Intrinsic Extended Kalman Filter on Lie Groups

D. H. S. Maithripala*

†Department of Mechanical Engineering, University of Peradeniya, KY 20400, Sri Lanka

Abstract

I. Introduction

[1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]

II. The Kalman Filter on $\mathbb{R}^n$

Consider a system modelled by a Markov process that takes the form

$$x_k = A_k\, x_{k-1} + G_k w_{k-1}, \tag{1}$$
$$y_k = H_k\, x_k + z_k. \tag{2}$$

We assume that $p(w_k) = \mathcal{N}(0, \Sigma_p)$, and $p(z_k) = \mathcal{N}(0, \Sigma_m)$ where $\mathcal{N}(\mu, \Sigma)$ denotes a multivariate gaussian with mean μ and covariance Σ.

Denote by Y_k the event that corresponds to a certain realization of the k-tuple of random variables $\{y_1, y_2, \cdots, y_k\}$ and the random variables $x_k^+ \triangleq x_k|Y_k$ and $x_k^- \triangleq x_k|Y_{k-1}$ that correspond to estimates of x_k given the events Y_k and Y_{k-1} respectively. Also define the estimated output $y_k^- \triangleq y_k|Y_{k-1}$. Motivated by (1) we assume the model

$$x_k^- = A_k\, x_{k-1}^+ + G_{k-1} w_{k-1}, \tag{3}$$
$$y_k^- = H_k\, x_k^- + z_k. \tag{4}$$

If we let $p(x_0^-) = \mathcal{N}(m_0, P_0)$ then since $p(w_k) = \mathcal{N}(0, \Sigma_p)$ the linear system (3) implies that $p(x_k \mid Y_{k-1}) = p(x_k^-) = \mathcal{N}(m_k^-, P_k^-)$ where the predicted mean, m_k^-, and the predicted covariance, P_k^-, satisfy

$$m_k^- = A_k\, m_{k-1}, \tag{5}$$
$$P_k^- = A_k P_{k-1} A_k^T + G_k \Sigma_p G_k^T. \tag{6}$$

Here we let $p(x_k \mid Y_k) = p(x_k^+) = \mathcal{N}(m_k, P_k)$.

Furthermore since $p(z_k) = \mathcal{N}(0, \Sigma_m)$ the linear output relationship (4) implies that $p(y_k|Y_{k-1}) = \mathcal{N}(H_k m_k^-, H_k P_k^- H_k^T + \Sigma_m)$. Thus we have the joint distribution

$$p\begin{pmatrix} x_k \mid Y_{k-1} \\ y_k \mid Y_{k-1} \end{pmatrix} = \mathcal{N}\left(\begin{bmatrix} m_k^- \\ H_k m_k^- \end{bmatrix}, \begin{bmatrix} P_k^- & P_k^- H_k^T \\ H_k P_k^- & H_k P_k^- H_k^T + \Sigma_m \end{bmatrix} \right).$$

Thus from the properties of multivariate normal distributions (please refer to [] for formulas for constructing the conditional probabilities of Gaussian distributions) we find that the conditional distribution $p(x_k \mid Y_k)$ is given by

$$p(x_k \mid Y_k) = \mathcal{N}\left(m_k^- + P_k^- H_k^T (H_k P_k^- H_k^T + \Sigma_m)^{-1}(y_k - H_k m_k^-), P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + \Sigma_m)^{-1} H_k P_k^- \right)$$

Thus we have that the updated mean and covariances are given by

$$K_k \triangleq P_k^- H_k^T (H_k P_k^- H_k^T + \Sigma_m)^{-1}, \tag{7}$$
$$m_k = m_k^- + K_k (y_k - H m_k^-), \tag{8}$$
$$P_k = (I - K_k H_k) P_k^-. \tag{9}$$

where $\Sigma_p = E(w_k w_k^T)$ and $\Sigma_m = E(z_k z_k^T)$. Define $e_k \triangleq x_k - m_k$ then we have

$$e_k = (I - K_k H_k) A_k e_{k-1} + (I - K_k H_k) w_{k-1} - K_k z_k$$
$$P_k = (I - K_k H_k)(A_k P_{k-1} A_k^T + \Sigma_p)$$

III. Discrete Time Pre-observers on Lie groups with Time Invariant Error Dynamics

Consider an n-dimensional Lie group G with Lie algebra $\mathcal{G}$. Let $(g,\zeta) \in G \times \mathcal{G}$ and $\phi : G \times M \mapsto M$ be a left or right-invariant action on a m-dimensional manifold M.

Consider the kinematic system that evolves according to

$$\dot{g} = g \cdot \zeta, \tag{10}$$
$$y = \phi_g(\gamma), \tag{11}$$

where $\zeta(t) \in \mathcal{G}$ is a known input and $\gamma \in M$ is a known constant. **In a typical inertial measurement system or more generally where inertial landmarks are observed in the body frame the outputs are right invariant.**

The discretized version of the system is given by

$$g_k = g_{k-1} \exp(\Delta T \zeta_{k-1}). \tag{12}$$

This discretization is not novel and has been used by many see for instance [7].

A. For left-invariant Outputs

Consider the pre-observer

$$\widetilde{g}_k^- = \widetilde{g}_{k-1} \exp(\Delta T(\zeta_{k-1})), \tag{13}$$
$$\widetilde{g}_k = \widetilde{g}_k^- \exp(\Delta T L(y_k, \widetilde{y}_k)), \tag{14}$$
$$\widetilde{y}_k = \phi_{\widetilde{g}_k}(\gamma) \tag{15}$$

where $L : M \times M \to \mathcal{G}$ is the Innovation term. This appears in [4]

Let $u_k \triangleq \exp(\Delta T \zeta_{k-1})$ and the erorr $e_k^- = (\widetilde{g}_k^-)^{-1} g_k$ and $e_k = \widetilde{g}_k^{-1} g_k$. Then the error dynamics are

$$e_k^- = (\widetilde{g}_k^-)^{-1} g_k = u_{k-1}^{-1} \widetilde{g}_{k-1}^{-1} g_{k-1} u_{k-1} = u_{k-1}^{-1} e_{k-1} u_{k-1},$$
$$e_k = \widetilde{g}_k^{-1} g_k = \exp(-\Delta T L(y_k, \widetilde{y}_k)) e_k^-.$$

