J Math Imaging Vis manuscript No.
The final publication is available at Springer via http://dx.doi.org/10.1007/s10851-014-0528-x

A compact formula for the derivative of a 3-D rotation in exponential coordinates

Guillermo Gallego · Anthony Yezzi

Received: date / Accepted: date

Abstract We present a compact formula for the derivative of a 3-D rotation matrix with respect to its exponential coordinates. A geometric interpretation of the resulting expression is provided, as well as its agreement with other less-compact but better-known formulas. To the best of our knowledge, this simpler formula does not appear anywhere in the literature. We hope by providing this more compact expression to alleviate the common pressure to reluctantly resort to alternative representations in various computational applications simply as a means to avoid the complexity of differential analysis in exponential coordinates.

Keywords Rotation · Lie group · exponential map · derivative of rotation · cross-product matrix · Rodrigues parameters · rotation vector.

G. Gallego is supported by the Marie Curie - COFUND Programme of the EU, as part of the Seventh Framework Programme (FP7). A. Yezzi is supported by National Science Foundation (NSF) Grants: CCF-1347191 and CMMI-1068624. This work has been partially supported by the Spanish Government under project TEC2010-20412 (Enhanced 3DTV), and by FP7's project 610691 (BRIDGET).

G. Gallego
was with Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid, E-28040, Spain. He is now with Robotics and Perception Group, AI Laboratory, University of Zürich, 8001 Zürich, Switzerland.
E-mail: guillermo.gallego@ifi.uzh.ch

A. Yezzi
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

1 Introduction

Three-dimensional rotations have numerous applications in many scientific areas, from quantum mechanics to stellar and planetary rotation, including the kinematics of rigid bodies. In particular, they are widespread in computer vision and robotics to describe the orientation of cameras and objects in the scene, as well as to describe the kinematics of wrists and other parts of a robot or a mobile computing device.

Space rotations have three degrees of freedom, and admit several ways to represent and operate with them. Each representation has advantages and disadvantages. Among the most common representations of rotations are Euler angles, axis-angle representation, exponential coordinates, unit quaternions, and rotation matrices. Euler angles [15, p. 31], axis-angle and exponential coordinates [15, p. 30] are very easy to visualize because they are directly related to world models; they are also compact representations, consisting of 3-4 real numbers. These representations are used as parametrizations of 3×3 rotation matrices [15, p. 23], which are easier to work with but require nine real numbers. Unit quaternions (also known as Euler-Rodrigues parameters) [2, 15, p. 33] are a less intuitive representation, but nevertheless more compact (4 real numbers) than 3×3 matrices, and are also easy to work with. Historical notes as well as additional references on the representations of rotations can be found in [14, p. 43], [3].

In many applications, it is not only necessary to know how to represent rotations and carry out simple group operations but also to be able to perform some differential analysis. This often requires the calculation of derivatives of the rotation matrix, for example, to find optimal rotations that control some process or that minimize some cost function (in cases where a closed form solution does not exist) [7,12]. Such is the case for the optimal pose estimation problem long studied within the computer vision and photogrammetry communities [16,9,6], as well as for other related problems [20,5,13,18].

Here we consider rotations parametrized by exponential coordinates using the well-known Euler-Rodrigues formula, and compute a compact expression, in matrix form, for the derivative of the parametrized rotation matrix. We also give a geometric interpretation of the formula in terms of the spatial decomposition given by the rotation axis. To the authors' knowledge, the result presented here has not been shown before and it fills a gap in the literature, at an intermediate point between numerical differentiation, the derivative at the identity for incremental rotations and general formulas for Lie groups. By providing this simpler, compact formula for the derivative of the rotation matrix we hope to alleviate the common pressure to reluctantly resort to alternative representations or the framework of incremental rotations (i.e., local charts) in various computational applications simply as a means to avoid the complexity of differential analysis in exponential coordinates.

The paper is organized as follows: Section 2 reviews the theory of 3-D rotations parametrized by exponential coordinates. Section 3 presents the main contributions of this paper, where proofs and secondary results have been moved to appendices for readers interested in technical details. Finally, conclusions are given in Section 4.

2 Parametrization of a rotation

In this section, we review the parametrization of space rotations using exponential coordinates, before proceeding to calculating derivatives.

A three-dimensional rotation is a circular movement of an object around an imaginary line called the rotation axis. The rotation angle measures the amount of circular displacement. Rotations preserve Euclidean distance and orientation. Algebraically, the rotation of a point $\mathbf{X} = (X, Y, Z)^\top$ to a point $\mathbf{X}' = (X', Y', Z')^\top$ can be expressed as $\mathbf{X}' = \mathtt{R}\mathbf{X}$, where the rotation matrix $\mathtt{R}$ is a 3×3 orthogonal matrix ($\mathtt{R}^\top\mathtt{R} = \mathtt{R}\mathtt{R}^\top = \mathrm{Id}$, the identity matrix) with determinant $\det(\mathtt{R}) = 1$.

The space of 3-D rotations is known as the matrix Lie group $\mathcal{SO}(3)$ (special orthogonal group of order three) [15, p. 24], and it is not isomorphic to $\mathbb{R}^3$ [19]. It has the structure of both a non-commutative group (under the composition of rotations) and a manifold for which the group operations are smooth. Since $\mathcal{SO}(3)$ is a differentiable manifold, each of its points (i.e., rotations) has a tangent space, and the corresponding vector space Skew_3 consists of all (real) 3×3 skew-symmetric matrices, which can be thought of as infinitesimal rotations [14, p. 25]. Moreover, the exponential map $\exp : \mathfrak{so}(3) \to \mathcal{SO}(3)$ can be defined, which allows one to recapture the local group structure of $\mathcal{SO}(3)$ from the Lie algebra $\mathfrak{so}(3)$, the latter consisting of Skew_3 together with the binary operation (Lie bracket or commutator) $[A, B] = AB - BA$, with $A, B \in \mathrm{Skew}_3$.

