Chapter 6

Pseudorange Estimation

The receiver position and its clock offset were determined from pseudoranges in Chapter 3. Estimating the pseudoranges by analyzing the received signals is the aim of the present chapter. This also requires that the Doppler shift be estimated. Section 6.1 parameterizes the signals in terms the propagation delay, Doppler-shift, and phase-offset and describes the channel model. The parameters mentioned represent a sufficient statistic for position estimation, which is exploited in Section 6.2 to derive the basic equation for estimating the pseudorange, the Doppler-shift, and the carrier phase offset. In the remaining sections we develop the determination of an initial estimate of these quantities and their subsequent tracking.

6.1 Channel Model and Received Signal

A channel model describes the modifications of the signal caused by propagation. This also includes modifications due to relative movement or differences in gravity at the location of the transmitter and receiver. In global navigation satellite systems, the interest is focussed on the delay caused by propagation. In TRANSIT and Tzikada, the interest was on the Doppler-shift. The latter is also present in GNS systems, but plays a secondary role in the determination of the position .

The signals from Chapter 5 have a common structure, described by Equation (5.18). They are the product of a carrier signal, a spreading signal and a navigation message:

$$s(t;\omega_c,\tau_0,\varphi_0,b) = \sqrt{2\mathcal{P}_0}\, b(t-\tau_0)\, c(t-\tau_0)\cos(\omega_0(t-\tau_0)+\varphi_0). \tag{6.1}$$

This equation includes a reference time τ_0. This time reference, as well as the carrier frequency ω_0 are derived from the satellite's local extremely stable atomic reference[1]. The variable ϕ_0 denotes the phase, and $\mathcal{P}_0$ the transmit power. The navigation signal $b(.)$ and the spreading signal $b(.)$ are described the Equations (5.17) and (5.19). All quantities depend on a satellite index, omitted for simplicity. Equation (6.1) describes the signal in the product representation. This will simplify the computation of correlations in the Section 6.3. The normalization $\sqrt{2\mathcal{P}_0}$ is related to the chip energy $\mathcal{E}_c$, by integrating the signal power over one chip period T_c :

$$\mathcal{E}_c = \int_0^{T_c} dt\, s^2(t;\omega_0,\tau_0,\varphi_0,b) = \mathcal{P}_0 T_c.$$

The signal $s(.)$ from Equation (6.1) propagates towards the receiver via a channel, typically described by a delay $\Delta\tau = \tau - \tau_0$, a phase rotation $\Delta\varphi = \varphi - \varphi_0$, and an attenuation $a = \sqrt{\mathcal{P}/\mathcal{P}_0}$, with $\mathcal{P}$ being the received power. Furthermore, the relative movement of the transmitter and

[1]The atomic clock is a extremely stable but still noisy and not perfect. Its behavior is described in a clock model. Parameters to correct the deterministic part of that model are provided in the navigation message.

© 2008 by Christoph Günther, all rights reserved.

receiver causes a Doppler shift $\Delta\omega = \omega - \omega_0$. The signal received from K satellites is thus described by

$$r(t) = \sum_{k=1}^{K} a^k s(t; \omega^k, \tau^k, \varphi^k, b^k) + n(t), \tag{6.2}$$

with $n(.)$ being zero-mean AWG noise with a covariance:

$$\mathcal{E}[n(t)n(t')] = \frac{\mathcal{N}_0}{2}\delta(t - t'). \tag{6.3}$$

If τ^k was measured with a clock synchronized to the clock of the k-th satellite, the difference $c\Delta\tau$ would simply be the range. Since both time offsets are measured using local clocks, the difference is related to the pseudorange. The time offsets τ^k and τ_0^k occur in an essential manner in two places: in the code signal and in the carrier signal. In an optimal receiver both components are used to estimate $\Delta\tau$. In most receivers this is not the case. Consumer receivers only use the code signal. In this case, the measurement provides the code pseudorange:

$$\tilde{\rho}^k = c\Delta\tau^k + \eta^k, \tag{6.4}$$

with η^k being the noise due to clocks, propagation, and the estimation of τ^k using the code component only of the noisy received signal $r(.)$. More sophisticated receiver also use the carrier phase, in which case the more accurate carrier phase dominates the accuracy of the estimate. In the absence of a Doppler shift and of a phase offset, the difference in the carrier phase of the received signal relates to $\Delta\tau$ as follows (see Equation (6.13)):

$$\lambda\tilde{\phi}^k = \left(c\Delta\tau^k + \epsilon^k\right) \bmod \lambda, \tag{6.5}$$

with ϵ^k being the noise based on estimating $\Delta\tau$ using the carrier phase signal only. The minus sign inherent in this definition is a convention. Unfortunately, this expression only provides the pseudorange modulo the wave length λ. This is due to the periodicity of the carrier phase signal. The resolution of the assocaited ambiguity is fundamental for the determination the pseudorange using carrier phase measurements. It is addressed in Chapter 16. The phase signal can also be used to determine the Doppler shift $\Delta\omega$. The latter is needed in the estimation process, as shall be seen below.

The AWG noise in Equation (6.2) describes noise components due to tropospheric and regular ionospheric propagation as well as diffuse multipath and mild forms of interference. In the presence of specular multipath, more general models need to be considered, e.g.:

$$r(t) = \sqrt{2\,\mathcal{P}} \sum_{k=1}^{K} \sum_{i=0}^{\ell-1} a_i^k s(t; \omega_i^k, \tau_i^k, \varphi_i^k, b^k) + n(t), \tag{6.6}$$

with i being the path index. Each path might have a different delay, phase, Doppler-shift, and attenuation. The attenuation is mostly due to fading, and often described as a Rice-distributed random variable. In the presence of a line of sight signal, the latter one corresponds to the path with the least delay. Sometimes the line of sight to one or several satellites might be obstructed, although reflections are still received. In such scenarios, advanced receivers are needed but not yet used. The situation is rather mitigated using tricks, like road matching, i.e. the property that cars stay on roads. Similar models to Equation (6.6) are also used in the presence of ionospheric scintillation. Strong interference finally requires additional components. The latter components significantly depend on the nature of the interference - a discussion is beyond the scope of this book.

© 2008 by Christoph Günther, all rights reserved.

6.2 Bayesian Formulation of the Estimation Problem

The detailed derivation will be included in a later version.

With the channel model described by Equation (6.2), finding the value of ρ_t that maximizes the a posteriori probability can thus be seen as finding the components of

$$\theta_t = (\omega_t, \tau_t, \varphi_t, b_t(.)),$$

(and a) for all satellites, which maximize

$$p(\theta_t | r_t(.)).$$

For simplicity, we will replace $r_t(.)$ by r, and also drop the index t on all other variables. In the absence of a priori information, this maximization is equivalent to the maximization of the likelihood function $p(r|\theta)$ or of its logarithm. This means determining the solution θ to the equation:

$$\frac{\partial \log p(r|\theta)}{\partial \theta} = 0.$$

This solution can be determined using a sampled version of the bandpass signal r, see Appendix B to Chapter 5. Assuming that the bandwidth is B, the minimum sampling rate is given by[2] $1/T_s = B$. The sampled version of Equation (6.2) becomes

$$r_i = as(iT_s, ; \theta) + n_i,$$

with a noise covariance

$$\mathcal{E}[n_i n_{i'}] = \frac{\mathcal{N}_0}{2} \frac{\delta_{i,i'}}{T_s}.$$

Since the variables n_i are independently Gaussian distributed, this implies that

$$p(r|\theta) = \prod_i \frac{1}{\sqrt{\pi \mathcal{N}_0 / T_s}} e^{-\frac{T_s}{\mathcal{N}_0} \left(r_i - \sum_k a^k s(iT_s; \theta^k) \right)^2}.$$

The log-likelihood then becomes

$$\log p(r|\theta) = -\frac{T_s}{\mathcal{N}_0} \sum_i \left(r_i - \sum_k a^k s(iT_s; \theta^k) \right)^2 + \text{const.}$$

and the maximization of the log-likelihood is found to be equivalent to the minimization of the quadratic error,

$$T_s \sum_i \left(r_i - \sum_k a^k s(iT_s; \theta^k) \right)^2 \tag{6.7}$$

which can again be expressed in continuous notations:

$$\hat{\theta} = \arg\min_\theta \int dt \left(r(t) - \sum_k a^k s(t; \theta^k) \right)^2. \tag{6.8}$$

In a strict sense, the equivalence of the latter two expressions is a consequence of the orthogonality[3] of:

$$\int_{-\infty}^{\infty} dt \, \text{sinc}(\pi(t-i)) \, \text{sinc}(\pi(t-j)) = \delta_{ij}. \tag{6.9}$$

[2] The Nyquist rate is also expressed in terms of the maximum frequency F present in the Fourier transform of the baseband signal. In this case $1/T_s = 2F$.

[3] This is the orthogonality of two non-overlapping rectangular pulses, expressed in the Fourier transformed domain.

© 2008 by Christoph Günther, all rights reserved.

The quadratic norm of $\sum_k a^k s(t;\theta^k)$ does not much depend on θ^k, since shifting in phase, time and frequency does not change the energy in the signal for signals with low crosscorrelations. Thus the minimization of the quadratic error is equivalent to the maximization of the correlation

$$\hat{\theta} = \arg\max_\theta \int dt\, r(t) \sum_k a^k s(t;\theta^k), \tag{6.10}$$

which due to the low crosscorrelations can be solved satellite by satellite

$$\hat{\theta^k} = \arg\max_{\theta^k} \int dt\, r(t) s(t;\theta^k). \tag{6.11}$$

The limits of these integrals were left unspecified. In a completely static environment, they would extend from $-\infty$ to t. This would provide the most accurate position estimates, and would correspond to an infinite signal-to-noise ratio. In real situations, the integral is limited to periods of time that are short enough for neglecting the relative movement of the user and satellites. The duration of the integration typically ranges from 1 ms in large classes of receivers to 100 ms in consumer products with indoor positioning capabilities. This is sufficient in most cases. In environments with extreme dynamics, as encountered in some military scenarios, inertial systems are used to account for changes in the movement of the receivers during the correlation.

Immediately after switching on, the receiver has no knowledge about the signal. In this phase, the receiver proceeds exactly, as described by Equation (6.10). The search is somewhat simplified by the very low data rate of the navigation message, and is further supported by the presence of a pilot signal in Galileo. Furthermore, the shape of the correlation, as derived in Section 6.3, allows to remove the phase by taking the absolute value. Thus the actual search reduces to a two dimensional joint search for the delay and frequency-offset. This is described in Section 6.4.

After initial acquisition, the receiver has prior knowledge about the signal. The optimal receiver should thus be derived from the maximum a posteriori probability principle. The present maximum likelihood is suboptimum but more tractable. The generic maximum likelihood tracking of parameters, under the assumption that some adequate a priori information is known, is developed in Section 6.5. It is applied to the tracking of delay τ in a Delay Locked Loop (DLL), of frequency ω in a Frequency Locked Loop (FLL), and potentially of phase ϕ in a Phase Locked Loop (PLL) in the subsequent sections. These loops are independent of each other, and separate for each satellite and signal. This is typically sufficient[4].

As mentioned the low bit-rate of the navigation message allows to eliminate the modulation by the navigation message. The latter is most disturbing during initial acquisition. Different techniques have been developed. They include short integration times, e.g. 1 ms, as well as the use of aiding data transmitted via mobile radio networks. However, even in the case of an integration over 10 ms, the sign changes due to the navigation bits affect at most[5] every 2nd integration in GPS. A very simple scheme thus is to perform a search on the odd and even correlation intervals in parallel. The only inconvenience is that this approach doubles the acquisition time. Since this method allows to somehow handle the data sequence $b(.)$ the navigation message will not be considered any further in the estimation of the other parameters below.

6.3 Computing Correlations

Without data modulation, the transmitted waveform becomes:

$$s(t) = \sqrt{2\mathcal{P}_0}\, c(t-\tau_0) \cos(\omega_0(t-\tau_0) + \varphi_0).$$

[4] In some situation, one has to resort to coupled loops, as obtained in the more general models cited above. Examples of such situations include maneuvering airplanes, that loose all satellites to the right during a left turn, and all satellites to the left in the following right turn, or car driving under trees that see a constantly changing set of satellites.

[5] A sign change only occurs if two consecutive navigation bits are different, which happens with probability 1/2.

© 2008 by Christoph Günther, all rights reserved.

Similarly the signal received from one single satellite over a single path channel becomes

$$r(t) = \sqrt{2\,\mathcal{P}}\, c_D(t-\tau)\cos\big(\omega(t-\tau)+\varphi\big) + n(t),$$

with $c_D(.)$ being the chip signal as modified by the relative movement of the receiver and the satellite, and $n(.)$ being AWG noise with a covariance given by Equation (6.3). In typical situations, the Doppler shift of the code is small and can be neglected for all conceivable durations of the integration[6]. Therefore the distinction between $c_D(.)$ and $c(.)$ does not need to be made below.

In a real receiver the incoming signal is first shifted in frequency to eliminate the underlying carrier mostly (down conversion). Subsequently the correlation is performed on the sampled signal in the baseband. This is mathematically equivalent to the computation of the correlation, as described in the previous section.