These error dynamics appear in [2], [3], [9], [11] as well as [4]

Let $L(y_k, \widetilde{y}_k)$ be G invariant. That is $L(\phi_g(y_1), \phi_g(y_2),) = L(y_1, y_2)$ for all $g \in G$. Thus with no noise

$$L(y_k, \widetilde{y}_k) = L(\phi_{g_k}(\gamma), \phi_{\widetilde{g}_k}(\gamma)) = L\left(\phi_{\widetilde{g}_k^{-1} g_k}(\gamma), \phi_{\widetilde{g}_k^{-1} \widetilde{g}_k}(\gamma)\right) = L(\phi_{e_k}(\gamma), \gamma)$$

Hence if $L(y_k, \widetilde{y}_k)$ is G invariant then the above error dynamics are autonomous.

B. For right-invariant outputs

Consider the pre-observer

$$\widetilde{g}_k^- = \widetilde{g}_{k-1} \exp(\Delta T(\zeta_{k-1})), \tag{16}$$
$$\widetilde{g}_k = \exp(\Delta T L(y_k, \widetilde{y}_k)) \widetilde{g}_k^-, \tag{17}$$
$$\widetilde{y}_k = \phi_{\widetilde{g}_k}(\gamma), \tag{18}$$

where $L : M \times M \to \mathcal{G}$ is the Innovation term.

Let $u_k \triangleq \exp(\Delta T \zeta_{k-1})$ and consider the right invariant error $e_k^- = g_k (\widetilde{g}_k^-)^{-1}$ and $e_k = g_k \widetilde{g}_k^{-1}$. Then the error dynamics are

$$e_k^- = g_k (\widetilde{g}_k^-)^{-1} = g_{k-1} u_{k-1} u_{k-1}^{-1} \widetilde{g}_{k-1}^{-1} = e_{k-1},$$
$$e_k = g_k \widetilde{g}_k^{-1} = g_{k-1} u_{k-1} (\widetilde{g}_k^-)^{-1} \exp(-\Delta T L(y_k, \widetilde{y}_k)) = e_k^- \exp(-\Delta T L(y_k, \widetilde{y}_k)).$$

Let $L(y_k, \widetilde{y}_k)$ be G invariant. That is $L(\phi_g(y_1), \phi_g(y_2),) = L(y_1, y_2)$ for all $g \in G$. Thus with no noise

$$L(y_k, \widetilde{y}_k) = L(\phi_{g_k}(\gamma), \phi_{\widetilde{g}_k}(\gamma)) = L\left(\phi_{g_k \widetilde{g}_k^{-1}}(\gamma), \phi_{\widetilde{g}_k \widetilde{g}_k^{-1}}(\gamma)\right) = L(\phi_{e_k}(\gamma), \gamma)$$

Hence if $L(y_k, \widetilde{y}_k)$ is G invariant then the above error dynamics are autonomous.

C. The IMU+GNSS Sensor Fusion Problem

From an application point of view estimating the orientation of a rigid body given the IMU+GPS measurements is of immense importance. The Gyroscope measures the angular velocity of the object in the body frame given by Ω. The accelerometers measure $A^m \triangleq R^T(\ddot{o} + ge_3)$ or in other words the external forces acting on the object represented in the body frame. In addition when GNSS measurements are available one also has the measurements $o, \dot{o}$. **Basically everything is measured except R. In practice this is a sensor fusion problem that has been tackled successfully by many.** The governing equations are given by:

$$\dot{R} = R\widehat{\Omega}, \tag{19}$$

$$\ddot{o} = \frac{f}{m} - ge_3. \tag{20}$$

A GPS would measure $o, \dot{o}$ while the gyroscopes measure Ω, and the accelerometers measure $A^m \triangleq f/m$.

Define $o^s(t) \triangleq o(t) + gt^2/2e_3$ then $v^s(t) \triangleq \dot{o}^s(t) = \dot{o}(t) + gte_3$ and $RA^m = \ddot{o}^s(t) = \ddot{o}(t) + ge_3$. In these notations we recast the problem as

$$\begin{aligned}
\dot{R} &= R\widehat{\Omega}, && (21)\\
\dot{v}^s &= RA^m, && (22)\\
\dot{o}^s &= v^s = RV^s, && (23)\\
y_o &= o^s, && (24)\\
y_v &= v^s && (25)
\end{aligned}$$

Let

$$X \triangleq \begin{bmatrix} R & v^s \\ 0 & 1 \end{bmatrix}, \qquad \zeta \triangleq \begin{bmatrix} \widehat{\Omega} & A^m \\ 0 & 0 \end{bmatrix}, \qquad \gamma_v = \begin{bmatrix} 0_{3\times 1} \\ 1 \end{bmatrix}$$

Then above equations take the form

$$\begin{aligned}
\dot{X} &= X\,\zeta, && (26)\\
y_v &= X\gamma_v. && (27)
\end{aligned}$$

$$\begin{aligned}
\dot{o}^s &= v^s = X\gamma_v, && (28)\\
y_o &= o^s, && (29)
\end{aligned}$$

The first equation (26) above when discretized takes the form

$$X_{k+1} = X_k \exp(\Delta t\, \zeta_k). \tag{30}$$

Note that

$$\exp(\Delta t\, \zeta_k) = \begin{bmatrix} \exp\left(\Delta t\, \widehat{\Omega}_k\right) & \Delta t\, {A^m}_k \\ 0 & 1 \end{bmatrix}$$

Then we have

$$\begin{aligned}
R_{k+1} &= R_k \exp\left(\Delta t\, \widehat{\Omega}_k\right),\\
v^s_{k+1} &= v^s_k + \Delta t\, R_k {A^m}_k.
\end{aligned}$$

From the equation (28) we have

$$o^s_{k+1} = o^s_k + \Delta t\, v^s_k.$$

In the original notations we then have the discretized equations

$$\begin{aligned}
R_{k+1} &\approx R_k \exp\left(\Delta t\, \widehat{\Omega}_k\right),\\
v_{k+1} &= v_k - g\,\Delta t\, e_3 + \Delta t\, R_k {A^m}_k,\\
o_{k+1} &= o_k + \Delta t\, v_k - \frac{g\,(\Delta t)^2}{2}\, e_3.
\end{aligned}$$

IV. Intrinsic Extended Kalman Filter on Lie Groups

Consider an n-dimensional Lie group G with Lie algebra $\mathcal{G}$. Let $(g, \zeta) \in G \times \mathcal{G}$ and $\phi : G \times M \mapsto M$ be a left-invariant linear action on a m-dimensional vector space M. That is for each $g \in G$ the mapping $\phi_g : M \to M$ is linear ($\phi_g \in GL(M)$).