The Euler-Rodrigues formula [4][15, p. 28] states that the rotation matrix representing a circular movement of angle θ (in radians) around a specified axis $\bar{\mathbf{v}} \in \mathbb{R}^3$ is given by

$$\mathtt{R} = \mathrm{Id} + \sin\theta\, [\bar{\mathbf{v}}]_\times + (1 - \cos\theta)\, [\bar{\mathbf{v}}]_\times^2\,, \tag{1}$$

where $\bar{\mathbf{v}}$ is a unit vector, and

$$[\mathbf{a}]_\times := \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \in \mathrm{Skew}_3 \tag{2}$$

is the cross product (skew-symmetric) matrix such that $[\mathbf{a}]_\times \mathbf{b} = \mathbf{a} \times \mathbf{b}$, for all $\mathbf{a} = (a_1, a_2, a_3)^\top, \mathbf{b} \in \mathbb{R}^3$.

An alternative formula for (1) is

$$\mathtt{R} = \cos\theta\, \mathrm{Id} + \sin\theta\, [\bar{\mathbf{v}}]_\times + (1 - \cos\theta) \bar{\mathbf{v}}\bar{\mathbf{v}}^\top \tag{3}$$

because any unit vector $\bar{\mathbf{v}}$ satisfies

$$[\bar{\mathbf{v}}]_\times^2 = \bar{\mathbf{v}}\bar{\mathbf{v}}^\top - \mathrm{Id}. \tag{4}$$

The exponential coordinates [15, p. 30] given by the rotation vector $\mathbf{v} := \theta\bar{\mathbf{v}}$ are a natural and compact representation of the rotation in terms of its geometric building blocks. They are also called the canonical coordinates of the rotation group. The Euler-Rodrigues rotation formula (1) is a closed form expression of the aforementioned exponential map [15, p. 29]

$$\mathtt{R} = \exp([\mathbf{v}]_\times) := \sum_{k=0}^{\infty} \frac{1}{k!} [\mathbf{v}]_\times^k = \sum_{k=0}^{\infty} \frac{\theta^k}{k!} [\bar{\mathbf{v}}]_\times^k\,. \tag{5}$$

Moreover, since the exponential map considered is surjective, every rotation matrix can be written as (5) for some coordinates $\mathbf{v}$, specifically, those with $\|\mathbf{v}\| \le \pi$, i.e., in the closed ball of radius π in $\mathbb{R}^3$. Hence, exponential coordinates can be used either locally (to represent incremental rotations between two nearby configurations) or globally (to represent total rotations with respect to a reference one) [17]. More observations of this parametrization can be found in [10, p. 624].

To retrieve the exponential coordinates or the axis-angle representation of a rotation matrix, we use the log map, $\log : \mathcal{SO}(3) \to \mathfrak{so}(3)$, given in [14, p. 27] by

$$\theta = \|\mathbf{v}\| = \arccos\left(\frac{\mathrm{trace}(\mathtt{R}) - 1}{2}\right)$$

and, if $\theta \neq 0$ and $\mathtt{R}_{ij}$ are the entries of $\mathtt{R}$,

$$\bar{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{2\sin\theta}\left(\mathtt{R}_{32} - \mathtt{R}_{23}, \mathtt{R}_{13} - \mathtt{R}_{31}, \mathtt{R}_{21} - \mathtt{R}_{12}\right)^\top.$$

Without loss of generality, if θ is the principal value of the inverse cosine function then $\mathbf{v}$ lies within the ball of radius π, as assumed from now on. In the identity case, $\mathtt{R} = \mathrm{Id}$, then $\theta = 0$ and $\bar{\mathbf{v}}$ can be chosen arbitrarily.

3 Derivative of a rotation

Given the exponential coordinate parametrization (5), we consider the calculation of the derivative of the Rotation matrix, which is a relevant topic on its own as well as due to its broad range of applications. Although formulas exist to express the derivative of the exponential map in general Lie groups [11, p. 95][8, p. 70], they are not computationally friendly. Instead, for the rotation group, researchers commonly resort to one of the following alternatives: numerical differentiation, using a complicated analytical formula for the derivative (see (7)) or reformulating the problem using incremental rotations so that formulas for the simplified case (linearization around the identity element) are used.

Here, we bridge the gap between the aforementioned general formulas and alternatives for the rotation group by providing a simple, analytical and computationally friendly formula to calculate the derivative of a rotation. We also give the geometric interpretation in terms of the spatial decomposition according to the rotation axis. After many false starts, we report the path that lead to the formula using well-known matrix identities. We conjecture that there is a way to obtain such formula from the general one for Lie groups, but so far we have not found it.

The incremental rotation approach has the following explanation. The Lie group framework allows $\mathcal{SO}(3)$ to be locally replaced by its linearized version, i.e., the Lie algebra $\mathfrak{so}(3)$, whose vector space is the tangent space of $\mathcal{SO}(3)$ at the identity element [14, p. 26]. This element plays a key role in differential analysis with the exponential map: it shows that rotations may be linearly approximated using three so-called group generators (the standard basis for Skew_3)

$$\mathtt{G}_i := \frac{\partial}{\partial v_i}\exp([\mathbf{v}]_\times)\Big|_{\mathbf{v}=\mathbf{0}} = [\mathbf{e}_i]_\times\,, \tag{6}$$

where $\mathbf{v} = (v_1, v_2, v_3)^\top$ and $\mathbf{e}_i$ is the i-th vector of the standard basis in $\mathbb{R}^3$. And it also provides a means to calculate derivatives of rotations as long as they are written in an incremental way, e.g., $\mathtt{R} = \exp([\mathbf{v}]_\times)\mathtt{R}_0$, so that derivatives are evaluated at $\mathbf{v} = \mathbf{0}$, as in (6).

