

On Cayley's Factorization with an Application to the Orthonormalization of Noisy Rotation Matrices

S. Sarabandi, A. Perez-Gracia and F. Thomas

Abstract. A real orthogonal matrix representing a rotation in four dimensions can be decomposed into the commutative product of a left- and a right-isoclinic rotation matrix. This operation, known as Cayley's factorization, directly provides the double quaternion representation of rotations in four dimensions. This factorization can be performed without divisions, thus avoiding the common numerical issues attributed to the computation of quaternions from rotation matrices. In this paper, it is shown how this result is particularly useful, when particularized to three dimensions, to re-orthonormalize a noisy rotation matrix by converting it to quaternion form and then obtaining back the corresponding proper rotation matrix. This re-orthonormalization method is commonly implemented using the Shepperd-Markley method, but the method derived here is shown to outperform it by returning results closer to those obtained using the Singular Value Decomposition which are known to be optimal in terms of the Frobenius norm.

Mathematics Subject Classification (2010). Primary 99Z99; Secondary 15A66.

Keywords. Rotation matrices, quaternions, double quaternions, Cayley's factorization, Shepperd-Markley method.

1. Introduction

Any rotation in four dimensions can be seen as the composition of two rotations in a pair of orthogonal two-dimensional subspaces [1]. When the absolute value of the rotated angles in these two subspaces are equal, the rotation

This work was partially supported by the Spanish Ministry of Economy and Competitiveness through the projects DPI2017-88282-P and MDM-2016-0656, and the USA National Science Foundation under Grant No. 1208385.

is said to be *isoclinic*. In 1855, Cayley proved that any rotation in four dimensions can be factored into the commutative composition of two isoclinic rotations [2]. The development of the first effective procedure for computing this factorization is attributed in [3] to Van Elfrinkhof [4]. Since this work, written in Dutch, remained unnoticed, other sources (see, for example, [5]) attribute it to Rosen [6]. The methods of Elfrinkhof and Rosen are equivalent (see [3, 12] for a detailed explanation). Although formally correct, this method was not designed taking into account numerical issues. Nevertheless, it has recently been proven that a slight variation on it leads to a division-free formulation which, when particularized to three dimensions, outperforms the celebrated Shepperd's method widely used in aerial navigation, computer graphics, and robotics [7].

In this paper, we further deepen the study of Cayley's factorization. We show, for example, that the left- and right-isoclinic rotations resulting from its application can be obtained as the kernel and the cokernel of a 4×4 matrix. This result, when particularized to three dimensions, coincides with a method proposed by Bar-Itzhack in [8] (see [7] for a simplified explanation) to obtain the quaternion corresponding to a 3×3 rotation matrix. Explicit formulas are also derived which are shown to be of particular interest to orthonormalize noisy rotation matrices based on the well-known technique of converting the noisy rotation matrix to quaternion form and recovering back the rotation matrix. The obtained results are better in terms of the Frobenius norm of the error than the commonly used Shepperd-Markley method [9]. The results are actually closer to the optimal results obtained using the Singular Value Decomposition (SVD).

This paper is organized as follows. Section 2 summarizes the basic facts about rotations in four dimensions that are used in Section 3 to derive a set of division-free formulas for obtaining the double quaternion representation of a rotation in four dimensions. Then, in Section 4, this result is particularized to rotations in three dimensions. Section 5 briefly reviews Shepperd's method, the standard method used in most applications to obtain the unit quaternion corresponding to a rotation matrix, and a variation on it introduced by Markley which makes it particularly useful to orthonormalizing noisy rotation matrices. Section 6 deals with the problem of orthonormalizing noisy rotation matrices. Section 7 presents a statistical analysis comparing the results obtained using the Shepperd-Markley method, the method derived in this paper, and those obtained using the SVD for variable levels of noise. Section 8 summarizes the main points.

2. Rotations in four dimensions

A rotation in four dimensions is defined by two hyperplanes and two angles of rotation, α_1 and α_2 , one for each plane, through which points in the planes rotate. All points not in the planes rotate through an angle between α_1 and

α_2 . See [10] for details on the geometric interpretation of rotations in four dimensions.

Isoclinic rotations are a particular case of rotations in four dimensions in which there are infinitely many invariant orthogonal planes, with same rotation angles, that is, $\alpha_1 = \pm\alpha_2$. These rotations can be left-isoclinic, when the rotation in both planes is the same ($\alpha_1 = \alpha_2$), or right-isoclinic, when the rotations in both planes have opposite signs ($\alpha_1 = -\alpha_2$).

Isoclinic rotations have several important properties: (a) the composition of two right- (left-) isoclinic rotations is a right- (left-) isoclinic rotation; (b) the composition of a right- and a left-isoclinic rotation is commutative; and (c) any rotation in four dimensions can be decomposed into the composition of a right- and a left-isoclinic rotations.