Recall that

$$g \exp(\zeta) g^{-1} = \exp(\mathrm{Ad}_g \cdot \zeta), \tag{31}$$

and that since $\phi : G \to GL(M)$ is a homomorphism it follows that

$$\phi_{\exp(\zeta)} = \exp(T_e\phi \circ \zeta). \tag{32}$$

The exponential map in the first of the above expression is $\exp : \mathcal{G} \to G$ while the exponential map in the right hand side of the second of the above expressions is $\exp : gl(M) \to GL(M)$. In particular if $M = \mathcal{G}$, $\phi = \mathrm{Ad}$ then $\mathrm{Ad}_{\exp(\zeta)} = \exp(\mathrm{ad}_\zeta)$.

Consider the left invariant Markovian stochastic processes $g(t) \in G$ and $y(t) \in M$ that evolve according to

$$\dot{g} = g \cdot (\zeta + n_\zeta), \tag{33}$$

$$y = \phi_{g^{-1}}(\gamma) + n, \tag{34}$$

where $\zeta(t) \in \mathcal{G}$ is a known input $\gamma \in M$ is a known constant, while $n_\zeta(t) \in \mathcal{G}$ and $n(t) \in \mathcal{G}^m$ are Gaussian white noise processes with zero mean and covariances Σ_g and Σ_y respectively.

Theorem 1: Define the matrices

$$A(t) = -\mathrm{ad}_{\zeta(t)}, \quad H(t) = -\phi_{\tilde{g}^{-1}}\left(T_e\phi \circ \mathrm{Ad}_{\tilde{g}}\, \eta_e(\gamma)\right).$$

Then if the pair $(A(t), H(t))$ is uniformly observable then the intrinsic Extended Kalman filter

$$\dot{\tilde{g}} = \tilde{g} \cdot (\zeta + K(t)(y - \tilde{y})), \tag{35}$$

$$\tilde{y} = \phi_{\tilde{g}^{-1}}(\gamma), \tag{36}$$

where the Kalman filter gain $K(t)$ is given by

$$\begin{aligned} \dot{P} &= A^T P + P A^T - P H^T \Sigma_y^{-1} H P + \Sigma_\zeta, \\ K &= P H^T \Sigma_y^{-1}. \end{aligned}$$

ensures that

$$\lim_{t\to\infty} E(e_g(t)) = I, \tag{37}$$

$$\lim_{t\to\infty} E(\eta_e \eta_e^T) = P_\infty. \tag{38}$$

for all e_g in some neighborhood $\mathcal{N}_\epsilon$ of the identity I of G where $e_g \triangleq \tilde{g}^{-1} g$ and $\eta_e \triangleq \log(e_g)$.

Proof of Theorem 1: **This is a sketch of the proof:** Let us define the estimation error stochastic processes $\{e_g(t)\}$ such that $e_g \triangleq \tilde{g}^{-1} g$ and the deviation output stochastic process $\{y_e(t)\}$ such that $y_e \triangleq (y - \tilde{y})$. From (33)–(34) and (35)–(36) we see that the stochastic processes $\{e_g(t)\}$ and $\{y_e(t)\}$ satisfy the Markov process

$$\dot{e}_g = e_g \cdot \Big((I - \mathrm{Ad}_{e_g^{-1}})\zeta - \mathrm{Ad}_{e_g^{-1}} K(t) y_e + n_\zeta \Big), \tag{39}$$

$$y_e = \phi_{\tilde{g}^{-1}}\Big((\phi_{\tilde{g} e_g^{-1} \tilde{g}^{-1}} - I)(\gamma) \Big) + n. \tag{40}$$

From (31) and (32) we have

$$(\phi_{\tilde{g} e_g^{-1} \tilde{g}^{-1}} - I) = (\phi_{\exp(-\mathrm{Ad}_{\tilde{g}} \cdot \eta_e)} - I) = \exp(-T_e\phi \circ \mathrm{Ad}_{\tilde{g}} \cdot \eta_e) - I$$

Let η_e be the exponential coordinate of e_g. That is let $e_g = \exp(\eta_e)$. Then $\{\eta_e(t)\}$ is a stochastic process on $\mathcal{G}$ with expectation $m_{\eta_e}(t) \triangleq E(\eta_e(t))$. Then from (39)–(40) we find that $\{\eta_e(t)\}$ and $\{y_{e_i}(t)\}$ evolve according to the Markovian processe on $\mathcal{G} \times \mathcal{G}^m$ that is given by

$$\dot{\eta}_e = -\mathrm{ad}_\zeta \eta_e - \left(\frac{\mathrm{ad}_{\eta_e}}{\exp(\mathrm{ad}_{\eta_e}) - I} \right) K(t) y_e + \left(\frac{\exp(\mathrm{ad}_{\eta_e}) \mathrm{ad}_{\eta_e}}{\exp(\mathrm{ad}_{\eta_e}) - I} \right) n_\zeta, \tag{41}$$

$$y_e = \phi_{\tilde{g}^{-1}}\left(\left(\exp(-T_e\phi \circ \mathrm{Ad}_{\tilde{g}} \cdot \eta_e) - I \right)(\gamma) \right) + n. \tag{42}$$

thus that these equations can be written down as

$$\dot{\eta}_e = (A(t) - K(t)H(t))\,\eta_e + n_\zeta - K(t)n + \left(\frac{1}{2}\mathrm{ad}_{\eta_e} + O(||\eta_e||^2)\right) n_\zeta + O(||\eta_e||^2)n, \tag{43}$$

$$y_{e_i} = H(t)\eta_e + n_i + O(||\eta_e||^2), \tag{44}$$

where

$$A(t) = -\mathrm{ad}_{\zeta(t)},$$

and

$$H(t)\eta_e = -\phi_{\tilde{g}^{-1}}\left(\left(T_e\phi \circ \mathrm{Ad}_{\tilde{g}}\,\eta_e\right)(\gamma)\right).$$

Note that $(T_e\phi \circ \mathrm{Ad}_{\tilde{g}})\,\eta_e \in gl(M)$ and hence the above expression is well defined.