Next, let us show a formula for the derivative of a rotation at an arbitrary element, not necessarily the simplified case of the identity element of the rotation group. First we show a well-known but complicated one and then our contribution.

Stemming from the Euler-Rodrigues formula (1), the derivative of a rotation $\mathtt{R}(\mathbf{v}) = \exp([\mathbf{v}]_\times)$ with respect to its exponential coordinates $\mathbf{v}$ is given by

$$\frac{\partial \mathtt{R}}{\partial v_i} = \cos\theta\,\bar{v}_i\,[\bar{\mathbf{v}}]_\times + \sin\theta\,\bar{v}_i\,[\bar{\mathbf{v}}]^2_\times + \frac{\sin\theta}{\theta}\,[\mathbf{e}_i - \bar{v}_i\bar{\mathbf{v}}]_\times + \frac{1-\cos\theta}{\theta}\left(\mathbf{e}_i\bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top - 2\bar{v}_i\bar{\mathbf{v}}\bar{\mathbf{v}}^\top\right), \tag{7}$$

where $\theta = \|\mathbf{v}\|$ and $\bar{\mathbf{v}} = (\bar{v}_1, \bar{v}_2, \bar{v}_3)^\top = \mathbf{v}/\|\mathbf{v}\|$. Formula (7) is used, for example, in the OpenCV library [1] (having more than 50 thousand people of user community and estimated number of downloads exceeding 7 million) if the rotation vector $\mathbf{v}$ is passed as argument to the appropriate function (`cvRodrigues`). The proof of (7) is given in Appendix E.

Here, however we follow a different approach and first compute the derivative of the product $\mathtt{R}\mathbf{u}$ where $\mathbf{u}$ is independent of the exponential coordinates $\mathbf{v}$. Once obtained a compact formula, it is used to compute the derivatives of the rotation matrix itself.

Result 1 *The derivative of* $\mathtt{R}(\mathbf{v})\mathbf{u} = \exp([\mathbf{v}]_\times)\mathbf{u}$ *with respect to the exponential coordinates* $\mathbf{v}$*, where* $\mathbf{u}$ *is independent of* $\mathbf{v}$*, is*

$$\frac{\partial \mathtt{R}(\mathbf{v})\mathbf{u}}{\partial \mathbf{v}} = -\mathtt{R}\,[\mathbf{u}]_\times \frac{\mathbf{v}\mathbf{v}^\top + (\mathtt{R}^\top - Id)\,[\mathbf{v}]_\times}{\|\mathbf{v}\|^2}. \tag{8}$$

The proof is given in Appendix B.

3.1 Geometric interpretation

Let the decomposition of a vector $\mathbf{b}$ onto the subspaces parallel and perpendicular components to the rotation axis $\bar{\mathbf{v}}$ be $\mathbf{b} = \mathbf{b}_\parallel + \mathbf{b}_\perp$, where $\mathbf{b}_\parallel \propto \bar{\mathbf{v}}$ is parallel to the rotation axis and $\mathbf{b}_\perp \perp \bar{\mathbf{v}}$ lies in the plane orthogonal to the rotation axis. Then, observe that formula (8) provides insight about the action of $\partial(\mathtt{R}\mathbf{u})/\partial\mathbf{v}$ on a vector $\mathbf{b}$. Such operation has two components according to the aforementioned decomposition along/orthogonal to the rotation axis,

$$\frac{\partial \mathtt{R}\mathbf{u}}{\partial \mathbf{v}}\,\mathbf{b} = -\mathtt{R}\,[\mathbf{u}]_\times\left((\mathbf{b}_\parallel\cdot\bar{\mathbf{v}})\bar{\mathbf{v}} + \frac{(\mathtt{R}^\top - \mathrm{Id})\,[\bar{\mathbf{v}}]_\times\,\mathbf{b}_\perp}{\|\mathbf{v}\|}\right),$$

and notice that both components scale differently: the first term $(\mathbf{b}_\parallel\cdot\bar{\mathbf{v}})\bar{\mathbf{v}}$ depends solely on $\mathbf{b}_\parallel$, whereas the second term involves $[\bar{\mathbf{v}}]_\times\,\mathbf{b}_\perp/\|\mathbf{v}\|$, which depends on both $\mathbf{b}_\perp$ and $\|\mathbf{v}\|$. This information is difficult to extract by using a formula like (7).

Another way to look at the geometric interpretation of our formula is through sensitivity analysis. The first order Taylor series approximation of the rotated point $\mathbf{u}' = \mathtt{R}(\mathbf{v})\mathbf{u}$ around $\mathbf{v}$ is

$$\mathbf{u}'(\mathbf{v}+\delta\mathbf{v}) \approx \mathbf{u}'(\mathbf{v}) + \frac{\partial \mathtt{R}\mathbf{u}}{\partial \mathbf{v}}\,\delta\mathbf{v}$$
$$= \mathbf{u}'(\mathbf{v}) - \mathtt{R}\,[\mathbf{u}]_\times \Big((\delta\mathbf{v}_\parallel \cdot \bar{\mathbf{v}})\bar{\mathbf{v}} + \frac{(\mathtt{R}^\top - \mathrm{Id})\,[\bar{\mathbf{v}}]_\times\, \delta\mathbf{v}_\perp}{\|\mathbf{v}\|} \Big),$$

where $\delta\mathbf{v} = \delta\mathbf{v}_\parallel + \delta\mathbf{v}_\perp$. As the rotation $\mathtt{R}(\mathbf{v})$ is perturbed, there are two different types of changes:

- If the perturbation $\delta\mathbf{v}$ is such that only the amount of rotation changes, but not the direction of rotation (rotation axis), i.e., $\delta\mathbf{v}_\perp = \mathbf{0}$, the rotated point becomes $\mathbf{u}'(\mathbf{v}+\delta\mathbf{v}) \approx \mathbf{u}'(\mathbf{v}) - \|\delta\mathbf{v}\|\,\mathtt{R}(\mathbf{u}\times\bar{\mathbf{v}})$, where the change is proportional to the rotation of $\mathbf{u}\times\bar{\mathbf{v}}$. Equivalently, using property (16) with $\mathtt{G}=\mathtt{R}$, $\mathtt{R}(\mathbf{u}\times\bar{\mathbf{v}}) = (\mathtt{R}\mathbf{u})\times(\mathtt{R}\bar{\mathbf{v}}) = \mathtt{R}\mathbf{u}\times\bar{\mathbf{v}}$, the change is perpendicular to both $\mathtt{R}\mathbf{u}$ and the rotation axis $\bar{\mathbf{v}}$, which is easy to visualize geometrically since the change is represented by the tangent vector to the circumference traced out by point $\mathbf{u}$ as it rotates around the fixed axis $\bar{\mathbf{v}}$, $(\mathbf{u}'(\mathbf{v}+\delta\mathbf{v}) - \mathbf{u}'(\mathbf{v}))/\|\delta\mathbf{v}\|$.
- If the perturbation $\delta\mathbf{v}$ is such that only the direction of the rotation changes, but not the amount of rotation, i.e., $\delta\mathbf{v}_\parallel = \mathbf{0}$, the rotated point becomes $\mathbf{u}'(\mathbf{v}+\delta\mathbf{v}) \approx \mathbf{u}'(\mathbf{v}) - \|\mathbf{v}\|^{-1}\mathtt{R}\,[\mathbf{u}]_\times\,(\mathtt{R}^\top - \mathrm{Id})\,[\bar{\mathbf{v}}]_\times\,\delta\mathbf{v}$. The scaling is different from previous case, since now the change in $\mathbf{u}'$ depends on both $\delta\mathbf{v}_\perp$ and $\|\mathbf{v}\|$.

For an arbitrary perturbation, the change on the rotated point has two components: one due to the part of the perturbation that modifies the amount of rotation, and another one due to the part of the perturbation that modifies the direction of the rotation.

3.2 Compact formula for the derivative of the rotation matrix

Next, we use Result 1 to compute the derivatives of the rotation matrix itself with respect to the exponential coordinates (5), without re-doing all calculations.

Result 2 *The derivative of* $\mathtt{R}(\mathbf{v}) = \exp([\mathbf{v}]_\times)$ *with respect to its exponential coordinates* $\mathbf{v} = (v_1, v_2, v_3)^\top$ *is*

$$\frac{\partial \mathtt{R}}{\partial v_i} = \frac{v_i\,[\mathbf{v}]_\times + [\mathbf{v}\times(Id - \mathtt{R})\mathbf{e}_i]_\times}{\|\mathbf{v}\|^2}\,\mathtt{R}, \tag{9}$$

where $\mathbf{e}_i$ *is the i-th vector of the standard basis in* $\mathbb{R}^3$.

The proof is given in Appendix C. To conclude, we also need to show that the compact formula (9) is consistent with (7). This is demonstrated in Appendix D.

3.3 Derivative at the identity.

Our result, evaluated at the identity element, agrees with the well-known result about the so-called generators $\mathtt{G}_i$ of the group (6). This can be shown by computing the limit as $\mathbf{v}\to 0$ of (9), and using the facts that $\lim_{\mathbf{v}\to\mathbf{0}} \mathtt{R} = \mathrm{Id}$ and $\lim_{\mathbf{v}\to\mathbf{0}} (\mathrm{Id}-\mathtt{R})/\|\mathbf{v}\| = -[\bar{\mathbf{v}}]_\times$,

$$\begin{aligned}
\lim_{\mathbf{v}\to\mathbf{0}} \frac{\partial \mathtt{R}}{\partial v_i} &\overset{(9)}{=} \lim_{\mathbf{v}\to\mathbf{0}} \Big(\big(\bar{v}_i\,[\bar{\mathbf{v}}]_\times + \frac{[\bar{\mathbf{v}}\times(\mathrm{Id}-\mathtt{R})\mathbf{e}_i]_\times}{\|\mathbf{v}\|} \big)\mathtt{R} \Big) \\
&= \bar{v}_i\,[\bar{\mathbf{v}}]_\times - \big[\bar{\mathbf{v}}\times([\bar{\mathbf{v}}]_\times\,\mathbf{e}_i)\big]_\times \\
&= \Big[\bar{v}_i\bar{\mathbf{v}} - [\bar{\mathbf{v}}]_\times^2\,\mathbf{e}_i\Big]_\times \\
&\overset{(4)}{=} [\mathbf{e}_i]_\times .
\end{aligned}$$

4 Conclusion

We have provided a compact formula for the derivative of a rotation matrix in exponential coordinates. The formula is not only simpler than existing ones but it also has an intuitive interpretation according to the geometric decomposition that it provides in terms of the amount of rotation and the direction of rotation. This, together with the Euler-Rodrigues formula and the fact that exponential coordinates provide a global chart of the rotation group are supporting arguments in favor of using such parametrization for the search of optimal rotations in first-order finite-dimensional optimization techniques. In addition, the formula can also provide a simple fix for numerical implementations that are based on the derivative of a linearization of the rotation matrix in exponential coordinates.