Right- and a left-isoclinic rotations form maximal and normal subgroups. Denote by S_R^3 the subgroups of right-isoclinic rotations, and by S_L^3 the subgroup of left-isoclinic rotations. The direct product $S_L^3 \times S_R^3$ is a double cover of the group $SO(4)$, as four-dimensional rotations can be seen as the composition of rotations of these two subgroups, and there are two expressions for each element of the group. The left- and right-isoclinic rotations can be represented by rotation matrices of the form

$$\mathbf{R}^L = \begin{pmatrix} l_0 & -l_3 & l_2 & -l_1 \\ l_3 & l_0 & -l_1 & -l_2 \\ -l_2 & l_1 & l_0 & -l_3 \\ l_1 & l_2 & l_3 & l_0 \end{pmatrix}, \quad (1)$$

and

$$\mathbf{R}^R = \begin{pmatrix} r_0 & -r_3 & r_2 & r_1 \\ r_3 & r_0 & -r_1 & r_2 \\ -r_2 & r_1 & r_0 & r_3 \\ -r_1 & -r_2 & -r_3 & r_0 \end{pmatrix}, \quad (2)$$

respectively, where

$$\mathbf{l} = \sigma(l_0 \ l_1 \ l_2 \ l_3)^T \quad \text{and} \quad \mathbf{r} = \sigma(r_0 \ r_1 \ r_2 \ r_3)^T \quad (3)$$

directly correspond, as we will see later, to their quaternion representation with $\sigma = \pm 1$.

Since (1) and (2) are rotation matrices, their rows and columns are unit vectors. As a consequence,

$$\mathbf{l} \cdot \mathbf{l}^T = 1, \quad \text{and} \quad \mathbf{r} \cdot \mathbf{r}^T = 1. \quad (4)$$

Without loss of generality, we have introduced some changes in the signs and indices of (1) and (2) with respect to the notation used by Cayley [2, 5] to provide a neat connection with the standard use of quaternions for representing rotations in three dimensions.

According to the above properties, a 4×4 rotation matrix, say \mathbf{R} , can be expressed as:

$$\mathbf{R} = \mathbf{R}^L \mathbf{R}^R = \mathbf{R}^R \mathbf{R}^L, \quad (5)$$

where

$$\mathbf{R}^L = l_0 \mathbf{I} + l_1 \mathbf{A}_1 + l_2 \mathbf{A}_2 + l_3 \mathbf{A}_3, \quad (6)$$

and

$$\mathbf{R}^R = r_0 \mathbf{I} + r_1 \mathbf{B}_1 + r_2 \mathbf{B}_2 + r_3 \mathbf{B}_3, \quad (7)$$

where \mathbf{I} stands for the 4×4 identity matrix and

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{B}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \mathbf{B}_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Now, it can be verified that

$$\mathbf{A}_1^2 = \mathbf{A}_2^2 = \mathbf{A}_3^2 = \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3 = -\mathbf{I}, \quad (8)$$

and

$$\mathbf{B}_1^2 = \mathbf{B}_2^2 = \mathbf{B}_3^2 = \mathbf{B}_1 \mathbf{B}_2 \mathbf{B}_3 = -\mathbf{I}. \quad (9)$$

Expression (8) determines all the possible products of \mathbf{A}_1 , \mathbf{A}_2 , and \mathbf{A}_3 resulting in

$$\begin{aligned} \mathbf{A}_1 \mathbf{A}_2 &= \mathbf{A}_3, & \mathbf{A}_2 \mathbf{A}_3 &= \mathbf{A}_1, & \mathbf{A}_3 \mathbf{A}_1 &= \mathbf{A}_2, \\ \mathbf{A}_2 \mathbf{A}_1 &= -\mathbf{A}_3, & \mathbf{A}_3 \mathbf{A}_2 &= -\mathbf{A}_1, & \mathbf{A}_1 \mathbf{A}_3 &= -\mathbf{A}_2. \end{aligned} \quad (10)$$

Likewise, all the possible products of \mathbf{B}_1 , \mathbf{B}_2 , and \mathbf{B}_3 can be derived from expression (9). Moreover, it can be verified that

$$\mathbf{A}_i \mathbf{B}_j = \mathbf{B}_j \mathbf{A}_i. \quad (11)$$

which is actually a consequence of the commutativity of left- and right-isoclinic rotations.

3. Cayley's factorization

The problem of factoring a 4×4 rotation matrix, say \mathbf{R} , into the product of a right- and a left- isoclinic rotation matrix consists in finding \mathbf{l} and \mathbf{r} that satisfy the following matrix equation:

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{pmatrix} = \begin{pmatrix} l_0 & -l_3 & l_2 & -l_1 \\ l_3 & l_0 & -l_1 & -l_2 \\ -l_2 & l_1 & l_0 & -l_3 \\ l_1 & l_2 & l_3 & l_0 \end{pmatrix} \begin{pmatrix} r_0 & -r_3 & r_2 & r_1 \\ r_3 & r_0 & -r_1 & r_2 \\ -r_2 & r_1 & r_0 & r_3 \\ -r_1 & -r_2 & -r_3 & r_0 \end{pmatrix}. \quad (12)$$

