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

Guillermo Gallego, Anthony Yezzi

***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.**

***Index Terms*—Rotation, rotation group, derivative of rotation, exponential map, cross-product matrix, rotation vector, Rodrigues parameters, optimization.**

I. INTRODUCTION

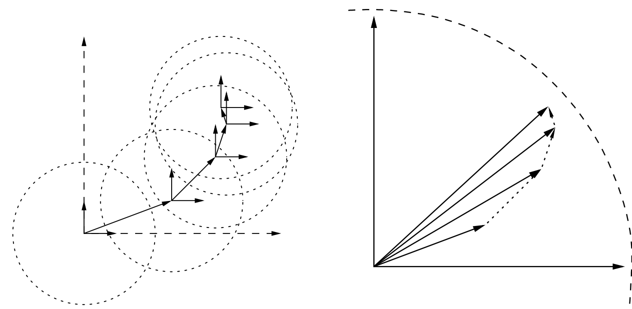
THREE-DIMENSIONAL rotations have numerous applications in many scientific areas, from quantum mechanics to stellar and planetary rotation, including the kinematics of rigid bodies. In particular, they are widespread in computer vision and robotics to describe the orientation of cameras and objects in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

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 [14, p. 31], axis-angle and exponential coordinates [14, p. 30] are very easy to visualize because they are directly related to world models; they are also compact representations, requiring 3-4 real numbers to represent rotations. These representations are used as parametrizations of 3×3 rotation matrices [14, p. 23], which are easier to work with but require nine real numbers. Unit quaternions (also known as Euler-Rodrigues parameters) [2], [14, 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 [13, 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], [10]. Such is the case for the optimal pose estimation problem long studied within the computer vision and photogrammetry communities [15], [9], [6], as well as for other related problems [19], [5], [12], [16].

G. Gallego is with the Grupo de Tratamiento de Imágenes, E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid, Madrid 28040, Spain, e-mail: ggb@gti.ssr.upm.es.

A. Yezzi is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. e-mail: ayezzi@ece.gatech.edu.



(a) Optimization by means of a sequence of local parametrizations (i.e., local charts, represented by circles centered at the corresponding chart origin and axes).

(b) Usual finite-dimensional optimization paradigm that uses a global chart to reach all valid locations of parameter space (interior of fictitious circle).

Figure I.1: Two optimization paradigms: local vs. global parametrizations of the search space.

The space of rotations, i.e., the Lie group $SO(3)$ or rotation group, has the structure of both a group and a manifold, and it is not isomorphic to $\mathbb{R}^3$ [17]. The usual approach to numerical optimization in the rotation group consists of constructing a sequence of local parametrizations (*charts* in the language of differential geometry), as illustrated in Fig. I.1a, rather than relying on a single global parametrization [18], [11], such as Euler angles, to avoid problems caused by singularities in the latter case [17]. This procedure of calculating incremental steps in the tangent space to the manifold relies on the fact that the rotation group has a natural parametrization based on the exponential operator associated with the Lie group.

The previous approach differs from most numerical finite-dimensional optimization paradigms, where the unknown parameters are assumed to lie in some vector space isomorphic to $\mathbb{R}^n$ and a global parametrization is defined, as depicted in Fig. I.1b. The exponential coordinates, however, can be used in this setting since it is a parametrization that covers the whole space of rotations, thus avoiding local charts, with an isolated and removable singularity at the origin.

Here we study 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 does not appear anywhere in the literature, and it will be shown that it agrees and simplifies the derivative computed in the OpenCV library [1], which has more than 50 thousand people of user community and estimated number of downloads exceeding 6 million. By providing this simpler, compact formula for the derivative of the rotation matrix we hope to alleviate the common pressure to reluctantly resort to 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 II reviews the theory of 3-D rotations parametrized by exponential coordinates and its use in iterative optimization methods as a motivation for the development of the main contributions of this paper, which are introduced in section III. Some proofs and secondary results are given in the form of appendices to prevent the reader from distracting from the main thread of the paper. Finally, conclusions are given in section IV.

II. 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 $\mathtt{R}$ is a 3×3 orthogonal matrix ($\mathtt{R}^\top\mathtt{R} = \mathtt{R}\mathtt{R}^\top = \mathrm{Id}$, the identity matrix) with determinant $\det(\mathtt{R}) = 1$ representing the rotation.

