Contents

The Cost Function

In this very first chapter, we derive the cost function for our PLL implementation. We will see, that the application of Gaussian message passing will come up very naturally, when it comes to minimizing this cost function. The system in which we try to track the phase can generally be described by an autonomus state space model

$$\begin{aligned} x_k &= \mathbf{A} x_{k-1} \\ y_k &= \mathbf{C} x_k + Z_k. \end{aligned} \tag{1}$$

The system parameters A, C and Z_k are given by

$$\begin{aligned} &\mathbf{A} = R(\omega_0) = \begin{bmatrix} \cos(\omega_0) & -\sin(\omega_0) \\ \sin(\omega_0) & \cos(\omega_0) \end{bmatrix}, \\ &\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \\ &Z_k \overset{i.i.d}{\sim} \mathcal{N}\left(0, \sigma^2\right). \end{aligned}$$

Now assume that at time K, we observed the outputs $y_1, y_2, \dots, y_K$. With these measurements, we want to find estimates $\tilde{y}, \tilde{y}_2, \dots, \tilde{y}_K$, that provide the best fit in terms of the squared error. Our cost function e^2 can therefore be described as follows

$$e^2 = \sum_{k=1}^{K} \left(y_k - \tilde{y}_k\right)^2. \tag{2}$$

This cost function is reasonable, because Z_k is modeled as white Gaussian noise. The estimates $\tilde{y}_k$ however, are subject to the physical laws imposed by the state space model in equation (1). Hence it is enough to compute the current state estimate denoted by $\hat{x}_K$. The estimated outputs are then given by

$$\tilde{y}_k = \mathbf{C}\mathbf{A}^{-(K-k)}\tilde{x}_K. \tag{3}$$

Using this relation in equation (2) yields

$$e^2\left(\tilde{x}_K\right) = \sum_{k=1}^{K}\left(y_k - \mathbf{C}\mathbf{A}^{-(K-k)}\tilde{x}_K\right)^2.$$

Note the dependence on $\tilde{x}_K$ only. We additionally introduce a so-called decay factor γ that puts less weight on past measurements. It is defined more formally in the next chapter. The final cost function then has the form

$$e^2\left(\tilde{x}_K\right) = \sum_{k=1}^{K}\gamma^{K-k}\left(y_k - \mathbf{C}\mathbf{A}^{-(K-k)}\tilde{x}_K\right)^2. \tag{4}$$

If $\tilde{x}_K$ is optimal, equation (4) is minimal. Hence $\tilde{x}_K$ can be found by minimizing e^2

$$\begin{aligned}
\tilde{x}_K &= \arg\min_x\left(\sum_{k=1}^{K}\gamma^{K-k}\left(y_k - \mathbf{C}\mathbf{A}^{-(K-k)}x\right)^2\right) \\
&= \arg\max_x\left(\prod_{k=1}^{K}e^{-\gamma^{K-k}\left(y_k-\mathbf{C}\mathbf{A}^{-(K-k)}x\right)^2}\right) \\
&= \arg\max_x\left(\prod_{k=1}^{K}e^{-\left(\sqrt{\gamma}^{K-k}y_k-\mathbf{C}\left(\sqrt{\gamma}\mathbf{A}\right)^{-(K-k)}x\right)^2}\right) \\
&= \arg\max_x\left(\prod_{k=1}^{K}e^{-\frac{z_k^2}{2\sigma^2}}\right),
\end{aligned} \tag{5}$$

where we used the noisy output relation of equation (1). On closer examination of equation (5), it turns out that our minimization problem can be formulated in terms of Gaussian sum-product message passing. The corresponding factor graph looks as depicted in Figure 1

With forward message passing on this factor graph, we can find the estimate $\hat{x}_K$ at time K that minimizes the squared error $e^2\left(\tilde{x}_K\right)$. Calculation of the involved forward messages is done in the next chapter.