A closed-form matrix solution for this problem can be found in [11]. Nevertheless, it involves four alternative mappings and the best one has to be chosen according to a voting scheme. This voting scheme, as we will see

later, leads to some inconveniences when dealing with noisy rotation matrices. An alternative matrix formulation can be obtained by first defining

$$\mathbf{P} = \mathbf{l} \cdot \mathbf{r}^T = \begin{pmatrix} l_0 r_0 & l_0 r_1 & l_0 r_2 & l_0 r_3 \\ l_1 r_0 & l_1 r_1 & l_1 r_2 & l_1 r_3 \\ l_2 r_0 & l_2 r_1 & l_2 r_2 & l_2 r_3 \\ l_3 r_0 & l_3 r_1 & l_3 r_2 & l_3 r_3 \end{pmatrix} \quad (13)$$

and

$$\mathbf{K} = \frac{1}{4} \begin{pmatrix} r_{11} + r_{22} + r_{33} + r_{44} & -r_{41} + r_{32} - r_{23} + r_{14} \\ r_{41} + r_{32} - r_{23} - r_{14} & r_{11} - r_{22} - r_{33} + r_{44} \\ -r_{31} + r_{42} + r_{13} - r_{24} & r_{21} + r_{12} - r_{43} - r_{34} \\ r_{21} - r_{12} + r_{43} - r_{34} & r_{31} + r_{42} + r_{13} + r_{24} \\ -r_{31} - r_{42} + r_{13} + r_{24} & r_{21} - r_{12} - r_{43} + r_{34} \\ r_{21} + r_{12} + r_{43} + r_{34} & r_{31} - r_{42} + r_{13} - r_{24} \\ -r_{11} + r_{22} - r_{33} + r_{44} & r_{41} + r_{32} + r_{23} + r_{14} \\ -r_{41} + r_{32} + r_{23} - r_{14} & -r_{11} - r_{22} + r_{33} + r_{44} \end{pmatrix}. \quad (14)$$

Now, using equation (12), it can be verified that¹

$$\mathbf{P} = \mathbf{K}. \quad (15)$$

If we right-multiply (15) by \mathbf{r} , we conclude that

$$\mathbf{l} = \mathbf{K}\mathbf{r}. \quad (16)$$

Likewise, if we transpose (15) and right-multiply it by \mathbf{l} , we have that

$$\mathbf{r} = \mathbf{K}^T \mathbf{l}. \quad (17)$$

Then, from (16) and (17), we conclude that

$$(\mathbf{K}\mathbf{K}^T - \mathbf{I})\mathbf{l} = 0 \quad \text{and} \quad \mathbf{r}(\mathbf{K}\mathbf{K}^T - \mathbf{I}) = 0. \quad (18)$$

Thus, the factorization into a left- and a right-isoclinic rotation can be obtained by computing the kernel and the cokernel, respectively, of $(\mathbf{K}\mathbf{K}^T - \mathbf{I})$.

An alternative to the above matrix approaches is obtained by a slight variation of the Elfrinkhof-Rosen method. The result is a division-free closed-form scalar formula for each element of the two quaternions. These formulas are obtained by simply observing that the norm of row i of \mathbf{P} is

$$+\sqrt{l_{i-1}^2(r_0^2 + r_1^2 + r_2^2 + r_3^2)} = |l_{i-1}|, \quad (19)$$

and the norm of column i is

$$+\sqrt{r_{i-1}^2(l_0^2 + l_1^2 + l_2^2 + l_3^2)} = |r_{i-1}|. \quad (20)$$

As a consequence, since $\mathbf{P} = \mathbf{K}$, the norms of the row and column vectors of \mathbf{K} gives us the absolute values of the elements of \mathbf{l} and \mathbf{r} , respectively. To assign a consistent set of signs to them, we can take any positive entry in \mathbf{K} , say the element (k, l) . Then, according to (13), l_{k-1} and r_{l-1} are both

¹The expression given in [12] for \mathbf{K} contains some typos.

As in the four-dimensional case, a closed-form scalar formula can be obtained for the elements of \mathbf{q} by computing the norms of the rows or the columns of (24). In this case, we have that

$$|q_0| = \frac{1}{4} \sqrt{(r_{11}+r_{22}+r_{33}+1)^2 + (r_{32}-r_{23})^2 + (r_{13}-r_{31})^2 + (r_{21}-r_{12})^2}, \quad (27)$$

$$|q_1| = \frac{1}{4} \sqrt{(r_{32}-r_{23})^2 + (r_{11}-r_{22}-r_{33}+1)^2 + (r_{21}+r_{12})^2 + (r_{31}+r_{13})^2}, \quad (28)$$

$$|q_2| = \frac{1}{4} \sqrt{(r_{13}-r_{31})^2 + (r_{21}+r_{12})^2 + (r_{22}-r_{11}-r_{33}+1)^2 + (r_{32}+r_{23})^2}, \quad (29)$$

$$|q_3| = \frac{1}{4} \sqrt{(r_{21}-r_{12})^2 + (r_{31}+r_{13})^2 + (r_{32}+r_{23})^2 + (r_{33}-r_{11}-r_{22}+1)^2}. \quad (30)$$