The space of 3-D rotations is known as the matrix Lie group $SO(3)$ (special orthogonal group) [14, p. 24]. 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. This implies that the exponential map $\exp : so(3) \to SO(3)$ can be defined, where the Lie algebra $so(3)$ associated to the rotation group $SO(3)$ consists of all skew-symmetric 3×3 matrices together with the binary operation (Lie bracket or commutator) $[A, B] = AB - BA$, with $A, B \in so(3)$.

The Euler-Rodrigues formula [4][14, 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

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

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

$$[\mathbf{a}]_\times := \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \tag{II.2}$$

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 (II.1) is

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

because for any unitary vector $\bar{\mathbf{v}}$ it is satisfied that

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

The exponential coordinates [14, p. 30] $\mathbf{v} := \theta\bar{\mathbf{v}}$ 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 (II.1) is a closed form expression of the aforementioned exponential map [14, p. 29]

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

since the powers of the skew-symmetric matrix $[\bar{\mathbf{v}}]_\times$ (with $\bar{\mathbf{v}}$ unitary) can be written in terms of Id and the first two powers, $[\bar{\mathbf{v}}]_\times$ and $[\bar{\mathbf{v}}]_\times^2$. All rotations may be represented by vectors $\|\mathbf{v}\| \leq \pi$, i.e., in the closed ball of radius π in $\mathbb{R}^3$. Hence, exponential coordinates are a global parametrization of the rotation group. More observations of this parametrization can be found in [8, p. 624].

To retrieve the exponential coordinates or the axis-angle representation of a rotation matrix, we use the log map, $\log : SO(3) \to so(3)$, given in [13, p. 27] by

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

error

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

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

If $\mathtt{R} = \mathrm{Id}$, then $\theta = 0$ and $\bar{\mathbf{v}}$ can be chosen arbitrarily.

A. Two optimization approaches: Local vs. Global parametrizations

Consider the common optimization problem consisting in the minimization of a cost function $f(\mathtt{R})$ with respect to some rotation $\mathtt{R}$. Using an iterative solver, the optimal rotation $\mathtt{R}_{\text{opt}}$ is obtained as the limit of a sequence of rotations: $\mathtt{R}_0, \mathtt{R}_1, \ldots, \mathtt{R}_m, \ldots \to \mathtt{R}_{\text{opt}}$. The exponential coordinates parametrization allows to consider two different approaches, which we compare next: local charts and global charts.

In the local charts approach, illustrated in Fig. I.1a, the rotation matrix update is encoded in an incremental way by $\mathtt{R}_m = \exp([\mathbf{w}_m]_\times)\mathtt{R}_{m-1}$, where $\exp([\mathbf{w}_m]_\times)$ is the (usually small) incremental rotation estimated at the m-th iteration and $\mathtt{R}_{m-1}$ accumulates all previous incremental rotations. Local parameter vector $\mathbf{w}_m$ is estimated by solving first order optimality conditions using a linearization of the cost function $g_m(\mathbf{w}) := f\big(\exp([\mathbf{w}]_\times)\mathtt{R}_{m-1}\big)$ around the origin ($\mathbf{w} = \mathbf{0}$). As iteration proceeds toward convergence, $\mathbf{w}_m \to \mathbf{0}$ so that the cumulative rotation $\mathtt{R}_{m-1} \to \mathtt{R}_{\text{opt}}$. This strategy requires the optimization algorithm to keep track of both $\mathbf{w}_m$ and $\mathtt{R}_{m-1}$. Conceptually, it is equivalent to the inconvenient situation of using a standard finite-dimensional optimization algorithm on a different (auxiliary) cost function per iteration, $g_m(\mathbf{w})$.

