Derivation of Update Rules for Kalman-Based PLL

In Figure 1 we can see a factor graph representation of a Kalman-based phase-locked loop. The matrices A, C and the observed measurements $\tilde{y}_k$ are given as follows

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

Since Z_k is a Gaussian random variable it follows that all random variables in the factor graph are Gaussian. The states X_k can therefore be described by its mean vector and covariance matrix.

In every step k, a new sample y_k is observed. Together with this sample and the previous state X_{k-1}, using Gaussian message passing, we can calculate the next state estimate X_k.

We begin the message passing algorithm by computing the message given by the observed samples $\tilde{y}_k$ [1]

$$\overrightarrow{m}_{Z_k} = 0 \qquad \overrightarrow{V}_{Z_k} = \sigma^2 \tag{1}$$
$$\overleftarrow{m}_{\tilde{Y}_k} = \tilde{y}_k \qquad \overleftarrow{V}_{\tilde{Y}_k} = 0. \tag{2}$$

In a next step, we add the noise in Eq. (1) to the observed sample in Eq. (2) to get the message at Y_k

$$\overleftarrow{m}_{Y_k} = \tilde{y}_k \qquad \overleftarrow{V}_{Y_k} = \sigma^2. \tag{3}$$

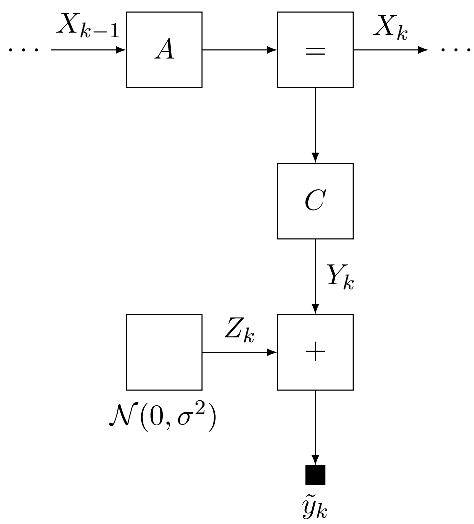

Figure 1: Factor graph representation of a Kalman filter's kth cell.

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

$$\overrightarrow{m}_{X''_k} = \mathbf{A}\overrightarrow{m}_{X_{k-1}} \qquad \overrightarrow{V}_{X''_k} = \mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}_T \tag{4}$$

$$\overleftarrow{W}_{X'_k} = \mathbf{C}^T\overleftarrow{W}_{Y_k}\mathbf{C} \qquad \overleftarrow{W}_{X'_k}\overleftarrow{m}_{X'_k} = \mathbf{C}^T\overleftarrow{W}_{Y_k}\overleftarrow{m}_{Y_k}, \tag{5}$$

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

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

and $\overleftarrow{W}_{X'_k}\overleftarrow{m}_{X'_k}$ denotes the weighted mean. The set of equations in (5) can be simplified by using the set of equations in (3)

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

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

$$
\begin{aligned}
\overrightarrow{W}_{X_k} &= \overrightarrow{W}_{X''_k} + \overleftarrow{W}_{X'_k} && (6)\\
&= \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1} + \mathbf{C}^T\frac{1}{\sigma^2}\mathbf{C} && (7)\\
&= \left(\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\right)^{-1} + \frac{1}{\sigma^2}\mathbf{C}^T\mathbf{C} && (8)\\
&= \mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1} + \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2} && (9)\\
&= \mathbf{A}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1} + \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2}, && (10)
\end{aligned}
$$

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

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

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

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 (16)

$$(\mathbf{B} + \mathbf{DEF})^{-1} = \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{D}\left(\mathbf{E}^{-1} + \mathbf{FB}^{-1}\mathbf{D}\right)^{-1}\mathbf{FB}^{-1}. \tag{16}$$

Thus we get the following assignments for the equation (6)

$$
\begin{aligned}
\mathbf{B} &= \mathbf{A}^{-T}\overrightarrow{W}_{X_{k-1}}\mathbf{A}^{-1}, && (17)\\
\mathbf{D} &= \mathbf{C}^T, && (18)\\
\mathbf{E} &= \frac{1}{\sigma^2}, && (19)\\
\mathbf{F} &= \mathbf{C}. && (20)
\end{aligned}
$$