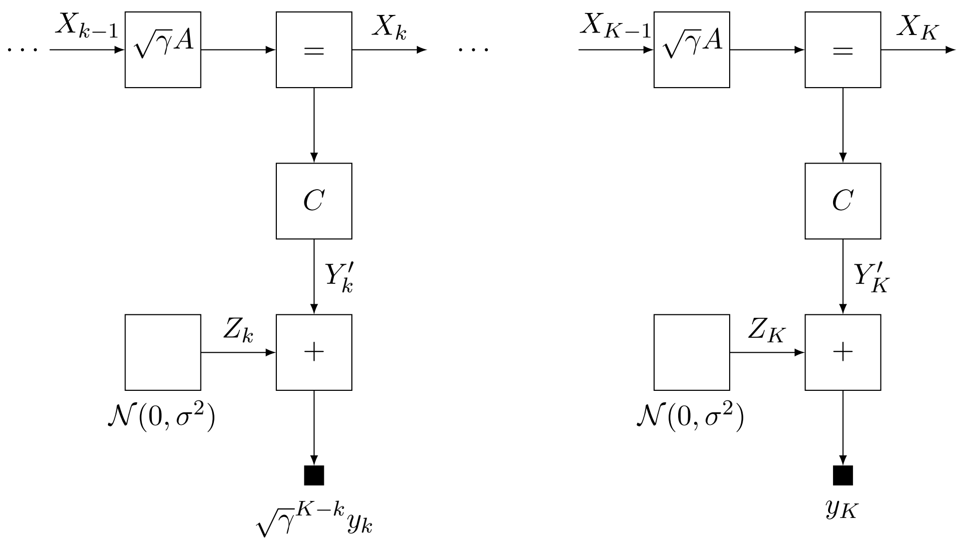

Figure 1: Factor graph representation of minimization problem in equation 4.

Derivation of Update Rules for Kalman-Based PLL

In order to facilitate the computations, we set $\gamma = 1$. It will be reintroduced later. In Figure 2 we can see a factor graph representation of a Kalman-based phase-locked loop where $\gamma = 1$.

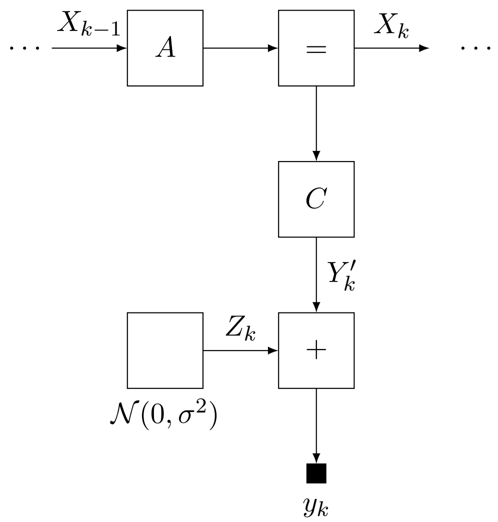

Figure 2: Factor graph representation of a Kalman filter's kth cell.

Since Z_k is a Gaussian random variable it follows that all random variables in the factor graph are Gaussian. As stated before, this gives rise to Gaussian message passing, i.e., all messages in the factor graph can be described by their mean vectors and covariance matrices.

We start the message passing algorithm by computing the message given by the observed samples y_k [1]

$$\overrightarrow{m}_{Z_k} = 0 \qquad \overrightarrow{V}_{Z_k} = \sigma^2 \tag{6}$$
$$\overleftarrow{m}_{Y_k} = y_k \qquad \overleftarrow{V}_{Y_k} = 0. \tag{7}$$

In a next step, we add the noise in Eq. (6) to the observed sample in Eq. (7) to get the message at Y'_k

$$\overleftarrow{m}_{Y'_k} = y_k \qquad \overleftarrow{V}_{Y'_k} = \sigma^2. \tag{8}$$

With these results, we can now compute the messages at the equality constraint

$$\overrightarrow{m}_{X''_k} = \mathbf{A}\overrightarrow{m}_{X_{k-1}} \qquad \overrightarrow{V}_{X''_k} = \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T \tag{9}$$
$$\overleftarrow{W}_{X'_k} = \mathbf{C}^T\overleftarrow{W}_{Y'_k}\mathbf{C} \qquad \overleftarrow{W}_{X'_k}\overleftarrow{m}_{X'_k} = \mathbf{C}^T\overleftarrow{W}_{Y'_k}\overleftarrow{m}_{Y'_k}, \tag{10}$$

where $\overleftarrow{W}_{X'_k}$ denotes the precision matrix with the following equality

$$\overrightarrow{W}^{-1}_{X_k} = \overrightarrow{V}_{X_k},$$

and $\overleftarrow{W}_{X'_k}\overleftarrow{m}_{X'_k}$ denotes the weighted mean. The set of equations in (10) can be simplified by using the set of equations in (8)

$$\overleftarrow{W}_{X'_k} = \mathbf{C}^T\frac{1}{\sigma^2}\mathbf{C} \qquad \overleftarrow{W}_{X'_k}\overleftarrow{m}_{X'_k} = \mathbf{C}^T\frac{1}{\sigma^2}y_k.$$