In contrast, in the global chart approach (Fig. I.1b), the rotation matrix is parametrized by $\mathtt{R}_m = \exp([\mathbf{v}_m]_\times)$, where parameter vector $\mathbf{w}_m$ is estimated by solving first order optimality conditions using a linearization of the cost function around $\mathbf{v}_{m-1}$. As iteration proceeds toward convergence, $\mathbf{v}_m \to \mathbf{v}_{\text{opt}}$ such that $\mathtt{R}_m \to \mathtt{R}_{\text{opt}} = \mathtt{R}(\mathbf{v}_{\text{opt}})$. Hence, only the usual cumulative parameter vector needs to be stored and standard optimization algorithms may be used to update $\mathtt{R}_m$ through $\mathbf{v}_m$ in the original cost function, $g(\mathbf{v}) := f\big(\mathtt{R}(\mathbf{v})\big)$. Thus, no auxiliary cost function needs to be defined and updates in the parameter vector are handled with simple addition, $\mathbf{v}_m = \mathbf{v}_{m-1} + \Delta\mathbf{v}$, like a regular optimization in $\mathbb{R}^3$. Finally, note that the exponentiation of such a simple addition is, in general, $\exp([\mathbf{a}+\mathbf{b}]_\times) \neq \exp([\mathbf{a}]_\times)\exp([\mathbf{b}]_\times)$, otherwise both local and global approaches would coincide.

B. *Change of coordinates*

The previous global parametrization can be given with respect to any rotation $\mathtt{R}_0$ (not necessarily the identity) by a suitable change of coordinates. For example, rotations of the form $\mathtt{R}(\mathbf{v}) = \exp([\mathbf{v}]_\times)\mathtt{R}_0$ are referenced with respect to a given rotation $\mathtt{R}_0$. Again, all rotations may be represented by vectors $\|\mathbf{v}\| \leq \pi$ and they represent incremental rotations with respect to the origin of rotations, $\mathtt{R}(\mathbf{0}) = \mathtt{R}_0$.

This can be useful in optimization problems, where an initial guess of the optimal rotation is given and the optimization algorithm (e.g., gradient descent) proceeds in a local search around that initial point. By translating the origin of rotations to that initial guess using the previous change of coordinates so that rotations are parameterized by $\mathtt{R}(\mathbf{v}) = \exp([\mathbf{v}]_\times)\mathtt{R}_0$, the optimization algorithm will focus on exploring rotations close to the origin, i.e., $\mathtt{R}_0$, thus avoiding the need to worry about *i*) whether antipodal rotations corresponding to $\mathtt{R}(\mathbf{v})$ with $\|\mathbf{v}\| = \pi$ (180 degrees) will be covered twice by the parametrization and *ii*) whether the explored rotations will be jumping around that topological boundary ($\|\mathbf{v}\| = \pi$) so that some renormalization [8, p. 624] is needed to ensure that the iteratively updated parameter vector $\mathbf{v}$ lies within the valid sphere $\|\mathbf{v}\| \leq \pi$. Under the hypothesis of a local search around the origin $\mathtt{R}_0$ both previous issues are avoided, and even though SO(3) is not isomorphic to $\mathbb{R}^3$, in this case one may consider that is actually carrying out a regular optimization in $\mathbb{R}^3$, where updates can be handled with simple addition.

Also, observe that a rotated vector $\mathbf{X}' = \exp([\mathbf{v}]_\times)\mathtt{R}_0\mathbf{X}$ may be expressed as $\mathbf{X}' = \exp([\mathbf{v}]_\times)\tilde{\mathbf{X}}$ with the reference rotation $\mathtt{R}_0$ subsumed in $\tilde{\mathbf{X}} = \mathtt{R}_0\mathbf{X}$, hence it is just necessary to study the behavior of the $\exp([\mathbf{v}]_\times)$ part regardless of the origin of rotations. According to the previous arguments, in the following we may assume, without loss of generality, that the parameter vector $\mathbf{v}$ is within the ball of radius π centered at some origin $\mathtt{R}_0$ possibly different from the identity.

III. Derivative of a rotation

The calculation of derivatives of a Rotation matrix is a relevant topic on its own due to its wide range of applications, but it is also needed in many optimization techniques that iteratively solve for optimal rotations. As mentioned in II-A, at each iteration of an optimization algorithm, the update on the parameter vector is computed using a linearization of the cost function, hence the importance of knowing how to compute and interpret derivatives of the rotation matrix.