A Some cross product relations

Let us use the dot notation for the Euclidean inner product $\mathbf{a}\cdot\mathbf{b} = \mathbf{a}^\top\mathbf{b}$. Also, let $\mathtt{G}$ be a 3×3 matrix, invertible when required so that it represents a change of coordinates in $\mathbb{R}^3$.

$$[\mathbf{a}]_\times\,\mathbf{a} = \mathbf{0} \tag{10}$$
$$[\mathbf{a}]_\times\,\mathbf{b} = -[\mathbf{b}]_\times\,\mathbf{a} \tag{11}$$
$$[\mathbf{a}]_\times\,[\mathbf{b}]_\times = \mathbf{b}\mathbf{a}^\top - (\mathbf{a}\cdot\mathbf{b})\mathrm{Id} \tag{12}$$
$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) \overset{(12)}{=} (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} \tag{13}$$
$$[\mathbf{a}\times\mathbf{b}]_\times = \mathbf{b}\mathbf{a}^\top - \mathbf{a}\mathbf{b}^\top \tag{14}$$
$$[\mathbf{a}\times\mathbf{b}]_\times = [\mathbf{a}]_\times\,[\mathbf{b}]_\times - [\mathbf{b}]_\times\,[\mathbf{a}]_\times \tag{15}$$
$$[(\mathtt{G}\mathbf{a})\times(\mathtt{G}\mathbf{b})]_\times = \mathtt{G}\,[\mathbf{a}\times\mathbf{b}]_\times\,\mathtt{G}^\top$$
$$(\mathtt{G}\mathbf{a})\times(\mathtt{G}\mathbf{b}) = \det(\mathtt{G})\mathtt{G}^{-\top}(\mathbf{a}\times\mathbf{b}) \tag{16}$$
$$[\mathbf{a}]_\times\,\mathtt{G} + \mathtt{G}^\top\,[\mathbf{a}]_\times = \mathrm{trace}(\mathtt{G})\,[\mathbf{a}]_\times - [\mathtt{G}\mathbf{a}]_\times \tag{17}$$
$$[\mathtt{G}\mathbf{a}]_\times = \det(\mathtt{G})\mathtt{G}^{-\top}\,[\mathbf{a}]_\times\,\mathtt{G}^{-1}$$

B Proof of Result 1

Proof Four terms result from applying the chain rule to (1) acting on vector $\mathbf{u}$. Let us use $\theta = \|\mathbf{v}\|$ and $\bar{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$, then

$$\frac{\partial \mathtt{R}\mathbf{u}}{\partial \mathbf{v}} = \sin\theta \frac{\partial [\bar{\mathbf{v}}]_\times \mathbf{u}}{\partial \mathbf{v}} + [\bar{\mathbf{v}}]_\times \mathbf{u} \frac{\partial \sin\theta}{\partial \mathbf{v}} + (1-\cos\theta)\frac{\partial [\bar{\mathbf{v}}]_\times^2 \mathbf{u}}{\partial \mathbf{v}} + [\bar{\mathbf{v}}]_\times^2 \mathbf{u} \frac{\partial (1-\cos\theta)}{\partial \mathbf{v}}.$$

The previous derivatives are computed next, using some of the cross product properties listed in Appendix A:

$$\frac{\partial [\bar{\mathbf{v}}]_\times \mathbf{u}}{\partial \mathbf{v}} \stackrel{(11)}{=} \frac{\partial (-[\mathbf{u}]_\times \bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}} \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}} = -[\mathbf{u}]_\times \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}},$$

with derivative of the unit rotation axis vector

$$\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}}\left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right) = \frac{1}{\theta}(\mathrm{Id} - \bar{\mathbf{v}}\bar{\mathbf{v}}^\top) \stackrel{(4)}{=} -\frac{1}{\theta}[\bar{\mathbf{v}}]_\times^2. \tag{18}$$

Also by the chain rule,

$$\frac{\partial \sin\theta}{\partial \mathbf{v}} = \frac{\partial \sin\theta}{\partial \theta}\frac{\partial \theta}{\partial \mathbf{v}} = \cos\theta\, \bar{\mathbf{v}}^\top,$$
$$\frac{\partial (1-\cos\theta)}{\partial \mathbf{v}} = -\frac{\partial \cos\theta}{\partial \theta}\frac{\partial \theta}{\partial \mathbf{v}} = \sin\theta\, \bar{\mathbf{v}}^\top,$$

and, applying the product rule twice,

$$\begin{aligned}\frac{\partial [\bar{\mathbf{v}}]_\times^2 \mathbf{u}}{\partial \mathbf{v}} &\stackrel{(4)}{=} \frac{\partial \bar{\mathbf{v}}(\bar{\mathbf{v}}^\top \mathbf{u})}{\partial \mathbf{v}} = \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}}(\bar{\mathbf{v}}^\top \mathbf{u}) + \bar{\mathbf{v}}\frac{\partial (\bar{\mathbf{v}}^\top \mathbf{u})}{\partial \mathbf{v}} \\ &= \left((\bar{\mathbf{v}}^\top \mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\right)\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}} \\ &\stackrel{(18)}{=} -\frac{1}{\theta}\left((\bar{\mathbf{v}}^\top \mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\right)[\bar{\mathbf{v}}]_\times^2,\end{aligned}$$

which can be rewritten as a sum of cross product matrix multiplications since

$$\begin{aligned}&\left((\bar{\mathbf{v}}^\top \mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\right)[\bar{\mathbf{v}}]_\times^2 \\ &\stackrel{(10)}{=} \left((\bar{\mathbf{v}}^\top \mathbf{u})\mathrm{Id} - \mathbf{u}\bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{u}^\top - \mathbf{u}\bar{\mathbf{v}}^\top\right)[\bar{\mathbf{v}}]_\times^2 \\ &\stackrel{(12),(14)}{=} -[\bar{\mathbf{v}}]_\times [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 + [\mathbf{u}\times\bar{\mathbf{v}}]_\times [\bar{\mathbf{v}}]_\times^2 \\ &\stackrel{(15),(4)}{=} -2[\bar{\mathbf{v}}]_\times [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 - [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times.\end{aligned}$$

Hence, so far the derivative of the rotated vector is

$$\frac{\partial \mathtt{R}\mathbf{u}}{\partial \mathbf{v}} = (\cos\theta [\bar{\mathbf{v}}]_\times + \sin\theta [\bar{\mathbf{v}}]_\times^2)\mathbf{u}\bar{\mathbf{v}}^\top + \frac{\sin\theta}{\theta}[\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 + \frac{1-\cos\theta}{\theta}(2[\bar{\mathbf{v}}]_\times [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 + [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times).$$