Hence we can characterize X_k by its precision matrix and its weighted mean (update rules). Note that $\mathbf{A}$ is an invertible matrix since

$$\begin{aligned}
\overrightarrow{W}_{X_k} &= \overrightarrow{W}_{X''_k} + \overleftarrow{W}_{X'_k} && (11)\\
&= \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1} + \mathbf{C}^T\frac{1}{\sigma^2}\mathbf{C} && (12)\\
&= \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1} + \frac{1}{\sigma^2}\mathbf{C}^T\mathbf{C} && (13)\\
&= \mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1} + \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2} && (14)\\
&= \mathbf{A}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1} + \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2}, && (15)
\end{aligned}$$

$$
\begin{aligned}
\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k} &= \overrightarrow{W}_{X''_k}\overrightarrow{m}_{X''_k} + \overrightarrow{W}_{X'_k}\overrightarrow{m}_{X'_k} = && (16)\\
&= \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1}\mathbf{A}\overrightarrow{m}_{X_{k-1}} + \mathbf{C}^T\frac{1}{\sigma^2}y_k && (17)\\
&= \left(\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1}\underbrace{\mathbf{A}^{-1}\mathbf{A}}_{\mathbf{I}}\overrightarrow{m}_{X_{k-1}} + \mathbf{C}^T\frac{1}{\sigma^2}y_k && (18)\\
&= \mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}\overrightarrow{m}_{X_{k-1}} + \frac{\mathbf{C}^T}{\sigma^2}y_k && (19)\\
&= \mathbf{A}\overrightarrow{W}_{X_{k-1}}\overrightarrow{m}_{X_{k-1}} + \frac{\mathbf{C}^T}{\sigma^2}y_k, && (20)
\end{aligned}
$$

where the last step in the particular update equations follows from the fact that $\mathbf{A}$ is the rotation matrix and therefore an orthogonal matrix, i.e., its transpose is equal to its inverse

$$\mathbf{A}^T = \mathbf{A}^{-1}.$$

What is still missing in this implementation, is the ability of our estimator to react to abrupt signal changes, that are not modeled with our state space model. Such changes serve the purpose to physically transmit information with a waveform (e.g., Phase Shift Keying (PSK)). This adaptive ability is modeled with the *decay factor* $\gamma \in (0,1)$ that we've already met in the previous chapter. In every iteration step k, $\overrightarrow{W}_{X_{k-1}}$ is scaled with γ which increases the uncertainty of past estimates and therefore puts more emphasis on new measurements. The new update rules are then given by

$$\overrightarrow{W}_{X_k} = \gamma\mathbf{A}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1} + \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2}, \tag{21}$$

and

$$\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k} = \gamma\mathbf{A}\overrightarrow{W}_{X_{k-1}}\overrightarrow{m}_{X_{k-1}} + \frac{\mathbf{C}^T}{\sigma^2}y_k. \tag{22}$$

In a next step we try to get to an expression for the covariance matrix and the mean vector by using the Matrix Inversion Lemma (23)

$$(\mathbf{B}+\mathbf{DEF})^{-1} = \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{D}\left(\mathbf{E}^{-1}+\mathbf{FB}^{-1}\mathbf{D}\right)^{-1}\mathbf{FB}^{-1}. \tag{23}$$

Thus we get the following assignments from equation (21)

$$
\begin{align}
\mathbf{B} &= \gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1}, \tag{24}\\
\mathbf{D} &= \mathbf{C}^{T}, \tag{25}\\
\mathbf{E} &= \frac{1}{\sigma^2}, \tag{26}\\
\mathbf{F} &= \mathbf{C}. \tag{27}
\end{align}
$$

Using these, the inverse of equation (11), i.e., the covariance matrix can be written as

$$
\begin{align}
\overrightarrow{V}_{X_k} &= \left[\gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} + \frac{\mathbf{C}^T \mathbf{C}}{\sigma^2}\right]^{-1} \tag{28}\\
&= \left(\gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1}\right)^{-1} \nonumber\\
&- \left(\gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1}\right)^{-1} \mathbf{C}^T \left(\sigma^2 + \mathbf{C} \left(\gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1}\right)^{-1} \mathbf{C}^T\right)^{-1} \mathbf{C} \left(\gamma \mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1}\right)^{-1} \tag{29}\\
&= \frac{1}{\gamma} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T - \frac{1}{\gamma^2} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T \left(\sigma^2 + \frac{1}{\gamma} \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T\right)^{-1} \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \nonumber\\
&= \frac{1}{\gamma} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T - \frac{1}{\gamma} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T \underbrace{\left(\gamma\sigma^2 + \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T\right)^{-1}}_{=:G} \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \tag{30}\\
&= \frac{1}{\gamma} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T - \frac{1}{\gamma} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T G \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \tag{31}
\end{align}
$$

where

$$
G := \left(\gamma\sigma^2 + \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T\right)^{-1}.
$$

