking and post-processing, all involved signals have been constructed at the intermediate frequency predefined by the used NordNav R30 Receiver. The samples that are output by the raw data grabber of the NordNav receiver are taken at a sampling frequency of about 16 MHz, and are stored as 4 bit signed integers, leaving a value range of only $[-127 : 127]$. An automatic gain control (AGC) mechanism has to assure that the sampled signal amplitudes fit sufficiently into this range to avoid a too strong quantization error.

5.1.1 Signal Generation

The following subsection describes the model of the implemented signal generator. The described signal generator is able to produce GNSS signal samples of one or more satellites received on multiple paths, modulated in the IF band, and sampled at a fixed sampling frequency of about 16 MHz. Each path's individual attenuation factor and its associated path delays are functions of time. A Doppler offset on the alternate path can therefore be generated by changing the NLOS path delay continually with respect to the LOS path's delay. Special attention has been drawn to assure that all signal paths retain continuity when signal parameters change dynamically, e.g. a path delay change which must not cause a phase jump of the corresponding signal carrier.

Signal parameters

Essentially, the signal generator constructs a multipath signal that follows the model described in Section 3.1. Thus, the generated signal consists of copies of the same PRN- and bit modulated carrier, each copy delayed and attenuated by its individual value, and shifted in frequency. The frequency shift corresponds to an increasing or decreasing path delay, as it was shown in Section 2.2. Therefore, it is sufficient to either parameterize the path delay as a function of time and derive the Doppler shift from that value, or to start with an initial path delay and alter only the Doppler shift during advance of simulation time.

In this implementation, the Doppler shift was designed to be an input function of time, where the path delay is only given as an initial value, $\tau_{0,k}$ for path k at time t_0. The effective path delay at time t can then be determined from

$$\tau(t) = \tau_0 - \frac{1}{f_c} \int_{t_0}^{t} \Delta f_D(t')dt' \ , \tag{5.1}$$

where the integrand $\Delta f_D(t)$ is the difference between the LOS path's Doppler shift and that of the particular NLOS path, also called the path Doppler.

The Doppler offset determines whether that the NLOS signal arrives slower (negative Doppler offset) or faster (positive Doppler offset) than the LOS component. Every path's signal, combined from the carrier oscillation, chips and bits, runs at its own speed, and the times when a chip and/or bit change occur may or may not be the same on all paths.

Intermediate-Frequency Signal Model

The generated signal is modulated at an intermediate frequency, making the required sampling rate more easy to process. The software receiver provided by the NordNav Receiver system was used for reference; so the nominal IF frequency and sampling frequency were decided to match the ones used by the NordNav RF frontend and the corresponding receiver software ($f_{IF} = 4.1304\text{MHz}$; $f_s = 16.6376\text{MHz}$).

To model the down-conversion correctly, the Doppler shift on the transmitted carrier frequency f_c must be applied unchanged onto the intermediate frequency f_{IF}. Thus, the generated carrier oscillates at $2\pi(f_{IF} + \Delta f_D)$. It is important to recognize that the phase relation between the Doppler-shifted IF carrier and the code varies due to their non-integer ratio. This means that the PRN chips and the data bits do not necessarily switch only at the zero-crossing of the carrier. In contrast to that, on the RF carrier, bit and chip switches only occur at zero phase, regardless of the Doppler shift.

To the IF carrier, the Doppler shift can be applied directly in the frequency domain by adding Δf_D to the IF frequency f_{IF}. The data bit stream, and the chip sequence, however, are usually denoted as functions of time where no data rate (and thus, frequency) can be derived directly. It is still possible to account for the effect of the Doppler shift by using a perturbation factor, which is applied to the time argument of these functions. According to [Psi01], the signal can be described by

$$s_r(t) = A\, b\big((1+\eta)(t-\tau)\big)\, c\big((1+\eta)(t-\tau)\big)\, \cos\big(2\pi(f_{IF}+\Delta f_D)t+\varphi\big) \ , \tag{5.2}$$

with the already well-known parameters A for the amplitude, $d(t)$ for the data bits, $c(t)$ for the PRN sequence, f_{IF} for the intermediate frequency, Δf_D for the Doppler shift, and φ for the initial carrier phase of the signal. The parameter η describes the mentioned perturbation factor that influences the bit rate and the chipping rate of the PRN sequence in the same way that

the Doppler shift does this for the carrier. Doppler shift and perturbation factor are related by

$$\eta = \frac{\Delta f_D}{f_c} \ , \tag{5.3}$$

which means that η measures the relative shift of the transmitted RF carrier frequency. As known from Section 2.2, the Doppler shift makes the signal run faster or slower, and this effect is not limited to the carrier only. It also affects the chip rate and the bit rate in the same relative magnitude.

With this representation, the underlying data bit and chip functions $d(t)$ and $c(t)$ have nominal bit rates of 50bps and 1023000 chips/s, respectively. The perturbing effect is constituted directly by preceding the time argument with the factor $(1 + \eta)$.

Concluding, the signal model for N satellites, received on K_n paths each, can be written as

$$\begin{aligned} r(t) \ &= \ \sum_{n=1}^{N} \sum_{k=1}^{K_n} A_{n,k} \ b_{n,k}\big((1+\eta)(t-\tau_{n,k})\big) \ c_{n,k}\big((1+\eta)(t-\tau_{n,k})\big) \cdot \\ &\quad \cdot \cos\big(2\pi(f_{IF} + \Delta f_{D\ n,k})t + \varphi_{n,k}\big) \ , \end{aligned} \tag{5.4}$$

where the subscripts $_{n,k}$ denote the parameter set for the n^{th} satellite, received on the k^{th} path.

In (5.4), the path amplitude and the path Doppler shift are denoted without their arguments of time for better readability only. In the signal generator implementation, these values are allowed to change and must be written as $A_{n,k}(t)$ and $\Delta f_{D\ n,k}(t)$, respectively.