Next, multiply on the left by $\mathtt{R}^\top$ and use

$$\mathtt{R}^\top [\bar{\mathbf{v}}]_\times \stackrel{(3),(10)}{=} \cos\theta\, [\bar{\mathbf{v}}]_\times - \sin\theta\, [\bar{\mathbf{v}}]_\times^2 \tag{19}$$

to simplify the first term of $\mathtt{R}^\top \partial(\mathtt{R}\mathbf{u})/\partial \mathbf{v}$,

$$\begin{aligned}&\mathtt{R}^\top(\cos\theta\, [\bar{\mathbf{v}}]_\times + \sin\theta\, [\bar{\mathbf{v}}]_\times^2)\mathbf{u}\bar{\mathbf{v}}^\top \\ &\stackrel{(19)}{=} (\cos^2\theta\, [\bar{\mathbf{v}}]_\times - \sin^2\theta\, [\bar{\mathbf{v}}]_\times^3)\mathbf{u}\bar{\mathbf{v}}^\top \\ &\stackrel{(4)}{=} (\cos^2\theta + \sin^2\theta)\, [\bar{\mathbf{v}}]_\times \mathbf{u}\bar{\mathbf{v}}^\top \\ &\stackrel{(11)}{=} -[\mathbf{u}]_\times \bar{\mathbf{v}}\bar{\mathbf{v}}^\top.\end{aligned} \tag{20}$$

For the remaining term of $\mathtt{R}^\top \partial(\mathtt{R}\mathbf{u})/\partial \mathbf{v}$, we use the transpose of (1) and apply $[\bar{\mathbf{v}}]_\times [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times \stackrel{(12)}{=} -(\mathbf{u}^\top \bar{\mathbf{v}})[\bar{\mathbf{v}}]_\times$ to simplify

$$\begin{aligned}&\left(\mathrm{Id} - \sin\theta\, [\bar{\mathbf{v}}]_\times + (1-\cos\theta)[\bar{\mathbf{v}}]_\times^2\right)\cdot\left(\sin\theta\, [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2\right. \\ &\quad \left. + (1-\cos\theta)(2[\bar{\mathbf{v}}]_\times [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 + [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times)\right) \\ &= \sin\theta\, [\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times^2 + (1-\cos\theta)[\mathbf{u}]_\times [\bar{\mathbf{v}}]_\times \\ &+ (\mathbf{u}^\top \bar{\mathbf{v}})\left(-2(1-\cos\theta) + \sin^2\theta + (1-\cos\theta)^2\right)[\bar{\mathbf{v}}]_\times \\ &= [\mathbf{u}]_\times\left(\sin\theta\, [\bar{\mathbf{v}}]_\times^2 - (1-\cos\theta)[\bar{\mathbf{v}}]_\times^3\right) \\ &= -[\mathbf{u}]_\times (\mathtt{R}^\top - \mathrm{Id})[\bar{\mathbf{v}}]_\times,\end{aligned}$$

where the term in $(\mathbf{u}^\top \bar{\mathbf{v}})$ vanished since $\sin^2\theta - 2(1-\cos\theta) + (1-\cos\theta)^2 = 0$. Collecting terms,

$$\mathtt{R}^\top \frac{\partial \mathtt{R}\mathbf{u}}{\partial \mathbf{v}} = -[\mathbf{u}]_\times \left(\bar{\mathbf{v}}\bar{\mathbf{v}}^\top + \frac{1}{\theta}(\mathtt{R}^\top - \mathrm{Id})[\bar{\mathbf{v}}]_\times\right). \tag{21}$$

Finally, multiply (21) on the left by $\mathtt{R}$ and use $\mathtt{R}\mathtt{R}^\top = \mathrm{Id}$, $\theta = \|\mathbf{v}\|$, $\bar{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$ to obtain (8).

C Proof of Result 2

Proof Stemming from (8), we show that it is possible to write

$$\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} = \mathtt{A}\mathbf{u} \tag{22}$$

for some matrix $\mathtt{A}$ and for all vector $\mathbf{u}$ independent of $\mathbf{v}$. Thus in this operator form, $\mathtt{A}$ is indeed the representation of $\partial \mathtt{R}/\partial v_i$. First, observe that

$$\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} = \frac{\partial}{\partial v_i}(\mathtt{R}\mathbf{u}) = \frac{\partial}{\partial \mathbf{v}}(\mathtt{R}\mathbf{u})\,\mathbf{e}_i,$$

then substitute (8) and simplify using the cross-product properties to arrive at (22):

$$\begin{aligned}\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} &= -\|\mathbf{v}\|^{-2}\mathtt{R}[\mathbf{u}]_\times \left(\mathbf{v}\mathbf{v}^\top + (\mathtt{R}^\top - \mathrm{Id})[\mathbf{v}]_\times\right)\mathbf{e}_i \\ &= -\|\mathbf{v}\|^{-2}\mathtt{R}[\mathbf{u}]_\times \left(\mathbf{v}\mathbf{v}^\top + [\mathbf{v}]_\times(\mathtt{R}^\top - \mathrm{Id})\right)\mathbf{e}_i \\ &= -\|\mathbf{v}\|^{-2}\mathtt{R}[\mathbf{u}]_\times \left(\mathbf{v}v_i + \left(\mathbf{v}\times(\mathtt{R}^\top - \mathrm{Id})\mathbf{e}_i\right)\right) \\ &\stackrel{(11)}{=} \|\mathbf{v}\|^{-2}\mathtt{R}\left[v_i\mathbf{v} + \left(\mathbf{v}\times(\mathtt{R}^\top - \mathrm{Id})\mathbf{e}_i\right)\right]_\times \mathbf{u}.\end{aligned} \tag{23}$$

After some manipulations,

$$\mathtt{R}\left[v_i\mathbf{v} + \left(\mathbf{v}\times(\mathtt{R}^\top - \mathrm{Id})\mathbf{e}_i\right)\right]_\times = \left[v_i\mathbf{v} + \left(\mathbf{v}\times(\mathrm{Id} - \mathtt{R})\mathbf{e}_i\right)\right]_\times \mathtt{R},$$

and so, substituting in (23) and using the linearity of the cross-product matrix (2), the desired formula (9) is obtained.