The mean vector $\overrightarrow{m}_{X_k}$ can be retrieved by multiplying the matrix $\overrightarrow{V}_{X_k}$ with (22)

$$
\begin{aligned}
\overrightarrow{m}_{X_k} &= \overrightarrow{V}_{X_k}\left(\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k}\right) && (32)\\
&= \frac{1}{\gamma}\left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T - \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right) && (33)\\
&\quad \cdot\left(\gamma\mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}\overrightarrow{m}_{X_{k-1}} + \frac{\mathbf{C}^T}{\sigma^2}y_k\right)\\
&= \mathbf{A}\underbrace{\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}}_{\mathbf{I}}\overrightarrow{m}_{X_{k-1}} - \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\mathbf{C}\mathbf{A}\underbrace{\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}}_{\mathbf{I}}\overrightarrow{m}_{X_{k-1}} && (34)\\
&\quad + \frac{1}{\gamma}\left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T - \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)\cdot\left(\frac{\mathbf{C}^T}{\sigma^2}y_k\right)\\
&= \mathbf{A}\overrightarrow{m}_{X_{k-1}} - \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\mathbf{C}\mathbf{A}\overrightarrow{m}_{X_{k-1}} && (35)\\
&\quad + \frac{1}{\gamma}\left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T - \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\cdot\left(\frac{1}{\sigma^2}\right)\cdot y_k\\
&= \mathbf{A}\overrightarrow{m}_{X_{k-1}} - \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\right)\mathbf{C}\mathbf{A}\overrightarrow{m}_{X_{k-1}} && (36)\\
&\quad + \frac{1}{\gamma}\left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\underbrace{\left(\mathbf{I} - G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\cdot\left(\frac{1}{\sigma^2}\right)}_{=:\lambda}\cdot y_k.
\end{aligned}
$$

The factor λ can further be simplified to

$$
\begin{aligned}
\lambda &:= \left(\mathbf{I} - G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\cdot\left(\frac{1}{\sigma^2}\right) && (37)\\
&= \left(GG^{-1} - G\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\cdot\left(\frac{1}{\sigma^2}\right) && (38)\\
&= G\left(G^{-1} - \mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)\cdot\left(\frac{1}{\sigma^2}\right) && (39)\\
&= G\left(\gamma\sigma^2 + \underbrace{\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T - \mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T}_{=0}\right)\cdot\left(\frac{1}{\sigma^2}\right) && (40)\\
&= G\left(\gamma\sigma^2\right)\cdot\left(\frac{1}{\sigma^2}\right) && (41)\\
&= \gamma G, && (42)
\end{aligned}
$$

where we used the fact, that $G^{-1} = \gamma\sigma^2 + \mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T$. So we finally find the following condensed expression for $\overrightarrow{m}_{X_k}$

$$
\begin{aligned}
\overrightarrow{m}_{X_k} &= \mathbf{A}\overrightarrow{m}_{X_{k-1}} - \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\right)\mathbf{C}\mathbf{A}\overrightarrow{m}_{X_{k-1}} + \frac{1}{\gamma}\left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\gamma G\right) y_k \quad (43)\\
&= \mathbf{A}\overrightarrow{m}_{X_{k-1}} + \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\left(y_k - \mathbf{C}\mathbf{A}\overrightarrow{m}_{X_{k-1}}\right). \quad (44)
\end{aligned}
$$

Amplitude and Phase Estimate and Cost Function

The predicted signal estimate given the measurements $y_1, ..., y_{k-1}$ can be written as

$$\tilde{y}_k = \mathbf{CA}\overrightarrow{m}_{X_{k-1}},$$

and the corrected estimate given the measurements $y_1, ..., y_k$ as

$$\hat{y}_k = \mathbf{C}\overrightarrow{m}_{X_{k-1}}.$$

The difference between the corrected and the predicted estimate can be deduced as

$$\begin{aligned}
\hat{y}_k - \tilde{y}_k &= \mathbf{C}\left(\overrightarrow{m}_{X_k} - \mathbf{A}\overrightarrow{m}_{X_{k-1}}\right) && (45)\\
&= \mathbf{C}\left(\overrightarrow{m}_{X_k} - \overrightarrow{m}_{X_k} + \underbrace{\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G}_{=:\mathbf{B}_k}\left(y_k - \mathbf{CA}\overrightarrow{m}_{X_{k-1}}\right)\right)\\
&= \mathbf{CB}_k\left(y_k - \tilde{y}_k\right) && (46)
\end{aligned}$$