5.1.2 Signal Measurement

Apart from the generated signals, measured GPS L1 signals have also been used to estimate multipath with the developed technique. Those signals have been obtained using the NordNav R30 receiver, which consists of a hardware frontend with a USB 2.0 interface, and a software component on a PC.

Receiver Hardware

The NordNav Receiver can grab IF data samples from the measured signals, while simultaneously performing acquisition and tracking on the fly. The receiver was equipped with a small patch antenna with an active, semi-spheric characteristic. The front end is responsible for the pre-amplification and an AGC, which is controlled from the software. The signal is down converted to an intermediate frequency of about 4 MHz, and sampled at

about 16 MHz sampling rate. The value- and time-discrete samples are then transferred via USB 2.0 to the host PC, where they can be written to disk, or processed in the receiver software.

Measurement Scenario

Measurements have been made in an urban area, where the antenna was placed outside the window of a building of Technische Universität München. The height above ground of the antenna was about 15 m, whereas the distance to the nearest reflectors was between 10 and 15 m. Two measurement locations were determined for IF sampling: One on the 3rd floor, heading west, with the housefront of a 4-storied building opposite the window, at a distance of about 15 m. The second location was set up on the 4th floor at the corner of the building, where both the opposite housefront (again at about 15m distance) and more distant possible reflectors were visible from the antenna position. Those reflectors included metal roofs of different pitches, and different other reflecting structures as stone or concrete walls and steel structures.

5.1.3 Signal Acquisition, Tracking and Post-Processing

Implementation of Acquisition and Tracking

To validate the proposed estimator, a GNSS receiver implementation that can provide the required output signals is needed. Johansson et al [JMTU98] have implemented a simple tracking loop in MATLAB, which was used as a basis for a software implementation of a GNSS receiver. Some modifications were made to meet the requirements of carefully post processing the receiver output data.

First, the desired output parameters had to be made available to the invoking function, including the carrier- and code phase and both I and Q channel prompt correlator outputs. To allow a visual inspection of the tracked signal, the cross correlation of the carrier demodulated signal and the prompt code is computed in the tracking loop, which can be examined later. Figure 5.1 shows a tracked satellite's cross correlation during 3 seconds, where the second axis denotes the code delay between the received signal and the locally generated prompt code, in chip lengths (T_c).

Multipath can be detected by visual inspection from such a plot by carefully comparing the shapes of the rising edge and the falling edge - if constructive multipath is active, the falling edge will be bulged upwards; if destructive multipath is active, the bulge will point downwards. The effect that a NLOS signal has on the correlator depends not only on the phase

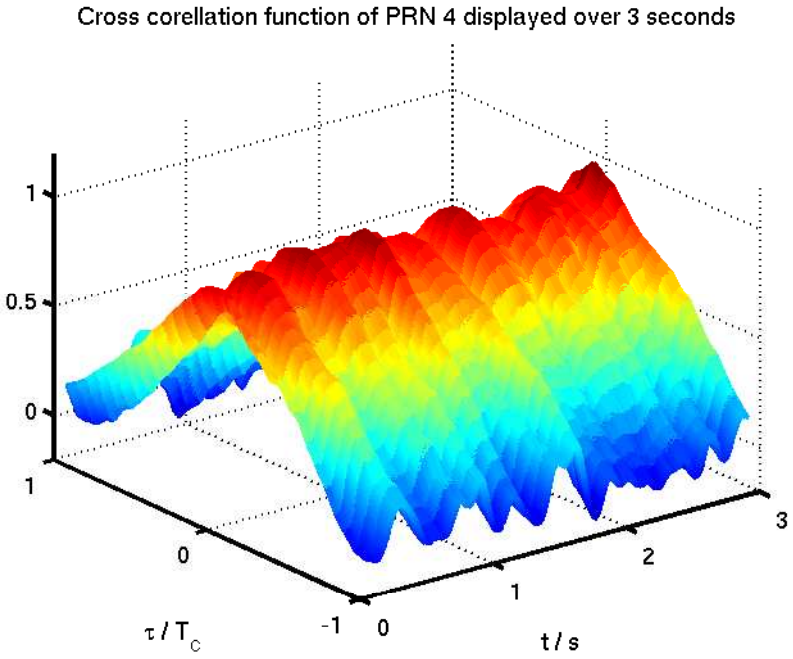


Figure 5.1: Exemplary cross correlation timeseries of a tracked satellite signal

relation between LOS and NLOS, but also on their amplitude ratio and the path delay that separates the reflected signal component from the direct path signal. In Figure 5.2, the possible deformations of the cross correlation function are shown for a number of different path delays. Each panel shows the same set of path delays, but with a specific reflection amplitude, or, when looked at in the base band, complex coefficient changing the reflection phase. This results either in constructive or destructive interference, and it can be seen that the effect of such an interference on the correlators depends on the path delay on the one hand, and on the correlator spacing on the other hand. If a narrow spacing is used, the distorting reflection has to occur earlier in order to touch the late correlator with the rising edge of its ACF. For a standard wide correlator spacing of $d = 1$, or $\tau_{\text{late}} - \tau_{\text{early}} = T_c$, the path length difference of $\Delta\rho = 450$m, shown in magenta, is the lower bound for a delayed signal copy which does not affect the correlator. This is valid provided that the correlator pair was locked on the direct path's peak before.

To automatize post processing of the tracking data, a tracking lock detector has been implemented in the loop. It evaluates the prompt correlator amplitude against a configurable threshold. If the low pass filtered amplitude goes below the threshold, tracking loss is indicated. This information can be used later to exclude tracking data where the signal was to weak to follow. When data is excluded from processing, it is very important to maintain signal continuity at the discontinued locations. Otherwise, the computed spectra may be distorted by jump discontinuities, which may occur especially in the carrier phase and the code phase.