Stemming from the Euler-Rodrigues formula (II.1), the derivative of a rotation $\mathtt{R}(\mathbf{v}) = \exp([\mathbf{v}]_\times)$ with respect to exponential coordinates $\mathbf{v} = (v_1, v_2, v_3)^\top$ can be computed as

$$\begin{aligned}\frac{\partial \mathtt{R}}{\partial v_i} \;=\;& \cos\theta\, \bar{v}_i\, [\bar{\mathbf{v}}]_\times + \sin\theta\, \bar{v}_i\, [\bar{\mathbf{v}}]_\times^2 + \frac{\sin\theta}{\theta}\, [\mathbf{e}_i - \bar{v}_i\bar{\mathbf{v}}]_\times \\ & + \frac{1-\cos\theta}{\theta}\left(\mathbf{e}_i\bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{e}_i^\top - 2\bar{v}_i\bar{\mathbf{v}}\bar{\mathbf{v}}^\top\right), \qquad \text{(III.1)}\end{aligned}$$

for $i = 1,2,3$, where $\theta = \|\mathbf{v}\|$, $\bar{\mathbf{v}} = (\bar{v}_1, \bar{v}_2, \bar{v}_3)^\top = \mathbf{v}/\|\mathbf{v}\|$, and $\mathbf{e}_i$ is the i-th vector of the canonical basis of $\mathbb{R}^3$. Formula (III.1) is used, for example, in the OpenCV library [1] if the rotation vector $\mathbf{v}$ (i.e., the exponential coordinates) is passed as argument to the appropriate function (`cvRodrigues`). For completeness, the proof of (III.1) is given in Appendix C.

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*

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

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

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

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

$$\frac{\partial\, [\bar{\mathbf{v}}]_\times \mathbf{u}}{\partial \mathbf{v}} \overset{\text{(A.2)}}{=} \frac{\partial (-\,[\mathbf{u}]_\times \bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}} \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}} = -\,[\mathbf{u}]_\times \frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}},$$

with derivative of the unitary rotation axis vector

$$\frac{\partial \bar{\mathbf{v}}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}}\left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right) = \frac{1}{\theta}(\mathrm{Id} - \bar{\mathbf{v}}\bar{\mathbf{v}}^\top) \overset{\text{(II.4)}}{=} -\frac{1}{\theta}\,[\bar{\mathbf{v}}]_\times^2. \qquad \text{(III.3)}$$

Also by the chain rule,

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

and, applying the product rule twice,

$$\begin{aligned}\frac{\partial\,[\bar{\mathbf{v}}]_\times^2\,\mathbf{u}}{\partial\mathbf{v}} &\overset{\text{(II.4)}}{=} \frac{\partial\bar{\mathbf{v}}(\bar{\mathbf{v}}^\top\mathbf{u})}{\partial\mathbf{v}} = \frac{\partial\bar{\mathbf{v}}}{\partial\mathbf{v}}(\bar{\mathbf{v}}^\top\mathbf{u}) + \bar{\mathbf{v}}\frac{\partial(\bar{\mathbf{v}}^\top\mathbf{u})}{\partial\mathbf{v}}\\ &= \big((\bar{\mathbf{v}}^\top\mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\big)\frac{\partial\bar{\mathbf{v}}}{\partial\mathbf{v}}\\ &\overset{\text{(III.3)}}{=} -\frac{1}{\theta}\big((\bar{\mathbf{v}}^\top\mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\big)\,[\bar{\mathbf{v}}]_\times^2\,,\end{aligned}$$

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

$$\begin{aligned}&\big((\bar{\mathbf{v}}^\top\mathbf{u})\mathrm{Id} + \bar{\mathbf{v}}\mathbf{u}^\top\big)\,[\bar{\mathbf{v}}]_\times^2\\ &\overset{\text{(A.1)}}{=} \big((\bar{\mathbf{v}}^\top\mathbf{u})\mathrm{Id} - \mathbf{u}\bar{\mathbf{v}}^\top + \bar{\mathbf{v}}\mathbf{u}^\top - \mathbf{u}\bar{\mathbf{v}}^\top\big)\,[\bar{\mathbf{v}}]_\times^2\\ &\overset{\text{(A.3),(A.5)}}{=} -[\bar{\mathbf{v}}]_\times\,[\mathbf{u}]_\times\,[\bar{\mathbf{v}}]_\times^2 + [\mathbf{u}\times\bar{\mathbf{v}}]_\times\,[\bar{\mathbf{v}}]_\times^2\\ &\overset{\text{(A.6),(II.4)}}{=} -2\,[\bar{\mathbf{v}}]_\times\,[\mathbf{u}]_\times\,[\bar{\mathbf{v}}]_\times^2 - [\mathbf{u}]_\times\,[\bar{\mathbf{v}}]_\times\,.\end{aligned}$$