The down conversion and I/Q demodulation is performed using the estimates ω_r, τ_r, ϕ_r of the frequency, delay, and phase respectively, to generate a local carrier replica $\cos\big(\omega_r(t-\tau_r)+\varphi_r\big)$. The values of ω_r, τ_r, ϕ_r are all referenced with respect to the receiver's "master clock". Multiplication with the above replica exposes the in-phase baseband component $x(t)$. Similarly $\sin\big(\omega_r(t-\tau_r)+\varphi_r\big)$, is used to extract the quadrature baseband component $y(t)$. During initialization a subset of values of ω_r and τ_r are searched, later these values are continuously tracked in loops. The baseband signal then becomes:

$$\begin{aligned}\begin{pmatrix} x \\ y \end{pmatrix}(t) &= 2\,r(t)\begin{pmatrix} \cos \\ \sin \end{pmatrix}(\omega_r(t-\tau_r)+\varphi_r) \\ &= \sqrt{2\,\mathcal{P}}\, c(t-\tau)\begin{pmatrix} \cos \\ \sin \end{pmatrix}(\Delta\omega(t-\tau_r) - \omega\Delta\tau + \Delta\varphi) + \begin{pmatrix} n_x \\ n_y \end{pmatrix}(t),\end{aligned}$$

with

$$\Delta\omega = \omega - \omega_r, \quad \Delta\tau = \tau - \tau_r \quad \text{and} \quad \Delta\varphi = \varphi - \varphi_r.$$

In order to perform the correlation, the result is multiplied by a reference waveform $c'(.)$ and integrated over the interval $[mT_i + \tau_r, (m+1)T_i + \tau_r)$ of duration $T_i = \kappa N T_c$, with κ being the number of periods of the spreading code used during the integration. Typical values of κ range from 1 to 100. The beginning of the interval was chosen to coincide with the locally generated reference signal

$$c'(t-\tau_r).$$

The code sequence c' might either coincide with c or be different. Different codes $c' \neq c$ imply that the signal comes from another satellite than the desired one. In this case, the Doppler shift will be substantial $\Delta\omega T_i \gg 1$, with a very high probability. The phase change induced by such a frequency offset makes the sequence of chips behave like a random sequence over the whole integration interval, which implies that the normalized correlation is $\mathrm{O}(1/\sqrt{\kappa N})$. We will thus focus on the case $c' = c$.

The result of the correlation is denoted by

$$\begin{pmatrix} C_x \\ C_y \end{pmatrix}_m,$$

and is a function of $(\Delta\omega, \Delta\tau, \Delta\varphi)$. Consider the x-component:

$$C_{x,m} = \sqrt{2\,\mathcal{P}} \int_{(m-1/2)T_i-\tau_r}^{(m+1/2)T_i-\tau_r} dt\, c(t-\tau_r)\, c(t-\tau)\cos\big(\Delta\omega(t-\tau_r) - \omega\Delta\tau + \Delta\varphi\big) + n_{x,m},$$

shifting the integration interval implies:

$$C_{x,m} = \sqrt{2\,\mathcal{P}} \int_{(m-1/2)T_i}^{(m+1/2)T_i} dt\, c(t)\, c(t-\Delta\tau)\cos\big(\Delta\omega t - \omega\Delta\tau + \Delta\varphi\big) + n_{x,m}. \tag{6.12}$$

[6]The worst case value of v/c for GPS is $3.1 \cdot 10^{-6}$, additionally, the Doppler shift of the code is corrected automatically whenever the receiver couples the DLL to either the FLL or the PLL.

© 2008 by Christoph Günther, all rights reserved.

A corresponding expression is obtained for the quadrature component $C_{y,m}$ by replacing the cosine function by the sine function. The same shift in the integration variable can also be performed in the noise component:

$$n_{x,m} = 2 \int_{(m-1/2)T_i}^{(m+1/2)T_i} dt\, n(t+\tau_r)c(t)\cos(\omega_r t + \varphi_r).$$

A corresponding expression is obtained for $n_{y,m}$. Both components stay zero-mean AWG and have a covariance:

$$\mathcal{E}[n_{x,m}n_{x,m'}] = \mathcal{E}[n_{y,m}n_{y,m'}] = \mathcal{N}_0 T_i \delta_{m,m'}, \qquad \mathcal{E}[n_{x,m}n_{y,m'}] = 0.$$

The in-phase and quadrature components are often combined using complex notations:

$$C_{z,m} = C_{x,m} + jC_{y,m}, \quad \text{and} \quad n_{z,m} = n_{x,m} + j\, n_{y,m},$$

with the covariance of the associated noise becoming:

$$\mathcal{E}[n_{z,m}n^*_{z,m'}] = 2\mathcal{N}_0 T_i \delta_{m,m'}, \qquad \mathcal{E}[n_{z,m}n_{z,m'}] = 0.$$

The resulting complex correlations can be computed in closed form in the case of no delay error $\Delta\tau = 0$ or of no frequency error $\Delta\omega = 0$. The general case is discussed in Appendix A, and leads to:

$$C_{z,m} = \alpha \left\{ R(\Delta\tau)\,\mathrm{sinc}\left(\frac{\Delta\omega T_i}{2}\right) e^{j(m\Delta\omega T_i - \omega\Delta\tau + \Delta\varphi)} + \mathrm{O}\left(\frac{1}{N^\beta}\right) + \mathrm{O}\left(\frac{\Delta\omega T_i}{\kappa}\right) \right\} + n_{z,m}. \quad (6.13)$$

In this expression, the following definitions were used:

$$\alpha = T_i\sqrt{2\mathcal{P}},$$

and

$$R(\Delta\tau) = \begin{cases} 1 - \dfrac{|\Delta\tau|}{T_c} & \text{if } |\Delta\tau| \le T_c \\ 0 & \text{otherwise.} \end{cases} \quad (6.14)$$

$R(\Delta\tau)$ is the autocorrelation function of a rectangular pulse. The term $\mathrm{O}(1/N^\beta) + \mathrm{O}(\Delta\omega T_i/\kappa)$ in Equation (6.13) will not be written out in subsequent sections. The exponent β is 1/2 for all reasonable spreading codes. In practice, the conditions N large, and $\Delta\omega T_i/\kappa$ small are usually met. In this case, the higher order terms can be omitted.

The function $R(\Delta\tau)$ is shown in the left side graph of Figure 6.1. The corresponding autocorrelation for a BOC(m,n) modulation is shown in the right side graph of the same figure. The shape of the latter autocorrelation can be understood as follows: for integer ratios m/n, each chip contains a full number of subarrier periods. Starting from the in-phase value, extrema of the correlation function occur when the reference code is shifted by a full tooth (half a period). The values of the extrema have alternating signs, and are identical with the corresponding values of the correlation of the rectangular waveform, since either all signs are equal or all opposite. The changes between the maxima are linear. This explains the correlation seen in Figure 6.1. The argument is valid for arbitrary integer m, and n. With $2m/n$ being the number of half-periods of the subcarrier contained in one chip, there is a total of $4m/n-1$ local extrema. In real-systems front-end filtering furthermore rounds off the corners of the rectangular pulse shape, and correspondingly smooths out the sharp edges of the correlation functions.

Equation (6.13) is the basis for acquisition and for all control loops. $R(\Delta\tau)$ is maximum when $\Delta\tau = 0$. The triangular shape of $R(\Delta\tau)$ is fundamental both for the initial acquisition and the tracking using a Delay Locked Loop (DLL). The maximum of the sinc-function is obtained for $\Delta\omega =$

© 2008 by Christoph Günther, all rights reserved.

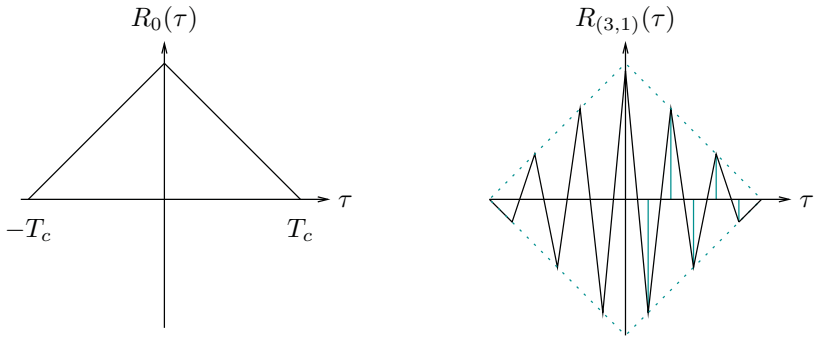

Figure 6.1: Correlation of the rectangular waveform $R_0(\tau)$ (left) and of a BOC(3,1) waveform $R_{(3,1)}(\tau)$ right.

0. The sinc-factor describes the degradation during initial acquisition. It is never used for estimating $\Delta\omega$. The latter estimate, as well as the estimate for frequency are obtained by maximizing the real part of the exponential function. A local maximum is obtained for $\Delta\tau = 0$, $\Delta\omega = 0$, and $\Delta\phi = 0$. The phase error $\Delta\phi$ only needs to be known modulo 2π. The frequency estimation is ambiguous, it is only determined modulo $2\pi/T_i$. This is resolved during initial acquisition. The phase pseudorange is extremely accurate, down to millimeter level. It is also ambiguous. The associated discussion is deferred until Chapter 16. The phase dependency on frequency, and phase are used in the Frequency Locked Loop (FLL) and Phase Locked Loop (PLL).

The signal to noise ratio of the correlation described by Equation (6.13) is given by

$$\mathrm{SNR} = \frac{|\mathcal{E}[C_{z,m}(\Delta\omega, \Delta\tau, \Delta\varphi)]|^2}{\mathrm{var}[C_{z,m}(\Delta\omega, \Delta\tau, \Delta\varphi)]} = \frac{\mathcal{P}T_i}{\mathcal{N}_0} R^2(\Delta\tau)\,\mathrm{sinc}^2\left(\frac{\Delta\omega T_i}{2}\right), \tag{6.15}$$

with $\mathcal{P}T_i$ being the energy accumulated during the correlation process. For zero frequency error $\Delta\omega = 0$, this expression is linear in the integration time, and one might choose a large integration time if the dynamics of the movement allows for that. For $\Delta\omega \neq 0$ the signal to noise ratio is degraded by the sinc-term, and an optimum value of T_i must be chosen. Methods for increasing the signal to noise ratio even in such cases are shortly addressed in Section 6.4.2.

6.4 Initial Acquisition

Equation (6.15), in the previous section, shows the degradation of the signal to noise ratio caused by an error in the delay or frequency error. If the delay error is larger than a chip or the frequency error $\Delta\omega$ larger than π/T_i, the signal to noise ratio essentially vanishes. As a consequence the signal acquisition must start with a two dimensional search process.

Equation (6.13) provides the information for selecting the grid in which to search. The spacing in delay is typically chosen to be $T_c/2$. This ensures that the factor $R(\Delta\tau)$ is at least $3/4$ for one of the samples. The search is performed from early to late in order to lock on the direct path rather than on any echo in the case of multipath propagation. The grid spacing for the frequency search is often chosen such that $\delta\Delta\omega\, T_i = 4\pi/3$. This ensures that $\mathrm{sinc}((\omega_r + i\delta\Delta\omega)T_i/2) \geq 0.41$ for at least one i. The maximum Doppler shift due to the satellite movement is ± 6000 Hz. The unaided search of the first satellite requires an extension of the window due to the finite stability of the receiver oscillator, which depends strongly on the quality of that oscillator or on the possibility to lock it onto the stable reference of a base station. For $T_i = m$ milliseconds the spacing in the frequency grid becomes $667/m$ Hz. This grid is sparse enough for $m = 1$, i.e., for an integration time of 1 ms, to search the span of possible frequencies effectively. In the case of 20 or 100 ms, the search

© 2008 by Christoph Günther, all rights reserved.

becomes increasingly cumbersome. The grid is searched, starting with the most likely values, e.g., starting from an estimate of ω one might have. The grid is then searched symmetrically from that value.

In the next step the receiver must be given a criterion for inferring that it has found the correct values of $\Delta\omega$, and $\Delta\tau$. In the maximum likelihood approach described so far the value that maximizes the correlation function is chosen. This is a near optimum criterion under circumstances described previously. They always assumed that the satellite, whose delay and frequency offset are searched are actually visible. In reality, this is not the case. Even if the satellite would be visible under clear sky conditions, it might hide beyond a building, or other obstacle. Taking the maximal correlation would give a random result. Similarly, a receiver that performs initial acquisition on life signals (rather than on stored ones) will experience fading and might thus choose the wrong values of $\Delta\omega$, and $\Delta\tau$.

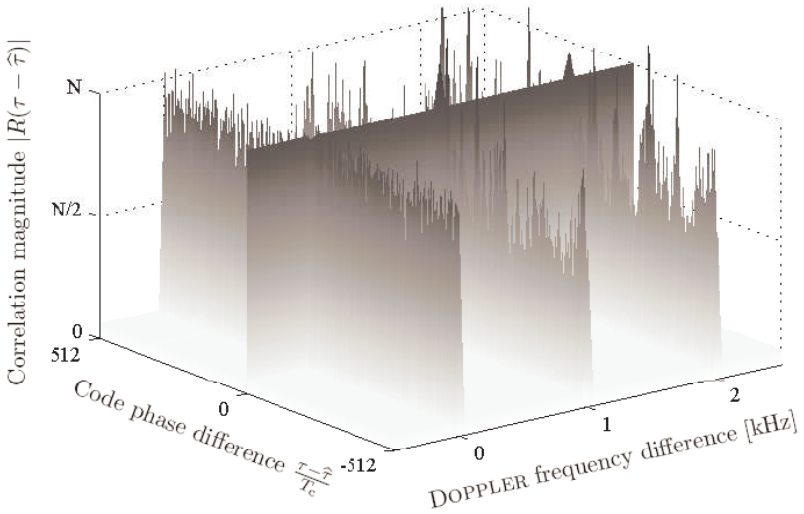

Figure 6.2: Correlation of a GPS signal in the presence of an interfering GPS signal with a 24dB higher power level in the case of $T_i = 20$ ms. The values are very low, except when the frequency offset is a multiple of 1 kHz. The matching must be as accurate as $\delta\omega T_i < 8\pi/3$. [Courtesy A. Schmid [4]]

6.4.1 Neyman-Pearson Criterion

The Neyman-Pearson criterion provides an alternative for finding the correct parameters $\Delta\omega$ and $\Delta\tau$, whenever $\mathcal{P}/\mathcal{N}_0$ can be estimated. The decisions are taken for each grid point individually. The method does thus not suffer from the deficiencies mentioned above. It decides whether a signal is present and whether the values $(\Delta\omega, \Delta\tau)$ are correct by comparing the correlation values with a threshold. Let H_1 denote the hypothesis that the signal is present and that $(\Delta\omega, \Delta\tau)$ is correct, and let H_0 be the probability that either of the conditions are not met, then the Neyman-Pearson criterion decides for

$$\begin{matrix} H_1 \\ H_0 \end{matrix} \quad \text{if} \quad |C_{z,m}| \begin{matrix} \geq \\ < \end{matrix} \gamma. \tag{6.16}$$

This decision can have four outcomes. Their names are listen in Table 6.1. The threshold is chosen to maximize the detection probability under the boundary condition that the probability of false alarm is small, typically $10^{-5}..10^{-3}$. The value 10^{-5} is used with snap-shot receivers, i.e. with

© 2008 by Christoph Günther, all rights reserved.