The tracking loop in the receiver code proposed by [JMTU98] does not integrate coherently along whole code sequences, but rather shifts the prompt, early and late code replicas in their phase. This results in integration intervals that cover the time of bit change, resulting in degraded correlation and noisy output. To fix this, the tracking loop has been enhanced to shift the input data stream according to the detected code phase, e.g. re-read or skip samples if a code phase shift has been detected. In order to have the best temporal resolution available, the tracking loop was programmed to integrate at every code sequence, resulting in 1000 measurements per second.

Acquisition is done using a full search algorithm, which covers the range of ±4kHz in bins of 50Hz, and the code phase at all sample steps, which is about 16 samples per chip length. In a second run, the Doppler is pulled in more exactly to provide a better starting value for the demodulator NCO in the tracking loop.

The carrier tracking PLL is a Costas loop [Cos56] with an inverse tangens ($\tan^{-1}(\frac{Q_p}{I_p})$) as discriminator function, and a first order low pass filter

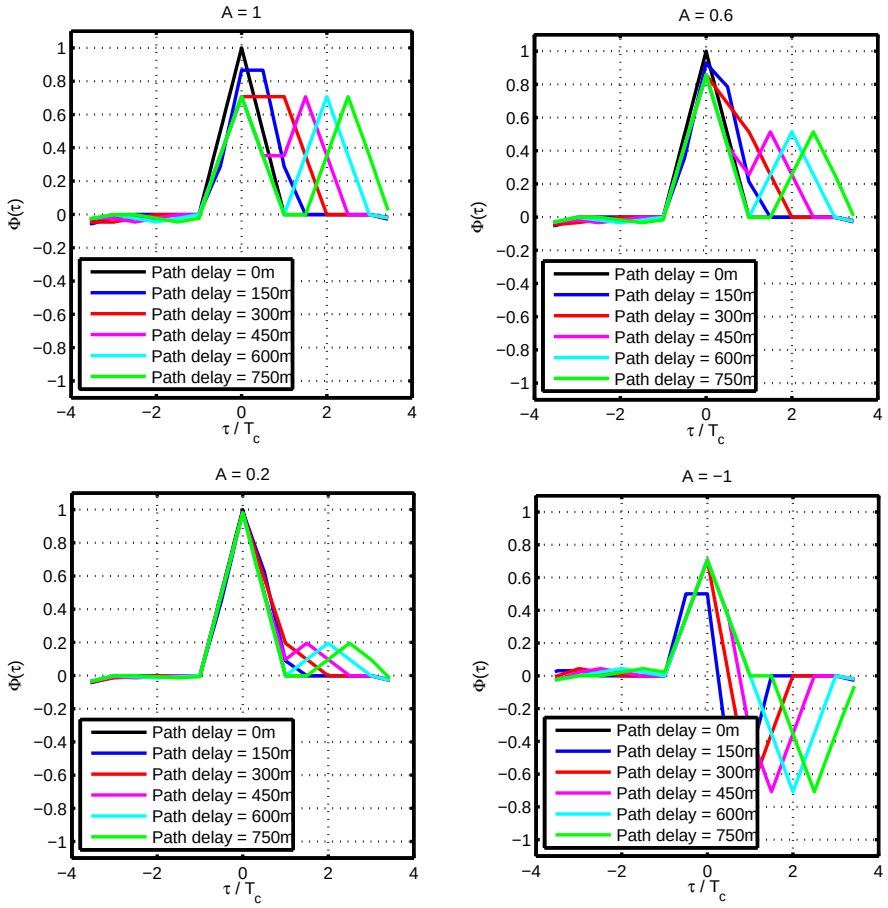


Figure 5.2: Cross correlation function distortion through multipath

to smoothen the carrier phase. Therefore, the frequency of the demodulator oscillator is fixed to an initially set value, and in order to have low phase corrections, a good estimate of the satellite Doppler from the acquisition stage is necessary. The filtering bandwidth has been set to different values between 1 and 50 Hz, where 20 Hz have resulted in the most flexible behavior. In general, a stationary receiver should be able to work with a much smaller PLL bandwidth, but since the generated signals have been parameterized to contain also high Doppler shifts with respect to those usually found in stationary received signals, the higher bandwidth has been chosen. Higher Doppler shifts shorten the minimum length of a generated signal that is needed to make the detection of the alternate paths possible. Signal generation is highly computational extensive in MATLAB, so shorter signals for testing purposes were preferred, which reduces also the needed space for data storage (one minute of samples equals to about one gigabyte of data). Although non-coherent tracking of the carrier phase has a slightly lower performance for small signal-to-noise ratios [ME01], it was decided that the performance degradation could be accepted in order to avoid programming the navigation processing, to remove data from the signal.

The DLL that tracks the code phase was implemented using the normalized early-minus-late envelope discriminator function [Kap96] which mitigates the effect of amplitude sensitivity. This discriminator function $D(e, l)$ is written as

$$D(e,l) = \frac{\sum \sqrt{I_E^2 + Q_E^2} - \sum \sqrt{I_L^2 + Q_L^2}}{\sum \sqrt{I_E^2 + Q_E^2} + \sum \sqrt{I_L^2 + Q_L^2}} \,, \tag{5.5}$$

and is evaluated against a threshold to initiate code shift in every measurement period. Again, non-coherent de-spreading is done here, because the data are not known. This approach squares also noise and thus degrades receiver performance, but for the present use case, this downside was accepted.

For the stationary receiver, the shifting distance was designed to be a constant value of one sample, which in fact limits the dynamics of the DLL to a code phase rate equivalent to one sample spacing per code sequence. However, under the assumption of a stationary receiver antenna, no higher pseudorange dynamics are to be expected. The correlator spacing can be set to different values, ranging from $d = 1.0$ to $d = \frac{1}{8}$, which is the lower limit for the chosen sampling rate.

Output Processing and Multipath Estimation

Of the receiver, only the tracking loop was implemented, which continuously follows code and carrier phase for a single satellite, as long as the signal can

be received strong enough. The recovery and post-processing of navigation data, and computation of a position solution from the data was not scope of the work. Instead, the loop outputs can be post-processed to obtain a Doppler spectrum of the satellite's received signal paths.