Hence, so far the derivative of the rotated vector is

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

Next, premultiply by $\mathtt{R}^\top$ and use

$$\mathtt{R}^\top\,[\bar{\mathbf{v}}]_\times \overset{\text{(II.3),(A.1)}}{=} \cos\theta\;[\bar{\mathbf{v}}]_\times - \sin\theta\;[\bar{\mathbf{v}}]_\times^2 \tag{III.4}$$

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

$$\begin{aligned}&\mathtt{R}^\top(\cos\theta\;[\bar{\mathbf{v}}]_\times + \sin\theta\;[\bar{\mathbf{v}}]_\times^2)\mathbf{u}\bar{\mathbf{v}}^\top\\ &= (\cos\theta\,\mathtt{R}^\top\,[\bar{\mathbf{v}}]_\times + \sin\theta\,\mathtt{R}^\top\,[\bar{\mathbf{v}}]_\times^2)\mathbf{u}\bar{\mathbf{v}}^\top\\ &\overset{\text{(III.4)}}{=} (\cos^2\theta\;[\bar{\mathbf{v}}]_\times - \sin^2\theta\;[\bar{\mathbf{v}}]_\times^3)\mathbf{u}\bar{\mathbf{v}}^\top\\ &\overset{\text{(II.4)}}{=} (\cos^2\theta\;[\bar{\mathbf{v}}]_\times + \sin^2\theta\;[\bar{\mathbf{v}}]_\times)\mathbf{u}\bar{\mathbf{v}}^\top\\ &= [\bar{\mathbf{v}}]_\times\,\mathbf{u}\bar{\mathbf{v}}^\top\\ &\overset{\text{(A.2)}}{=} -[\mathbf{u}]_\times\,\bar{\mathbf{v}}\bar{\mathbf{v}}^\top.\end{aligned} \tag{III.5}$$

For the remaining term of $\mathtt{R}^\top\partial(\mathtt{R}\mathbf{u})/\partial\mathbf{v}$, we use the transpose of (II.1), and apply $[\bar{\mathbf{v}}]_\times\,[\mathbf{u}]_\times\,[\bar{\mathbf{v}}]_\times \overset{\text{(A.3)}}{=} -(\mathbf{u}^\top\bar{\mathbf{v}})\,[\bar{\mathbf{v}}]_\times$ to simplify

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

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,

$$\mathtt{R}^\top\frac{\partial\mathtt{R}\mathbf{u}}{\partial\mathbf{v}} = -\,[\mathbf{u}]_\times\,\big(\bar{\mathbf{v}}\bar{\mathbf{v}}^\top + \frac{1}{\theta}(\mathtt{R}^\top - \mathrm{Id})\,[\bar{\mathbf{v}}]_\times\big). \tag{III.6}$$

Finally, premultiply (III.6) by $\mathtt{R}$ (recall that $\mathtt{R}\mathtt{R}^\top = \mathrm{Id}$) to get

$$\frac{\partial\mathtt{R}\mathbf{u}}{\partial\mathbf{v}} = -\mathtt{R}\,[\mathbf{u}]_\times\,\big(\bar{\mathbf{v}}\bar{\mathbf{v}}^\top + \frac{1}{\theta}(\mathtt{R}^\top - \mathrm{Id})\,[\bar{\mathbf{v}}]_\times\big),$$

which coincides with (III.2) after substituting $\theta = \|\mathbf{v}\|$ and $\bar{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$. ■

A. 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 (III.2) provides some 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,

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

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 (III.1).

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

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

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 (A.7) 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.