Test on $\lvert C_{z,m}\rvert$	Signal	Outcome
$\geq \gamma$	present	correct detection
$\geq \gamma$	not present	false alarm
$< \gamma$	present	missed detection
$< \gamma$	not present	correct non-detection

Table 6.1: Classification of the outcomes of the Neyman-Pearson test.

receivers that only perform acquisition, and no subsequent tracking. The receivers in consumer devices operate in this mode in order to minimize power consumption. The value 10^{-3} is used in tracking receivers, which discover erroneous choices in the latter tracking phase. Most professional receivers, e.g. for survey, aeronautical application, farm automation, and the like are of this type. The Neyman-Pearson approach is very effective since it stops the search for a satellite, whenever it found a value above the threshold γ.

The threshold γ needs to discriminate the signal against cross-correlations, and off-peak autocorrelations. The cross-correlations have usually a significant Doppler offset and typically average out to very small values. The periodic code structure leads to a line spectrum of crosscorrelations with a linewidth inversely proportional to T_i. If the Doppler shift is chosen accordingly, the cross-correlations reach a maximum, which is of the same order as the off-peak autocorrelations (see Figure 6.2). Thus, the further discussion can be limited to this size. For a fixed value of $\Delta\omega$, $\Delta\tau$, and $\Delta\varphi$, Equation (6.13) can be rewritten in the form

$$C = \zeta + n_z$$

with ζ being the deterministic average and n_z being the AWG noise contribution. Without loss of generality, assume that ζ is rotated to lie on the positive real axis, i.e. $\Im(\zeta) = 0$, then the probability of obtaining the correlation C given ζ is:

$$p(C = re^{j\Psi}|\zeta) = \frac{1}{2\pi\mathcal{N}_0 T_i} e^{-((r\cos\Psi-\zeta)^2+(r\sin\Psi)^2)/(2\mathcal{N}_0 T_i)}.$$

The probability of a particular value of the amplitude r of C becomes

$$\begin{aligned} p(|C| = r|\zeta) &= \int_{-\pi}^{\pi} r\, d\Psi\, p(C = re^{j\Psi}|\zeta) \\ &= \frac{r}{\mathcal{N}_0 T_i}\, e^{-(r^2+\zeta^2)/(2\mathcal{N}_0 T_i)} \frac{1}{2\pi} \int_{-\pi}^{\pi} d\Psi\, e^{r\zeta\cos\Psi/(\mathcal{N}_0 T_i)} \\ &= \frac{r}{\mathcal{N}_0 T_i}\, e^{-(r^2+\zeta^2)/(2\mathcal{N}_0 T_i)} I_0\left(\frac{r\zeta}{\mathcal{N}_0 T_i}\right) \end{aligned} \tag{6.17}$$

with I_0 being the 0-th order modified Bessel function of the first kind [2]. In the case $\zeta = 0$, the distribution becomes a Rayleigh distribution. For $\zeta \neq 0$ the distribution is a Rice distribution.

The modified Bessel functions of the first kind I_α have a power series expansion of the form

$$I_\alpha(r) = \sum_{k=0}^{\infty} \frac{(r/2)^{\alpha+2k}}{k!\Gamma(\alpha+k+1)},$$

which converges for $r > 0$.

Let γ be the decision threshold, and let

$$\zeta_1 = \alpha R(\Delta\tau)|_{\Delta\tau=T_c/2}\, \mathrm{sinc}\left(\frac{\Delta\omega T_i}{2}\right)\Big|_{\Delta\omega=4\pi/(3T_i)}$$

© 2008 by Christoph Günther, all rights reserved.

be the minimum amplitude of the in-phase autocorrelation function sampled by searching the grid, and let

$$\zeta_0 = \frac{\alpha}{\nu N} \max_{c' \neq c \text{ or } n \neq 0} |C_n^e(c', c)|$$

be the maximum amplitude of the out-of-phase autocorrelation or crosscorrelation, then the criterion described by Equation (6.16) leads to a probability $P_{d,\min}$ of detecting the smallest signal that might occur under regular conditions (signal above threshold):

$$P_d \geq P_{d,\min} = \int_\gamma^\infty dr\, p(|C| = r \big| \zeta_1), \tag{6.18}$$

and a probability of false alarm $P_{f,\max}$, i.e. the probability that noise and cross-correlations combine in such a manner that they cross the threshold (noise and interference above threshold):

$$P_f \leq P_{f,\max} = \int_\gamma^\infty dr\, p(|C| = r \big| \zeta_0). \tag{6.19}$$

The integrals (6.18) and (6.19) are generalized Q-functions, called Marcum's Q-function, see [3]:

$$Q\left(\frac{\zeta}{\sqrt{\mathcal{N}_0 T_i}}, \frac{\gamma}{\sqrt{\mathcal{N}_0 T_i}}\right) = \int_\gamma^\infty dr\, p(|C| = r \big| \zeta).$$

The function Q is monotonic decreasing in γ. It can be inverted for $0 < P_f < 1$ and has a power series expansion

$$Q(x, y) = e^{-(x^2+y^2)/2} \sum_{k=0}^{\infty} \left(\frac{x}{y}\right)^k I_k(xy).$$

The threshold γ is determined from Equation (6.19) by specifying the probability of false detection P_f, e.g. 10^{-3} or 10^{-5}. Equation (6.18) then defines P_d. Figure 6.3 shows results. Two variants of the probability of false alarm are plotted: one for the minimum and one for the maximum crosscorrelation. The parameters used to generate the left plot were a signal-to-noise ratio of 45 dBHz, an integration time of $T_i = 10$ ms, a spacing of the samples in delay space of $T_c/2$, and a corresponding spacing in frequency space of $4\pi/(3T_i)$. The result in this left plot is not really satisfactory. It can be improved in different ways. The right plot shows results for twice as many samples in the frequency search, and otherwise identical parameters. This solves all problems. Alternatively, one might choose a large probability of detection, e.g. $P_d = 0.98$, which leads to a large probability of false alarm. Thus, each time the decision is positive, this decision must be questioned, and verified, either by increasing T_i, and thus the signal-to-noise ratio or by tightening the grid around that value.

6.4.2 Enhancement of the Signal to Noise Ratio during Initial Acquisition

GNS Systems are designed to allow for an acquisition under clear sky conditions with short T_i. GNSS receivers are, however, also used in more demanding environments, under foliage, in moderate indoor environments, as well as in the presence of interference. As a consequence, the signal-to-noise ratio obtained by coherent integration, as described above might not be sufficient for acquiring the signal. From the previous discussion of Equation (6.13), it is clear that increasing the coherent integration time does not lead to a solution, since it excessively increases the computational complexity of the acquisition process.

Different alternatives have been conceived, they include incoherent accumulation

$$\Gamma_k = \sum_{m=\kappa M}^{\kappa(M+1)-1} |C_{z,m}|^2,$$

© 2008 by Christoph Günther, all rights reserved.

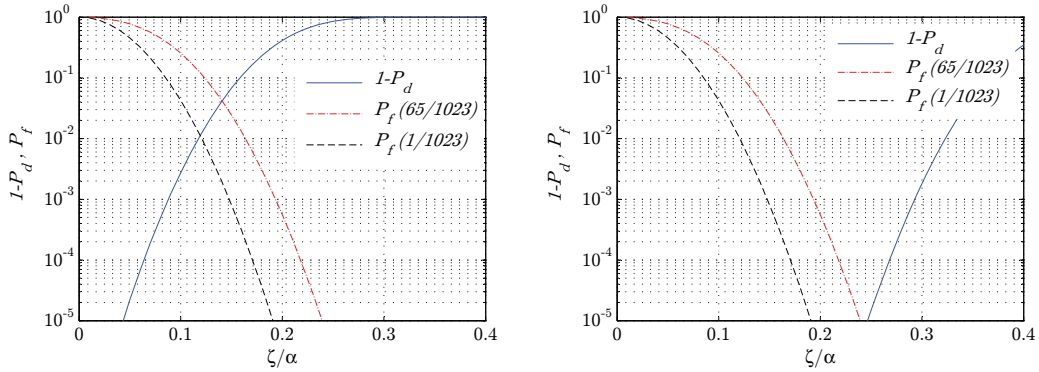

Figure 6.3: The graphs show the probability of missed detection $1-P_d$ as a function of the threshold γ. The probability of false alarm is represented in two different curves, one for the minimum and one for the maximum crosscorrelations of Gold sequences ($\zeta_0 = 1/1023$ and $65/1023$). The left graph corresponds to $\mathcal{P}/\mathcal{N}_0 = 45$ dBHz, and an integration time of 10 ms, and a search space as described in the text. The right graph shows the result, when the number samples in the frequency domain is doubled.

and more recently differential correlation [5]

$$\Gamma_k = \sum_{m=\kappa M}^{\kappa(M+1)-1} C_{z,m+1} C_{z,m}^*.$$

These techniques are used in snap shot receivers in deep urban and indoor environments, i.e., with receivers that do not track the signal.

6.5 Generic Parameter Tracking

The search performed during acquisition leads to a first estimation of the delay τ, and the frequency ω. This knowledge can be used to track the parameters as a function of time, and to refine the first estimate. Equation (6.10), i.e.:

$$\theta_{\max} = \arg\max_\theta (C_{z,m}(\theta))$$

is the basis for tracking. It is equivalent to the maximization of

$$F(\theta) = f(C_{z,m}(\theta))$$

for any function that is monotonic around $C_{z,m}(\theta_{\max})$. The function is preferably also differentiable, so that the local maximization can be reformulated in differential form:

$$F'(\theta_{\max}) = \frac{\partial}{\partial\theta} F(\theta)\Big|_{\theta=\theta_{\max}} = 0.$$

The task of searching zeros can be solved iteratively, e.g. using the Newton algorithm (see Appendix B). In this case, successive approximations $\hat{\theta}_0, \hat{\theta}_1, \ldots$ for finding the zero $\theta_{\max}$ of $F'(\theta_{\max}) = 0$ are obtained by:

$$\hat{\theta}_{m+1} = \hat{\theta}_m - \frac{F'(\hat{\theta}_m)}{F''(\hat{\theta}_m)}. \tag{6.20}$$

© 2008 by Christoph Günther, all rights reserved.

The function

$$D(\hat{\theta}_m) = \frac{F'(\hat{\theta}_m)}{F''(\hat{\theta}_m)} \tag{6.21}$$

is called discriminator. The analysis of the iteration is simplest whenever the function $D(.)$ is linear or nearly linear. Consider the case $f(C_{z,m}) = |C_{z,m}|$, and assume that the frequency error is small, then in the case of a BPSK signal, as used in GPS:

$$F(\Delta\tau) = f(C_{z,m}(\theta)) = |C_{z,m}| \simeq \alpha R(\Delta\tau).$$

Approximating the derivatives of F by finite differences leads to the following expressions valid for $|\Delta\tau| < dT_c/2$:

$$\begin{aligned} F'(\Delta\tau) &= \alpha \frac{1}{dT_c}\left(R(\Delta\tau + \frac{dT_c}{2}) - R(\Delta\tau - \frac{dT_c}{2})\right) \\ &= -\frac{2\alpha}{dT_c^2}\Delta\tau, \end{aligned} \tag{6.22}$$

and

$$\begin{aligned} F''(\Delta\tau) &= \alpha \frac{1}{(dT_c)^2}\left(R(\Delta\tau + dT_c) - 2R(\Delta\tau) + R(\Delta\tau - dT_c)\right) \\ &= -\frac{2\alpha}{dT_c^2}\left(1 - \frac{|\Delta\tau|}{dT_c}\right). \end{aligned}$$

At least in a neighborhood of $\Delta\tau = 0$, this provides the desired functional dependency of the discriminator, and also a negative second derivative showing that the extremum is indeed a maximum:

$$D(\Delta\tau) = \frac{\Delta\tau}{1 - \frac{|\Delta\tau|}{dTc}} \simeq \Delta\tau.$$

This discriminator is called envelope discriminator. The result can also be interpreted as using the definition

$$D(\hat{\theta}_m) = \frac{F'(\hat{\theta}_m)}{F''(\theta_{\max})}$$

instead of Equation (6.21). In later sections, choices that further depart from Equation (6.21) will be used. The associated freedom leads to solutions that "work." The present developments provide the motivation for that. An important criterium that the function $D(.)$ should fulfil is to be equal to $\theta - \theta_{\max}$, whenever this quantity is small. Tracking is first used for refining the initial estimates of the delay, frequency, and phase, and later to follow their change over time. The latter change during one integration period $[mT_i, (m+1)T_i)$ must be small as compared to the desired tracking accuracy. Thus, there is an optimum integration time T_i, which must be large enough for reducing noise and small enough for accommodating the dynamics of the receiver movement. Examples of discriminators are provided in the Figures 6.10, 6.11, and 6.12.

A comprehensive review of sampled parameter tracking was given by Lindsey and Chie in 1981. The present section somewhat follows their presentation. Figure 6.4 shows the generic form of the loop for the parameter θ. The received signal $r(.)$, parameterized by θ is discriminated against the waveform $s(., \hat{\theta})$ generated using the estimate $\hat{\theta}$. The discriminator is typically chosen to be linear with a slope of 1. Its output includes a deterministic part:

$$x_m = \theta_m - \hat{\theta}_m, \tag{6.23}$$

and a superimposed noise contribution

$$n_m.$$

© 2008 by Christoph Günther, all rights reserved.

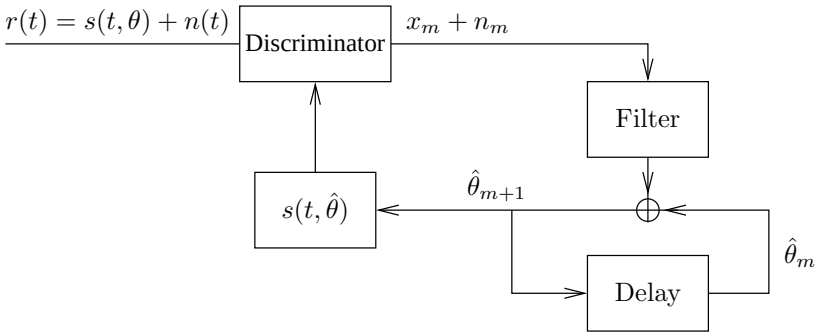

Figure 6.4: Generic form of a tracking loop for the parameter θ.