Finally it follows the update rules for the covariance matrix and the mean vector

$$
\begin{aligned}
\overrightarrow{V}_{X_k} &= \left[\mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} + \frac{\mathbf{C}^T \mathbf{C}}{\sigma^2} \right]^{-1} && (21) \\
&= \left(\mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} \right)^{-1} \\
&- \left(\mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} \right)^{-1} \mathbf{C}^T \left(\sigma^2 + \mathbf{C} \left(\mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} \right)^{-1} \mathbf{C}^T \right)^{-1} \mathbf{C} \left(\mathbf{A}^{-T} \overrightarrow{W}_{X_{k-1}} \mathbf{A}^{-1} \right)^{-1} && (22) \\
&= \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T - \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T \underbrace{\left(\sigma^2 + \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T \right)^{-1}}_{=:G} \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \\
&= \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T - \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T G \mathbf{C} \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T && (23)
\end{aligned}
$$

$$
\begin{aligned}
\overrightarrow{m}_{X_k} &= \ldots && (24) \\
&= \ldots \\
&= \mathbf{A} \overrightarrow{m}_{X_{k-1}} + \mathbf{A} \overrightarrow{V}_{X_{k-1}} \mathbf{A}^T \mathbf{C}^T G \left(y_k - \mathbf{C} \mathbf{A} \overrightarrow{m}_{X_{k-1}} \right) && (25)
\end{aligned}
$$

with matrix $\mathbf{G}$ given as follows

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

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

$$\begin{aligned}
\hat{y}_k - \tilde{y}_k &= \mathbf{C}\left(\overrightarrow{m}_{X_k} - \mathbf{A}\overrightarrow{m}_{X_{k-1}}\right) && (26)\\
&= \mathbf{C}\left(\overrightarrow{m}_{X_k} - \overrightarrow{m}_{X_k} + \underbrace{\mathbf{A}\overrightarrow{V}_{X_{k-1}}\mathbf{A}^T\mathbf{C}^T G\left(y_k - \mathbf{CA}\overrightarrow{m}_{X_{k-1}}\right)}_{=:\mathbf{B}_k}\right) && (27)\\
&= && (28)
\end{aligned}$$

Independence of State Update from Noise

By induction and for a time index $k > 0$, the precision matrix update rule in eq. (6) can be written as follows

$$\overrightarrow{W}_{X_k} = \sum_{j=0}^{k-1} \left(\mathbf{A}^{-T}\right)^j \frac{\mathbf{C}^T\mathbf{C}}{\sigma^2} \left(\mathbf{A}^{-1}\right)^j . \tag{29}$$

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

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H, \tag{30}$$

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

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

In a next step, the new expression for $\mathbf{A}$ can be inserted into equation (29) [2]

$$
\begin{aligned}
\overrightarrow{W}_{X_k} &= \sum_{j=0}^{k-1}\left(\left(\left(\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H\right)^H\right)^{-j}\frac{\mathbf{C}^T\mathbf{C}}{\sigma^2}\left(\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H\right)^{-j}\right)\\
&= \frac{1}{\sigma^2}\sum_{j=0}^{k-1}\left(\mathbf{Q}\left(\bar{\Lambda}\right)^{-j}\mathbf{Q}^H\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\mathbf{Q}\boldsymbol{\Lambda}^{-j}\mathbf{Q}^H\right) \qquad (31)\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\mathbf{Q}\left(\bar{\Lambda}\right)^{-j}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\boldsymbol{\Lambda}^{-j}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}\exp(i\omega_0 j) & 0\\ 0 & \exp(-i\omega_0 j)\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\exp(-i\omega_0 j) & 0\\ 0 & \exp(i\omega_0 j)\end{bmatrix}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}\exp(i\omega_0 j) & \exp(i\omega_0 j)\\ \exp(-i\omega_0 j) & \exp(-i\omega_0 j)\end{bmatrix}\begin{bmatrix}\exp(-i\omega_0 j) & 0\\ 0 & \exp(i\omega_0 j)\end{bmatrix}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}\exp(i\omega_0 j) & 0\\ 0 & \exp(-i\omega_0 j)\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\exp(-i\omega_0 j) & 0\\ 0 & \exp(i\omega_0 j)\end{bmatrix}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}1 & \exp(2i\omega_0 j)\\ \exp(-2i\omega_0 j) & 1\end{bmatrix}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\begin{bmatrix}0 & \exp(2i\omega_0 j)\\ \exp(-2i\omega_0 j) & 0\end{bmatrix}\right)\mathbf{Q}^H\\
&= \frac{1}{\sigma^2}\left(\frac{k}{2}\mathbf{I}_2+\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix}0 & \exp(2i\omega_0 j)\\ \exp(-2i\omega_0 j) & 0\end{bmatrix}\right)\mathbf{Q}^H\right) \qquad (32)
\end{aligned}
$$

