Notes on Jacobian of the PSIMU measurement model with local velocity difference

Creation Date: 2022-July-27
Previous Edit: 2022-August-21
Author: MT

Preliminary

In this set of notes, we follow a similar notation scheme as the one used in [1]. This document never means to be rigorous in terms of narration, but the key idea is to keep a record of the derivation of the Jacobian for the PSIMU measurement model.

1 Notation

1.1 General Notation

Symbol	Description
$\mathbf{R}$	3 x 3 rotation matrix on $SO(3)$
ϕ or ψ	3 x 1 axis-angle representation of rotation
$\mathbf{v}$	3 x 1 translational velocity vector
$(\cdot)^\wedge$	skew-symmetric operator, $\boldsymbol{a}^\wedge = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}^\wedge = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}_{3\times 3} \in \mathfrak{so}(3)$
$\exp(\cdot)$	matrix exponential map, where, $\mathbf{R} = \exp(\phi^\wedge)$

1.2 Frame-specific Notation

Symbol	Description
$\mathbf{R}_{iv_s}$	rotation from the **starting** vehicle frame among a sliding window to inertial frame
$\mathbf{v}_i^{v_s i}$	translational velocity of the **starting** vehicle frame among a sliding window with respect to inertial frame, expressed in inertial frame
$\mathbf{v}_i^{v_f i}$	velocity of the **final** vehicle frame among a sliding window with respect to inertial frame, expressed in inertial frame
$\gamma_{v_s}^{v_f v_s}$	velocity difference of the final frame with respect to the starting frame, expressed in the starting frame among a sliding window. It is defined as $\gamma_{v_s}^{v_f v_s} := \mathbf{R}_{iv_s}^T(\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i)$
$\mathbf{y}_{v_s}^{v_f v_s}$	measurement of velocity difference provided by the PSIMU model; $\mathbf{y}_{v_s}^{v_f v_s} = \gamma_{v_s}^{v_f v_s}$, ideally. However, equality does not hold due to noise and biases.
$\mathbf{g}_i$	gravitational acceleration in inertial frame. Assumed to be constant as $[0, 0, -9.8x]^T$
Δt	time span of a sliding window, i.e., the wall-clock time of the data sequence

2 Jacobian of the PSIMU measurement model

In the pseudo-IMU (PSIMU) setup, the model measurement is obtained by sending a sequence (i.e., a sliding window) of body-frame IMU readings into a convolutional neural network (CNN). The CNN outputs a 3-vector representing the velocity difference between the last and first timesteps of the sequence, expressed in the vehicle frame at the first timestep. For a given sliding window, we denote the measurement as $\mathbf{y}_{v_s}^{v_f v_s}$, which is a 3×1 vector in $\mathbb{R}^3$.

Ideally, we have the equation

$$\mathbf{y}_{v_s}^{v_f v_s} = \gamma_{v_s}^{v_f v_s} = \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t), \tag{1}$$

where the definitions of terms are given in Section 1.2; and $\mathbf{R}_{iv_s}$, $\mathbf{v}_i^{v_f i}$, and $\mathbf{v}_i^{v_s i}$ are vehicle state that we aim to estimate. Nonetheless, the first equality does not hold in practice due to the existence of noises and biases. Thus, we can define an error term for each given sliding window as

$$\mathbf{e}_{v_s} = \mathbf{y}_{v_s}^{v_f v_s} - \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t), \tag{2}$$

where $\mathbf{e}_{v_s}$ is a 3×1 vector representing the error of the velocity difference, expressed in the starting vehicle frame among the sliding window.

We now want to take the derivative of $\mathbf{e}_{v_s}$ with respect to the vehicle state involved in Eq. (2). There are several ways to perform this task, where each achieves similar or the same results. In this note, we show two ways in the following sections that achieve the same result.