Then the expectation $E(\eta_e(t)) \triangleq m_{\eta_e}(t)$ and the covariance $E(\eta_e {\eta_e}^T) \triangleq P(t)$ evolve according to

$$\dot{m}_{\eta_e} = (A(t) - K(t)H(t))\,m_{\eta_e} + E(O(||\eta_e||^2)), \tag{45}$$

$$\dot{P} = (A(t) - K(t)H(t))P + P(A(t) - K(t)H(t))^T - PH^T\Sigma_y^{-1}HP + \Sigma_\zeta + \frac{1}{4}E\left(\mathrm{ad}_{\eta_e}\Sigma_\zeta\mathrm{ad}^T_{\eta_e}\right) + E(O(||\eta_e||^3)) \tag{46}$$

If the pair $(A(t), H(t))$ is uniformly observable and $K(t)$ is the Kalman gain for the linear system $(A(t), H(t), \Sigma_p, \Sigma)$ then $\lim_{t\to\infty} m_{\eta_e}(t) = 0$. Define the mean zero stochastic variable $\epsilon \triangleq \eta_e - m_{\eta_e}$. The Kalman filter property ensures that that the covariance of $\epsilon(t)$ converges to a constant and is small. Thus since $e_g = e^{\eta_e} = e^{(m_{\eta_e}+\epsilon)} \approx e^{m_{\eta_e}}e^{\epsilon}$ we have that $\lim_{t\to\infty} e_g(t) = e^{\epsilon}$ and hence that

$$\lim_{t\to\infty} \tilde{g} = g(t)e^{-\epsilon(t)}. \tag{47}$$

A. Discretized EKF on Lie Groups

Defining the matrices

$$A_k = \left(I - \Delta T\mathrm{ad}_{\zeta_k}\right), \quad H_k = \begin{bmatrix} -\phi_{\widetilde{g}_k^{-1}} \circ \left(T_e\phi \circ \mathrm{Ad}_{\widetilde{g}_k}(\cdot)\right)(\gamma_1) \\ -\phi_{\widetilde{g}_k^{-1}} \circ \left(T_e\phi \circ \mathrm{Ad}_{\widetilde{g}_k}(\cdot)\right)(\gamma_2) \\ \vdots \\ -\phi_{\widetilde{g}_k^{-1}} \circ \left(T_e\phi \circ \mathrm{Ad}_{\widetilde{g}_k}(\cdot)\right)(\gamma_m) \end{bmatrix}, \quad G_k = \sqrt{\Delta T}\,I$$

and $\Sigma_m = E(n_k n_k^T)$ and $\Sigma_p = E(n_{\zeta_k} n_{\zeta_k}^T)$. The discretized version of the Kalman filter (35)–(36) is given by

$$\widetilde{g}_k^- = \widetilde{g}_{k-1}\exp\left(\Delta T(\zeta_{k-1})\right), \tag{48}$$

$$\widetilde{g}_k = \widetilde{g}_k^-\exp\left(\Delta T K_k(y_k - \widetilde{y}_k)\right), \tag{49}$$

$$\tag{50}$$

where the Kalman gain K_k and the error covariance are computed by

$$P_k^- = A_{k-1}P_{k-1}A_{k-1}^T + G_k\Sigma_q G_k^T, \tag{51}$$

$$K_k \triangleq P_k^- H_k^T(H_k P_k^- H_k^T + \Sigma_m)^{-1}, \tag{52}$$

$$P_k = (I - K_k H_k)P_k^-. \tag{53}$$

V. Application to Rigid Body Motion

Rigid body kinematics are defined by

$$\dot{g} = g \cdot \zeta \tag{54}$$

where $g \in SE(3) = SO(3)\otimes_{\mathrm{s}}\mathbb{R}^3$ and $\zeta \in se(3)$. Explicitly written down

$$g = \begin{bmatrix} R & o \\ 0 & 1 \end{bmatrix}, \quad \zeta = \begin{bmatrix} \widehat{\Omega} & V \\ 0 & 0 \end{bmatrix},$$

where $R \in SO(3)$, $o, V \in \mathbb{R}^3$ and $\widehat{\Omega} \in so(3)$ and the $\cdot$ is simply matrix multiplication. Here $g \in SE(3)$ represents a change of frame between a fixed frame $\mathbf{e}$ and a moving frame $\mathbf{b}$ fixed on the body.

Also if $\zeta = (\Omega, V) \in se(3)$, $\xi = (\Phi, U) \in se(3)$ then

$$\mathrm{ad}_{\zeta}\xi = (\Omega \times \Phi\ ,\ \Omega \times U - \Phi \times V)$$

and hence

$$\mathrm{ad}_{\zeta} = \begin{bmatrix} \widehat{\Omega} & 0 \\ \widehat{V} & \widehat{\Omega} \end{bmatrix}.$$

The Euclidean motion group $SE(3)$ acts on the left on $\mathbb{R}^4$, $\phi : SE(3) \times \mathbb{R}^3 \to \mathbb{R}^3$, by multiplication where

$$\phi_g(x) = \begin{bmatrix} R & o \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ \alpha \end{bmatrix},$$

for $g = (R, o) \in SO(3) \otimes_{\mathrm{s}} \mathbb{R}^3 = SE(3)$, $x \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$. This action is transitive and proper. Its orbit on points of the form $\gamma = [x \quad 1]^T$ represents different perspectives of a physical point in space.