The tracking loop provides accumulated I and Q samples for each tracking period from the early, prompt, and late correlators. Additionally, the carrier PLL feedback $PNCO$ and the code phase are available. Since tracking is done code-word-coherently for every complete code sequence of 1 ms duration, 1000 datasets per second of input sampling data are available. The Doppler spectrum is determined as already described, from the prompt I and Q outputs, and the carrier feedback.

5.2 Estimation with Simulated Signals

5.2.1 Signals from MATLAB Signal Generator

The first simulated signals were used to validate the estimator algorithm, and were therefore generated without noise. Although the IF samples provide a noise-free signal, the resulting tracking characteristics contain noise, e.g. on the carrier phase. This is due to the quantization at a quite low resolution (4 bit), and the methods of discrete signal processing used in the tracking loop. This self-noise, however, is very low, so the phasor traces (see Figure 4.17) are clearly visible. The tracking loop was then modified such as to allow the ad-hoc addition of white gaussian noise to the input samples, where the noise power can be adjusted dynamically for each tracking run. With this configuration, it is possible to use the same generated signal twice for tracking, but change the SNR, in order to empirically determine performance limits of the tracking loop.

A noise-free example is given in Figure 5.3. Here, a satellite signal with two reflections has been used as a basis for the detection algorithm using the sub-optimal amplitude spectrum. The reflection parameters in this signal are: Path Dopplers 15Hz and 18Hz with respect to the LOS path, Amplitudes of the reflections $A_2 = 0.75A_1$ and $A_3 = 0.6A_1$; Path delays $\Delta\tau_2 = 0.5T_C; \Delta\tau_3 = 0.6T_c$. These values represent relatively strong and also fast changing reflections which are unlikely to occur in a real measurement on a stationary site. However, fast Dopplers decrease the necessary observation time, and thus the amount of simulated IF data samples that would have been mandatory to be created. By measuring longer, also Dopplers that are much lower can be detected, if the channel is stable enough during the measurement interval.

The plots in the first row show the reconstructed phase and the amplitude of the signal. The phase is determined from the phase feedback, and

the I/Q outputs by means of their quotient's inverse tangens, $\tan^{-1}(Q/I)$. The amplitude is the root of the squared sum of in-phase and quadrature output, $\sqrt{I^2+Q^2}$. In both plots, the double modulation caused by the two reflection components can be seen clearly. The second row shows the effect of normalization and filtering of the reconstructed phase. In this example, normalization has been done using a polynomial approximation which was subtracted from the phase signal. The polynomial is also hinted in the first plot as a dashed line. The filtering done in the fourth plot is a low pass filtering by means of windowing in the time domain with a hamming window. Its main objective is to remove noise as seen at the maximum values of the phase, which occur due to the very small amplitude at those times (compare plot 2 and 4).

In the third row, the phasor trace of the carrier, and the resulting amplitude spectrum are given. The phasor trace shows that, after removal of the DC part of the phase, the carrier phasor follows a path which is fully determined by the influence of the two reflections. It can be seen that for such high reflection amplitudes, the phasor spreads over the full I/Q plane, which would result in erroneous decoding of the data if this situation occured in a real measurement.

5.2.2 SPIRENT Simulator

Figure 5.4 is a plot of the Doppler spectrum resulting from three measurements done with the SPIRENT Simulator at DLR, Oberpfaffenhofen. The measurements are equivalent to measurements from real satellites, with the important difference that the CIR can be determined exactly and thus, the exact number, amplitudes and path delays for each echo are known. In the case of these measurements, a single reflection at a path Doppler of 0.1 Hertz has been applied to the LOS path with -3 dB and -12dB attenuation. A third measurement shows the Doppler spectrum without the influence of the reflection. Although the theoretically optimal complex Doppler spectrum has been used here, harmonics at multiples of 0.1 Hz can be seen both for the -3dB and -12dB spectra. Those harmonics are believed to result from the inertia of the phase feedback, which is low pass filtered and has therefore a different response depending on the direction of phase change when a common Doppler shift is superimposed to the phase distortion. This common shift results from the error in the initial Doppler estimation during acquisition, and makes the PLL steer the phase constantly into one direction to keep up with the slightly different frequency of the received signal. A superimposed oscillation of the phase due to multipath is then not filtered exacltly synchronously, and therefore the complex phasor sum does not perform a perfect circle, resulting in harmonics on the spectrum. Compared