2.1 Taking Derivative w.r.t. Vehicle State Separately

The derivatives w.r.t. the velocity terms are more convenient, thus we will start with them. For $\mathbf{v}_i^{v_f i}$, we have

$$\begin{aligned}
\frac{\partial \mathbf{e}_{v_s}}{\partial \mathbf{v}_i^{v_f i\,T}} &= \frac{\partial \mathbf{y}_{v_s}^{v_f v_s}}{\partial \mathbf{v}_i^{v_f i\,T}} - \frac{\partial \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t)}{\partial \mathbf{v}_i^{v_f i\,T}} && \text{(3a)} \\
&= \mathbf{0}_{3\times3} - \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_f i}}{\partial \mathbf{v}_i^{v_f i\,T}} + \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_s i}}{\partial \mathbf{v}_i^{v_f i\,T}} + \mathbf{R}_{iv_s}^T \frac{\mathbf{g}_i \, \Delta t}{\partial \mathbf{v}_i^{v_f i\,T}} && \text{(3b)} \\
&= \mathbf{0}_{3\times3} - \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_f i}}{\partial \mathbf{v}_i^{v_f i\,T}} + \mathbf{0}_{3\times3} + \mathbf{0}_{3\times3} && \text{(3c)} \\
&= -\mathbf{R}_{iv_s}^T \, \mathbf{I}_{3\times3} && \text{(3d)} \\
&= -\mathbf{R}_{iv_s}^T. && \text{(3e)}
\end{aligned}$$

Note that there is a transpose at the denominator of the partial derivative, such that the numerator is a 3×1 vector, and the denominator is a 1×3 vector, resulting in a 3×3 matrix as the Jacobian. This notation convention follows the one used in [2, Appendix B] and the typical machine learning literature. Other than explicitly showing the transpose at the denominator, one can follow the notation convention, for instance, in [1] or in Section 3.3.5 of [2]. Gao and Zhang in [2, Appendix B] also provide a discussion on the notation of matrix derivatives. One should keep in mind that the different notation convention means the same derivative, despite the slight variation in human-readable symbolic terms. The choice of notation here is solely a preference of the author to keep completeness at the cost of verboseness.

Similarly, for $\mathbf{v}_i^{v_s i}$, we have

$$
\begin{aligned}
\frac{\partial \mathbf{e}_{v_s}}{\partial \mathbf{v}_i^{v_s i^T}} &= \frac{\partial \mathbf{y}_{v_s}^{v_f v_s}}{\partial \mathbf{v}_i^{v_s i^T}} - \frac{\partial \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t)}{\partial \mathbf{v}_i^{v_s i^T}} && \text{(4a)} \\
&= \mathbf{0}_{3\times 3} - \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_f i}}{\partial \mathbf{v}_i^{v_s i^T}} + \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_s i}}{\partial \mathbf{v}_i^{v_s i^T}} + \mathbf{R}_{iv_s}^T \frac{\mathbf{g}_i \, \Delta t}{\partial \mathbf{v}_i^{v_s i^T}} && \text{(4b)} \\
&= \mathbf{0}_{3\times 3} - \mathbf{0}_{3\times 3} + \mathbf{R}_{iv_s}^T \frac{\partial \mathbf{v}_i^{v_s i}}{\partial \mathbf{v}_i^{v_s i^T}} + \mathbf{0}_{3\times 3} && \text{(4c)} \\
&= \mathbf{R}_{iv_s}^T \, \mathbf{I}_{3\times 3} && \text{(4d)} \\
&= \mathbf{R}_{iv_s}^T. && \text{(4e)}
\end{aligned}
$$

To take the derivative w.r.t. the rotation, one can proceed with two different methods. The first method refers to as the "Perturbation Model", while the second refers to as the "Derivative Model" in [2]. Although the "Perturbation Model" is more commonly used in practice, we will also present the latter. The two methods result in similar but slightly different results. But in practice, using both methods in an optimization problem with the same initial condition can often result in a similar point.