Since both the input signal and estimate is known to be sinusoidal the mean vector at time index k can be written as [2]

$$\overrightarrow{m}_{X_k} = \hat{a}_k \begin{bmatrix} \cos\left(\omega_0 k + \hat{\phi}_k\right) \\ \sin\left(\omega_0 k + \hat{\phi}_k\right) \end{bmatrix}$$

and subsequently

$$y_k = a\cos(\omega_0 k + \phi) + Z_k, \qquad Z_k \overset{i.i.d}{\sim} \mathcal{N}\left(0, \sigma^2\right), \tag{47}$$

$$\tilde{y}_k = \hat{a}_{k-1}\cos\left(\omega_0 k + \hat{\phi}_{k-1}\right), \tag{48}$$

$$\hat{y}_k = \hat{a}_k\cos\left(\omega_0 k + \hat{\phi}_k\right) \tag{49}$$

Now we consider the scenario where only the phase is tracked and the amplitude is assumed to be $\hat{a}_k = \hat{a}_{k-1} = a_k$. If the phase estimation error is small such that $\left|\hat{\phi}_k - \phi\right| \ll 1$ and $\left|\hat{\phi}_{k-1} - \phi\right| \ll 1$, equation (46) can be reduced to

$$a\left(\cos\left(\omega_0 k + \hat{\phi}_k\right) - \cos\left(\omega_0 k + \hat{\phi}_{k-1}\right)\right) = a\alpha_k\left(\cos\left(\omega_0 k + \hat{\phi}_k\right) - \cos\left(\omega_0 k + \hat{\phi}_{k-1}\right)\right) \tag{50}$$

$$\sin\left(\frac{\hat{\phi}_k - \hat{\phi}_{k-1}}{2}\right) \approx \alpha_k \sin\left(\frac{\phi - \hat{\phi}_{k-1}}{2}\right) \tag{51}$$

$$\hat{\phi}_k - \hat{\phi}_{k-1} \approx \alpha_k\phi - \hat{\phi}_{k-1}. \tag{52}$$

Thus follows from equation (52) that the phase update can be approximated as

$$\hat{\phi}_k \approx \alpha_k\phi + (1 - \alpha_k)\,\hat{\phi}_{k-1}. \tag{53}$$

In a similar way, the amplitudes can be extracted under the condition that $\hat{\phi}_k = \hat{\phi}_{k-1} = \phi$

$$\hat{a}_k\cos(\omega_0 k + \phi) - \hat{a}_{k-1}\cos(\omega_o k + \phi) = \alpha_k\left(a\cos(\omega_0 k + \phi) - \hat{a}_{k-1}\cos(\omega_o k + \phi)\right) \tag{54}$$

$$\hat{a}_k - \hat{a}_{k-1} = \alpha_k\left(a - \hat{a}_{k-1}\right) \tag{55}$$

$$\hat{a}_k = \alpha_k a + (1 - \alpha_k)\,\hat{a}_{k-1}. \tag{56}$$

The cost function for the PLL then can be written as follows

$$J = \mathbb{E}\left[\right] \tag{57}$$

Independence of State Update from Noise

The update rules in Equations (21) and (22) are still iterative methods to compute the precision matrix and the weighted mean vector. Equations (58) and (59) provide explicit formulas to compute the $\overrightarrow{W}_{X_k}$ and $\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k}$ respectively

$$\overrightarrow{W}_{X_k} = \frac{1}{\sigma^2}\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T\mathbf{C}\left(\mathbf{A}^{-1}\right)^j, \tag{58}$$

$$\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k} = \frac{1}{\sigma^2}\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T y_{k-j}. \tag{59}$$