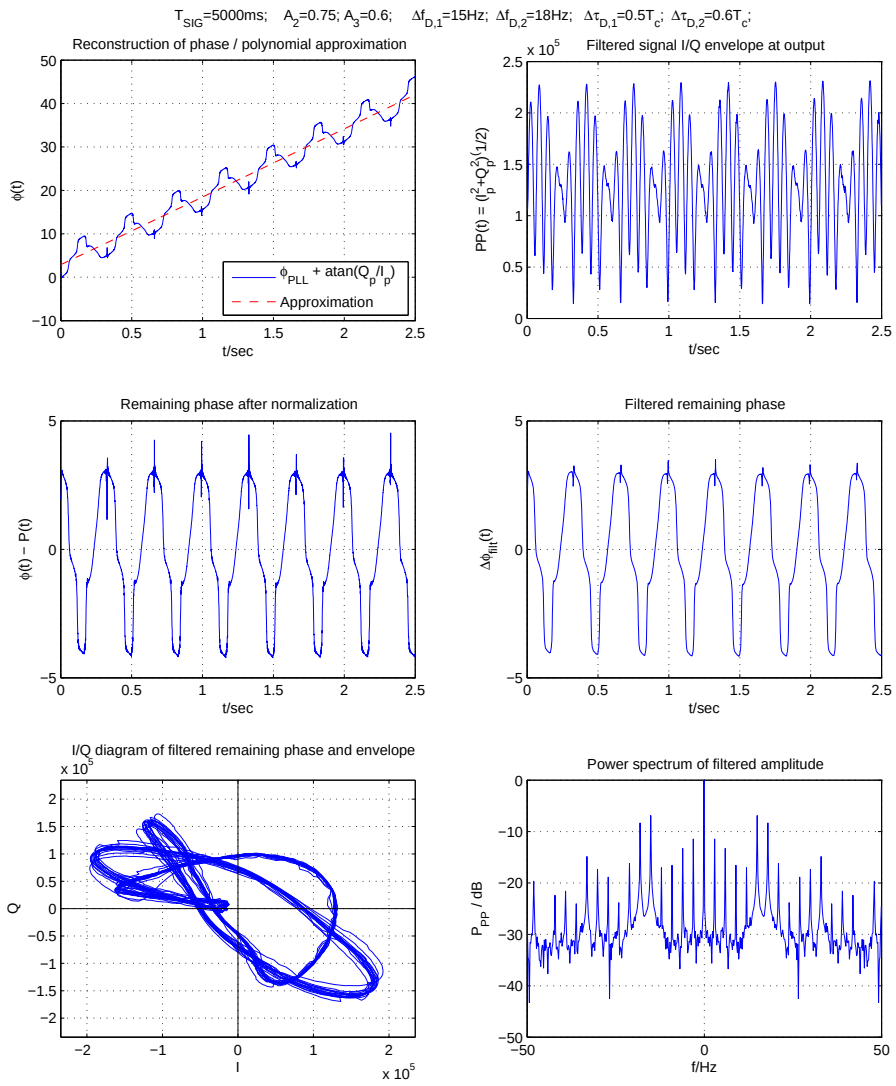


Figure 5.3: Detection Result from Generated Signal with 2 Reflections

with the harmonics from Figure 4.15, however, this disturbance is much weaker.

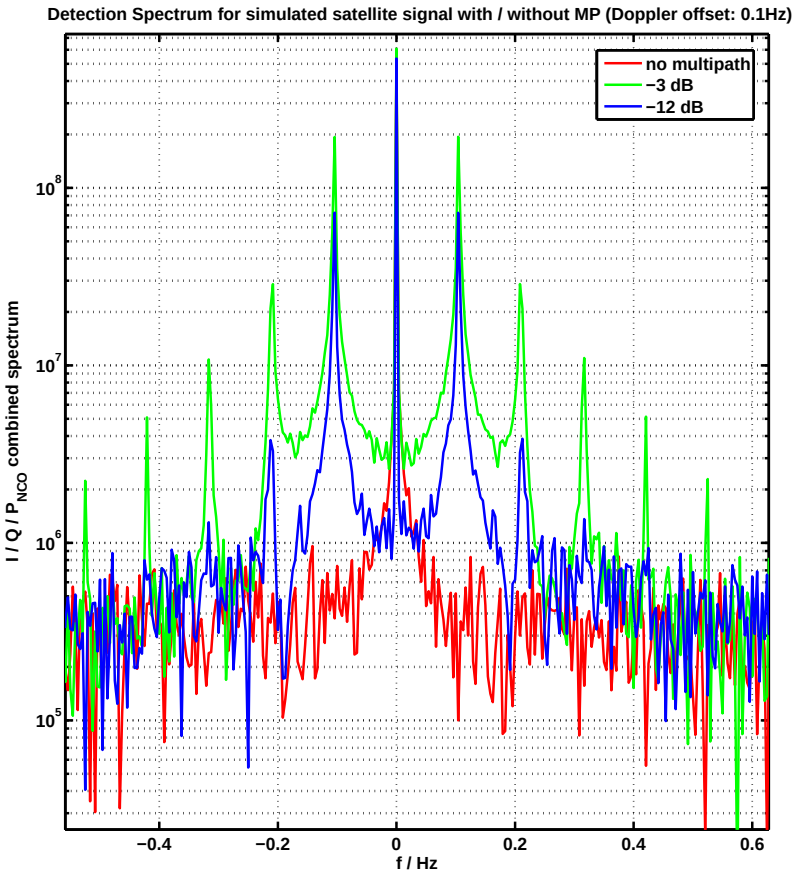


Figure 5.4: Complex Spectrum with and without Multipath

In Figure 5.5, a Doppler spectrum resulting from a Spirent-generated signal with a multitude of reflections has been analyzed. The symmetrical spectrum is shown here only in one half, to obtain a better readability. The path Dopplers used in these measurements were $\Delta f_k = [0.2; 0.85; 1.3; 1.5; 2.1]$Hz. In the plot, all these paths appear as distinct peak; however, one peak at about 1.9Hz is visible in the spectrum which could not be assigned to a generated echo. Because this signal contains the usual noise as it would be the case in a real measurement, the spectrum peaks are much less clear than in the analysis that was shown in Figure 5.3.