2.1.1 Derivative w.r.t. rotation using the Perturbation Model

To take the derivative of $\mathbf{e}_{v_s}$ w.r.t. the rotation $\mathbf{R}_{iv_s}$ using the perturbation method, we follow the idea described in Section 3.3.4 of [2], but a right perturbation is used rather than the originally left perturbation. The usage of right perturbation is to account for the fact that we often keep the perturbations on the "vehicle" side of the pose, and we are working with $\mathbf{R}_{iv_s}$ as this is the most appropriate way to express the IMU motion model [1, Section 9.4.3]. Note that right perturbation is commonly used in the robotic literature when dealing with IMU-based motion models, such as Forster et al. in [3], Brossard et al. in [4], Sola in [5], and Barfoot in Section 9.4 of [1]. Barfoot in [1, Section 9.4.9] provides a discussion of the relations between the left and right perturbations.

To proceed with the derivative, we first define

$$
\gamma_i^{v_f v_s} := \mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t, \tag{5}
$$

which is the velocity difference of the final frame with respect to the starting frame among the sliding window, expressed in **inertial frame**. In the perturbation model, we adopt the $SO(3)$-sensitive perturbation scheme to introduce a right perturbation, $\exp(\psi_{v_s}^{\wedge})$, over the rotation matrix, $\mathbf{R}_{iv_s} = \exp(\phi_{v_s}^{\wedge})$, to see the change of the result relative to the disturbance [2]. In this way, the derivative is given as:

$$
\begin{aligned}
\frac{\partial \mathbf{e}_{v_s}}{\partial {\psi_{v_s}}^T} &= \frac{\partial \mathbf{y}_{v_s}^{v_f v_s}}{\partial {\psi_{v_s}}^T} - \frac{\partial \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t)}{\partial {\psi_{v_s}}^T} && \text{(6a)} \\
&= \mathbf{0}_{3\times 3} - \frac{\partial \mathbf{R}_{iv_s}^T}{\partial {\psi_{v_s}}^T} (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i \, \Delta t) && \text{(6b)}
\end{aligned}
$$

$$= -\frac{\partial \mathbf{R}_{iv_s}^T}{\partial \boldsymbol{\psi}_{v_s}{}^T}\, \boldsymbol{\gamma}_i^{v_f v_s} \tag{6c}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\left(\mathbf{R}_{iv_s} \exp\left(\boldsymbol{\psi}_{v_s}^{\wedge}\right)\right)^T - \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6d}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\exp\left(\boldsymbol{\psi}_{v_s}^{\wedge}\right)^T \mathbf{R}_{iv_s}^T - \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6e}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\exp\left(-\boldsymbol{\psi}_{v_s}^{\wedge}\right) \mathbf{R}_{iv_s}^T - \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6f}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\left(\mathbf{I}_{3\times 3} - \boldsymbol{\psi}_{v_s}^{\wedge}\right) \mathbf{R}_{iv_s}^T - \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6g}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\mathbf{R}_{iv_s}^T - \boldsymbol{\psi}_{v_s}^{\wedge} \mathbf{R}_{iv_s}^T - \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6h}$$

$$= -\left(\lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{-\boldsymbol{\psi}_{v_s}^{\wedge} \mathbf{R}_{iv_s}^T}{\boldsymbol{\psi}_{v_s}{}^T} \right) \boldsymbol{\gamma}_i^{v_f v_s} \tag{6i}$$

$$= - \lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{-\boldsymbol{\psi}_{v_s}^{\wedge} \mathbf{R}_{iv_s}^T \boldsymbol{\gamma}_i^{v_f v_s}}{\boldsymbol{\psi}_{v_s}{}^T} \tag{6j}$$

$$= - \lim_{\boldsymbol{\psi} \to \mathbf{0}} \frac{\left(\mathbf{R}_{iv_s}^T \boldsymbol{\gamma}_i^{v_f v_s}\right)^{\wedge} \boldsymbol{\psi}_{v_s}}{\boldsymbol{\psi}_{v_s}{}^T} \tag{6k}$$