This superposition is filtered:

$$y_m = T_i \sum_{k=0}^{\infty} f_k (x_{m-k} + n_{m-k}). \tag{6.24}$$

and used to update the estimate. In the present case, successive time instants are separated by T_i seconds. Thus the sum is a Riemann sum, which would converges towards an integral if $T_i \to 0$. The scaling implies that the coefficients $f_.$ only weakly depend on T_i. The estimate $\hat{\theta}_m$ is updated using the filtered correction:

$$\hat{\theta}_{m+1} = \hat{\theta}_m + y_m. \tag{6.25}$$

Equations (6.23)-(6.25) define a difference equation, either in the sequence $\hat{\theta}_.$ or $x_.$. The latter choice shall typically be preferred, since $x_.$ is small whenever the loop tracks the parameter, while $\hat{\theta}_.$ is supposed to follow $\theta_.$ and thus includes a non-zero statistical mean. Equation (6.25) then becomes:

$$x_{m+1} - x_m = \theta_{m+1} - \theta_m - y_m. \tag{6.26}$$

The solution and discussion of the difference equation is simplified by introducing the z-transform. The z-transform allows to compute the solution to the above set of difference equations (6.24)-(6.26) in closed algebraic form. For a sequence x_m, it is defined by the Laurent-series:

$$X(z) = \sum_{m=-\infty}^{\infty} x_m z^{-m}, \tag{6.27}$$

which can also be expressed as a sum of Taylor-series in z, and $1/z$:

$$X(z) = \sum_{m=0}^{\infty} x_{-m} z^m + \sum_{m=1}^{\infty} x_m \left(\frac{1}{z}\right)^m.$$

Both series typically have a positive convergence radius R_+ and R_-, respectively. The Laurent-series thus defines a complex analytical function in the region:

$$R_- < |z| < R_+,$$

see Figure 6.5. Analytical functions have properties, which are greatly supportive in studying tracking loops. Some of these properties will be detailed below, see also [10], [11], and other text books.

© 2008 by Christoph Günther, all rights reserved.

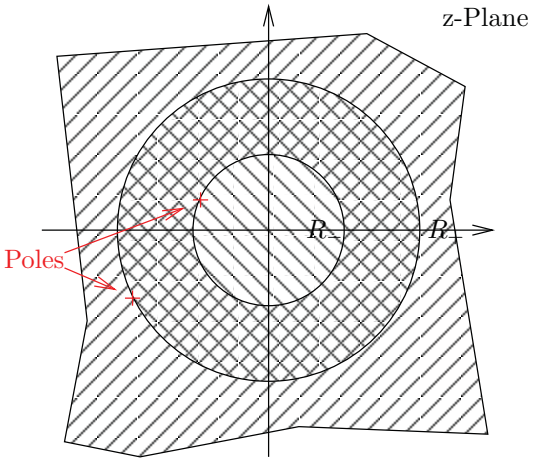

Figure 6.5: The area of convergence of a Laurent-Series is limited by the singularities of the associated analytical function. The series with negative exponents converges in the region $|z| > R_-$ shaded with a pattern ///, and the series with positive exponents converges in the region $|z| < R_+$, which is shaded in the reverse direction. The Laurent-series converges in the intersection of both regions.

Laurent-series and difference equations are related through the observation that shifting in time corresponds to a multiplication by z[7]: let $\tilde{y}_m = x_{m+1}$, then $\tilde{Y}(z) = zX(z)$. Similarly, integration $\bar{y}_m = \bar{y}_{m-1} + T_i x_m$ corresponds to:

$$\bar{Y}(z) = T_i \frac{1}{1 - z^{-1}} X(z), \tag{6.28}$$

furthermore the z-transform of the convolution of $x_\cdot$ by a filter function $f_\cdot$, as defined in Equation (6.24) becomes:

$$Y(z) = T_i F(z)[X(z) + N(z)].$$

Equation (6.23) can be rewritten as:

$$X(z) = \Theta(z) - \hat{\Theta}(z).$$

Together with the Equations (6.24) and (6.26), this leads to the following equation for estimating $X(z)$:

$$(z-1)X(z) = (z-1)\Theta(z) - T_i F(z)[X(z) + N(z)],$$

which can be solved for $X(z)$:

$$X(z) = [1 - H(z)]\Theta(z) - H(z)N(z), \tag{6.29}$$

with:

$$H(z) = \frac{T_i F(z)}{z - 1 + T_i F(z)}. \tag{6.30}$$

The loop described in the parameters θ, and x, as well as using the z-transform of the integrator and the filter looks slightly simpler. It is shown in Figure 6.6.

[7]This is the reason for using z^{-m} rather than z^m in Equation (6.27).

© 2008 by Christoph Günther, all rights reserved.

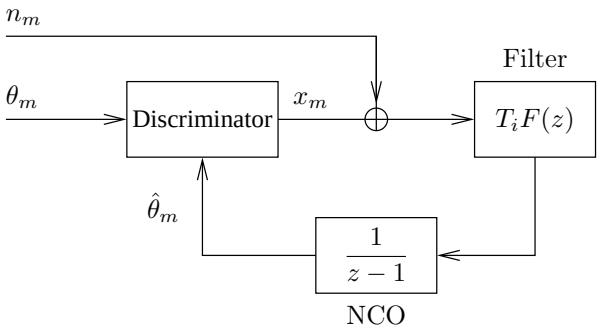

Figure 6.6: Tracking loop described by z-Transforms.

Order ν	G_0	G_1	G_2
1st	ω_0		
2nd	$a\omega_0$	ω_0^2	
3rd	$b\omega_0$	$a\omega_0^2$	ω_0^3

Table 6.2: Correspondence of the parameters of the loop filter.

Typical loop filters have the form:

$$F(z) = \sum_{n=0}^{\nu-1} \frac{G_n T_i^n}{(1-z^{-1})^n} = \sum_{n=0}^{\nu-1} \frac{G_n T_i^n z^n}{(z-1)^n} \tag{6.31}$$

with ν denoting the order of the filter. The terms in the sum are nested integrators as can be seen from Equation (6.28). Figure 6.7 shows a representation of such filters for low degrees as used by Ward [13]. The correspondence of the coefficients is provided in Table 6.2. Its general form reads

$$G_n = a_{\nu-n-1}\,\omega_0^{n+1}, \qquad n \in \{0, 1, \ldots \nu-1\},$$

with

$$a_0 = 1,\ a_1 = a,\ a_2 = b,\ \ldots$$

It leads to the following expression for the filter:

$$T_iF(z) = \omega_0 T_i \sum_{n=0}^{\nu-1} a_{\nu-1-n} \left(\frac{\omega_0 T_i z}{z-1}\right)^n.$$

The ratio of analytical functions implicit in the Equations (6.29) and (6.30) defines another analytical function. The zeros of the denominator become poles of $X(z)$, unless they are canceled by zeros of the numerator.

The coefficients of the Laurent-series from Equation (6.27), i.e. the elements of the sequence, can be recovered by computing the contour integral

$$x_m = \frac{1}{2\pi j} \oint_\gamma dz\, X(z) z^{m+1}, \tag{6.32}$$

© 2008 by Christoph Günther, all rights reserved.

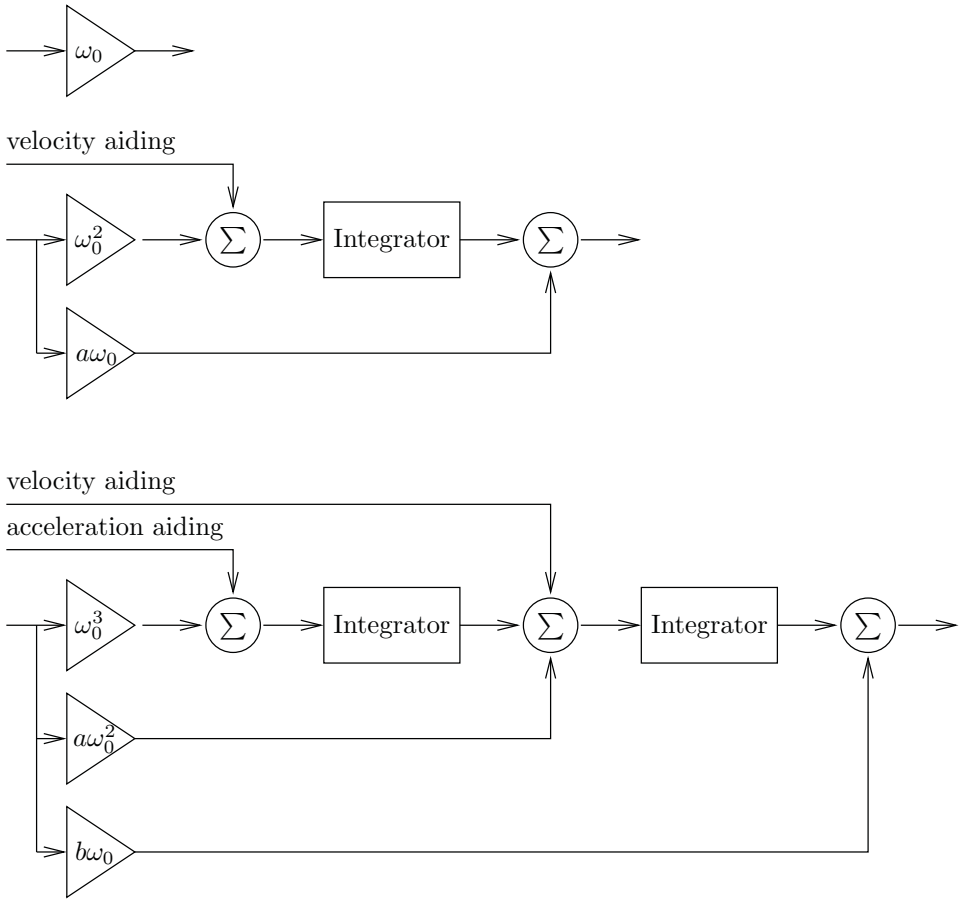

Figure 6.7: Loop filters of order 1, 2, and 3. The order counts the total number of integrators including the one modeling the delay generator. The notation follows Ward [13].

with γ being a contour that circles once around the origin staying completely in the area of convergence of the function $X(z)$. These integrals are typically computed using residuals at the individual poles (see [11] for example):

$$\frac{1}{2\pi j}\oint_\gamma dz\, f(z) = \sum_{\text{poles } z_0 \text{ inside } \gamma} \text{Res}(f(z)|z_0).$$

Let

$$f(z) = \frac{g(z)}{(z-z_0)^s}$$

with $g(z)$ analytic in z_0, then the residual can itself be computed from

$$\text{Res}(f(z)|z_0) = \frac{1}{(s-1)!}\frac{d^{s-1}g(z)}{dz^{s-1}}\bigg|_{z=z_0},$$

i.e. the elements of the sequence x can be characterized by derivatives at the poles of $X(.)$ inside some domain. This domain will be further characterized in the next section.

© 2008 by Christoph Günther, all rights reserved.

6.5.1 Loop Response without Noise

This is the simplest situation to study. In this case $N(z) = 0$. Typical input errors are steps in the parameter, its first derivative, or its second derivative. In the case of the phase locked loop, which is the most sensitive loop, the steps in the parameter and the two derivatives are steps in the phase, in the frequency, and in the derivative of the frequency. A step is a mathematical idealization for an abrupt change. Physically such a change might be caused by the dynamics of the receiver, of the ionosphere or by surrounding scatterers, for example. The z-transformation of various steps is listed in Table 6.3.

A linear convolutional system

$$y_m = \sum_{k=-\infty}^{\infty} \eta_k x_{m-k}$$

is defined as *causal* if $\eta_k = 0$ for all $k < 0$. The associated Laurent-series

$$H(z) = \sum_{k=0}^{\infty} \eta_k z^{-k}$$

is singular in 0. The convergence region of the series is given by $\{z \in \mathbb{C} \big| \, |z| > |z_0|\}$ with z_0 being the pole with the maximum absolute value.

A convolutional system (causal or acausal) is said to be *stable* if every bounded input $|x_.| \leq \mu < \infty$ produces a bounded output $|y_.| \leq \mu' < \infty$. Such systems are called bounded input bounded output (BIBO) stable. The system is stable if and only if

$$\sum_{k=-\infty}^{\infty} |\eta_k| < \infty.$$

The inequality

$$|y_m| \leq \sum_{k=-\infty}^{\infty} |\eta_k| \max_m |x_m| = \mu \sum_{k=-\infty}^{\infty} |\eta_k|$$

proves that the condition is sufficient. The choice $x_k = \eta_k^*/|\eta_k|$ implies $|x_k| = 1$, which is definitively bounded, and since

$$y_0 = \sum_{k=-\infty}^{\infty} |\eta_k|, \quad \text{and} \quad |y_0| \leq \mu'$$

this also proves the necessity of the condition.

Next, consider the inequality

$$|H(z)| = |\sum_{k=-\infty}^{\infty} \eta_k z^{-k}| \leq \sum_{k=-\infty}^{\infty} |\eta_k|.$$

It implies that a system cannot be stable if its z-transform has a pole on the unit circle.

Step functions are themselves singular. If a filter applied to a step function is to provide a bounded output, the poles of the step function must be canceled by zeros of the function $K(z) = 1 - H(z)$, which is the factor multiplying the step function in Equation (6.29). Specifically, a step of order ν, requires a filter of order ν. Consider

$$\begin{aligned} K(z) &= \frac{(z-1)^\nu}{(z-1)^\nu + (z-1)^{\nu-1} T_i F(z)} \\ &= \frac{(z-1)^\nu}{(z-1)^\nu + \omega_0 T_i \sum_{n=0}^{\nu-1} a_{\nu-1-n} (\omega_0 T_i z)^n (z-1)^{\nu-1-n}} \\ &= \frac{(z-1)^\nu}{P(z)}, \end{aligned}$$

© 2008 by Christoph Günther, all rights reserved.