B. 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, without re-doing all calculations.

Result 2. *The derivative of* $\mathtt{R}$ *in (II.5) with respect to its exponential coordinates* $\mathbf{v}$ *is*

$$\frac{\partial \mathtt{R}}{\partial v_i} = \frac{v_i\,[\mathbf{v}]_\times + [\mathbf{v}\times(Id-\mathtt{R})\mathbf{e}_i]_\times}{\|\mathbf{v}\|^2}\,\mathtt{R}. \tag{III.7}$$

Proof: Stemming from (III.2), we will show that it is possible to write

$$\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} = \mathtt{A}\mathbf{u} \tag{III.8}$$

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 (III.2) and simplify using the cross-product properties to arrive at (III.8):

$$\begin{aligned}
\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} &= -\|\mathbf{v}\|^{-2}\,\mathtt{R}\,[\mathbf{u}]_\times\big(\mathbf{v}\mathbf{v}^\top + (\mathtt{R}^\top-\mathrm{Id})\,[\mathbf{v}]_\times\big)\mathbf{e}_i\\
&= -\|\mathbf{v}\|^{-2}\,\mathtt{R}\,[\mathbf{u}]_\times\big(\mathbf{v}\mathbf{v}^\top + [\mathbf{v}]_\times(\mathtt{R}^\top-\mathrm{Id})\big)\mathbf{e}_i\\
&= -\|\mathbf{v}\|^{-2}\,\mathtt{R}\,[\mathbf{u}]_\times\big(\mathbf{v}v_i + \big(\mathbf{v}\times(\mathtt{R}^\top-\mathrm{Id})\mathbf{e}_i\big)\big)\\
&\overset{(A.2)}{=} \|\mathbf{v}\|^{-2}\,\mathtt{R}\,\big[v_i\mathbf{v} + \big(\mathbf{v}\times(\mathtt{R}^\top-\mathrm{Id})\mathbf{e}_i\big)\big]_\times\,\mathbf{u}.
\end{aligned}$$

After some manipulations, $\mathtt{R}\,\big[v_i\mathbf{v} + \big(\mathbf{v}\times(\mathtt{R}^\top-\mathrm{Id})\mathbf{e}_i\big)\big]_\times = \big[v_i\mathbf{v} + \big(\mathbf{v}\times(\mathrm{Id}-\mathtt{R})\mathbf{e}_i\big)\big]_\times\mathtt{R}$, and so

$$\begin{aligned}
\frac{\partial \mathtt{R}}{\partial v_i}\mathbf{u} &= \mathtt{R}\frac{\big[v_i\mathbf{v} + \big(\mathbf{v}\times(\mathtt{R}^\top-\mathrm{Id})\mathbf{e}_i\big)\big]_\times}{\|\mathbf{v}\|^2}\mathbf{u}\\
&= \frac{\big[v_i\mathbf{v} + \big(\mathbf{v}\times(\mathrm{Id}-\mathtt{R})\mathbf{e}_i\big)\big]_\times}{\|\mathbf{v}\|^2}\mathtt{R}\mathbf{u},
\end{aligned}$$

which is the desired formula (III.7) due to the linearity of the cross-product matrix (II.2). ■

To conclude, we also need to show that the compact formula (III.7) is consistent with (III.1). This is elaborated in Appendix B.

C. Derivative at the identity.

Our result agrees with the well-known theory of matrix Lie groups stating that

$$\frac{\partial}{\partial v_i}\exp([\mathbf{v}]_\times)\Big|_{\mathbf{v}=\mathbf{0}} = [\mathbf{e}_i]_\times\,. \tag{III.9}$$

This can be shown by computing the limit as $\mathbf{v}\to 0$ of (III.7), and using the facts that $\lim_{\mathbf{v}\to\mathbf{0}}\mathtt{R} = \mathrm{Id}$ and $\lim_{\mathbf{v}\to\mathbf{0}}(\mathrm{Id}-\mathtt{R})/\|\mathbf{v}\| = -\,[\bar{\mathbf{v}}]_\times$,