$$= -\left(\mathbf{R}_{iv_s}^T \boldsymbol{\gamma}_i^{v_f v_s}\right)^{\wedge} \mathbf{I}_{3\times 3} \tag{6l}$$

$$= -\left(\mathbf{R}_{iv_s}^T \boldsymbol{\gamma}_i^{v_f v_s}\right)^{\wedge} \tag{6m}$$

$$= -\left(\mathbf{R}_{iv_s}^T \left(\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i\, \Delta t\right)\right)^{\wedge}. \tag{6n}$$

In Eq. (6), the third equality (6c) is obtained by substituting (5) into (6b); the fourth equality (6d) is obtained by the definition of the derivative [2, Section 3.3.3]; the sixth equality (6f) uses the identity of [3, Section V.A]

$$\exp\left(\boldsymbol{\varphi}^{\wedge}\right)^T \equiv \exp\left(-\boldsymbol{\varphi}^{\wedge}\right), \tag{7}$$

where $\boldsymbol{\varphi}$ is a 3×1 vector in $\mathbb{R}^3$, and $\boldsymbol{\varphi}^{\wedge} \in \mathfrak{so}(3)$; the seventh equality (6g) uses the relationship of [3, Section III.A]

$$\exp\left(\boldsymbol{\varphi}^{\wedge}\right) \approx \mathbf{I}_{3\times 3} + \boldsymbol{\varphi}^{\wedge}, \tag{8}$$

where the relationship holds when $\boldsymbol{\varphi}$ is small, and the approximation approaches equality because the limit is taken [2, Section 3.3.3]; the eleventh equality (6k) uses the identity of [3, Section III.A] [1, Section 7.2.5]

$$\mathbf{b}^{\wedge}\, \mathbf{c} \equiv -\mathbf{c}^{\wedge}\, \mathbf{b}, \quad \forall\, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3; \tag{9}$$

and the last equality (6n) is obtained by substituting (5) into $\boldsymbol{\gamma}_i^{v_f v_s}$.

At this point, we have acquired the derivatives of the measurement error, $\mathbf{e}_{v_s}$, w.r.t. the state $\mathbf{R}_{iv_s}$, $\mathbf{v}_i^{v_s i}$, and $\mathbf{v}_i^{v_f i}$ as (6), (4), and (3), respectively. In (6), the derivative is obtained using the perturbation model. Next, we also include the derivative w.r.t. the rotation using the Derivative

Model (Section 3.3.3 in [2]) for the sake of completeness. However, one should note that we generally use the (6), (4), and (3) in code implementation.

2.1.2 Derivative w.r.t. rotation using the Derivative Model

To take the derivative of $\mathbf{e}_{v_s}$ w.r.t. the rotation $\mathbf{R}_{iv_s}$ using the derivative model, we follow the idea described in Section 3.3.3 of [2], but a right perturbation is used rather than the originally left perturbation, similar to Section 2.1.1.

One should note that the Derivative Model (described in [2]) is closely related to the so-called middle perturbation described in [1]. Barfoot in [1, Section 8.3.1] provides a discussion on the different perturbation options (left, middle, right), and one may also read [6] for a short notes on the topic.

Recall that we have $\mathbf{R}_{iv_s} = \exp(\phi_{v_s}^{\wedge})$, where $\phi_{v_s}^{\wedge}$ and ϕ_{v_s} are the corresponding lie algebra and tangential vector of $\mathbf{R}_{iv_s}$, respectively. Then, the derivative of $\mathbf{e}_{v_s}$ w.r.t. the rotation $\mathbf{R}_{iv_s}$ is given as [2, Section 3.3.3]:

$$
\begin{aligned}
\frac{\partial \mathbf{e}_{v_s}}{\partial {\phi_{v_s}}^T} &= \frac{\partial \mathbf{y}_{v_s}^{v_f v_s}}{\partial {\phi_{v_s}}^T} - \frac{\partial \mathbf{R}_{iv_s}^T (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i\,\Delta t)}{\partial {\phi_{v_s}}^T} && \text{(10a)}\\
&= \mathbf{0}_{3\times 3} - \frac{\partial \exp(\phi_{v_s}^{\wedge})^T \gamma_i^{v_f v_s}}{\partial {\phi_{v_s}}^T} && \text{(10b)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\exp((\phi_{v_s} + \delta\phi_{v_s})^{\wedge})^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10c)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\left(\exp(\phi_{v_s}^{\wedge})\exp((\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge})\right)^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10d)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\exp((\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge})^T \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10e)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\exp(-(\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge}) \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10f)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{(\mathbf{I}_{3\times 3} - (\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge}) \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10g)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\exp(\phi_{v_s}^{\wedge})^T \gamma_i^{v_f v_s} - (\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge} \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s} - \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10h)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{-(\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge} \exp(\phi_{v_s}^{\wedge})^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10i)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{-(\,\mathrm{J_r}\,\delta\phi_{v_s})^{\wedge}\ \mathbf{R}_{iv_s}^T\ \gamma_i^{v_f v_s}}{\delta\phi_{v_s}^T} && \text{(10j)}\\
&= -\lim_{\delta\phi\to\mathbf{0}} \frac{\left(\mathbf{R}_{iv_s}^T\ \gamma_i^{v_f v_s}\right)^{\wedge}\ \mathrm{J_r}\,\delta\phi_{v_s}}{\delta\phi_{v_s}^T} && \text{(10k)}\\
&= -\left(\mathbf{R}_{iv_s}^T\ \gamma_i^{v_f v_s}\right)^{\wedge}\ \mathrm{J_r}\,\mathbf{I}_{3\times 3} \quad = \quad -\left(\mathbf{R}_{iv_s}^T\ (\mathbf{v}_i^{v_f i} - \mathbf{v}_i^{v_s i} - \mathbf{g}_i\,\Delta t)\right)^{\wedge}\ \mathrm{J_r} && \text{(10l)}
\end{aligned}
$$

where $\mathrm{J_r} \coloneqq \mathrm{J_r}(\phi_{v_s})$ is the 3×3 right Jacobian of the tangential vector, ϕ_{v_s}. The expression of right Jacobian is given as ([3, Section III.A] and [1, Section 8.1]):

$$\mathrm{J}_r(\boldsymbol{\varphi}) = \frac{\sin\varphi}{\varphi}\,\mathbf{I}_{3\times3} + (1 - \frac{\sin\varphi}{\varphi})\,\mathbf{a}\mathbf{a}^T - \frac{1-\cos\varphi}{\varphi}\,\mathbf{a}^\wedge \tag{11a}$$

$$\equiv \mathbf{I}_{3\times3} - \frac{1-\cos\varphi}{\varphi^2}\,\boldsymbol{\varphi}^\wedge + \frac{\varphi - \sin\varphi}{\varphi^3}(\boldsymbol{\varphi}^\wedge)^2, \tag{11b}$$

where $\boldsymbol{\varphi}$ is a 3×1 tangential vector corresponding to a rotation; $\varphi = \|\boldsymbol{\varphi}\|$ is a scalar denoting the angle of rotation; and $\mathbf{a} = \frac{\boldsymbol{\varphi}}{\varphi} = \frac{\boldsymbol{\varphi}}{\|\boldsymbol{\varphi}\|}$ is a 3×1 unit vector representing the axis of rotation.