If we assume that q_0 is positive, we can give a consistent set of signs to the other elements of the quaternion by simply assigning the signs of $(r_{32} - r_{23})$, $(r_{13} - r_{31})$, and $(r_{21} - r_{12})$, to q_1 , q_2 , and q_3 , respectively. Alternatively, if we assume that q_1 is positive, a consistent set of signs to the other elements of the quaternion result from assigning the signs of $(r_{32} - r_{23})$, $(r_{21} + r_{12})$, and $(r_{1,3} + r_{3,1})$ to q_0 , q_2 , and q_3 , respectively (see [7] for details). The following table summarizes the four possible alternatives:

sign(q_0)	+	sign($r_{32} - r_{23}$)	sign($r_{13} - r_{31}$)	sign($r_{21} - r_{12}$)
sign(q_1)	sign($r_{32} - r_{23}$)	+	sign($r_{21} + r_{12}$)	sign($r_{13} + r_{31}$)
sign(q_2)	sign($r_{13} - r_{31}$)	sign($r_{21} + r_{12}$)	+	sign($r_{32} + r_{23}$)
sign(q_3)	sign($r_{21} - r_{12}$)	sign($r_{13} + r_{31}$)	sign($r_{32} + r_{23}$)	+

Any of these four alternatives gives a correct consistent set of signs. Nevertheless, if \mathbf{R}' is noisy, the use of some of these sign assignments, without further considerations, might lead to erroneous results. This is better understood through an example.

Let us consider the following rotation matrix in three dimensions

$$\mathbf{R}'' = \begin{pmatrix} -0.88614058 & 0.23685074 & 0.39831731 \\ 0.23723170 & -0.50650954 & 0.82895672 \\ 0.39809051 & 0.82906568 & 0.39265025 \end{pmatrix} \quad (31)$$

which after adding some noise becomes

$$\hat{\mathbf{R}} = \begin{pmatrix} -0.88607281 & 0.23738025 & 0.39857802 \\ 0.23662227 & -0.50746065 & 0.82897574 \\ 0.39732188 & 0.82870960 & 0.39185813 \end{pmatrix}. \quad (32)$$

The quaternions corresponding to \mathbf{R}'' and $\hat{\mathbf{R}}$, obtained using the procedure presented in this section, and assuming that q_0 is positive, are

$$(0.0001 \quad 0.2386 \quad 0.4967 \quad 0.8345)^T$$

and

$$(0.0006 \quad -0.2386 \quad 0.4966 \quad -0.8344)^T,$$

respectively. The second one is clearly incorrect as it should be equal to the first one up to a small perturbation. Thus, assuming that an element

of the quaternion close to zero is positive is a bad choice because a small perturbation in the rotation matrix might induce a chance of its sign leading to an inconsistency with the rest of signs. A better option is to assume that the largest component of the quaternion is positive and then to assign the other component signs according to the table given above.

5. Shepperd-Markley method

Since it was first proposed in [13], Shepperd's method remains as one of the most popular methods for computing the quaternion corresponding to a 3×3 rotation matrix. It improves on Hughes' method [14] via a voting scheme. In Shepperd's method, there are four different formulas for computing the quaternion as a function of the elements of the rotation matrix, all of them formally equivalent, provided that the matrix is orthogonal. Numerically, however, these four formulas are not identical and, depending on the rotation matrix, one of them is numerically better conditioned than the others. These four formulas are:

$$\mathbf{u}_1 = \frac{1}{2} \begin{pmatrix} (1+r_{11}+r_{22}+r_{33})^{\frac{1}{2}} \\ (r_{32}-r_{23})/(1+r_{11}+r_{22}+r_{33})^{\frac{1}{2}} \\ (r_{13}-r_{31})/(1+r_{11}+r_{22}+r_{33})^{\frac{1}{2}} \\ (r_{21}-r_{12})/(1+r_{11}+r_{22}+r_{33})^{\frac{1}{2}} \end{pmatrix}, \quad (33)$$

$$\mathbf{u}_2 = \frac{1}{2} \begin{pmatrix} (r_{32}-r_{23})/(1+r_{11}-r_{22}-r_{33})^{\frac{1}{2}} \\ (1+r_{11}-r_{22}-r_{33})^{\frac{1}{2}} \\ (r_{12}+r_{21})/(1+r_{11}-r_{22}-r_{33})^{\frac{1}{2}} \\ (r_{31}+r_{13})/(1+r_{11}-r_{22}-r_{33})^{\frac{1}{2}} \end{pmatrix}, \quad (34)$$

$$\mathbf{u}_3 = \frac{1}{2} \begin{pmatrix} (r_{13}-r_{31})/(1-r_{11}+r_{22}-r_{33})^{\frac{1}{2}} \\ (r_{12}+r_{21})/(1-r_{11}+r_{22}-r_{33})^{\frac{1}{2}} \\ (1-r_{11}+r_{22}-r_{33})^{\frac{1}{2}} \\ (r_{23}+r_{32})/(1-r_{11}+r_{22}-r_{33})^{\frac{1}{2}} \end{pmatrix}, \quad (35)$$

$$\mathbf{u}_4 = \frac{1}{2} \begin{pmatrix} (r_{21}-r_{12})/(1-r_{11}-r_{22}+r_{33})^{\frac{1}{2}} \\ (r_{31}+r_{13})/(1-r_{11}-r_{22}+r_{33})^{\frac{1}{2}} \\ (r_{32}+r_{23})/(1-r_{11}-r_{22}+r_{33})^{\frac{1}{2}} \\ (1-r_{11}-r_{22}+r_{33})^{\frac{1}{2}} \end{pmatrix}. \quad (36)$$

The best of the four above alternatives, from the numerical point of view, is the one that involves the largest denominator in its computation.