Step in	$\theta(t)$ (for $t \geq 0$)	$\Theta(z)$
parameter	1	$\frac{z}{z-1}$
1st derivative	t	$T_i \frac{z}{(z-1)^2}$
2nd derivative	t^2	$T_i^2 \frac{z(z+1)}{(z-1)^3}$
3rd derivative	t^3	$T_i^3 \frac{z(z^2+4z+1)}{(z-1)^4}$
k-th derivative	t^k	$\left(-T_i z \frac{d}{dz}\right)^k \frac{z}{z-1}$

Table 6.3: The table shows different filter inputs $\theta(t)$ with $\theta(t) = 0$ for $t < 0$, and functions for positive t listed in the table. The time variable t is incremented in steps of T_i, i.e.: $t = nT_i$. The main purpose of the table is to provide the associated z-transforms $\Theta(z)$.

with

$$P(z) = (z-1)^\nu + \omega_0 T_i \sum_{n=0}^{\nu-1} a_{\nu-n-1} (\omega_0 T_i z)^n (z-1)^{\nu-1-n}.$$

Since

$$P(z)|_{z=1} = \sum_{n=0}^{\nu-1} a_{\nu-1-n} (\omega_0 T_i z)^n (z-1)^{\nu-1-n} \Big|_{z=1} = a_0 (\omega_0 T_i)^{\nu-1} = (\omega_0 T_i)^{\nu-1},$$

the above factor removes all poles in 1 up to order ν, and thus leads to a stable system. No filter of lower order can achieve this.

In the case of a causal and stable system $K(z)$, the following inequality holds for $|z| \geq 1$:

$$|K(z)| \leq \sum_{n=0}^{\infty} |\kappa_k| |z|^{-k} \leq \sum_{n=0}^{\infty} |\kappa_k| < \infty.$$

This implies that all poles must lie strictly inside the unit circle, and thus that the contour γ in the integral in Equation (6.32) can be chosen to be the unit circle or somewhat smaller. The latter is desirable for multiplying the numerator and denominator by powers of $(z-1)$. The requirement on the poles of $K(z)$ implies that the zeros of $P(z)$ must lie inside the unit circle. For a first order loop, the solution of $P(z) = 0$ reads

$$z = 1 - \omega_0 T_i = 0 \tag{6.33}$$

and the conditions $|z| < 1$ and $\omega_0 T_i$ real-valued and positive imply:

$$\omega_0 T_i < 2.$$

De facto for the loop to reduce the noise, this product shall be significantly smaller than 1. For a second order loop, the equation becomes:

$$z^2 - (2 - a\omega_0 T_i - \omega_0^2 T_i^2) z + 1 - a\omega_0 T_i = 0.$$

Its solution can be indicated in closed form:

$$z = 1 - \frac{\omega_0 T_i}{2} \left(a + \omega_0 T_i \pm \sqrt{(a + \omega_0 T_i)^2 - 4} \right), \tag{6.34}$$

and analyzed. Equation (6.38) in the next section shows that the noise reduction is maximized by a loop of second order if both $\omega_0 T_i \ll 1$ and $a \sim 1$. Thus $|a + \omega_0 T_i| < 2$, and the expression in the square root is negative, i.e. the root has only an imaginary part. Thus the condition $|z| < 1$ becomes:

$$\big(2 - \omega_0 T_i (a + \omega_0 T_i)\big)^2 + (\omega_0 T_i)^2 \big(a + \omega_0 T_i)^2 - 4\big) < 4.$$

© 2008 by Christoph Günther, all rights reserved.

A little algebra leads to the the equivalent condition

$$\omega_0 T_i\big((a+\omega_0 T_i)^2-4\big) < 2a,$$

which is fulfilled under the assumption that $a > 0$ and $|a+\omega_0 T_i| < 2$. Thus the only condition on a is that

$$0 < a < 2 - \omega_0 T_i. \tag{6.35}$$

6.5.2 Tracking in the Presence of Noise

For the study of the tracking performance in the presence of noise, Equation (6.29) is reformulated as a difference equation:

$$x_m = \xi_m - \sum_{k=-\infty}^{\infty} h_k n_{m-k}.$$

If $n_\cdot$ is zero mean, which should obviously be the case within the current model, the expectation becomes:

$$\mathcal{E}[x_m] = \xi_m.$$

In this case, the covariance matrix Λ of the correction of the loop parameter $x_\cdot$ can be expressed in terms of the covariance of the noise $n_\cdot$, which is assumed to be time independent (stationary noise):

$$\begin{aligned} Q_{m,p} &= \mathcal{E}[(x_m - \mathcal{E}[x_m])(x_{m+p} - \mathcal{E}[x_{m+p}])] \\ &= \sum_{k,l} h_k h_l \mathcal{E}[n_{m-k} n_{m+p-l}] \\ &= Q_{0,p}, \qquad \text{i.e. time independent.} \end{aligned}$$

The z-transform of this expression becomes:

$$\begin{aligned} Q_{x_\cdot}(z) &= \sum_p Q_{0,p} z^{-p} \\ &= \sum_{n,k,l} h_k \left(\frac{1}{z}\right)^{-k} \cdot h_l\, z^{-l} \cdot \mathcal{E}[n_{m-k} n_{m+p-l}]\, z^{-(n-l+k)} \\ &= H\left(\frac{1}{z}\right) H(z)\, Q_{n_\cdot}(z), \end{aligned}$$

with $Q_{x_\cdot}(z)$, and $Q_{n_\cdot}(z)$, denoting the z-transform of the corresponding sequences. These choices were made in order to reduce the potential for confusion. If the input noise $n_\cdot$ is white, the following holds:

$$Q_{n_\cdot}(z) = \sigma_n^2.$$

A causal and stable system has a z-transform $H(z)$ that is bounded on the unit circle. Since the transformation $1/z$ maps the unit circle on itself, the same applies to $H(1/z)$. Thus $H(z)$ is analytical in some neighborhood of the unit circle and the zero-th Laurent-coefficient $Q_{x_\cdot,0}$ can be computed from the integral

$$\sigma_x^2 = \mathcal{E}[(x_m - \xi_m)^2] = Q_{x_\cdot,0} = \sigma_n^2 \frac{1}{2\pi j} \oint_\gamma dz z^{-1} H\left(\frac{1}{z}\right) H(z) \tag{6.36}$$

with the integration being performed over the unit circle. Using that all coefficients of $H(.)$ are real values, the integral can be expressed in the form:

$$\frac{\sigma_x^2}{\sigma_n^2} = \frac{1}{2\pi} \int_0^{2\pi} d\varphi |H(e^{j\varphi})|^2.$$

© 2008 by Christoph Günther, all rights reserved.

Since the noise is white, it can be fully characterized by the bandwidth of the filter to which it is subjected. Consider an equivalent box-shaped filter, and let B_L denote the one sided bandwidth of the loop filter, then the ratio of the noise variances is given by:

$$\frac{\sigma_x^2}{\sigma_n^2} = \frac{2B_L}{1/T_i},$$

with T_i being the sampling rate of the input.

The key relation for determining the actual value of this quantity is given by Equation (6.36). The contour integral in this equation can be computed explicitly. To this end, the poles of the integrand inside the unit circle C_1 must be determined: they are the poles of $1/z$ and of $H(z)$. $H(1/z)$ has no poles inside C_1 because $H(z)$ is regular outside C_1, and because $1/z$ exchanges inside and outside. $H(z)$ can be expressed in terms of $P(z)$:

$$H(z) = \frac{P(z) - (z-1)^\nu}{P(z)}.$$

Since $H(0) = a_{\nu-1}\omega_0 T_i\big((-1)^\nu + a_{\nu-1}\omega_0 T_i\big)^{-1}$ is finite and non-zero, $H(z)$ does neither cancel the pole in $z = 0$, nor does it increase its order. Furthermore, since $\deg\big(P(z)-(z-1)^\nu\big) < \deg(P(z))$:

$$\lim_{z\to 0} H(\frac{1}{z}) = \lim_{z\to\infty} H(z) = 0.$$

Thus the residuum at 0 is zero, and the poles of $H(z)$, i.e. the zeros of $P(z)$, are the only ones to be considered. Let them be denoted by z_m, and assume that they have multiplicity 1, i.e. $z_m \neq z_{m'}$ for all $m \neq m'$. The generalization to higher multiplicities would be straight forward. Since the coefficient of z^ν in $P(z)$ is equal to 1, $P(z)$ can be represented in the form:

$$P(z) = \prod_{m=0}^{\nu-1}(z - z_m),$$

which finally leads to

$$\frac{1}{2\pi j}\oint_{C_1} dz \frac{1}{z} H\left(\frac{1}{z}\right) H(z) = \sum_{n=0}^{\nu-1} \frac{-(z_n-1)^\nu H(1/z_n)}{z_n \prod_{(m\neq n)}(z_n - z_m)}. \tag{6.37}$$

This is an algebraic function, which can be evaluated whenever, the roots of $P(z)$ are known. For $\nu = 1$, and $\nu = 2$, the roots were determined by Equation (6.33) and (6.34).

For the first order filter ($\nu = 1$), the evaluation of the above integral is straight forward:

$$2B_L^{(1)} T_i = \frac{\omega_0 T_i}{2 - \omega_0 T_i} \simeq \frac{\omega_0 T_i}{2},$$

with the last approximation being valid for small values of $\omega_0 T_i$. For the second order filter ($\nu = 2$), the evaluation is more involved and a symbolic computation program such as Maxima [14] becomes helpful:

$$2B_L^{(2)} T_i = \frac{\omega_0 T_i}{a} \cdot \frac{2 + 2a^2 + a\omega_0 T_i}{4 - 2a\omega_0 T_i - (\omega_0 T_i)^2} \simeq \frac{\omega_0 T_i}{2}\frac{1+a^2}{a}. \tag{6.38}$$

The same assumption was used for the approximation in the second order expression. The interesting property of Equation (6.37) is that it can be evaluated for arbitrary loop orders. The zeros must be determined numerically in such cases. If they all have multiplicity 1, the right hand side of Equation (6.37) is then evaluated, otherwise this is done using a generalization of that expression to roots with multiplicities.

As a final comment, on notes that the parameter a that leads to the smallest inflation of the second order result with respect to the first order result is $a = 1$, which is well inside the region of stability for frequency steps described by Equation (6.35). The penalty in noise inflation paid for the increased stability is thus $(1 + a^2)/a = 5/2$.

© 2008 by Christoph Günther, all rights reserved.

6.6 Hierarchy of Control Loops

Having studied the performance of a generic control loop, the control architecture is now introduced. In principle, the parameter $\theta = (\Delta\omega, \Delta\tau, \Delta\varphi)$ must be controlled to zero. This would be a difficult task, if it was to be done at once. Fortunately, the correlation result from Equation (6.13), which is the basis for constructing discriminators essentially has a near product form. More precisely,

$$|\mathcal{E}[C_{z,m}]| = \alpha R(\Delta\tau)\,\mathrm{sinc}\left(\frac{\Delta\omega T_i}{2}\right)$$

mostly depends on $\Delta\tau$, when $\Delta\omega$ is small, and thus allows to control the delay, while

$$\mathcal{E}[C_{z,m+1}C^*_{z,m}] = \alpha^2 R^2(\Delta\tau)\, e^{j\Delta\omega T_i}\mathrm{sinc}\left(\frac{\Delta\omega T_i}{2}\right)$$

allows to control the frequency, if the delay and frequency errors are kept near zero. For simplicity, we considered $|\mathcal{E}[C_{z,m}]|$, which is not realizable in practice. The true discriminator will be described in the next section.

With such loops in mind, it is natural to transit from initial acquisition to a second stage with delay and frequency tracking. The loop used to track delay is called "Delay Locked Loop (DLL)," and, the loop, which tracks frequency is called "Frequency Locked Loop (FLL)." The lock of these loops is determined by analyzing the variance in delay and frequency, and performing a Neyman-Pearson type test. When tracking is initiated, the oscillators used in the DLL and the FLL are adjusted individually. Once lock is established, the oscillators are kept synchronous, since the change in delay and frequency have as a common origin: the relative movement of the receiver and satellite. This coupling also compensates for any Doppler in the code waveform, since the delay oscillator generates the local copy of the code waveform.

If lock is not reached, the receiver returns to the initial acquisition state. Once lock has been achieved, the carrier phase loop is started. The various state transitions are shown in Figure 6.8. The carrier phase loop is called "Phase Locked Loop (PLL)." It is special in so far as the tracking of the phase, predominantly implies a tracking of delay and frequency. Assuming that the delay and frequency error are small enough:

$$R(\Delta\tau) \simeq 1, \quad \text{and} \quad \mathrm{sinc}\left(\frac{\Delta\omega T_i}{2}\right) \simeq 1$$

implies that:

$$\mathcal{E}[C_{z,m}] = \alpha e^{j(\Delta\omega m T_i - \omega\Delta\tau + \Delta\varphi)}$$

After some initialization, $\Delta\varphi$ should not change anymore. Thus the loop will compensate changes in delay and frequency. Again if the variance increases too much, PLL tracking is abandoned and DLL/FLL tracking is resumed. The DLL is never disrupted since it ensures the robustness of tracking. The signal-to-noise ratio of the correlation outcome should always stay large.

In the case of perfect PLL tracking, the quadrature signal vanishes. In such a case, it is sufficient to consider the x-component, this reduced the noise by a factor of two. In practice this possibility is not used, however, since PLL tracking of the delay improves the accuracy by three orders of magnitude, already.

The next two sections describe the discriminators used for the DLL and for the FLL and DLL, respectively, and analyze the noise performance of these discriminators. This is the input for evaluating the noise performance of the DLL, FLL, and PLL, using the results about the generic loop performance from the previous section.

All of the above tracking loops are operated independently for each satellite and each signal. Clearly, all adjustments have some common origin in the mentioned relative movement or changes in the propagation environments. As already mentioned, the coupling of these loops is increasingly considered [7]-[**?**].

© 2008 by Christoph Günther, all rights reserved.