Since $\phi_g(\gamma)$ is simple matrix multiplication of γ by g we have that for $\zeta = (\Omega, V)$

$$(T_e\phi \circ \zeta)(\gamma) = \begin{bmatrix} \widehat{\Omega} & V \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}.$$

Then

$$\begin{aligned}
\phi_{\tilde{g}^{-1}}\left(\left(T_e\phi \circ \mathrm{Ad}_{\tilde{g}}\,\zeta\right)(\gamma)\right) &= \tilde{g}^{-1}\tilde{g}\zeta\tilde{g}^{-1}\gamma = \zeta\tilde{g}^{-1}\gamma \\
&= \begin{bmatrix} \widehat{\Omega} & V \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{R}^T & -\tilde{R}^T\tilde{o} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \widehat{\Omega} & V \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{R}^T(x - \tilde{o}) \\ 1 \end{bmatrix} = \begin{bmatrix} \widehat{\Omega}\tilde{R}^T(x - \alpha\tilde{o}) + V \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -\tilde{R}^T\widehat{(x - \tilde{o})}\tilde{R}\Omega + V \\ 0 \end{bmatrix} = \begin{bmatrix} -\tilde{R}^T\widehat{(x - \tilde{o})}\tilde{R} & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Omega \\ V \end{bmatrix}
\end{aligned}$$

A. *SLAM*

Let $\mathbf{e}$ be some fixed frame in space and p_i for $i = 1, 2, \cdots, N$ be some set of fixed points in space (landmarks). Let $\gamma_i = [x_i \quad 1]^T \in \mathbb{R}^4$ where $x_i \in \mathbb{R}^3$ is the representation of the physical point $p_i \in \mathbb{E}^3$ in the fixed frame $\mathbf{e}$. If X_i denotes the representation of p_i in the frame fixed on a rigid body moving in space whose relationship to the fixed frame $\mathbf{e}$ is given by $g \in SE(3)$ then the SLAM problem equivalent to

$$\begin{aligned}
\dot{g} &= g \cdot (\zeta + n_p)\,, \\
y_i &= \phi_{g^{-1}}(\gamma_i) + n_m,
\end{aligned}$$

where $y_i = [X_i \quad 1]^T$, $n_p \sim \mathcal{N}(0, \Sigma_p)$, $n_m \sim \mathcal{N}(0, \Sigma_m)$ and ϕ_g is simply matrix multiplication by g. In this formulation there is no bias in the velocity measurements but can be included easily.

Then from the above results we have

$$A_k = \left(I_{6\times 6} - \Delta T \begin{bmatrix} \widehat{\Omega}_k & 0 \\ \widehat{V}_k & \widehat{\Omega}_k \end{bmatrix}\right), \quad H_k = \begin{bmatrix} \widetilde{R}_k^T\widehat{(x_1 - \widetilde{o}_k)}\widetilde{R}_k & -I_{3\times 3} \\ \widetilde{R}_k^T\widehat{(x_2 - \widetilde{o}_k)}\widetilde{R}_k & -I_{3\times 3} \\ \vdots & \\ \widetilde{R}_k^T\widehat{(x_m - \widetilde{o}_k)}\widetilde{R}_k & -I_{3\times 3} \end{bmatrix}, \quad G_k = \sqrt{\Delta T}\, I$$

B. Attitude Estimation from IMUs

In the attitude estimation problem $G = SO(3)$, $M = \mathcal{G} = so(3) \simeq \mathbb{R}^3$, and ϕ is the adjoint representation $\mathrm{Ad} : SO(3) \times \mathbb{R}^3 \to \mathbb{R}^3$ that is explictly given by $\mathrm{Ad}_R x = Rx$ for $R \in SO(3)$ and $x \in \mathbb{R}^3$. Then $\mathrm{ad}_\Omega \Phi = \widehat{\Omega}\Phi = \Omega \times \Phi$ for $\Omega, \Phi \in \mathbb{R}^3$ and the output is

$$y_k = \begin{bmatrix} R_k^T e_1 \\ R_k^T e_2 \end{bmatrix}$$

where $e_1, e_2 \in \mathbb{R}^3$. Then we have

$$(\phi_{R^T} \circ T_e\phi \circ \mathrm{Ad}_R \cdot (\Omega))\, x = R^T \mathrm{ad}_{\mathrm{Ad}_R \Omega} x = R^T R \widehat{\Omega} R^T x = -(R^T \widehat{x} R)\, \Omega$$

For this system we have that the discrete Kalman filter given by (35)–(36) is is explicitly given by

$$\tilde{R}_k^- = \tilde{R}_{k-1} \exp(\Delta T \widehat{\Omega}_{k-1}), \tag{55}$$
$$\widetilde{R}_k = \widetilde{R}_k^- \exp(\Delta T K_k (y_k - \widetilde{y}_k)), \tag{56}$$

where the Kalman gain K_k and the error covariance are computed by

$$P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + G_k \Sigma_q G_k^T, \tag{57}$$
$$K_k \triangleq P_k^- H_k^T (H_k P_k^- H_k^T + \Sigma_m)^{-1}, \tag{58}$$
$$P_k = (I - K_k H_k) P_k^-. \tag{59}$$

where

$$A_k = \left(I_{3\times3} - \Delta T\, \widehat{\Omega}_k\right), \quad H_k = \begin{bmatrix} \widetilde{R}_k^T \widehat{x}_1 \widetilde{R}_k \\ \\ \widetilde{R}_k^T \widehat{x}_1 \widetilde{R}_k \end{bmatrix}, \quad G_k = \sqrt{\Delta T}\, I_{3\times3}$$

1) Simulation Results for Attitude Estimation from IMUs: Figure-1

Appendix

A. Lie Group Stuff

Let $\phi : G \mapsto H$ be a Lie-group homomorphism.