In 2008, Markley introduced a simple variation on this method [9]. Instead of using (33)-(36), he defined:

$$\mathbf{u}'_1 = \begin{pmatrix} 1+r_{1,1}+r_{2,2}+r_{3,3} \\ r_{3,2}-r_{2,3} \\ r_{1,3}-r_{3,1} \\ r_{2,1}-r_{1,2} \end{pmatrix}, \quad (37)$$

$$\mathbf{u}'_2 = \begin{pmatrix} r_{3,2} - r_{2,3} \\ 1 + r_{1,1} - r_{2,2} - r_{3,3} \\ r_{1,2} + r_{2,1} \\ r_{3,1} + r_{1,3} \end{pmatrix}, \quad (38)$$

$$\mathbf{u}'_3 = \begin{pmatrix} r_{1,3} - r_{3,1} \\ r_{1,2} + r_{2,1} \\ 1 - r_{1,1} + r_{2,2} - r_{3,3} \\ r_{2,3} + r_{3,2} \end{pmatrix}, \quad (39)$$

$$\mathbf{u}'_4 = \begin{pmatrix} r_{2,1} - r_{1,2} \\ r_{3,1} + r_{1,3} \\ r_{2,3} + r_{3,2} \\ 1 - r_{1,1} - r_{2,2} + r_{3,3} \end{pmatrix}. \quad (40)$$

Since $\mathbf{u}_i = \mathbf{u}'_i / \|\mathbf{u}'_i\|$, $i = 1, \dots, 4$, \mathbf{u}'_i is a non-unit quaternion proportional to the unit quaternion \mathbf{u}_i . Considering that in most applications it is not strictly necessary to operate with unit quaternions (quaternions can actually be treated as a vector of homogeneous coordinates [15]), this is an interesting alternative because it avoids the computation of square roots and divisions. Even if we need to compute the rotation matrix corresponding to a non-unit quaternion, we do not need to previously normalize it. Indeed, it is not difficult to prove that the rotation matrix corresponding to a non-unit quaternion is given by

$$\mathbf{R}(\mathbf{q}) = \frac{1}{q_0^2 + q_2^2 + q_3^2 + q_4^2} \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & q_0^2 - q_1^2 - q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}, \quad (41)$$

which does not involve any square root.

Now, it can be checked that both, the method derived in Section 4 and the Shepperd-Markley method, are well-defined even if the input is an arbitrary 3×3 matrix. Nevertheless, contrarily to the method presented in Section 4, Shepperd-Markley uses four mappings that are algebraically equivalent provided that the input matrix is orthogonal. If not, each mapping, in general, gives a different result. As a consequence, a slight variation in the input matrix might lead to a discontinuity in the obtained quaternion. This non-smooth behavior is not mentioned in the literature [9], but is clearly undesirable. In this context, the method derived in Section 4 seems particularly useful to compute the quaternion associated with a noisy rotation matrix, and, as a result, to orthonormalize it to the *nearest* rotation matrix, as explained in the next section.

6. Orthonormalization of noisy rotation matrices

Due to floating-point errors: (a) the cumulative multiplication of rotation matrices does not result in an orthogonal matrix [18], and (b) the integration of angular velocity differential equations in \mathbb{E}^3 leads to a matrix that departs

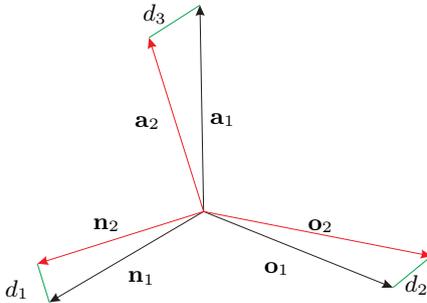


FIGURE 1. Following the standard robotics notation, the column vectors of the rotation matrix \mathbf{R}_i can be made explicit as $\mathbf{R}_i = (\mathbf{n}_i \ \mathbf{o}_i \ \mathbf{a}_i)$ [17]. These vectors determine a reference frame. Then, in this figure, the Frobenius norm of the difference between \mathbf{R}_1 and \mathbf{R}_2 is equal to $\sqrt{d_1^2 + d_2^2 + d_3^2}$.

from being orthogonal as time increases [19]. At least in these two situations, we have to restore the orthogonality of a rotation matrix, say $\mathring{\mathbf{R}}$, by finding the *closest* orthogonal matrix, say $\hat{\mathbf{R}}$. The first problem we face when trying to solve this problem is how to define a measure of closeness endowed with physical meaning.

The Euclidean or Frobenius norm (denoted as $\|\cdot\|_F$) of $\mathring{\mathbf{R}} - \hat{\mathbf{R}}$ (*i.e.*, the sum of squares of elements of $\mathring{\mathbf{R}} - \hat{\mathbf{R}}$) is commonly used to solve the stated problem because

- It has a simple geometric interpretation (see Fig. 1).
- Its derivatives can be easily obtained. This allows to derive a closed-form solution. If we would use the 2-norm, the closeness measure would be given by the largest singular value of $\mathring{\mathbf{R}} - \hat{\mathbf{R}}$ which is not easy to handle.
- It leads to a unique solution. If we would use the 2-norm, the solution is not necessarily unique [20].