$$\begin{aligned}
\lim_{\mathbf{v}\to\mathbf{0}}\frac{\partial \mathtt{R}}{\partial v_i} &\overset{(III.7)}{=} \lim_{\mathbf{v}\to\mathbf{0}}\Big(\Big(\bar{v}_i\,[\bar{\mathbf{v}}]_\times + \frac{[\bar{\mathbf{v}}\times(\mathrm{Id}-\mathtt{R})\mathbf{e}_i]_\times}{\|\mathbf{v}\|}\Big)\mathtt{R}\Big)\\
&= \bar{v}_i\,[\bar{\mathbf{v}}]_\times - \big[\bar{\mathbf{v}}\times([\bar{\mathbf{v}}]_\times\,\mathbf{e}_i)\big]_\times\\
&= \Big[\bar{v}_i\bar{\mathbf{v}} - [\bar{\mathbf{v}}]_\times^2\,\mathbf{e}_i\Big]_\times\\
&\overset{(II.4)}{=} [\mathbf{e}_i]_\times\,.
\end{aligned}$$

IV. 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 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 codes that are based on the derivative of a linearization of the rotation matrix that uses exponential coordinates.

Appendix 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 $\mathbb{R}^3$.

$$\begin{aligned}
[\mathbf{a}]_\times\,\mathbf{a} &= \mathbf{0} && \text{(A.1)}\\
[\mathbf{a}]_\times\,\mathbf{b} &= -\,[\mathbf{b}]_\times\,\mathbf{a} && \text{(A.2)}\\
[\mathbf{a}]_\times\,[\mathbf{b}]_\times &= \mathbf{b}\mathbf{a}^\top - (\mathbf{a}\cdot\mathbf{b})\mathrm{Id} && \text{(A.3)}\\
\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) &\overset{(A.3)}{=} (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} && \text{(A.4)}\\
[\mathbf{a}\times\mathbf{b}]_\times &= \mathbf{b}\mathbf{a}^\top - \mathbf{a}\mathbf{b}^\top && \text{(A.5)}\\
[\mathbf{a}\times\mathbf{b}]_\times &= [\mathbf{a}]_\times\,[\mathbf{b}]_\times - [\mathbf{b}]_\times\,[\mathbf{a}]_\times && \text{(A.6)}\\
[(\mathtt{G}\mathbf{a})\times(\mathtt{G}\mathbf{b})]_\times &= \mathtt{G}\,[\mathbf{a}\times\mathbf{b}]_\times\,\mathtt{G}^\top &&\\
(\mathtt{G}\mathbf{a})\times(\mathtt{G}\mathbf{b}) &= \det(\mathtt{G})\mathtt{G}^{-\top}(\mathbf{a}\times\mathbf{b}) && \text{(A.7)}\\
[\mathbf{a}]_\times\,\mathtt{G} + \mathtt{G}^\top\,[\mathbf{a}]_\times &= \mathrm{trace}(\mathtt{G})\,[\mathbf{a}]_\times - [\mathtt{G}\mathbf{a}]_\times && \text{(A.8)}\\
[\mathtt{G}\mathbf{a}]_\times &= \det(\mathtt{G})\mathtt{G}^{-\top}\,[\mathbf{a}]_\times\,\mathtt{G}^{-1} &&
\end{aligned}$$

Appendix B
Agreement between derivative formulas

Here we show the agreement between (III.1) and (III.7). First, use $\theta = \|\mathbf{v}\|$ and the definition of the unit vector $\bar{\mathbf{v}} = \mathbf{v}/\theta$, so that (III.7) becomes

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

Since the rotation matrix is (II.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 $[\bar{\mathbf{v}}\times\mathbf{u}]_\times = \mathbf{u}\bar{\mathbf{v}}^\top - \bar{\mathbf{v}}\mathbf{u}^\top$ and $\mathtt{R}^\top\bar{\mathbf{v}} = \bar{\mathbf{v}}$, we have that

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

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

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

Using the property (A.8) with $\mathbf{a}=\bar{\mathbf{v}}$, $\mathtt{G}=\mathbf{e}_i\bar{\mathbf{v}}^\top$ we have that

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

Finally, substituting previous results in (B.1), we obtain the desired result (III.1):

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

APPENDIX C
DERIVATIVE FORMULA WITH SINES AND COSINES

Here, we show how to obtain formula (III.1). First, differentiate the Euler-Rodrigues rotation formula (II.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 (II.3), gives