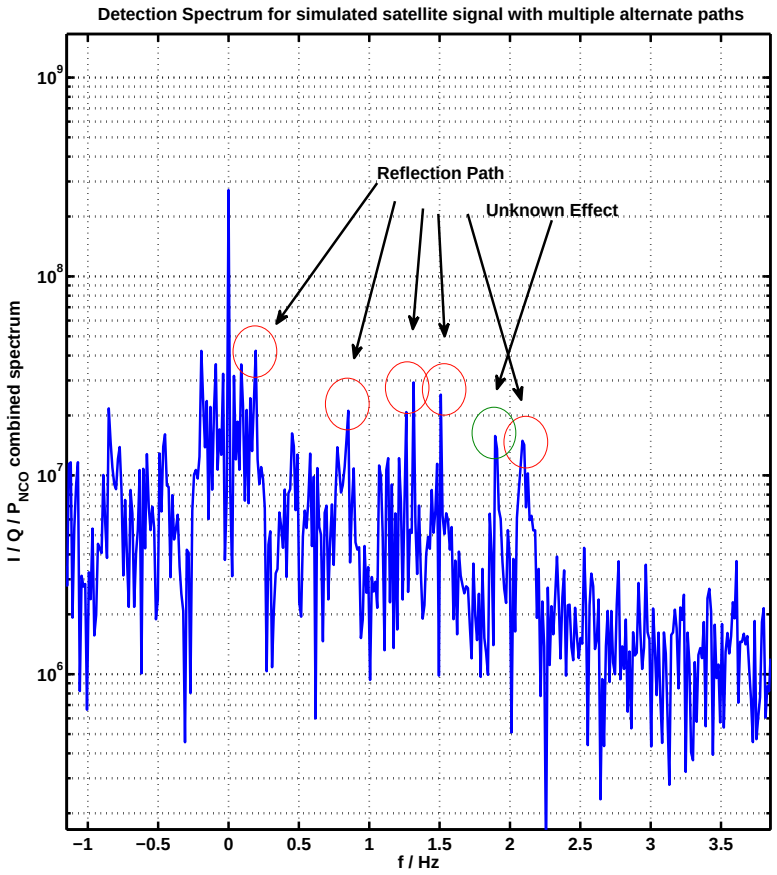


Figure 5.5: Spectrum with multiple reflections visible

5.3 Estimation of Measured GPS Signals

Although the estimation of a measured signal can only be validated if the channel impulse response, and therefore the parameters of all reflections, would be known, some measured data has been fed into the tracking loop as well, and applied to the post-processing stage which tries to detect multipath.

5.3.1 Measurements Description

The measurements were done in the city center of Munich, on the campus of Technische Universität München. The location of the receiver antenna was in front of the window, in the fourth floor of a four-storied building surrounded by similar buildings in close vicinity. The location had been chosen for its supposed multipath characteristics: A housefront directly vis-a-vis of the measurement site at a distance of about 15 meters made up most of the visible surroundings. Apart from that, some more distant reflecting structures besides the close housefront are likely to cause multipath. As discussed in Chapter 3, the distance of the reflector has a strong impact on the path Dopplers that can be expected. A large distance implies faster changing path lengths, and is therefore easier to detect for limited observation times.

5.3.2 Measurement Results

Using a program to monitor the satellite constellation in real time, and aerial pictures of the scenery, some careful estimates about expectable reflections could be made. With the NordNav Receiver running, it was possible to track also NLOS satellites in a few cases. The NLOS condition could be validated by comparing the satellite positition computed by the receiver with local obstacle configurations. For all satellites in view, attempts to track them with the NordNav Receiver have been made before the MATLAB tracking loop was used to generate the data for the Doppler spectrum. Figures 5.6, 5.7 and 5.8 show two determined Doppler spectra where peaks similar to those obtained with the Spirent simulator are visible. It it therefore possible that the signals of those three satellites contained multiple reflections. The satellite PRNs 16 and 25 had been seen in the sky at a position of $E = 20°; A = 45°$(PRN25) and $E = 14°; A = 53°$(PRN16), so they appeared very close to each other. Both these positions, measured from the north-eastern corner of the building in the 4th floor, have been free from obstacles and in direct view from the antenna position, so a direct component is believed to have been received in the measurement. Below the line of sight, a large metal roof with lots of metal structures like chimney encasings is situated at a receiver distance of about 70 meters (measures taken from

an aerial photograph), which could be the reason for the observed Doppler peaks.

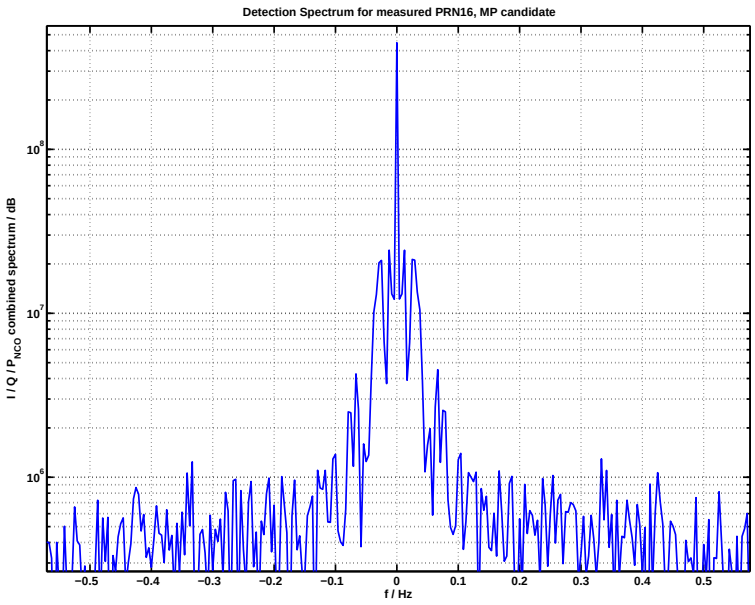


Figure 5.6: Measured PRN 16 with strong Doppler peaks

Satellite PRN 23 at $E = 47°; A = 64°$ was seen under a significantly higher elevation, and from a slightly different azimuth angle. The scenery for this satellite is not much different to the one of SVs 16 and 25; however the higher elevation results probably in completely different reflectors, e.g. the ground, which would have been a street in this case. The resulting spectrum contains a lot of peaks between zero and 0.2 Hz.

The peaks in the spectra are altogether in a frequency range below 0.2 Hertz, e.g. 0.03 Hz in Figure 5.6. The corresponding path delay would have changed at a rate of about 6 millimeters per second, which is a credible value when compared with the typical path delay change rates derived in Section 3.2. However, it is not possible to clearly assign this peak to a multipath effect, or even to a specific, observable reflection because we don't know the channel characteristics of this measurement.