Thus, using the Frobenius norm, the problem in algebraic terms can be stated as follows: given the noisy rotation matrix $\mathring{\mathbf{R}}$, the problem consists in finding $\hat{\mathbf{R}}$ that minimizes

$$\left\| \hat{\mathbf{R}} - \mathring{\mathbf{R}} \right\|_F^2 = \text{Trace} \left((\hat{\mathbf{R}} - \mathring{\mathbf{R}})(\hat{\mathbf{R}} - \mathring{\mathbf{R}})^T \right), \quad (42)$$

subject to $\hat{\mathbf{R}}^T \hat{\mathbf{R}} = \mathbf{I}$.

Using Lagrange multipliers, it can be proved that the optimal solution to this constrained optimization problem is given by (see, for example, [16])

$$\hat{\mathbf{R}} = \mathbf{R} (\mathbf{R}^T \mathbf{R})^{-\frac{1}{2}}. \quad (43)$$

It is easy to verify that $\hat{\mathbf{R}}$ thus obtained is orthonormal, *i.e.* $\hat{\mathbf{R}}^T \hat{\mathbf{R}} = \mathbf{I}$. However, there is no guarantee that $\det(\hat{\mathbf{R}}) = +1$. To represent a rotation,

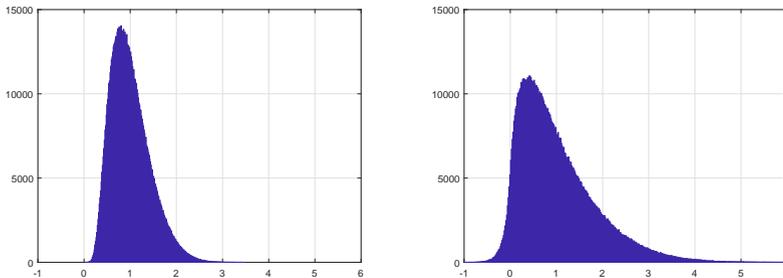


FIGURE 2. Histograms of the value of the determinants of 10^6 randomly generated 3×3 rotation matrices. While no matrices whose determinant is negative arise with an additive noise uniformly distributed in the interval of $[-0.4, 0.4]$ (left), only 3.7% of the matrices have a negative determinant even if the error is increased to $[-0.7, 0.7]$ (right).

the orthonormal matrix $\hat{\mathbf{R}}$ has to satisfy this condition as well. Otherwise it represents a reflection, not a rotation. There is no easy way to enforce this condition, and with highly noisy rotation matrices, we may have that $\det(\hat{\mathbf{R}}) = -1$. Nevertheless, the noise should be very high to induce an inversion of sign. For example, if we randomly generate 10^6 rotation matrices and a uniformly distributed noise in the interval $[-0.4, 0.4]$, which is an extremely high level of noise, no sign change arises (see Fig. 2-left). We have to increase the noise to $[-0.7, 0.7]$ to get an inversion in only 3.7% of the cases (see Fig. 2-right).

To avoid the computation of the square root of a matrix, the optimal solution using the Frobenius norm can alternatively be obtained using the SVD of \mathbf{R} [21, 22], which can be expressed as

$$\mathbf{R} = \mathbf{U}\mathbf{D}\mathbf{V}^T, \quad (44)$$

where $\mathbf{D} = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ is the diagonal matrix of singular values. Then,

$$\hat{\mathbf{R}} = \mathbf{U}\mathbf{V}^T. \quad (45)$$

To reduce the computational burden required to obtain the optimal solution using the SVD, a simpler method is commonly used. It consists in two steps: (a) converting the noisy rotation matrix to quaternion form (since the matrix is not orthogonal, the result is not a unit quaternion), and (b) converting back the quaternion to a proper orthogonal rotation matrix using (41). This approach has two disadvantages:

- No particular norm is apparently minimized.
- It contains a flaw: the derivation of the formulas for obtaining the quaternion assume that the quaternion has unit norm. For example, in our case, this assumption is used in (19) and (20). To the best of our knowledge, this point is not mentioned in the literature.

Despite these drawbacks, this approach is preferred in many applications because of the following reasons [9]:

- The Shepperd-Markley method is much less expensive than the SVD. It simply boils down to the evaluation of a rational algebraic formula in terms of the elements of the noisy rotation matrix.
- The difference between the use of the SVD or the Shepperd-Markley method is negligible for the level of expected numerical errors in most applications.
- The solution obtained using the SVD is not optimal, after all, unless the errors are isotropic [22, 23].

Despite that the above advantages are commonly accepted, the difference between the use of the SVD and the Shepperd-Markley method has not been previously quantified to the best of our knowledge. This analysis, including the performance of the alternative to the Shepperd-Markley proposed in this paper, is conducted in the next section.