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

In the previous formula, the term

$$\frac{\partial\,[\bar{\mathbf{v}}]_\times}{\partial v_i}=\frac{\partial}{\partial v_i}\left[\frac{\mathbf{v}}{\|\mathbf{v}\|}\right]_\times$$

can be computed using

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

where $\theta=\|\mathbf{v}\|$, resulting the concise formula

$$\frac{\partial\,[\bar{\mathbf{v}}]_\times}{\partial v_i}=\frac{1}{\theta}\,[\mathbf{e}_i-\bar{v}_i\bar{\mathbf{v}}]_\times\,.$$

The last term in (C.1) requires the calculation of

$$\frac{\partial\bar{\mathbf{v}}}{\partial v_i}=\frac{\partial}{\partial v_i}\left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right)=\frac{1}{\theta}(\mathrm{Id}-\bar{\mathbf{v}}\bar{\mathbf{v}}^\top)\mathbf{e}_i,\tag{C.2}$$

as we see in

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

Substituting the previous expressions in (C.1), gives (III.1). A similar proof is outlined in [12] using Einstein summation index notation.

ACKNOWLEDGMENT

G. Gallego is supported by the Marie Curie - COFUND Programme of the EU, UNITE Contract 246565, as part of the Seventh Framework Programme (FP7). This work has been partially supported by the Ministerio de Ciencia e Innovación of the Spanish Government under project TEC2010-20412 (Enhanced 3DTV).

REFERENCES

[1] Open Source Computer Vision (OpenCV). http://opencv.org.
[2] S.L. Altmann. *Rotations, Quaternions, And Double Groups*. Clarendon Press, Oxford, UK, 1986.
[3] O.A. Bauchau and J.-Y. Choi. The Vector Parameterization of Motion. *Nonlinear Dynamics*, 33(2):165–188, 2003.
[4] H. Cheng and K.C. Gupta. An Historical Note on Finite Rotations. *J. Appl. Mech.*, 56(1):139–145, 1989.
[5] S. Dambreville, R. Sandhu, A. Yezzi, and A. Tannenbaum. 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, March 2010.
[6] O. Enqvist and F. Kahl. Robust Optimal Pose Estimation. In *Proc. 10th European Conf. on Computer Vision*, ECCV '08, pages 141–153, Berlin, Heidelberg, 2008. Springer-Verlag.
[7] C. Gurwitz and M.L. Overton. 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] R Hartley and A Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2003.
[9] R.I. Hartley and F. Kahl. Global Optimization through Searching Rotation Space and Optimal Estimation of the Essential Matrix. In *Proc. IEEE 11th Int. Conf. on Computer Vision*, ICCV 2007, pages 1–8, 2007.
[10] B. K. P. Horn. Relative Orientation. *Int. J. Computer Vision*, 4(1):59–78, January 1990.
[11] A. Ibrahimbegovic. On the choice of finite rotation parameters. *Computer Methods in Applied Mech. and Eng.*, 149(1–4):49–71, 1997.
[12] F. O. Kuehnel. On the minimization over SO(3) manifolds. Technical Report TR03.12, Research Institute for Advanced Computer Science (RIACS). NASA Ames Research Center, USA, 2003.
[13] Y. Ma, S. Soatto, J. Kosecká, and S.S. Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer, 2003.
[14] R.M. Murray, Z. Li, and S.S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1994.
[15] C. Olsson, F. Kahl, and M. Oskarsson. Optimal Estimation of Perspective Camera Pose. In *Proc. 18th Int. Conf. on Pattern Recognition*, volume 2 of *ICPR 2006*, pages 5–8, 2006.
[16] R. Sandhu, S. Dambreville, and A. Tannenbaum. Point Set Registration via Particle Filtering and Stochastic Dynamics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(8):1459–1473, 2010.
[17] J. Stuelpnagel. On the Parametrization of the Three-Dimensional Rotation Group. *SIAM Review*, 6(4):422–430, October 1964.
[18] C. J. Taylor and D. J. Kriegman. Minimization on the Lie Group SO(3) and Related Manifolds. Technical Report 9405, Yale University, New Haven, CT, April 1994.
[19] G. B. Unal, A. J. Yezzi, S. Soatto, and G. G. Slabaugh. A Variational Approach to Problems in Calibration of Multiple Cameras. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(8):1322–1338, 2007.