The equality of (58) and (21) as well as the equality of (59) and (22) can easily be verified by induction. The independence of $\overrightarrow{m}_{X_k}$ from the output noise variance σ^2 can now be proved by combining the two equations, which yields

$$\begin{aligned}\overrightarrow{m}_{X_k} &= \left(\overrightarrow{W}_{X_k}\right)^{-1}\left(\overrightarrow{W}_{X_k}\overrightarrow{m}_{X_k}\right) && (60)\\ &= \left(\frac{1}{\sigma^2}\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T\mathbf{C}\left(\mathbf{A}^{-1}\right)^j,\right)^{-1}\cdot\left(\frac{1}{\sigma^2}\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T y_{k-j}\right) && (61)\\ &= \left(\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T\mathbf{C}\left(\mathbf{A}^{-1}\right)^j,\right)^{-1}\cdot\left(\sum_{j=0}^{k-1}\gamma^j\left(\mathbf{A}^{-T}\right)^j\mathbf{C}^T y_{k-j}\right) && (62)\end{aligned}$$

As we can see, the noise σ^2 cancels out in the calculation above, i.e. our estimator does not need any information about the Gaussian noise process.

Steady State Precision Matrix

1^{st} case: Decay factor $0 < \gamma < 1$

Since $\mathbf{A}$ is a rotation matrix it can be diagonalized in an orthonormal basis in $\mathcal{C}$ such that

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H, \tag{63}$$

where $\mathbf{Q}$ is a unitary matrix and $\mathbf{Q}^H$ denotes its Hermitian transpose. The decomposition of the matrix $\mathbf{A}$ can be done as follows

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix},$$
$$\boldsymbol{\Lambda} = \begin{bmatrix} \exp(i\omega_0) & 0 \\ 0 & \exp(-i\omega_0) \end{bmatrix}.$$

In a next step, the new expression for $\mathbf{A}$ can be inserted into equation (58) [2]

$$
\begin{aligned}
\overrightarrow{W}_{X_k} &= \sum_{l=0}^{k-1}\left(\gamma^l\left(\left(\mathbf{Q\Lambda Q}^H\right)^H\right)^{-l}\frac{\mathbf{C}^T\mathbf{C}}{\sigma^2}\left(\mathbf{Q\Lambda Q}^H\right)^{-l}\right) \\
&= \frac{1}{\sigma^2}\sum_{l=0}^{k-1}\left(\gamma^l\mathbf{Q}\left(\bar{\mathbf{\Lambda}}\right)^{-l}\mathbf{Q}^H\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\mathbf{Q\Lambda}^{-l}\mathbf{Q}^H\right) \qquad (64)\\
&= \frac{1}{2\sigma^2}\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\left(\bar{\mathbf{\Lambda}}\right)^{-l}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\mathbf{\Lambda}^{-l}\right)\mathbf{Q}^H \\
&= \frac{1}{2\sigma^2}\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\begin{bmatrix}\exp\left(i\omega_0 l\right) & 0\\ 0 & \exp\left(-i\omega_0 l\right)\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}\exp\left(-i\omega_0 l\right) & 0\\ 0 & \exp\left(i\omega_0 l\right)\end{bmatrix}\right)\mathbf{Q}^H \\
&= \frac{1}{2\sigma^2}\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\begin{bmatrix}\exp\left(i\omega_0 l\right) & \exp\left(i\omega_0 l\right)\\ \exp\left(-i\omega_0 l\right) & \exp\left(-i\omega_0 l\right)\end{bmatrix}\begin{bmatrix}\exp\left(-i\omega_0 l\right) & 0\\ 0 & \exp\left(i\omega_0 l\right)\end{bmatrix}\right)\mathbf{Q}^H \\
&= \frac{1}{2\sigma^2}\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\begin{bmatrix}1 & \exp\left(2i\omega_0 l\right)\\ \exp\left(-2i\omega_0 l\right) & 1\end{bmatrix}\right)\mathbf{Q}^H \\
&= \frac{1}{2\sigma^2}\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\gamma^l\begin{bmatrix}0 & \exp\left(2i\omega_0 l\right)\\ \exp\left(-2i\omega_0 l\right) & 0\end{bmatrix}\right)\mathbf{Q}^H \\
&= \frac{1}{2\sigma^2}\left(\frac{1-\gamma^k}{1-\gamma}\mathbf{I}_2+\mathbf{Q}\sum_{l=0}^{k-1}\left(\gamma^l\begin{bmatrix}0 & \exp\left(2i\omega_0 l\right)\\ \exp\left(-2i\omega_0 l\right) & 0\end{bmatrix}\right)\mathbf{Q}^H\right), \qquad (65)
\end{aligned}
$$

where equation (64) follows from the unitary property of $\mathbf{Q}$. Moreover, we used the geometric series in equation (65)

$$
\sum_{n=0}^{n-1} ar^k = a\frac{1-r^n}{1-r},
$$

where we implicitly assumed that $0 < \gamma < 1$. Now we rewrite equation (65) as follows