7. Statistical analysis

Let us consider a rotation matrix, say \mathbf{R} , to which an error randomly and uniformly distributed in the interval $[-0.5, 0.5]$ is added to each of its elements to obtain a noisy rotation matrix, say $\mathring{\mathbf{R}}$. As an example, let us take

$$\mathbf{R} = \begin{pmatrix} 0.5865 & -0.2210 & 0.7792 \\ -0.0493 & -0.9700 & -0.2380 \\ 0.8085 & 0.1011 & -0.5798 \end{pmatrix} \text{ and } \mathring{\mathbf{R}} = \begin{pmatrix} 0.3879 & -0.1819 & 0.4574 \\ 0.1518 & -0.7719 & -0.6100 \\ 0.9748 & 0.2676 & -0.0807 \end{pmatrix}.$$

Then, let us compute the corresponding quaternion using the method proposed in this paper and the Shepperd-Markley method and the result is transformed to a rotation matrix using (41). The results are, respectively,

$$\hat{\mathbf{R}}_1 = \begin{pmatrix} 0.3596 & -0.6072 & 0.7085 \\ 0.0933 & -0.7321 & -0.6747 \\ 0.9284 & 0.3087 & -0.2066 \end{pmatrix} \text{ and } \hat{\mathbf{R}}_2 = \begin{pmatrix} 0.4767 & -0.3378 & 0.8116 \\ 0.3033 & -0.8033 & -0.5125 \\ 0.8251 & 0.4905 & -0.2805 \end{pmatrix}.$$

Let us also compute the optimal solution using the SVD. In this case, the result is

$$\hat{\mathbf{R}}_3 = \begin{pmatrix} 0.3855 & -0.4814 & 0.7872 \\ 0.1722 & -0.8006 & -0.5739 \\ 0.9065 & 0.3568 & -0.2258 \end{pmatrix}.$$

The obtained errors are

$$\left\| \hat{\mathbf{R}}_1 - \mathring{\mathbf{R}} \right\|_F = 0.5231, \quad \left\| \hat{\mathbf{R}}_2 - \mathring{\mathbf{R}} \right\|_F = 0.5505, \quad \left\| \hat{\mathbf{R}}_3 - \mathring{\mathbf{R}} \right\|_F = 0.4844. \quad (46)$$

As expected, the result using the SVD is the best one because it corresponds to the optimum. The other two methods provide results that are not far from the optimum despite the high level of noise, but that obtained using the method derived in this paper is better. This observation is for a particular example. Thus, to provide a correct assessment of the quality of the proposed

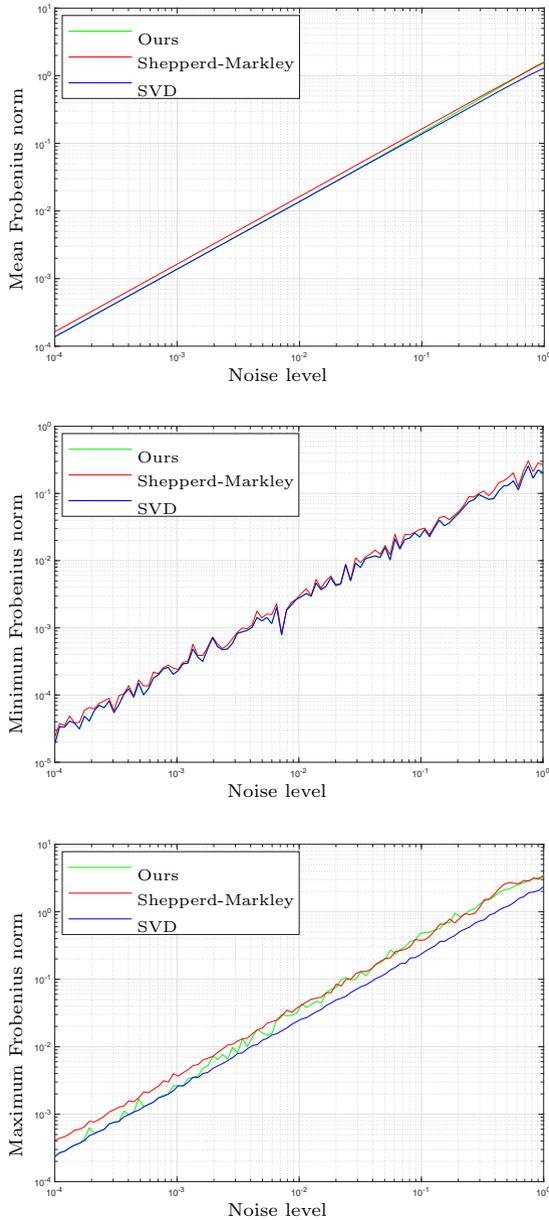


FIGURE 3. Mean, minimum and maximum Frobenius norm as a function of the input error for the three analyzed methods. 10,000 random noisy rotation matrices have been taken for each value of the input error.