$$\phi(\exp(t\zeta)) = \exp(t T_e\phi \cdot \zeta). \tag{60}$$

In the special case where $\phi = \mathrm{I}_g : G \mapsto G$ for each g, we have that

$$h(t) = g \exp(\zeta t) g^{-1} = \exp(t \mathrm{Ad}_g \zeta) \tag{61}$$

In the special case where $\phi = \mathrm{Ad} : G \mapsto GL(\mathcal{G})$, we have that

$$h(t) = \mathrm{Ad}_{\exp(\zeta t)} = \exp(t \mathrm{ad}_\zeta) = I + t\mathrm{ad}_\zeta + \frac{t^2}{2!}\mathrm{ad}_\zeta^2 + \cdots \tag{62}$$

Define

$$\Gamma(s,t) \triangleq e^{-s\zeta(t)} \frac{\partial}{\partial t} e^{s\zeta(t)}$$

Then

$$\Gamma(1,t) = \int_0^1 \frac{\partial \Gamma(s,t)}{\partial s} ds = e^{-s\zeta(t)} \frac{\partial}{\partial t} e^{s\zeta(t)} = \int_0^1 e^{-s\mathrm{ad}_\zeta} \frac{d\zeta}{dt}\, ds = \left(\frac{I - \mathrm{ad}_\zeta}{\mathrm{ad}_\zeta}\right) \frac{d\zeta}{dt}.$$

From which we get

$$\frac{d}{dt} e^\zeta = e^\zeta \Gamma(1,t) = e^\zeta \left(\frac{I - e^{-\mathrm{ad}_\zeta}}{\mathrm{ad}_\zeta}\right) \frac{d\zeta}{dt} \tag{63}$$

Let $e^{\chi(t)} = e^\zeta e^{t\eta}$. Then

$$\eta = e^{-\chi(t)} \frac{d}{dt} e^{\chi(t)} = \left(\frac{I - e^{-\mathrm{ad}_\chi}}{\mathrm{ad}_\chi}\right) \frac{d\chi(t)}{dt}.$$

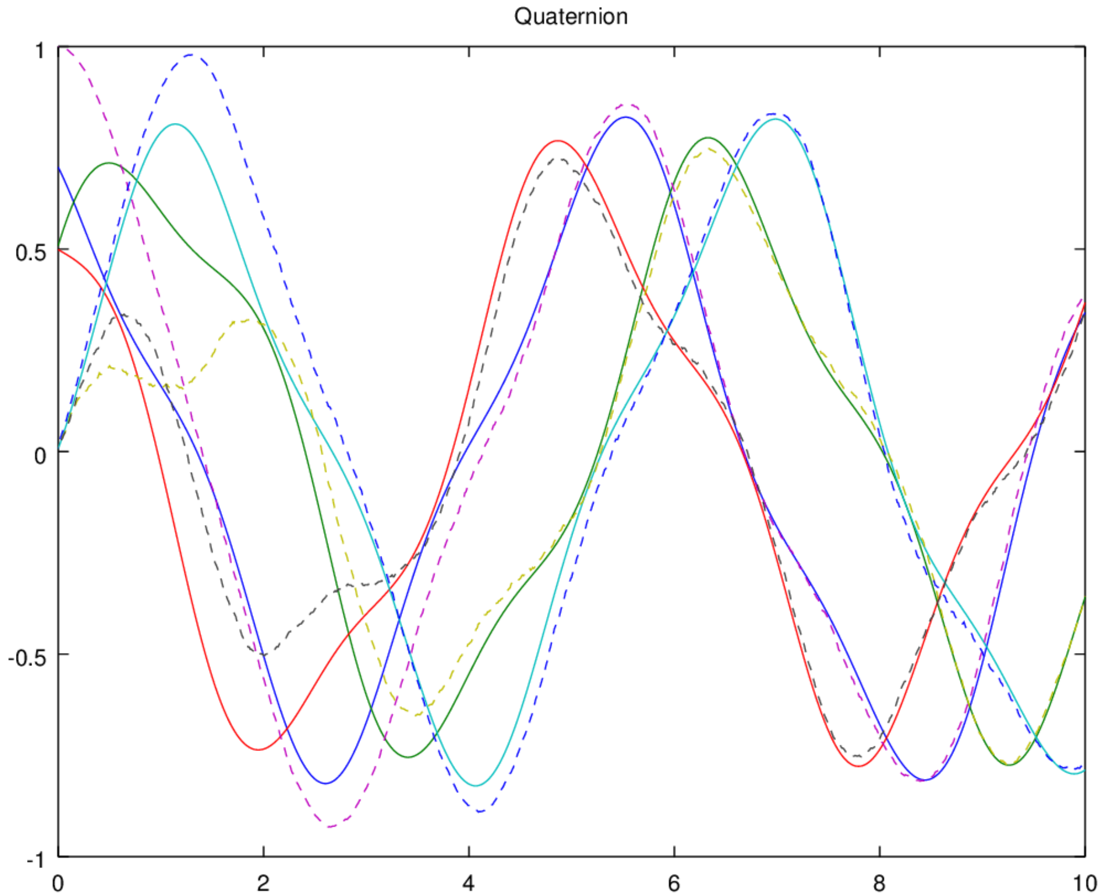

Fig. 1. Estimated and Actual Quaternions

Thus formally

$$\frac{d\chi(t)}{dt} = \left(\frac{\mathrm{ad}_\chi}{I - e^{-\mathrm{ad}_\chi}}\right)\eta.$$

Hence integrating from 0 to 1 with respect to t we have

$$\log\left(e^{\zeta}e^{\eta}\right) = \chi(1) = \zeta + \left(\int_0^1 \frac{\mathrm{ad}_\chi}{I - e^{-\mathrm{ad}_\chi}}\, dt\right)\eta.$$

The formal series in the integral is defined as

$$\frac{\mathrm{ad}_\chi}{I - e^{-\mathrm{ad}_\chi}} = \psi(e^{\mathrm{ad}_\chi}) = \psi(\mathrm{Ad}_{e^\chi}) = \psi(\mathrm{Ad}_{e^\zeta e^{t\eta}}) = \psi(e^{\mathrm{ad}_\zeta} e^{t\mathrm{ad}_\eta})$$

where

$$\psi(w) \triangleq \frac{w \log w}{w-1} = 1 + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m(m+1)}(w-1)^m, \quad ||w|| < 1.$$