D Agreement between derivative formulas

Here we show the agreement between (7) and (9). First, use $\theta = \|\mathbf{v}\|$ and the definition of the unit vector $\bar{\mathbf{v}} = \mathbf{v}/\theta$, to write (9) as

$$\frac{\partial \mathtt{R}}{\partial v_i} = \bar{v}_i [\bar{\mathbf{v}}]_\times \mathtt{R} + \frac{1}{\theta}[\bar{\mathbf{v}}\times(\mathrm{Id} - \mathtt{R})\mathbf{e}_i]_\times \mathtt{R}. \tag{24}$$

Using (3) and $[\bar{\mathbf{v}}]_\times \bar{\mathbf{v}} = \mathbf{0}$, it follows that

$$[\bar{\mathbf{v}}]_\times \mathtt{R} = \cos\theta\, [\bar{\mathbf{v}}]_\times + \sin\theta\, [\bar{\mathbf{v}}]_\times^2.$$

Also, since $[\bar{\mathbf{v}} \times \mathbf{b}]_\times = \mathbf{b}\bar{\mathbf{v}}^\top - \bar{\mathbf{v}}\mathbf{b}^\top$ and $\mathtt{R}^\top \bar{\mathbf{v}} = \bar{\mathbf{v}}$, it yields

$$\begin{aligned}[\bar{\mathbf{v}} \times (\mathrm{Id} - \mathtt{R})\mathbf{e}_i]_\times \mathtt{R} &= (\mathrm{Id} - \mathtt{R})\mathbf{e}_i \bar{\mathbf{v}}^\top \mathtt{R} - \bar{\mathbf{v}}\mathbf{e}_i^\top (\mathrm{Id} - \mathtt{R}^\top)\mathtt{R} \\ &= (\mathrm{Id} - \mathtt{R})\mathbf{e}_i \bar{\mathbf{v}}^\top - \bar{\mathbf{v}}\mathbf{e}_i^\top (\mathtt{R} - \mathrm{Id}) \\ &= \mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top - (\mathtt{R}\mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top \mathtt{R}),\end{aligned}$$

and expanding $\mathtt{R}\mathbf{e}_i$ and $\mathbf{e}_i^\top \mathtt{R}$ in the previous formula by means of (3), we obtain

$$\begin{aligned}&[\bar{\mathbf{v}} \times (\mathrm{Id} - \mathtt{R})\mathbf{e}_i]_\times \mathtt{R} \\ &= \mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top \\ &\quad - \left(\cos\theta\, \mathbf{e}_i \bar{\mathbf{v}}^\top + \sin\theta\, [\bar{\mathbf{v}}]_\times \mathbf{e}_i \bar{\mathbf{v}}^\top + (1 - \cos\theta)\bar{v}_i \bar{\mathbf{v}}\bar{\mathbf{v}}^\top\right) \\ &\quad - \left(\cos\theta\, \bar{\mathbf{v}}\mathbf{e}_i^\top + \sin\theta\, \bar{\mathbf{v}}\mathbf{e}_i^\top [\bar{\mathbf{v}}]_\times + (1 - \cos\theta)\bar{v}_i \bar{\mathbf{v}}\bar{\mathbf{v}}^\top\right) \\ &= (1 - \cos\theta)(\mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top - 2\bar{v}_i \bar{\mathbf{v}}\bar{\mathbf{v}}^\top) \\ &\quad - \sin\theta\left([\bar{\mathbf{v}}]_\times \mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top [\bar{\mathbf{v}}]_\times\right).\end{aligned}$$

Using property (17) with $\mathbf{a} = \bar{\mathbf{v}}$, $\mathtt{G} = \mathbf{e}_i \bar{\mathbf{v}}^\top$ we have that

$$\begin{aligned}[\bar{\mathbf{v}}]_\times \mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top [\bar{\mathbf{v}}]_\times &= \mathrm{trace}(\mathbf{e}_i \bar{\mathbf{v}}^\top)\, [\bar{\mathbf{v}}]_\times - \left[\mathbf{e}_i \bar{\mathbf{v}}^\top \bar{\mathbf{v}}\right]_\times, \\ &= \mathrm{trace}(\bar{\mathbf{v}}^\top \mathbf{e}_i)\, [\bar{\mathbf{v}}]_\times - \left[\mathbf{e}_i \|\bar{\mathbf{v}}\|^2\right]_\times \\ &= \bar{v}_i\, [\bar{\mathbf{v}}]_\times - [\mathbf{e}_i]_\times \\ &= [\bar{v}_i \bar{\mathbf{v}} - \mathbf{e}_i]_\times .\end{aligned}$$

Finally, substituting previous results in (24), the desired result (7) is obtained.