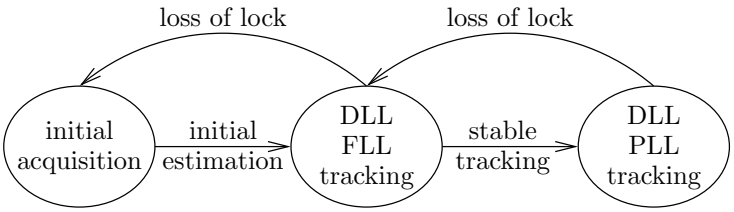

Figure 6.8: States of the acquisition and tracking process, as well as the associated transitions.

6.7 Code Tracking - Delay Locked Loop

The Delay Locked Loop is considered first. Figure 6.9 shows a typical implementation with the three key elements:

1. the discriminator, which maps the correlation results into a measure for the delay offset τ,
2. the loop filter, which ensures the stability of the loop and reduces the noise
3. the reference signal generator, consisting of the code generator and a Numerically Controlled Oscillator (NCO)[8], which adapts the timing of the reference signal.

The discriminator shown correlates an early and late version of the received signal with the local code, and takes the difference "early minus late"[9]. Equivalently, the receiver correlates the received signal with a late version of the local code $c(t - dT_c/2)$ and with a early version of the waveform $c(t + dT_c/2)$, and takes the difference. As described below, this convention provides an estimate of the delay τ. The reversed convention is also used sometimes with early and late referring to the local code. All these differences lead to an estimation of τ.

A prompt signal is also generated as an input to the FLL and PLL loops, that are operating around the DLL, i.e. they use the output of the DLL, while adapting the frequency used in the down-conversion between the antenna and the DLL.

6.7.1 DLL Discriminators and their Performance

In the stationary case, the frequency are locked, i.e., $\Delta\omega \simeq 0$, and the delay is small as well: $\Delta\tau \simeq 0$. One would therefore be inclined to use the coherent discriminator, described above:

$$C_E(\Delta\tau) - C_L(\Delta\tau), \tag{6.39}$$

with

$$\begin{aligned} C_{\substack{E\\L}}(\Delta\tau) &= C_{z,m}\left(\Delta\tau \mp \frac{dT_c}{2}, \Delta\omega, \Delta\varphi \mp \omega\frac{dT_c}{2}\right) \\ &= \alpha\, R(\Delta\tau \mp \frac{dT_c}{2})\, e^{j\left(\Delta\omega m T_i - \omega\Delta\tau + \Delta\varphi\right)} + n_{\substack{E\\L}} \end{aligned}$$

and

$$n_{\substack{E\\L},m} = \int_{(m-1/2)T_i \mp \frac{dT_c}{2}}^{(m+1/2)T_i \mp \frac{dT_c}{2}} dt\, c(t \mp \frac{dT_c}{2}) n(t) e^{j\omega_r t}.$$

[8] An NCO is an oscillator, which adapts the frequency of its output according to the amplitude of its numerical input.

[9] "Early minus late" means that the sign is reversed with respect to the discrete implementation of the derivation in Equation (6.22). This is necessary to accommodate the negative sign of $F''(\Delta\tau)$.

© 2008 by Christoph Günther, all rights reserved.

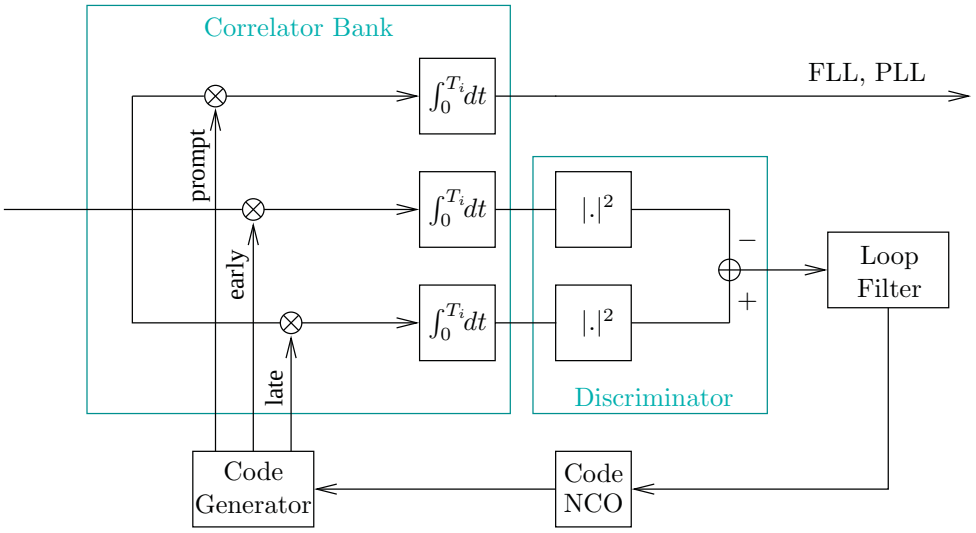

Figure 6.9: Delay Locked Loop (DLL) consisting of a discriminator, a loop filter, and an NCO.

Since the DLL processes the different integration intervals subsequently, the index m will dropped below. Using the coherent discriminator, described by Equation (6.39), would require that the phase error is small, which cannot be guaranteed in early tracking stages. The functional behavior of this discriminator is shown in Figure 6.10. It is indeed linear in the interval $[-\frac{dT_c}{2}, \frac{dT_c}{2}]$. If the phase error was small so-called cycle slips could occur. They are phase jumps by π (or 2π depending on the phase discriminator). They might flip the sign of the function and might thus destabilize the DLL. Cycle slips are caused by fades in the signal-to-noise ratio, which allow the phase to wander to another local equilibrium of the phase discriminator. With this considerations in mind, the difference $|C_E(\Delta\tau)| - |C_L(\Delta\tau)|$, i.e. the envelope discriminator, considered in Section 6.5, is more suited. It has the same characteristic as the coherent discriminator with an ideal phase estimate. Unfortunately, it could not be analyzed in the presence of noise so far.

Discriminators commonly considered are non-linear, with a piece of nearly linear characteristic around $\Delta\tau = 0$. Examples of noiseless discriminator responses are shown in Figure 6.10. They include the the dot discriminator $C^*(\Delta\tau)(C_E(\Delta\tau) - C_L(\Delta\tau))$, and the power discriminator. The latter is defined by

$$D(\Delta\tau) = |C_E(\Delta\tau)|^2 - |C_L(\Delta\tau)|^2.$$

It has the advantages of being easy to implement, and most importantly of having a strictly linear slope in the interval $[-\frac{dT_c}{2}, \frac{dT_c}{2}]$. It thus unifies robustness and linearity, and can furthermore be analyzed to a significant extent, even in the presence of noise. A minor inconvenience is that the squaring process enhances the noise at low signal-to-noise ratios. The impact is small in the usual modes of operations, however.

The control signal generated by this discriminator in the presence of noise is characterized by computing the expectation value. The following notation is introduced in order to keep the expressions simple:

$$R_{\substack{E \\ L}} = R(\Delta\tau \mp \delta) \quad \text{with} \quad \delta = \frac{dT_c}{2}.$$

The loop needs to keep the delay error inside the linear range of the discriminator $\Delta\tau \in [-\delta, +\delta]$. In this range, the following applies:

$$R_{\substack{E \\ L}} = 1 - \frac{\delta \mp \Delta\tau}{T_c},$$

© 2008 by Christoph Günther, all rights reserved.

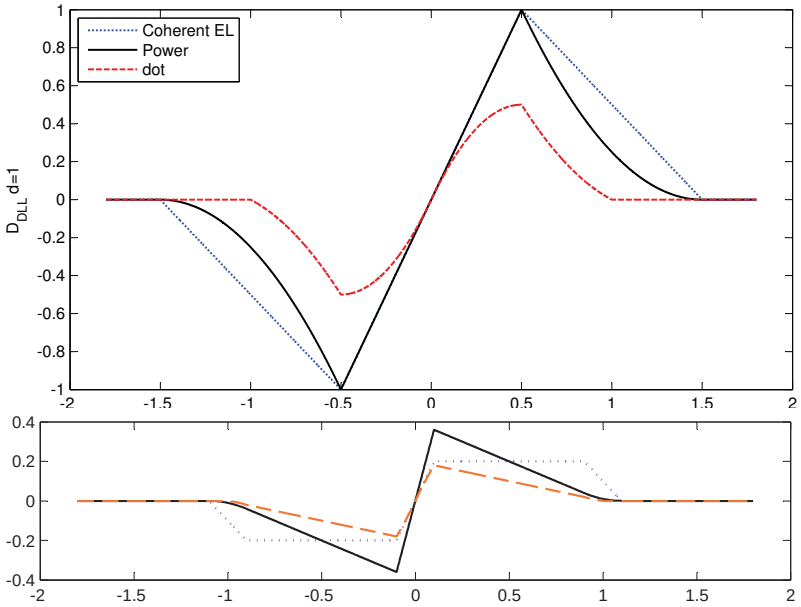

Figure 6.10: Discriminator functions with $d = 1$ in the upper graph and $d = 0.2$ in the lower graph. The coherent and envelope discriminators have the same characteristic. The dot discriminator has a non-linear characteristic in the control region around $\Delta\tau = 0$.

and

$$(R_{\substack{E\\L}})^2 = R^2(\delta) \pm 2R(\delta)\frac{\Delta\tau}{T_c} + \left(\frac{\Delta\tau}{T_c}\right)^2,$$

with

$$R(\delta) = 1 - \frac{\delta}{T_c} = 1 - \frac{d}{2}.$$

Since the noise is zero-mean with a time invariant statistic:

$$\mathcal{E}[|n_E|^2] = \mathcal{E}[|n_L|^2] = \mathcal{E}[|n|^2] = 2\mathcal{N}_0 T_i,$$

the computation of $\mathcal{E}[D(\Delta\tau)]$ becomes straight forward:

$$\begin{aligned}\mathcal{E}[D(\Delta\tau)] &= \alpha^2(R_E^2 - R_L^2) + \mathcal{E}[|n_E|^2] - \mathcal{E}[|n_L|^2] \\ &= 4\alpha^2 R(\delta)\frac{\Delta\tau}{T_c} \\ &= 2\alpha^2(2-d)\frac{\Delta\tau}{T_c}.\end{aligned}$$

On this basis, the expression

$$\frac{\hat{\Delta\tau}}{T_c} = \frac{1}{2\alpha^2(2-d)}\,D(\Delta\tau), \tag{6.40}$$

looks like a good estimator for the tracking error $\Delta\tau$, as long as $\Delta\tau$ stays small. The expression also depends on d. The original choice of d was $d = 1$, i.e., the early and late correlators where spaced by one chip. Dierendonck, Fenton and Ford [15] discovered that smaller spacings lead to lower estimation errors, and developed the theory presented in the present section. It is immediate that the characteristic of the discriminator does not provide a significant gain for small d. The origin of the gain rather lies in the statistical dependence of the noise, which increases as the early

© 2008 by Christoph Günther, all rights reserved.

and late waveforms overlap. The other factors should not be interpreted at this stage, they are the consequence of a particular choice of the normalization.

The discriminator function is further characterized by its noise variance:

$$\mathrm{var}[D(\Delta\tau)] = \mathcal{E}\left[|D(\Delta\tau)|^2\right] - \left|\mathcal{E}[D(\Delta\tau)]\right|^2.$$

The second term is the square of the expectation computed above. The first one is expanded into

$$\begin{aligned}
\mathcal{E}\left[|D(\Delta\tau)|^2\right] &= \mathcal{E}\Big[\big\{(\alpha^2|R_E|^2 + \alpha\,(R_E^* n_E + R_E n_E^*) + |n_E|^2) \\
&\quad -(\alpha^2|R_L|^2 + \alpha\,(R_L^* n_L + R_L n_L^*) + |n_L|^2)\big\}^2\Big] \\
&= \mathcal{E}\Big[\alpha^4\left(|R_E|^4 - 2|R_E|^2|R_L|^2 + |R_L|^4\right) + \\
&\quad +2\alpha^2\left(|R_E|^2(2|n_E|^2 - |n_L|^2) - 2\Re(R_E^* n_E R_L n_L^*) + |R_L|^2(2|n_L|^2 - |n_E|^2)\right) + \\
&\quad +|n_E|^4 - 2|n_E|^2|n_L|^2 + |n_L|^4\Big].
\end{aligned} \tag{6.41}$$

Products with an odd number of terms vanish because the noise is Gaussian and zero-mean. Furthermore: $\mathcal{E}[n_E n_L] = \mathcal{E}[n_E^* n_L^*] = \mathcal{E}[n_E^2] - \mathcal{E}[n_L^2] = 0$. The missing noise expectations need to be computed. By substituting the definitions, one obtains

$$\mathcal{E}[n_E n_L^*] = 2\mathcal{N}_0 T_i R(2\delta),$$

and by using the standard expression for the moments of a multivariate Gaussian distribution:

$$\begin{aligned}
\mathcal{E}\left[|n_E|^2|n_L|^2\right] &= \int dt'_E\,dt_E\,dt'_L\,dt_L\,c(t'_E+\delta)c(t_E+\delta)c(t'_L-\delta)c(t_L-\delta) \\
&\qquad\qquad \cdot\mathcal{E}[n^*(t'_E)n(t_E)n^*(t'_L)n(t_L)] \\
&= \int dt'_E\,dt_E\,dt'_L\,dt_L\,c(t'_E+\delta)c(t_E+\delta)c(t'_L-\delta)c(t_L-\delta) \\
&\quad \cdot\left(\mathcal{E}[n^*(t'_E)n(t_E)]\,\mathcal{E}[n^*(t'_L)n(t_L)] + \mathcal{E}[n^*(t'_E)n(t_L)]\,\mathcal{E}[n^*(t'_L)n(t_E)]\right) \\
&= \mathcal{N}_0^2\left(\int dt_E\,dt_L\,c^2(t_E+\delta)c^2(t_L-\delta) + \right. \\
&\quad \left. + \int dt_E\,dt_L\,c(t_L+\delta)c(t_E+\delta)c(t_E-\delta)c(t_L-\delta)\right) \\
&= 4\mathcal{N}_0^2 T_i^2\left(1 + R^2(2\delta)\right).
\end{aligned}$$

and similarly:

$$\mathcal{E}\left[|n_E|^4\right] = \mathcal{E}\left[|n_L|^4\right] = 8\mathcal{N}_0^2 T_i^2.$$

Using these results in Equation (6.41) leads to the following expression for the variance:

$$\begin{aligned}
\mathrm{var}[D(\Delta\tau)] &= 8\alpha^2\mathcal{N}_0 T_i R^2(\delta)\big(1 - R(2\delta)\big) \\
&\quad + 8\mathcal{N}_0^2 T_i^2\big(1 - R^2(2\delta)\big) + 8\alpha^2\mathcal{N}_0 T_i\left(\frac{\Delta\tau}{T_c}\right)^2\big(1 + R(2\delta)\big),
\end{aligned}$$

and by replacing $R(2\delta) = 1 - d$:

$$\mathrm{var}[D(\Delta\tau)] = 2\mathcal{N}_0 T_i(2-d)\left(\alpha^2 d(2-d) + 4\mathcal{N}_0 T_i d + 4\alpha^2\left(\frac{\Delta\tau}{T_c}\right)^2\right).$$

© 2008 by Christoph Günther, all rights reserved.