In Eq. (10), the third equality (10c) is obtained by the definition of the derivative; the fourth equality (10d) is obtained through the first order Baker-Campbell-Hausdorff (BCH) approximation ([2, Section 3.3.3], [3, Section III.A], [1, Section 8.1]):

$$\exp\left((\boldsymbol{\varphi} + \delta\boldsymbol{\varphi})^\wedge\right) \approx \exp\left(\boldsymbol{\varphi}^\wedge\right)\,\exp\Big((\,\mathrm{J_r}(\boldsymbol{\varphi})\,\delta\boldsymbol{\varphi})^\wedge\Big), \tag{12}$$

where the relationship holds when $\delta\boldsymbol{\varphi}$ is small, and the approximation reaches equality because the limit is taken; the sixth equality (10f) is obtained through the identity described in (7); the seventh equality (10g) uses the relationship described in (8), which is a first-order Taylor-series approximation omitting the higher-order terms, and the equality holds because the limit is taken [2, Section 3.3.3]; the eleventh equality (10k) is obtained by using the identity described in (9); and the last equality is achieved by substituting (5) into $\gamma_i^{v_f v_s}$.

When comparing (6) to (10), we can observe that the result derived from the Derivative Model requires the computation of an additional term of $\mathrm{J_r}(\phi_{v_s})$, while the remaining parts stay the same. In practice, it is more common to use the derivative obtained through the Perturbation Model, i.e., (6), which has a simpler form and requires fewer computation [2, Section 3.3.3 - 3.3.4].

In the next subsection, we present another way to derive the Jacobian, and the final result will be consistent with the ones given by (3), (4), and (6).

2.2 Computing Derivatives w.r.t. Vehicle State by Perturbation

Another way to approach the problem is to treat the state as Gaussian random variables, where each state variable comprises a 'large', noise-free component and a 'small' noisy component that is zero-mean [1, Section 8.3]. In this way, for $SO(3)$, a random variable, $\mathbf{R}$, will have the form

$$\mathbf{R} = \overline{\mathbf{R}}\ \exp\left(\boldsymbol{\psi}^\wedge\right), \tag{13}$$

where $\overline{\mathbf{R}} \in SO(3)$ is a 'large', noise-free, nominal rotation; and $\boldsymbol{\psi} \in \mathbb{R}^3$ is a 'small', noisy component (i.e., a random variable from a vector space) [1, Section 8.3]. Notice that (13) has a similar form as the Perturbation Model introduced in Section 2.1.1, so we can treat $\exp(\boldsymbol{\psi}^\wedge)$ as a perturbation applied to a nominal rotation, $\overline{\mathbf{R}}$.

Similarly, for velocities in $\mathbb{R}^3$, we have

$$\mathbf{v} = \overline{\mathbf{v}} + \delta\mathbf{v} \tag{14}$$

where $\overline{\mathbf{v}}$ is a noise-free nominal velocity, and $\delta\mathbf{v}$ is a zero-mean Gaussian noise (treated as a perturbation). In this way, we can formulate our state variables as

$$\mathbf{R}_{iv_s} = \overline{\mathbf{R}}_{iv_s}\ \exp\left(\boldsymbol{\psi}_{v_s}^\wedge\right) \tag{15a}$$

$$\mathbf{v}_i^{v_f i} = \overline{\mathbf{v}}_i^{v_f i} + \delta \mathbf{v}_i^{v_f i} \tag{15b}$$

$$\mathbf{v}_i^{v_s i} = \overline{\mathbf{v}}_i^{v_s i} + \delta \mathbf{v}_i^{v_s i} \, , \tag{15c}$$

where in practice, we often use our current estimate of the state as the nominal value (called the operating point), and the perturbations are the changes of values that will be used to update the state estimates.