$$
\begin{aligned}
\sum_{l=0}^{k-1} \gamma^l \exp\left(2i\omega_0 l\right) &= \sum_{l=0}^{k-1} \left(\gamma \exp\left(2i\omega_0\right)\right)^l && (66)\\
&= \frac{1-\gamma^k \exp\left(2i\omega_0 k\right)}{1-\gamma \exp\left(2i\omega_0\right)}\\
&= \frac{\left(1-\gamma^k\right) + \gamma^k - \gamma^k \exp\left(2i\omega_0 k\right)}{(1-\gamma) + \gamma - \gamma \exp\left(2i\omega_0\right)}\\
&= \frac{\left(1-\gamma^k\right) - \gamma^k \exp\left(i\omega_0 k\right)\left(\exp\left(i\omega_0 k\right) - \exp\left(-i\omega_0 k\right)\right)}{(1-\gamma) - \gamma \exp\left(i\omega_0\right)\left(\exp\left(i\omega_0\right) - \exp\left(-i\omega_0\right)\right)}\\
&= \frac{\left(1-\gamma^k\right) - 2\gamma^k \exp\left(i\omega_0 k\right)\sin\left(\omega_0 k\right)}{(1-\gamma) - 2\gamma \exp\left(i\omega_0\right)\sin\left(\omega_0\right)} && (67)
\end{aligned}
$$

where we made again use of the geometric series and the following identity for the sine (derived from Euler's formula) (see equation (67))

$$
\sin(x) = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right).
$$

Similarly we get

$$
\begin{aligned}
\sum_{j=0}^{k-1} \gamma^l \exp\left(-2i\omega_0 j\right) &= \frac{1-\gamma^k \exp\left(-2i\omega_0 k\right)}{1-\gamma \exp\left(-2i\omega_0\right)} && (68)\\
&= \frac{\left(1-\gamma^k\right) + 2\gamma^k \exp\left(-i\omega_0 k\right)\sin\left(\omega_0 k\right)}{(1-\gamma) + 2\gamma \exp\left(-i\omega_0\right)\sin\left(\omega_0\right)}. && (69)
\end{aligned}
$$

We are now interested in the long-term behavior of the precision matrix $\overrightarrow{W}_{X_k}$, i.e., in the special case where $k \to \infty$. We will call the resulting matrix $\overrightarrow{W}_{ss}$ steady state precision matrix. It is formally defined as follows

$$
\overrightarrow{W}_{ss} := \lim_{k\to\infty}\left(\overrightarrow{W}_{X_k}\right).
$$

As k tends to infinity, the second term of equation (65) simplifies to

$$
\lim_{k\to\infty} \mathbf{Q} \sum_{l=0}^{k-1} \left(\gamma^l \begin{bmatrix} 0 & \exp(2i\omega_0 l) \\ \exp(-2i\omega_0 l) & 0 \end{bmatrix} \right) \mathbf{Q}^H \tag{70}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \begin{bmatrix} 0 & (1-\gamma\exp(2i\omega_0))^{-1} \\ (1-\gamma\exp(-2i\omega_0))^{-1} & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} (1-\gamma\exp(-2i\omega_0))^{-1} & (1-\gamma\exp(2i\omega_0))^{-1} \\ i(1-\gamma\exp(-2i\omega_0))^{-1} & -i(1-\gamma\exp(2i\omega_0))^{-1} \end{bmatrix} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \\
&= \frac{1}{2} \frac{1}{(1-\gamma\exp(2i\omega_0))(1-\gamma\exp(-2i\omega_0))} \\
&\cdot \begin{bmatrix} 2-\gamma(\exp(2i\omega_0)+\exp(-2i\omega_0)) & -i\gamma(\exp(2i\omega_0)-\exp(-2i\omega_0)) \\ -i\gamma(\exp(2i\omega_0)-\exp(-2i\omega_0)) & -2+\gamma(\exp(2i\omega_0)+\exp(-2i\omega_0)) \end{bmatrix} \\
&= \frac{1}{2} \frac{1}{1+\gamma^2-2\gamma\cos(2\omega_0)} \begin{bmatrix} 2(1-\gamma\cos(2\omega_0)) & 2\gamma\sin(2\omega_0) \\ 2\gamma\sin(2\omega_0) & -2(1-\gamma\cos(2\omega_0)) \end{bmatrix}.
\end{aligned} \tag{71}
$$