Figure 5.9 has a spectrum where no strong peaks are visible. This satellite - PRN 20 at $E = 17°; A = 120°$ - was also clearly visible without any obstruction on the direct path. However, because its azimuth made it appear directly above the rooftop of the opposite housefront, it is very unlikely

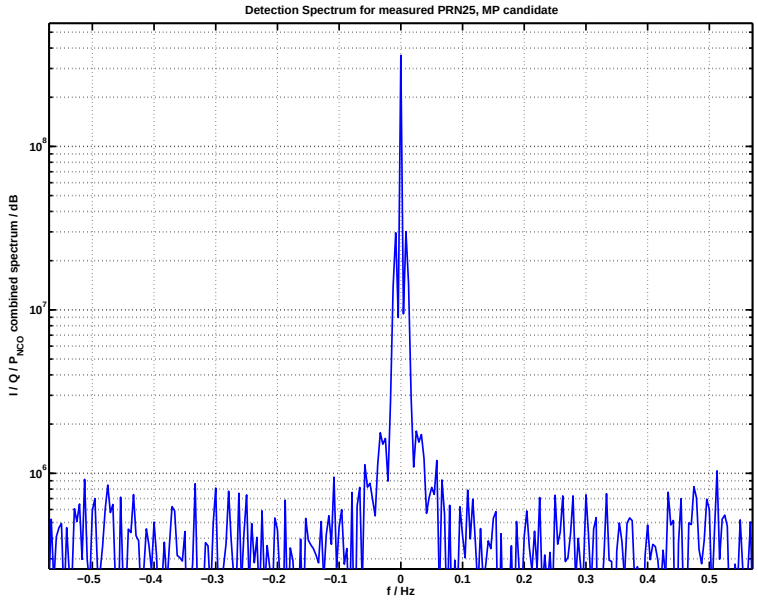


Figure 5.7: Measured PRN 25 with strong Doppler peaks

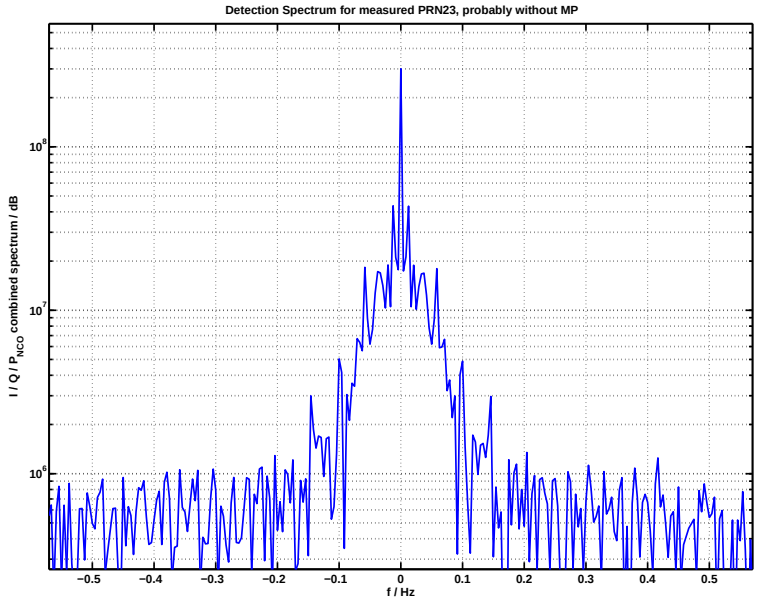


Figure 5.8: Measured PRN 23 with strong Doppler peaks

that a reflection from around the LOS axis occured here; however reflections from the side would be still possible.

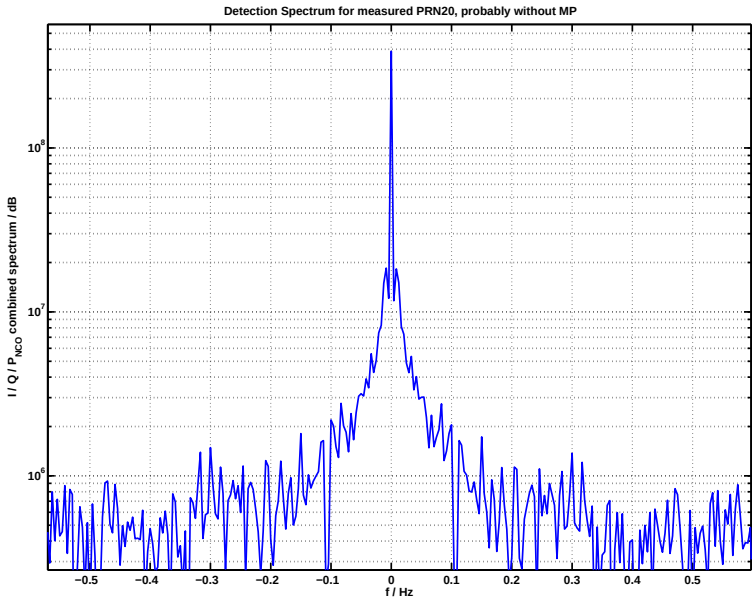


Figure 5.9: Measured PRN 20, no strong Doppler peaks visible

Although no definite statement about the existence or non-existence of multipath effect can be made in this case, the generated Doppler spectra from the measurements show at least that, although all of the satellites were clearly visible and at similar positions from the user point of view, their spectra show significant differences - three of them contain distinguishable peaks in the Doppler spectrum, whereas one satellite has a very smooth spectrum with only small noise components. Considering that all the analyses have been made from the same set of IF data, and therefore from the same measurement with equal equipment and under similar geometry, this gives at least a hint that the Doppler spectrum can provide information about existing reflections also in reality.

Chapter 6

Conclusion and Future Work

The Diploma Thesis showed that it is possible to detect individual reflection paths from a GPS signal by analyzing the spectrum of a reconstructed complex carrier phasor. The Doppler shift differences between the paths, which are the basis for this detection, are determined by the reflection scenario, which was analyzed in detail for the case of a ground reflection and a reflection at a housefront or a similar reflector. Typical Doppler shifts have been determined by simulating the satellite constellation in combination with typical reflector distances and positions.