Thus we have the Baker-Campbell-Hausdorff formula

$$\log\left(e^X e^Y\right) = X + \left(\int_0^1 \psi(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})\, dt\right) Y \tag{64}$$

$$= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\left([X,[X,Y]] + [Y,[Y,X]]\right) \tag{65}$$

$$
\begin{aligned}
&- \frac{1}{24}[Y,[X,[X,Y]]] && (66)\\
&- \frac{1}{720}\left([Y,[Y,[Y,[Y,X]]]] + [X,[X,[X,[X,Y]]]]\right) && (67)\\
&+ \frac{1}{360}\left([X,[Y,[Y,[Y,X]]]] + [Y,[X,[X,[X,Y]]]]\right) && (68)\\
&+ \frac{1}{120}\left([Y,[X,[Y,[X,Y]]]] + [X,[Y,[X,[Y,X]]]]\right) + \cdots . && (69)
\end{aligned}
$$

Alternatively

$$
\log(e^X e^Y) = X + \frac{\mathrm{ad}_X \; e^{\mathrm{ad}_X}}{e^{\mathrm{ad}_X} - I} \; Y + O(Y^2). \tag{70}
$$

Let

$$
\dot{g} = g\zeta \tag{71}
$$

and η be such that $\dot{g} = e^{\eta}$ Discretization of this gives

$$
g_k = g_{k-1} \exp\left(\Delta T \zeta_{k-1}\right). \tag{72}
$$

Define η_k such that $g_k = e^{\eta_k}$. Then we have from the above that

$$
\eta_k = \eta_{k-1} + \Delta T \frac{\mathrm{ad}_{\eta_{k-1}} \; e^{\mathrm{ad}_{\eta_{k-1}}}}{e^{\mathrm{ad}_{\eta_{k-1}}} - I} \; \zeta_{k-1} + O(\Delta T^2). \tag{73}
$$

Which gives us that

$$
\begin{aligned}
\dot{\eta} &= \left(\frac{\mathrm{ad}_\eta \; e^{\mathrm{ad}_\eta}}{e^{\mathrm{ad}_\eta} - I} \right) \; \zeta(t), && (74)\\
&= \left(I + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m(m+1)} (e^{\mathrm{ad}_\eta} - I)^m \right) \; \zeta(t) && (75)
\end{aligned}
$$

The quaternion representation of $\dot{R} = R\widehat{\Omega}$.

$$
\begin{bmatrix} \dot{q}_0 \\ \dot{w} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -\Omega \cdot w \\ q_0 \Omega - \Omega \times w \end{bmatrix}
$$

$$
\dot{q} = \frac{1}{2} \begin{bmatrix} 0 & -\Omega^T \\ \Omega & -\widehat{\Omega} \end{bmatrix} q
$$

$$
\exp\left(\widehat{\Omega}\right) = I + \frac{\sin ||\Omega||}{||\Omega||} \widehat{\Omega} + \frac{1}{2} \left(\frac{\sin \frac{||\Omega||}{2}}{\frac{||\Omega||}{2}} \right)^2 \widehat{\Omega}^2. \tag{76}
$$

B. Exponential Convergence of the 1D Kalman Filter

Lemma 1: Consider the discrete time Markov process given by

$$
\begin{aligned}
\mu_t &= \alpha_{t-1} \mu_{t-1} + n_\mu \\
X_t &= h_t \mu_t + n_x,
\end{aligned}
$$

where, $p(n_\mu) \sim \mathcal{N}(0, \sigma_\mu^2)$ and $p(n_x) \sim \mathcal{N}(0, \sigma_x^2)$ and the deterministic sequences $\{h_t\}$, $\{\alpha_t\}$ satisfies $0 < h_{\min} < h_t < h_{\max}$, $0 < \alpha_{\min} < \alpha_t < \alpha_{\max}$ for all t and the Gaussian estimator

$$
\begin{aligned}
K_t &= \frac{h_t \alpha_t^2 P_{t-1}}{h_t^2(\alpha_t^2 P_{t-1} + \sigma_\mu^2) + \sigma_m^2} \\
\widehat{\mu}^t &= \alpha_{t-1} \widehat{\mu}_{t-1} + K_t (X_t - h_t \alpha_{t-1} \widehat{\mu}_{t-1}), \\
P_t &= (\alpha_{t-1})^2 (1 - K_t h_t) P_{t-1} + \sigma_\mu^2 = \frac{h_t^2 \sigma_\mu^2 + \sigma_m^2}{h_t^2(\alpha_t^2 P_{t-1} + \sigma_\mu^2) + \sigma_m^2} P_{t-1} + \sigma_\mu^2.
\end{aligned}
$$

Proof of Lemma 1: Note that when $\sigma_\mu > 0$

$$P_t = P_0 \prod_{j=1}^{t}(1-h_jK_j)(\alpha_{j-1})^2 + \sigma_\mu^2 \sum_{j=1}^{t}\prod_{r=0}^{j-1}(1-h_{t-r}K_{t-r})(\alpha_{t-r-1})^2.$$

where $S_m = \inf_{t\geq 0} S_t$. Let $\beta > 0$ be a constant such that $\beta = \inf_{t\geq 0} h_t K_t = \frac{h_{\min}^2(\alpha_{\min}^2 S_m + \sigma_\mu^2)}{h_{\min}^2(\alpha_{\min}^2 S_m+\sigma_\mu^2)+\sigma_x^2}$. Then we have,

$$S_i^t \leq (1-\beta)^t \alpha^2 S_0 + \left(\frac{1-\beta}{\beta}\right)\alpha^2\sigma_x^2$$

Note that if $\sigma_\mu > 0$ then $\beta > 0$ and thus $\lim_{t\to\infty} S_t = S_\infty < \infty$ when $\sigma_\mu > 0$. Let λ be the constant such that $-\log(1-\beta) = \lambda$. Then we have,

$$S_t \leq \alpha^2 S_0 \exp(-\lambda t) + \eta$$

where $\eta = \sigma_x^2\sigma_x^2/(\alpha_{\min}^2 S_m + \sigma_\mu^2)$. Thus proving the exponential convergence of the variance of $\widehat{\mu}_t = E_\mu(\mu_t|X^t)$ as long as $\sigma_\mu > 0$. Note that this proof also shows that any constant $K > 0$ such that $0 < Kh_{\min} < Kh_{\max} < 1$ also ensures exponential convergence of S_t even when $\sigma_\mu = 0$.