Using the relation between the discriminator output $D(\Delta\tau)$ and the estimate $\hat{\Delta}\tau$ of the delay from Equation (6.40), this allows finally to indicate the desired variance of the estimator:

$$\begin{aligned}\frac{\sigma^2_{\Delta\tau}}{T_c^2} &= \frac{\mathrm{var}[D(\Delta\tau)^2]}{4\alpha^4(2-d)^2} \\ &= \frac{d}{4\mathcal{E}_i/\mathcal{N}_0}\left(1+\frac{2}{\mathcal{E}_i/\mathcal{N}_0(2-d)}+\frac{4}{d(2-d)}\left(\frac{\Delta\tau}{T_c}\right)^2\right), \end{aligned} \tag{6.42}$$

with $\mathcal{E}_i = \mathcal{P}T_i$. This important relation characterizes the noise at the output of one of the most widely used discriminators. The variance is expressed relative to the chip duration. Thus a shorter chip duration, i.e., a wider bandwidth leads to a reduction of the noise in the delay estimation. Furthermore, the standard deviation σ_τ is inversely proportional to the square root of the signal-to-noise ratio. The signal-to-noise ratio must therefore be increased by a factor of four for reducing the standard deviation by a factor of 2. The same holds for the correlator spacing. The above equations were derived without including the influence of the bandwidth B_{fe} of the front-end. Betz and Kolodziejski [16] derived a more general result, which shows that Equation (6.42) holds for front-end bandwidths, which satisfy the condition

$$B_{\mathrm{fe}} \geq \frac{\pi}{dT_c}.$$

The first term in Equation (6.42) is also present with a coherent discriminator. The second term is additional and is called "squaring loss." This loss scales with the signal-to-noise ratio. It is only significant when the latter ratio is small. The last term vanishes at the operational point of the loop $\Delta\tau = 0$.

6.7.2 DLL Loop Performance

The generic performance of a loop was studied in detail in Section 6.5. The results can be applied immediately, and leads to the famous expression

$$\frac{\sigma^2_{\hat{\tau}}}{T_c^2} = \frac{dB_L}{2\mathcal{P}/\mathcal{N}_0}\left(1+\frac{2}{\mathcal{E}_i/\mathcal{N}_0(2-d)}\right). \tag{6.43}$$

for the noise of the delay estimation, with B_L denoting either the bandwidth of the first or second order loop. The loop introduces an new degree of freedom. It allows to reduce T_i, without degrading the performance much, as long as B_L is kept fixed. Filtering allows to limit the sensitivity of the integration to frequency errors. The squaring loss stays inversely proportional to T_i, which means that T_i should not be reduced too much. Values of $2B_LT_i$ in the order of 1/10 are quite common.

6.8 Carrier Tracking - Phase and Frequency Locked Loops

In the case of correct delay tracking, the prompt signal in Figure 6.9 is kept at the maximum of the autocorrelation function. This signal thus has the maximum signal to noise ratio. It is used for frequency and phase tracking. Evaluating Equation (6.13) with a delay controlled to $\Delta\tau \sim 0$ and a reasonable initial frequency estimate $\mathrm{sinc}(\Delta\omega\frac{T_i}{2}) \sim 1$ implies

$$C_{z,m} = \alpha e^{j\left(\Delta\omega\, m\, T_i - \omega\Delta\tau + \Delta\varphi\right)} + n_m. \tag{6.44}$$

The approach for constructing the PLL and FLL loops is very similar to the DLL case: discriminators are first defined, and analyzed. These results are used together with the analysis of generic parameter tracking for providing the results on the loop performance.

© 2008 by Christoph Günther, all rights reserved.

6.8.1 Discriminators for the PLL

The PLL estimates and controls the phase $\Phi_m = \Delta\omega\, m\, T_i - \omega\Delta\tau + \Delta\varphi$ of the signal. The estimation of the phase does not relate results associated with different integration periods. Therefore, the index m is not written in the present section. In this case, the pre-detection result reads:

$$C(\Phi) = \alpha e^{j\,\Phi} + n.$$

The first and most immediate discriminator is the argument of C:

$$D_{\mathrm{PLL},a}(\Phi) = \arg(\frac{C(\Phi)}{\alpha}).$$

This discriminator directly generates Φ over the range $\Phi \in [-\pi, \pi]$. Its weakness is that it requires the correct estimate of the sign of the navigation bit. If this sign is erroneous, the phase jumps by π and causes a cycle slip. Other discriminators with similar properties have been considered, like the real part divided by the magnitude. This discriminator has the characteristic $\sin\Phi$, and thus a very similar behavior at large signal to noise ratios.

In order to eliminate the sensitivity to the estimation of the navigation bits, non-linear discriminators are typically used. The most important such discriminator is:

$$\begin{aligned} D_{\mathrm{PLL},C}(\Phi) &= \frac{1}{\alpha^2}\Re[C(\Phi)]\Im[C(\Phi)] \\ &= \frac{1}{\alpha^2}(\alpha\cos\Phi + n_x)(\alpha\sin\Phi + n_y) \end{aligned} \tag{6.45}$$

It was proposed by Costas and is accessible to analysis. The expression for $D_{\mathrm{PLL},C}$ is expanded into:

$$D_{\mathrm{PLL},C} = \frac{1}{2}\sin(2\Phi) + \frac{1}{\alpha}\left(\cos(\Phi)n_y + \sin(\Phi)n_x\right) + \frac{1}{\alpha^2}n_x n_y.$$

Since the noise is zero mean $\mathcal{E}[D_{\mathrm{PLL},C}] = \sin(2\Phi)/2 \simeq \Phi$. Furthermore the variance can be evaluated in a similar fashion as for the DLL:

$$\begin{aligned} \mathrm{var}[D_{\mathrm{PLL},C}] &= \mathcal{E}[(D_{\mathrm{PLL},C} - \mathcal{E}[D_{\mathrm{PLL},C}])^2] \\ &= \frac{1}{\alpha^2}\left(\cos^2\Phi\,\mathcal{E}[n_y^2] + \sin^2\Phi\,\mathcal{E}[n_x^2]\right) + \frac{1}{\alpha^4}\mathcal{E}[n_x n_y n_x n_y] && (6.46) \\ &= \frac{T_i\mathcal{N}_0}{\alpha^2} + \frac{T_i^2\mathcal{N}_0^2}{\alpha^4}, && (6.47) \\ &= \frac{1}{2\mathcal{E}_i/\mathcal{N}_0}\left(1 + \frac{1}{2\mathcal{E}_i/\mathcal{N}_0}\right). && (6.48) \end{aligned}$$

Strictly speaking, this is the variance around the approximation $\sin(2\Phi)/2$ of Φ. The approximation is exact in $\Phi = 0$. This discriminator shows again a squaring loss. The periodicity of this discriminator determines the amplitude of the cycle slip, which is π for the Costas discriminator. With the exception of the argument discriminator, which has a periodicity of 2π, this property is shared by all other discriminators considered in the literature. The latter additionally include: $\Re(C)\mathrm{sign}(\Im(C))$, and $\mathrm{arctg}(\Re(C)/\Im(C))$. All of them are insensitive to data modulation, and more or less linear around $\Delta\phi = k\pi$, (see Figure 6.11). The Costas discriminator is easiest to analyze.

6.8.2 Discriminators for the FLL

The FLL controls the frequency during the initial tracking stage, before the PLL is activated. The FLL uses the correlation in the form of Equation (6.44). The discriminators are based on functions of $C_{z,m+1}C^*_{z,m}$, which has the expectation value:

$$\mathcal{E}[C_{z,m+1}C^*_{z,m}] = \alpha^2 e^{j\Delta\omega T_i}.$$

© 2008 by Christoph Günther, all rights reserved.

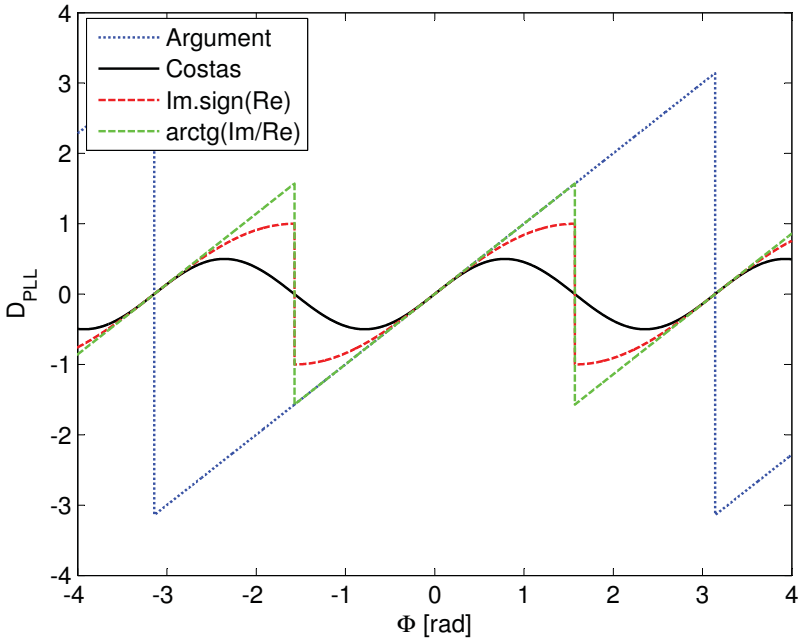

Figure 6.11: PLL-Discriminator functions in the interval $[-\pi, \pi]$, as a function of the phase offset. The argument discriminator has the broadest pull-in range and linear range.

The argument of this quantity is a natural function to consider:

$$D_{\mathrm{FLL},a}(\Delta\omega) = \frac{1}{T_i} \arg\left(\frac{C_{z,m+1}C^*_{z,m}}{\alpha^2}\right).$$

This discriminator is linear and generates corrections towards $\Delta\omega = 0$ for $\Delta\omega \in [-\pi/T_i, \pi/T_i]$. Note that this pull-in range cannot be used in practice, because the correlation result would be too much degraded, due to the presence of the sinc-factor in the complete expression for the correlation. A simpler discriminator that is often used, is again a Costas discriminator:

$$\begin{aligned} D_{\mathrm{FLL},C}(\Delta\omega) &= \frac{1}{\alpha^2 T_i}\Im\left(C_{z,m+1}C^*_{z,m}\right) \\ &= \frac{1}{T_i}\sin(\Delta\omega T_i) \\ &\quad + \frac{1}{\alpha T_i}\Big(-\cos\left(\Delta\omega(m+1)T_i + \Phi'\right) n_{m,y} + \sin\left(\Delta\omega(m+1)T_i + \Phi'\right) n_{m,x} \\ &\qquad\qquad -n_{m+1,x}\sin\left(\Delta\omega\, mT_i + \Phi'\right) + n_{m+1,y}\cos\left(\Delta\omega\, mT_i + \Phi'\right)\Big) \\ &\quad + \frac{1}{\alpha^2 T_i}\big(-n_{m+1,x}n_{m,y} + n_{m+1,y}n_{m,x}\big), \end{aligned}$$

with $\Phi' = \Delta\omega T_i/2 - \omega\Delta\tau + \Delta\varphi$. The range of linearity is smaller, but large enough to cover all frequency offsets that might still be present after initial acquisition. Most importantly, it is again easy to analyze. In a similar manner as before, one computes the normalized variance of the frequency estimate:

$$\left(\frac{\sigma_{\Delta f}}{1/T_i}\right)^2 = \frac{\mathrm{var}[D_{\mathrm{FLL},C}]}{1/T_i^2} = \frac{1}{\mathcal{E}_i/\mathcal{N}_0}\left(1 + \frac{1}{2\mathcal{E}_i/\mathcal{N}_0}\right).$$

© 2008 by Christoph Günther, all rights reserved.

The standard deviation of the frequency error scales like $1/T_i$, i.e.: long integration intervals lead to a smaller noise in the frequency estimate. A squaring loss is again present.

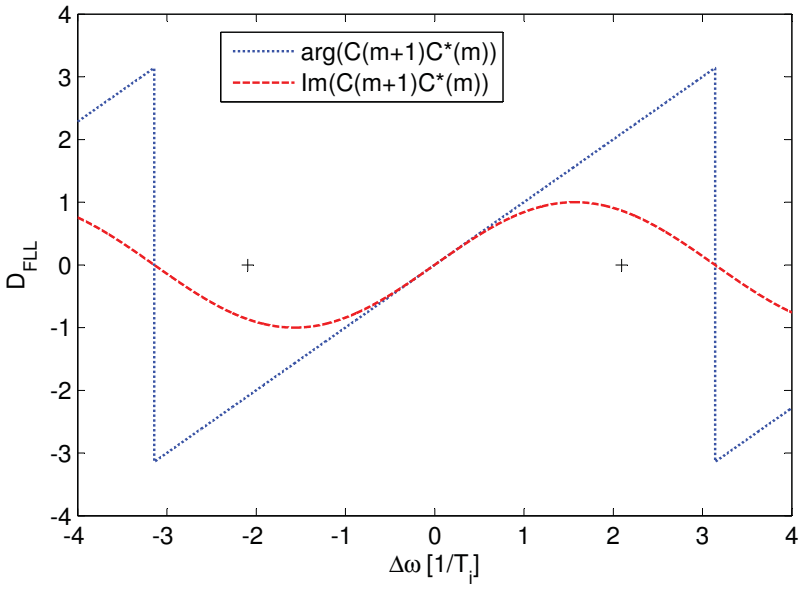

Figure 6.12: FLL-Discriminator functions as a function of the frequency offset in the interval $[-\pi/T_i, \pi/T_i]$. The argument discriminator has the broadest pull-in and linear range. The acquisition search ensures that the initial frequency is between the two +-signs.