The Doppler spectrum - which can be obtained easily by combination of already present receiver output signals - provides an estimate on the lower bound of reflection paths. This was validated in simulations, where the GNSS signal as received on the ground was reproduced completely, including noise and errors from the well-known error sources. The simulation provided a means to test the reliability of the information contained in the Doppler spectrum, because the channel impulse response, in particular the number and parameters of the echoes, were known for the generated signal.

To obtain a spectrum at a reasonably high resolution, very long observation times are necessary. As a rule of thumb, any path Doppler can only be seen in the spectrum if the observation lasts at least as long as the Doppler's period. But to be able to distinguish between paths with very similar frequencies, an even longer measurement must be taken. With the simulator, the channel stability allowed reconstruction of paths which had a Doppler of only 0.2 Hz, but investigations on the real-world GNSS channel show that the channel changes very fast, also for a stationary user [SL04]. Since path Dopplers of very near reflections are especially low, those reflections which are almost fully processed in the correlation stage are still a major threat.

The technique does not provide a means to remove the influence of the detected multipath - it can not estimate its path delay, but only the change rate of the path delay and the path amplitude. Still, it can be useful where

long term measurements are taken, and it is desireable to exclude measurements with strong multipath from further processing. Examples for this may be a GBAS reference station where the particular satellite could be processed using a different antenna, if multipath was detected. In surveying, the rover could exclude the measurement if enough satellites are available, or trigger different smoothing techniques that cope with the determined periodic properties of the code and carrier phase distortions.

Another use case could be the assessment of possible antenna locations with respect to multipath occurence. If several locations for a fixed reference station are possible, the multipath reception characteristics of each site could be evaluated, and the best location can be chosen.

Proposals for Future Work

Validation with measured scenarios

Further work on the topic might include more real-life measurements, where it could be tried to estimate the channel characteristics with respect to individual reflections prior to measuring. Especially, the stability of the channel during the measurement is still unclear.

Adaptive filtering

If the frequency compounds of the distortive influence of multipath can be estimated correctly, the PLL filter can be adjusted to mitigate the phase distortion immediately.

Mutual influence of DLL and PLL tracking

Because the PLL and DLL tracking is done numerically with discrete values, both code phase and carrier phase estimates can necessarily not always be equal to the values of the received signal. In particular, in the present implementation, the DLL code phase has a granularity of only about $1/16$ chip length, because the reference code sequence was generated at the same sampling rate than the received samples. Therefore, the code phase estimate oscillates around the true code phase constantly, which also influences the I and Q outputs. The carrier phase distortion again influences the DLL, because the demodulation of the true LOS component is partially inhibited when the phase error is high. The mutual influence of both loops in the common system could be investigated closer to help understanding various side-effects that occured during the analysis done in this work.

Cartographic application

With parameters about the satellite position and movement given, detected Doppler shifts between paths can be used to estimate the position and distance of a reflector. If the measurements are done long enough, satellites from many directions can be included into the analysis, so that a "reflector cartography" of the receiver antenna neighborhood can be made. With this information, a priori knowledge about the channel for a satellite at a given position could be collected, which can then be used in signal processing for a reference station application. This technique would require a "training phase" for a new reference site, and would return better results when the station is in regular operation.

Chapter 7

Bibliography

[Cos56] John P Costas. Synchronous communications. In *Proceedings of the IRE*. IEEE, 1956.

[Gü05] Christoph Günther. Lecture "satellite navigation", tu münchen. 2005.

[Hag03] Joachim Hagenauer. Mobile communications. Lecture Notes, 2003.

[HGI+] Guenter W Hein, Jeremie Godet, Jean-Luc Issler, Jean-Christophe Martin, Philippe Erhard, Rafael Lucas-Rodriguez, and Tony Pratt. Status of galileo frequency and signal design. Galileo Signal Task Force of the European Commission, Brussels.

[JMTU98] Fredrik Johansson, Rahman Mollaei, Jonas Thor, and Jörgen Uusitalo. Gps satellite signal acquisition and tracking. Master's thesis, Luleå University of Technology, 1998.

[Kap96] Eliott D Kaplan. *Understanding GPS - Principles and Applications*. Artech House Publishers, 1996.

[Lan02] Manfred Lang. Signaldarstellung. Lecture Notes, Institute for Human-Machine Communication, TU München, 2002.

[LS05] Andreas Lehner and Alexander Steingass. A novel channel model for land mobile satellite navigation. In *GNSS 2005 Proceedings*, 2005.

[Lyo04] Richard G Lyons. *Understanding Digital Signal Processing*. Prentice Hall, 2004.

[ME01] Pratap Misra and Per Enge. *Global Positioning System: Signals, Measurements and Performance*. Ganga-Jamuna Press, 2001. ISBN: 0970954409.

[ns04] No names specified. Is-gps-200 rev d - space segment/navigation user interfaces. Technical report, 2004.

[PS96] Bradford W Parkinson and James J Spilker, editors. *Global Positioning System: Theory & Applications (Volume One) (Progress in Astronautics and Aeronautics)*. AIAA (American Institute of Aeronautics & Ast, 1st edition, 1996. ISBN: 156347106X.

[Psi01] Mark L Psiaki. Block acquisition of weak gps signals in a software receiver. In *ION GPS 91*. Cornell University, 2001.

[RW00] Lennart Rade and Bertil Westergren. *Springers Mathematische Formeln. Taschenbuch für Ingenieure, Naturwissenschaftler, Informatiker, Wirtschaftswissenschaftler. 3., durchges. Aufl.* Springer, Berlin, 2000.

[SL04] Alexander Steingass and Andreas Lehner. Measuring the navigation multipath channel - a statistical analysis. In *GNSS 2004 Proceedings*, 2004.

[vDFF92] A. J van Dierendonck, Pat Fenton, and Tom Ford. Theory and performance of narrow correlator spacing in a gps receiver. In *Navigation: The Journal of The Institute of Navigation*, volume 39. ION, 1992.