At this point, we can substitute Eq. (15) into (2) and manipulate as

$$\begin{aligned}
\mathbf{e}_{v_s} &= \mathbf{y}_{v_s}^{v_f v_s} - \left(\overline{\mathbf{R}}_{iv_s} \; \exp\left(\boldsymbol{\psi}_{v_s}^{\wedge}\right)\right)^T \left((\overline{\mathbf{v}}_i^{v_f i} + \delta \mathbf{v}_i^{v_f i}) - (\overline{\mathbf{v}}_i^{v_s i} + \delta \mathbf{v}_i^{v_s i}) - \mathbf{g}_i \, \Delta t \right) && \text{(16a)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - \exp\left(\boldsymbol{\psi}_{v_s}^{\wedge}\right)^T \overline{\mathbf{R}}_{iv_s}^T \left((\, \overline{\mathbf{v}}_i^{v_f i} - \overline{\mathbf{v}}_i^{v_s i} - \mathbf{g}_i \, \Delta t) + \delta \mathbf{v}_i^{v_f i} - \delta \mathbf{v}_i^{v_s i} \right) && \text{(16b)} \\
&\approx \mathbf{y}_{v_s}^{v_f v_s} - (\mathbf{I}_{3\times 3} + \boldsymbol{\psi}_{v_s}^{\wedge})^T \, \overline{\mathbf{R}}_{iv_s}^T \left((\, \overline{\mathbf{v}}_i^{v_f i} - \overline{\mathbf{v}}_i^{v_s i} - \mathbf{g}_i \, \Delta t) + \delta \mathbf{v}_i^{v_f i} - \delta \mathbf{v}_i^{v_s i} \right) && \text{(16c)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - (\mathbf{I}_{3\times 3} - \boldsymbol{\psi}_{v_s}^{\wedge}) \, \overline{\mathbf{R}}_{iv_s}^T \left((\, \overline{\mathbf{v}}_i^{v_f i} - \overline{\mathbf{v}}_i^{v_s i} - \mathbf{g}_i \, \Delta t) + \delta \mathbf{v}_i^{v_f i} - \delta \mathbf{v}_i^{v_s i} \right) && \text{(16d)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - \left(\overline{\mathbf{R}}_{iv_s}^T - \boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \right) \left((\, \overline{\mathbf{v}}_i^{v_f i} - \overline{\mathbf{v}}_i^{v_s i} - \mathbf{g}_i \, \Delta t) + \delta \mathbf{v}_i^{v_f i} - \delta \mathbf{v}_i^{v_s i} \right) && \text{(16e)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - \left(\overline{\mathbf{R}}_{iv_s}^T - \boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \right) \left(\overline{\boldsymbol{\gamma}}_i^{v_f v_s} + \delta \mathbf{v}_i^{v_f i} - \delta \mathbf{v}_i^{v_s i} \right) && \text{(16f)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_f i} + \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_s i} + \boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} + \underbrace{\boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_f i} - \boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_s i}}_{\text{omit higher order perturbation terms}} && \text{(16g)} \\
&\approx \mathbf{y}_{v_s}^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_f i} + \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_s i} + \boldsymbol{\psi}_{v_s}^{\wedge} \, \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} && \text{(16h)} \\
&= \mathbf{y}_{v_s}^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_f i} + \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_s i} - \left(\overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} \right)^{\wedge} \boldsymbol{\psi}_{v_s} && \text{(16i)} \\
&= \; \overline{\mathbf{e}}_{v_s} - \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_f i} + \overline{\mathbf{R}}_{iv_s}^T \delta \mathbf{v}_i^{v_s i} - \left(\overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} \right)^{\wedge} \boldsymbol{\psi}_{v_s} && \text{(16j)} \\
&= \; \overline{\mathbf{e}}_{v_s} + \begin{bmatrix} \overline{\mathbf{R}}_{iv_s}^T & -\overline{\mathbf{R}}_{iv_s}^T & -\left(\overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} \right)^{\wedge} \end{bmatrix} \begin{bmatrix} \delta \mathbf{v}_i^{v_s i} \\ \delta \mathbf{v}_i^{v_f i} \\ \boldsymbol{\psi}_{v_s} \end{bmatrix} && \text{(16k)} \\
&= \; \overline{\mathbf{e}}_{v_s} + \mathbf{F} \, \delta \mathbf{x} \, , && \text{(16l)}
\end{aligned}$$