E Derivative formula with sines and cosines

Here, we show how to obtain formula (7). First, differentiate the Euler-Rodrigues rotation formula (3) with respect to the i-th component of the parametrizing vector $\mathbf{v} = \theta\bar{\mathbf{v}}$ and take into account that

$$\theta^2 = \|\mathbf{v}\|^2 \implies \frac{\partial\theta}{\partial v_i} = \frac{v_i}{\theta} =: \bar{v}_i.$$

Applying the chain rule to (3), gives

$$\begin{aligned}\frac{\partial \mathtt{R}}{\partial v_i} &= -\sin\theta\, \bar{v}_i \mathrm{Id} + \cos\theta\, \bar{v}_i\, [\bar{\mathbf{v}}]_\times + \sin\theta\, \bar{v}_i \bar{\mathbf{v}}\bar{\mathbf{v}}^\top \\ &\quad + \sin\theta\, \frac{\partial\, [\bar{\mathbf{v}}]_\times}{\partial v_i} + (1 - \cos\theta)\frac{\partial(\bar{\mathbf{v}}\bar{\mathbf{v}}^\top)}{\partial v_i}.\end{aligned} \tag{25}$$

Next, we use

$$\frac{\partial}{\partial v_i}\left(\frac{v_j}{\|\mathbf{v}\|}\right) = \begin{cases} -\frac{1}{\theta}\bar{v}_i \bar{v}_j & i \neq j \\ \frac{1}{\theta}(1 - \bar{v}_i^2) & i = j \end{cases}, \tag{26}$$

to simplify one of the terms in (25),

$$\frac{\partial\, [\bar{\mathbf{v}}]_\times}{\partial v_i} = \frac{\partial}{\partial v_i}\left[\frac{\mathbf{v}}{\|\mathbf{v}\|}\right]_\times \overset{(26)}{=} \frac{1}{\theta}\, [\mathbf{e}_i - \bar{v}_i \bar{\mathbf{v}}]_\times .$$

Applying the product rule to the last term in (25) and using

$$\frac{\partial \bar{\mathbf{v}}}{\partial v_i} = \frac{\partial}{\partial v_i}\left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right) = \frac{1}{\theta}(\mathrm{Id} - \bar{\mathbf{v}}\bar{\mathbf{v}}^\top)\mathbf{e}_i, \tag{27}$$

gives

$$\frac{\partial(\bar{\mathbf{v}}\bar{\mathbf{v}}^\top)}{\partial v_i} \overset{(27)}{=} \frac{1}{\theta}\left(\mathbf{e}_i \bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top - 2\bar{v}_i \bar{\mathbf{v}}\bar{\mathbf{v}}^\top\right).$$

Finally, substituting the previous expressions in (25), yields (7). A similar proof is outlined in [13] using Einstein summation index notation.

Acknowledgements The authors thank the anonymous reviewers for their helpful comments and suggestions.

References

1. Open Source Computer Vision (OpenCV). http://opencv.org
2. Altmann, S.: Rotations, Quaternions, And Double Groups. Clarendon Press, Oxford, UK (1986)
3. Bauchau, O., Choi, J.Y.: The Vector Parameterization of Motion. Nonlinear Dynamics **33**(2), 165–188 (2003)
4. Cheng, H., Gupta, K.: An Historical Note on Finite Rotations. J. Appl. Mech. **56**(1), 139–145 (1989)
5. Dambreville, S., Sandhu, R., Yezzi, A., Tannenbaum, A.: A Geometric Approach to Joint 2D Region-Based Segmentation and 3D Pose Estimation Using a 3D Shape Prior. SIAM J. Img. Sci. **3**(1), 110–132 (2010)
6. Enqvist, O., Kahl, F.: Robust Optimal Pose Estimation. In: Proc. 10th European Conf. on Computer Vision, ECCV '08, pp. 141–153. Springer-Verlag, Berlin, Heidelberg (2008)
7. Gurwitz, C., Overton, M.: A globally convergent algorithm for minimizing over the rotation group of quadratic forms. IEEE Trans. Pattern Anal. Mach. Intell. **11**(11), 1228–1232 (1989)
8. Hall, B.: Lie Groups, Lie algebras and representations. Graduate Text in Mathematics. Springer (2004)
9. Hartley, R., Kahl, F.: Global Optimization through Searching Rotation Space and Optimal Estimation of the Essential Matrix. In: Proc. IEEE 11th Int. Conf. on Computer Vision, ICCV 2007, pp. 1–8 (2007)
10. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press (2003)
11. Helgason, S.: Differential geometry and symmetric spaces. Pure and Applied Mathematics. Elsevier Science (1962)
12. Horn, B.K.P.: Relative Orientation. Int. J. Computer Vision **4**(1), 59–78 (1990)
13. Kuehnel, F.O.: On the minimization over SO(3) manifolds. Tech. Rep. 03.12, Research Institute for Advanced Computer Science (RIACS). NASA Ames Research Center, USA (2003)
14. Ma, Y., Soatto, S., Kosecká, J., Sastry, S.: An Invitation to 3-D Vision: from images to geometric models. Springer (2003)
15. Murray, R., Li, Z., Sastry, S.: A Mathematical Introduction to Robotic Manipulation. CRC Press (1994)
16. Olsson, C., Kahl, F., Oskarsson, M.: Optimal Estimation of Perspective Camera Pose. In: Proc. 18th Int. Conf. on Pattern Recognition, *ICPR 2006*, vol. 2, pp. 5–8 (2006)
17. Ritto-Corrêa, M., Camotim, D.: On the differentiation of the Rodrigues formula and its significance for the vector-like parameterization of Reissner–Simo beam theory. Int. J. Numerical Methods in Engineering **55**(9), 1005–1032 (2002)
18. Sandhu, R., Dambreville, S., Tannenbaum, A.: Point Set Registration via Particle Filtering and Stochastic Dynamics. IEEE Trans. Pattern Anal. Mach. Intell. **32**(8), 1459–1473 (2010)
19. Stuelpnagel, J.: On the Parametrization of the Three-Dimensional Rotation Group. SIAM Review **6**(4), 422–430 (1964)
20. Unal, G.B., Yezzi, A.J., Soatto, S., Slabaugh, G.G.: A Variational Approach to Problems in Calibration of Multiple Cameras. IEEE Trans. Pattern Anal. Mach. Intell. **29**(8), 1322–1338 (2007)