The steady state precision matrix is therefore given by

$$
\overrightarrow{W}_{ss} = \frac{1}{2\sigma^2} \left(\frac{1}{1-\gamma} \mathbf{I}_2 + \frac{1}{1+\gamma^2-2\gamma\cos(2\omega_0)} \begin{bmatrix} 1-\gamma\cos(2\omega_0) & \gamma\sin(2\omega_0) \\ \gamma\sin(2\omega_0) & -1+\gamma\cos(2\omega_0) \end{bmatrix} \right).
$$

2nd case: Decay factor $\gamma = 1$

For the special case where $\gamma = 1$, equations (67) and (68) simplify to

$$
\sum_{l=0}^{k-1} \exp(2i\omega_0 l) = \exp(i\omega_0(k-1)) \frac{\sin(\omega_0 k)}{\sin(\omega_0)}
$$

and

$$
\sum_{l=0}^{k-1} \exp(2i\omega_0 l) = \exp(-i\omega_0(k-1)) \frac{\sin(\omega_0 k)}{\sin(\omega_0)}
$$

respectively. The identical calculation as in 70 then yields

$$
\begin{aligned}
&\mathbf{Q}\sum_{l=0}^{k-1}\left(\begin{bmatrix}0 & \exp\left(2i\omega_0 l\right)\\ \exp\left(-2i\omega_0 l\right) & 0\end{bmatrix}\right)\mathbf{Q}^H\\
&=\frac{\sin\left(\omega_0 k\right)}{2\sin\left(\omega_0\right)}\begin{bmatrix}1 & 1\\ -i & i\end{bmatrix}\begin{bmatrix}0 & \exp\left(i\omega_0\left(k-1\right)\right)\\ \exp\left(-i\omega_0\left(k-1\right)\right) & 0\end{bmatrix}\begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}\\
&=\frac{\sin\left(\omega_0 k\right)}{2\sin\left(\omega_0\right)}\begin{bmatrix}\exp\left(-i\omega_0\left(k-1\right)\right) & \exp\left(i\omega_0\left(k-1\right)\right)\\ i\exp\left(-i\omega_0\left(k-1\right)\right) & -i\exp\left(i\omega_0\left(k-1\right)\right)\end{bmatrix}\begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}\\
&=\frac{\sin\left(\omega_0 k\right)}{2\sin\left(\omega_0\right)}\begin{bmatrix}2\cos\left(\omega_0\left(k-1\right)\right) & 2\sin\left(\omega_0\left(k-1\right)\right)\\ 2\sin\left(\omega_0\left(k-1\right)\right) & -2\cos\left(\omega_0\left(k-1\right)\right)\end{bmatrix}\\
&=\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\mathbf{A}^{k-1}\mathbf{S} \qquad (72)
\end{aligned}
$$

where $\mathbf{S}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$. The precision matrix is now reduced to

$$
\overrightarrow{W}_{X_k}=\frac{1}{2\sigma^2}\left(k\mathbf{I}_2+\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\mathbf{A}^{k-1}\mathbf{S}\right).
$$

As we can see, $\overrightarrow{W}_{X_k}$ will never end up in a steady state, i.e., for $k\to\infty$, the precision matrix will get bigger with linear divergence speed. The corresponding covariance matrix looks as follows

$$
\begin{aligned}
\overrightarrow{V}_{X_k}&=\left(\overrightarrow{W}_{X_k}\right)^{-1} && (73)\\
&=\frac{2\sigma^2}{k^2-\left(\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\right)^2}\left(k\mathbf{I}_2-\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\mathbf{A}^{k-1}\mathbf{S}\right). && (74)
\end{aligned}
$$

The Kalman gain is then defined as

$$
\begin{aligned}
\alpha_k&=\mathbf{C}\mathbf{B}_k && (75)\\
&=\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G && (76)\\
&=\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\left(\sigma^2+\mathbf{C}\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T\right)^{-1} && (77)\\
&=\frac{2\left(k-\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\cos\left(\omega_0\left(k+1\right)\right)\right)}{k^2-\left(\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\right)^2+2\left(k-\frac{\sin\left(\omega_0 k\right)}{\sin\left(\omega_0\right)}\cos\left(\omega_0\left(k+1\right)\right)\right)}. && (78)
\end{aligned}
$$

Bibliography

[1] H.-A. Loeliger. *Signal and Information Processing: Modeling, Filtering, Learning.* Signal and Information Processing Laboratory, ETH Zurich, 2013.

[2] Erik Hampus Malmberg. Kalman filter-based phase-locked loops for harmonic signals. semester thesis. ISI, ETH Zurich, 2013.