where equation (31) follows from the unitary property of $\mathbf{Q}$. Then we rewrite the equation (32) as follows

$$
\begin{aligned}
\sum_{j=0}^{k-1}\exp(2i\omega_0 j) &= \frac{1-\exp(2i\omega_0 k)}{1-\exp(2i\omega_0)} \qquad (33)\\
&= \frac{\exp(i\omega_0 k)\left(\exp(i\omega_0 k)-\exp(-i\omega_0 k)\right)}{\exp(i\omega_0)\left(\exp(i\omega_0)-\exp(-i\omega_0)\right)}\\
&= \exp\left(i\omega_0\left(k-1\right)\right)\frac{\sin(\omega_0 k)}{\sin(\omega_0)} \qquad (34)
\end{aligned}
$$

where we used the geometric series in equation (33)

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

and the following identity for the sine (derived from Euler's formula) (see equation (34))

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

In a similar way we get the following equation

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

Thus the second term in equation (32) can be written as

$$\begin{aligned}
&\mathbf{Q}\sum_{j=0}^{k-1}\left(\begin{bmatrix} 0 & \exp(2i\omega_0 j) \\ \exp(-2i\omega_0 j) & 0 \end{bmatrix}\right)\mathbf{Q}^H \\
&= \frac{\sin(\omega_0 k)}{2\sin(\omega_0)}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}\begin{bmatrix} 0 & \exp(i\omega_0(k-1)) \\ \exp(-i\omega_0(k-1)) & 0 \end{bmatrix}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \\
&= \frac{\sin(\omega_0 k)}{2\sin(\omega_0)}\begin{bmatrix} \exp(-i\omega_0(k-1)) & \exp(i\omega_0(k-1)) \\ i\exp(-i\omega_0(k-1)) & -i\exp(i\omega_0(k-1)) \end{bmatrix}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \\
&= \frac{\sin(\omega_0 k)}{2\sin(\omega_0)}\begin{bmatrix} 2\cos(\omega_0(k-1)) & 2\sin(\omega_0(k-1)) \\ 2\sin(\omega_0(k-1)) & -2\cos(\omega_0(k-1)) \end{bmatrix} \\
&= \frac{\sin(\omega_0 k)}{\sin(\omega_0)}\mathbf{A}^{k-1}\mathbf{S} \qquad (35)
\end{aligned}$$

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

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

$$\begin{aligned}
\overrightarrow{V}_{X_k} &= \left(\overrightarrow{W}_{X_k}\right)^{-1} && (36)\\
&= \frac{2\sigma^2}{k^2 - \left(\frac{\sin(\omega_0 k)}{\sin(\omega_0)}\right)^2}\left(k\mathbf{I}_2 - \frac{\sin(\omega_0 k)}{\sin(\omega_0)}\mathbf{A}^{k-1}\mathbf{S}\right) && (37)
\end{aligned}$$

$$\alpha_k = \mathbf{C}\mathbf{B}_k = \frac{2\left(k - \frac{\sin(\omega_0 k)}{\sin(\omega_0)}\cos(\omega_0(k+1))\right)}{k^2 - \left(\frac{\sin(\omega_0 k)}{\sin(\omega_0)}\right)^2 + 2\left(k - \frac{\sin(\omega_0 k)}{\sin(\omega_0)}\cos(\omega_0(k+1))\right)}$$

Cost function

[1]

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.