where we have defined

$$\begin{aligned}
\overline{\boldsymbol{\gamma}}_i^{v_f v_s} &:= \overline{\mathbf{v}}_i^{v_f i} - \overline{\mathbf{v}}_i^{v_s i} - \mathbf{g}_i \, \Delta t \, , && \text{(17a)} \\
\overline{\mathbf{e}}_{v_s} &:= \mathbf{y}_{v_s}^{v_f v_s} - \overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} \, , && \text{(17b)} \\
\mathbf{F} &:= \begin{bmatrix} \overline{\mathbf{R}}_{iv_s}^T & -\overline{\mathbf{R}}_{iv_s}^T & -\left(\overline{\mathbf{R}}_{iv_s}^T \, \overline{\boldsymbol{\gamma}}_i^{v_f v_s} \right)^{\wedge} \end{bmatrix}_{3\times 9} \, , && \text{(17c)} \\
\delta \mathbf{x} &:= \begin{bmatrix} \delta \mathbf{v}_i^{v_s i} \\ \delta \mathbf{v}_i^{v_f i} \\ \boldsymbol{\psi}_{v_s} \end{bmatrix}_{9\times 1} \, . && \text{(17d)}
\end{aligned}$$

In Eq. (16), the approximation (16c) is obtained by using (8); the fourth equality (16d) is obtained through the identity [1]:

$$\mathbf{c}^{\wedge^T} \equiv -\mathbf{c}^{\wedge}, \quad \forall \; \mathbf{c} \in \mathbb{R}^3; \tag{18}$$

the sixth equality (16f) is obtained by using the definition (17a); the approximation (16h) is caused by omitting the higher-order perturbation terms; the ninth equality (16i) is obtained through the identity (9); the tenth equality (16j) substitutes in the definition (17b); and the last two equalities (16k) and (16l) are obtained by arranging the coefficient matrices and state variables into a stacked form as (17c) and (17d), respectively.

We observe that the stacked coefficient matrix (17c) is consistent with the results that we have acquired in (3), (4), and (6). We should note that Eq. (16) has the form of a 'first-order' Taylor-series approximation as

$$\mathbf{e}(\overline{\mathbf{x}} + \delta\mathbf{x}) \;\approx\; \mathbf{e}(\overline{\mathbf{x}}) + \left.\frac{\partial \mathbf{e}(\mathbf{x})}{\partial \mathbf{x}^T}\right|_{\overline{\mathbf{x}}} \delta\mathbf{x} \,,$$

where $\mathbf{x}$, $\overline{\mathbf{x}}$, and $\delta\mathbf{x}$ are the stacked state variable, nominal component of the state variable, and perturbation component of the state variable, respectively [1].

It is worthwhile to reiterate that this set of notes only documents two possible ways to derive the Jacobian with first-order approximations used in several steps. There are other forms of the Jacobian (e.g., using a higher-order approximation) which can be used in practice (and even give a more accurate result).

References

[1] Timothy D Barfoot. *State estimation for robotics*. Cambridge University Press, 2 edition, 2022.

[2] Xiang Gao and Tao Zhang. *Introduction to Visual SLAM: From Theory to Practice*. Springer Nature, 2021.

[3] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza. On-manifold preintegration for real-time visual-inertial odometry. *IEEE Transactions on Robotics*, 33(1):1–21, 2017.

[4] Martin Brossard, Axel Barrau, Paul Chauchat, and Silvère Bonnabel. Associating uncertainty to extended poses for on lie group imu preintegration with rotating earth. *IEEE Transactions on Robotics*, 38(2):998–1015, 2021.

[5] Joan Sola. Quaternion kinematics for the error-state kalman filter. *arXiv preprint arXiv:1711.02508*, 2017.

[6] M.T. Side note: Perturbation scheme on so(3) rotation. https://www.overleaf.com/read/zggdmvkjgbyv, June 03, 2022 [Online].