6.8.3 FLL and PLL Loop Performance

With the results from Section 6.5, the FLL performance becomes

$$\left(\frac{\sigma_f}{1/T_i}\right)^2 = \frac{2B_L}{\mathcal{P}/\mathcal{N}_0}\left(1+\frac{1}{2\mathcal{E}_i/\mathcal{N}_0}\right).$$

After initialization, the mode of the loop is switched from FLL to PLL. From this time on, the PLL must track both the phase and the frequency. Enabling this implies that the PLL must at least be of second order.

$$(\sigma_\Phi)^2 = \frac{B_L^{(2)}}{\mathcal{P}/\mathcal{N}_0}\left(1+\frac{1}{2\mathcal{E}_i/\mathcal{N}_0}\right). \tag{6.49}$$

Comparing this latter result with the expression for the DLL in Equation (6.43), shows that both have essentially the same relative standard deviation. "Relative" means referred to the chip duration in the case of the DLL and referred to the wavelength in the case of the PLL. The chip duration (300 m) is, however, roughly 1500 times longer than a carrier cycle (20cm). The loop bandwidth B_L is slightly different for the various loops. Typical values are 0.2 Hz for a DLL, 2 Hz for a FLL, 10 Hz and for a PLL, which implies that one cannot realize the full gain. The ratio is still around a factor 300, which leads from ca. 70 cm to 2 mm. In order to take advantage of this potential, the ambiguity of the carrier phase estimation must be resolved and all other error sources must be controlled to millimeter level as well. This is discussed in Chapter 11 and 16.

© 2008 by Christoph Günther, all rights reserved.

Appendix A Computation of the Correlation

The purpose of the present appendix is to prove Equation (6.13). Consider Equation (6.12), the corresponding equation for $C_{y,m}$, and write the result in the form

$$C_{z,m} = \sqrt{2\,\mathcal{P}} \int_{(m-1/2)T_i}^{(m+1/2)T_i} dt\, c(t)\, c(t-\Delta\tau)\, e^{j\left(\Delta\omega t - \omega\Delta\tau + \Delta\varphi\right)} + n_{z,m}.$$

This can be reformulated using the following representation of t, and $\Delta\tau$:

$$\begin{array}{lll} t = s + nT_c + mT_i, & \text{with} & s \in [-\frac{T_c}{2}, \frac{T_c}{2}) \\ \Delta\tau = \sigma + \nu T_c, & \text{with} & \sigma \in [-\frac{T_c}{2}, \frac{T_c}{2}), \end{array}$$

which leads to the decomposition of the integral into:

$$\int_{(m-1/2)T_i}^{(m+1/2)T_i} dt = \sum_{n=-L}^{L} \int_{-T_c/2}^{T_c/2} ds,$$

with[10] $L = \frac{\kappa N - 1}{2}$, and the following reformulation of the expression for the complex correlation $C_{z,m}$:

$$\begin{aligned} C_{z,m} \;=\; & \sqrt{2\,\mathcal{P}} \sum_{n=-L}^{L} \int_{-T_c/2}^{T_c/2} ds\, e^{j\Delta\omega s} \sum_{n'=-\infty}^{\infty} \sum_{n''=-\infty}^{\infty} c_{n'} c_{n''}\, p\big(s + (n-n')T_c\big) \\ & \cdot p\big(s - \sigma + (n - n'' - \nu)T_c\big)\, e^{j\Delta\omega n T_c}\, e^{j(\Delta\omega m T_i - \omega\Delta\tau + \Delta\varphi)} + n_{z,m} \end{aligned}$$

Since the waveform $p(.)$ is zero outside the interval $[-T_c/2, T_c/2)$ (see Equation (5.17), the first factor $p(.)$ implies that $n' = n$, while the second together with the range restrictions for the variables leads to four cases. The first two cases assume $\sigma \geq 0$. Under this assumption, $s - \sigma \in [-T_c, T_c/2]$, and therefore $n - n'' - \nu = 0$ or 1, which leads to the following distinctions:

- $n'' = n - \nu$ implies that $-T_c/2 \leq s - \sigma \leq T_c/2$, and together with the integration limits $\sigma - T_c/2 \leq s \leq T_c/2$.
- $n'' = n - \nu - 1$ implies $-T_c/2 \leq s - \sigma + T_c \leq T_c/2$, and with the same consideration $-T_c/2 \leq s \leq \sigma - T_c/2$.

With a similar reasoning one obtains the following contributions in the case $\sigma < 0$:

- $n'' = n - \nu$, and $-T_c/2 \leq s \leq \sigma + T_c/2$.
- $n'' = n - \nu + 1$, and $\sigma + T_c/2 \leq s \leq T_c/2$.

For $\sigma > 0$, the expression for $C_{z,m}$ can be rewritten in the form:

$$\begin{aligned} C_{z,m} \;=\; & \sqrt{2\,\mathcal{P}} \left(\int_{\sigma - T_c/2}^{T_c/2} ds\, e^{j\Delta\omega s} \sum_{n=-L}^{L} c_n c_{n-\nu}\, e^{j\Delta\omega n T_c} \right. \\ & \left. + \int_{-T_c/2}^{\sigma - T_c/2} ds\, e^{j\Delta\omega s} \sum_{n=-L}^{L} c_n c_{n-\nu-1}\, e^{j\Delta\omega n T_c} \right) e^{j(\Delta\omega m T_i - \omega\Delta\tau + \Delta\varphi)} + n_{z,m}. \end{aligned}$$

A similar expression is obtained for $\sigma < 0$. In both cases:

$$\mathcal{E}[n_{z,m} n^*_{z,m'}] = 2\mathcal{N}_0 T_i\, \delta_{mm'}, \qquad \mathcal{E}[n_{z,m} n_{z,m'}] = \mathcal{E}[n^*_{z,m} n^*_{z,m'}] = 0.$$

[10] This assumes that κN is odd. The corresponding expressions for κN even are easy to derive.

© 2008 by Christoph Günther, all rights reserved.

Clearly, the in-phase correlation $\Delta\tau = 0$, leads to the maximal correlation. In that case,

$$\begin{aligned}\mathcal{E}[C_{z,m}] &= \sqrt{2P}\int_{-T_c/2}^{T_c/2} ds\, e^{j\Delta\omega s} \sum_{n=-L}^{L} e^{j\Delta\omega n T_c} e^{j(\Delta\omega m T_i - \Delta\phi)} \\ &= \sqrt{2P}\, T_i \operatorname{sinc}\left(\frac{\Delta\omega T_i}{2}\right) e^{j(\Delta\omega m T_i - \Delta\phi)},\end{aligned}$$

which is small unless $\Delta\omega T_i \lesssim 1$. This means that the phase should not rotate by more than a fraction of 2π over the integration interval. The correlation would be otherwise destroyed. The out-off code-phase correlations $\Delta\tau \neq 0$ can be estimated as follows:

$$\begin{aligned}|\mathcal{E}[C_{z,m}]| &= \left|\sqrt{2P}\int ds\, e^{j\Delta\omega s} \sum_{g=-(\kappa-1)/2}^{(\kappa-1)/2} e^{j\Delta\omega g N T_c} \sum_{\tilde{n}=-(N-1)/2}^{(N-1)/2} c_{\tilde{n}} c_{\tilde{n}-\nu}\, e^{j\Delta\omega n T_c}\right| \\ &\leq \sqrt{2P}\, T_c \cdot \kappa \cdot \left(\mathrm{O}\left(N^{1-\beta}\right) + \mathrm{O}\left(\Delta\omega N^2 T_c\right)\right) \qquad (6.50) \\ &= T_i\left(\mathrm{O}\left(\frac{1}{N^\beta}\right) + \mathrm{O}\left(\frac{\Delta\omega T_i}{\kappa}\right)\right).\end{aligned}$$

This is valid for all possible limits of the integral over s that are spaced by less than T_c, and all values of $\nu \neq 0$. Thus the only really interesting case is $\nu = 0$. first consider $\sigma > 0$:

$$\begin{aligned}C_{z,m} = \sqrt{2P}\Bigg(&\int_{\sigma-T_c/2}^{T_c/2} ds\, e^{j\Delta\omega s} \sum_{n=-L}^{L} e^{j\Delta\omega n T_c} \\ &+ \int_{-T_c/2}^{\sigma-T_c/2} ds\, e^{j\Delta\omega s} \sum_{n=-L}^{L} c_n c_{n-1}\, e^{j\Delta\omega n T_c}\Bigg) e^{j(\Delta\omega m T_i - \omega\Delta\tau + \Delta\varphi)} + n_{z,m}\end{aligned}$$

The second summand can be treated in a manner similar to Equation (6.50). The first summand must be evaluated. To this end, one computes:

$$\begin{aligned}\sum_{n=-L}^{L} e^{j\Delta\omega n T_c} &= \frac{\sin(\Delta\omega T_i/2)}{\sin(\Delta\omega T_c/2)} \\ &= \kappa N \frac{\operatorname{sinc}(\Delta\omega T_i/2)}{\operatorname{sinc}(\Delta\omega T_c/2)},\end{aligned}$$

and

$$\begin{aligned}\int_{\sigma-Tc/2}^{T_c/2} ds\, e^{j\Delta\omega s} &= T_c\, e^{j\Delta\omega\sigma/2}\left(1-\frac{\sigma}{T_c}\right)\operatorname{sinc}\left(\left(1-\frac{\sigma}{T_c}\right)\frac{\Delta\omega T_c}{2}\right) \\ &= T_c\left\{\left(1-\frac{\sigma}{T_c}\right)\operatorname{sinc}\left(\frac{\Delta\omega T_c}{2}\right) + \mathrm{O}\left(\Delta\omega T_c\right)\right\},\end{aligned}$$

as well as

$$\int_{-Tc/2}^{\sigma+T_c/2} ds\, e^{j\Delta\omega s} = T_c\left\{\left(1-\frac{-\sigma}{T_c}\right)\operatorname{sinc}\left(\frac{\Delta\omega T_c}{2}\right) + \mathrm{O}\left(\Delta\omega T_c\right)\right\}$$

for the case $\sigma < 0$. The combination of all terms then finally leads to

$$C_{z,m} = \alpha\left\{R(\Delta\tau)\operatorname{sinc}\left(\frac{\Delta\omega T_i}{2}\right) e^{j(m\Delta\omega T_i - \omega\Delta\tau + \Delta\varphi)} + \mathrm{O}\left(\frac{1}{N^\beta}\right) + \mathrm{O}(\frac{\Delta\omega T_i}{\kappa})\right\} + n_{z,m}.$$

This completes the proof of Equation (6.13).

© 2008 by Christoph Günther, all rights reserved.

Bibliography

[1] M. Lentmaier, B. Krach, P. Robertson, T. Thiasiriphet, "Dynamic Multipath Estimation by Sequential Monte-Calrlo Methods," *ION GNSS'07,* pp. 1712-1721, 2007.

[2] M. Abramowitz, I.A. Stegun, (Ed.) *Handbook of Mathematical Functions,* Ninth Ed., Dover Publ., Inc., New York, 1970.

[3] Marcum, J. I., "A Statistical Theory of Target Detection by Pulsed Radar: Mathematical Appendix," *RAND Corporation,* Santa Monica, CA, Research Memorandum RM-753, July 1, 1948. Reprinted in *IRE Trans. on Inform. Theory,* vol. IT-6, pp. 59267, Apr. 1960.

[4] A. Schmid, "Enhanced Sensitivity for Galileo and GPS Receivers," *PhD-Thesis*, TU-München, 2007.

[5] A. Schmid, A. Neubauer, "Performance Evaluation of Differential Correlation for Single Shot Measurement Positioning," *Proc. ION GNSS*, pp. 1998-2009, Sept. 2004.

[6] J.J. Spilker, "Fundamentals of Signal Tracking Theory," in *Global Positioning System, Theory and Applications,* vol. I, pp. 245-327, AIAA, Washington DC, 1996.

[7] M. Zhodzishsky, S. Yudanov, V. Veitsel, J. Ashjaee, "Co-Op Tracking for carrier phase," *Proc. of ION-GPS98,* pp. 653-664, Sep. 1998.

[8] K. Giger, P. Henkel, C. Günther, "Position-Domain Joint Satellite Tracking," *Proc. Navitec'10, ESA, Nordwijk*, pp. 1-8, Dec. 2010.

[9] K. Giger, P. Henkel, C. Günther, "Joint Satellite Code and Carrier Tracking," *Proc. ION ITM'10*, pp. 636-645, Jan. 2010.

[10] E.T. Whittaker, G.N. Watson, *A Course of Modern Analysis,* 4th Ed., Cambridge Univ. Press, 1927.

[11] L.V. Ahlfors, *Complex Analysis,* McGraw Hill, New York, 1953.

[12] W.C Lindsey, C.M. Chie, "A Survey of Digital Phase-Locked Loops," *Proc. IEEE,* vol. 69, pp. 410-431, 1981.

[13] Ph.W. Ward, "Performance Comparisons Between FLL, PLL and a Novel FLL-Assisted-PLL Carrier Tracking Loop under RF Interference Conditions, *Proc. ION GPS'98,* pp. 783-795, 1998.

[14] `http:/\!/maxima.sourceforge.net`

[15] A.J. Van Dierendonck, P. Fenton, T. Ford, "Theory and Performance of Narrow Correlator Spacing in a GPS Receiver," *J. Inst. of Navigation,* vol. 39, pp. 115-124, Fall 1992.

[16] J.W. Betz, K.R. Kolodziejski, "Extended Theory of Early-Late Code Tracking for a Bandlimited GPS Receiver,"

© 2008 by Christoph Günther, all rights reserved.