C. Gaussian Processes Regression

Let G be some set. Consider a one dimensional Gaussian process $X_g \sim \mathcal{N}(\mu_g, \kappa(g,g))$ where $g \in G$ and $\kappa : G\times G \mapsto \mathbb{R}$ is the covariance $E((X_g-\mu_g)(X_{g'}-\mu_{g'})) = \kappa(g,g')$. Let Y_g be another stochastic process that satisfies

$$Y_g = X_g + \nu,$$

where $\nu \sim \mathcal{N}(0,\sigma_m)$. This represents a noisy sampling of X_g. If one samples the stochastic process Y_g at different g and have the set of observations $y_n \triangleq [y_{g_1}, y_{g_2}, \cdots, y_{g_n}]^T$, the Gaussian Process (GP) regression is the problem of estimating Y_g given the observations y_n.

Consider the multidimensional random variable $\mathcal{Y}_n \triangleq [Y_{g_1}, Y_{g_2}, \cdots, Y_{g_n}]^T$ and $\mathcal{X}_n \triangleq [X_{g_1}, X_{g_2}, \cdots, X_{g_n}]^T$. Then $\mathcal{Y}_n \sim \mathcal{N}(\mu_n, K_n + \sigma_m^2 I)$ where, $\mu_n \triangleq [\mu(g_1), \mu(g_2), \cdots, \mu(g_n)]^T$ and K_n is the $n\times n$ auto-correlation matrix

$$K_n \triangleq E\left((\mathcal{X}_n-\mu_n)(\mathcal{X}_n-\mu_n)^T\right) = \begin{bmatrix} \kappa(g_1,g_1) & \kappa(g_1,g_2) & \cdots & \kappa(g_1,g_n) \\ \vdots & \vdots & \vdots & \vdots \\ \kappa(g_n,g_1) & \kappa(g_n,g_2) & \cdots & \kappa(g_n,g_n) \end{bmatrix}.$$

Thus the Gaussian Process (GP) regression problem boils down to finding $p(Y_g|\mathcal{Y}_n = y_n)$. We see that

$$p\begin{pmatrix} Y_g \\ \mathcal{Y}_n \end{pmatrix} = \mathcal{N}\left(\begin{bmatrix}\mu(g) \\ \mu_n\end{bmatrix}, \begin{bmatrix}\kappa(g,g) & k^T \\ k & K_n\end{bmatrix} + \sigma_m^2 I\right).$$

where $k = [\kappa(g_1,g), \kappa(g_2,g), \cdots, \kappa(g_n,g)]^T$. Let $p(Y_g \mid \mathcal{Y}_n = y_n) \sim \mathcal{N}(\widehat{\mu}_g, \sigma_g^2)$. From the properties of normal distributions (please refer to [] for formulas for constructing the conditional probabilities of Gaussian distributions) we find that

$$\widehat{\mu}_g = \mu(g) + k^T(K_n + \sigma_m^2 I)^{-1}(y_n - \mu_n), \tag{77}$$

$$\sigma_g^2 = \sigma_m^2 + \kappa(g,g) - k^T(K_n + \sigma_m^2 I)^{-1}k. \tag{78}$$

Thus the GP regression problem boils down to finding the appropriate auto-correlation function $\kappa(g,g')$.

References

[1] K. C. Wolfe, M. Mashner, and G. S. Chirikjian, "Bayesian fusion on lie groups," *Journal Of Algebraic Statistics*, vol. 2, no. 1, pp. 75–97, 2011.

[2] S. Bonnable, P. Martin, and E. Salan, "Invariant extended kalman filter: theory and application to a velocity-aided attitude estimation problem," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, Dec 2009, pp. 1297–1304.

[3] S. Bonnabel, "Left-invariant extended kalman filter and attitude estimation," in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 1027–1032.

[4] G. Bourmaud, R. Mgret, A. Giremus, and Y. Berthoumieu, "Discrete extended kalman filter on lie groups," in *21st European Signal Processing Conference (EUSIPCO 2013)*, Sept 2013, pp. 1–5.

[5] G. S. Chirikjian, "Information theory on lie groups and mobile robotics applications," in *2010 IEEE International Conference on Robotics and Automation*, May 2010, pp. 2751–2757.

[6] Y. Wang and G. S. Chirikjian, "Error propagation on the euclidean group with applications to manipulator kinematics," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, Aug 2006.

[7] M. J. Piggott and V. Solo, "Stochastic numerical analysis for brownian motion on so(3)," in *53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 3420–3425.

[8] O. Tuzel, F. Porikli, and P. Meer, "Learning on lie groups for invariant detection and tracking," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.

[9] A. Barrau and S. Bonnabel, "Intrinsic filtering on lie groups with applications to attitude estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, Feb 2015.

[10] S. Bonnabel and A. Barrau, "An intrinsic cramar-rao bound on so(3) for (dynamic) attitude filtering," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 2158–2163.

[11] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, April 2017.

[12] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a lie group," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 367–377, Feb 2010.

[13] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetry-preserving observers on lie groups," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1709–1713, July 2009.

[14] M. Izadi and A. K. Sanyal, "Rigid body attitude estimation based on the lagrange–d'alembert principle," *Automatica*, vol. 50, no. 10, pp. 2570 – 2577, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109814003112

[15] S. Bonnabel and J. J. Slotine, "A contraction theory-based analysis of the stability of the deterministic extended kalman filter," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 565–569, Feb 2015.