method, we have repeated this experiment for 10,000 random rotation matrices with a variable level of noise ranging from -0.0001 to 1 (*i.e.*, a randomly uniformly distributed noise in intervals ranging from $[-0.0001, 0.0001]$ to $[-1, 1]$ is added to the elements of the matrix). These random matrices have been obtained from random quaternions generated using the algorithm described in [24], which permits to generate uniformly distributed points in \mathbb{S}^3 . The program has been implemented in MATLAB[®]. The results appear as double-logarithmic plots in Fig. 3. Predictably, the plots for the mean Frobenius norm (see Fig. 3-top) obtained using the method proposed in this paper (in green) and the Shepperd-Markey method (in red) lie above the optimal values obtained using the SVD (in blue). Nevertheless, while the results using the Shepperd-Markey method departs a constant amount from the optimal, the values obtained with the proposed method practically coincide with the optimal values for noise levels below 0.01 and they are still clearly better than those obtained using the Shepperd-Markey method for noise levels as high as 0.1. Moreover, the plot of the minimum Frobenius norm reached using the proposed method overlaps the plot obtained using the SVD (Fig. 3-center). In other words, the solution obtained using the proposed method coincides, at the chosen resolution level, with the optimal one at least in one of the 10,000 samples. This is not the case for the Shepperd-Markey method. Finally, it is also worth noting that the plot for the maximum Frobenius norm reached using the proposed method (Fig. 3-bottom) lies below that of the the Shepperd-Markey for noise levels below 0.01. Summing up, the method derived in this paper for orthonormalizing noisy rotation matrices is a serious alternative to the Shepperd-Markey method with an average behavior very close to optimal for moderate levels of noise.

8. Conclusion

We have deepened the study of Cayley's factorization both in its matrix and scalar versions. In the matrix version, we have presented a new method consisting in the computation of the kernel and the cokernel of a 4×4 matrix expressed in terms of the elements of the rotation matrix. This method, when particularized to three dimensions, coincides with a method proposed by Bar-Itzhack. In the scalar version, we have presented a division-free variation on the Elfrinkhof-Rosen method. This method, when particularized to three dimensions, has been shown to be of particular interest for the orthonormalization of noisy rotation matrices. It actually outperforms the Shepperd-Markey method, commonly used in most applications. The ultimate reason for the good behavior of this method is that it involves a single mapping that includes all the elements of the rotation matrix in the computation of each element of the quaternion.

References

- [1] J.L. Weiner and G.R. Wilkens, "Quaternions and rotations in E^4 ," *The American Mathematical Monthly*, Vol. 112, No. 1, pp. 69-76, 2005.
- [2] A. Cayley, "Recherches ultérieures sur les déterminants gauches," *The Collected Mathematical Papers Of Arthur Cayley*, article 137, Cambridge University Press, pp. 202-215, 1891.
- [3] J.E. Mebius, *Applications of Quaternions to Dynamical Simulation, Computer Graphics and Biomechanics*, Ph.D. Thesis, Delft University of Technology, 1994.
- [4] L. van Elfrinkhof, "Eene eigenschap van de orthogonale substitutie van de vierde orde," *Handelingen van het zesde Nederlandsch Natuurkundig en Geneeskundig Congres (Acts of the sixth Dutch physical and medical congress)*, Delft, pp. 237-240, 1897.
- [5] F.L. Hitchcock, "Analysis of rotations in euclidean four-space by sedenions," *Journal of Mathematics and Physics*, Vol. 9, No. 3, pp. 188-193, 1930.
- [6] N. Rosen, "Note on the general Lorentz transformation," *Journal of Mathematics and Physics*, Vol. 9, pp. 181-187, 1930.
- [7] S. Sarabandi and F. Thomas, "A Survey on the computation of quaternions from rotation matrices," *ASME Journal of Mechanisms and Robotics*, to appear, 2019.
- [8] I. Bar-Itzhack, "New method for extracting the quaternion from a rotation matrix," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 6, pp. 1085-1087, 2000.
- [9] F.L. Markley, "Unit quaternion from rotation matrix," *Journal of Guidance, Control and Dynamics*, Vol. 31, No. 2, pp. 440-442, 2008.
- [10] L. Pertti, *Clifford algebras and spinors*, Cambridge University Press, 2001.
- [11] A. Perez-Gracia and F. Thomas, "On Cayley's factorization of 4D rotations and applications," *Advances in Applied Clifford Algebras*, Vol. 27, No. 1, pp 523-538, 2017.
- [12] F. Thomas, "Approaching dual quaternions from matrix algebra," *IEEE Transactions on Robotics*, Vol. 30, No. 5, pp. 1037-1048, 2014.
- [13] S.W. Shepperd, "Quaternion from rotation matrix," *Journal of Guidance and Control*, Vol. 1, No. 3, pp. 223-224, 1978.
- [14] P.C. Hughes, *Spacecraft Attitude Dynamics*, John Wiley & Sons, Inc, 1986.
- [15] A. Purwar and Q. J. Ge, "On the effect of dual weights in Computer Aided Design of rational motions," *ASME Journal of Mechanical Design*, Vol. 127, No. 5, pp. 967-972, 2005.
- [16] C.R. Giardina, R. Bronson, and L. Wallen, "An optimal normalization scheme," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-11, No. 4, pp. 443-446, 1975.
- [17] R.P. Paul, *Robot Manipulators: Mathematics, Programming, and Control*, MIT Press, Cambridge, Massachusetts, 1981.
- [18] C. Rucker, "Integrating rotations using nonunit quaternions," *IEEE Robotics and Automation Letters*, Vol. 3, No. 4, pp. 2979-2986, 2018.
- [19] D. Herbison-Evans and D.S. Richardson, "Control of round-off propagation in articulating the human figure," *Computer Graphics and Image Processing*, Vol. 17, No. 4, pp. 386-393, 1981.

- [20] W. Kahan, "The nearest orthogonal or unitary matrix," University of California at Berkeley, available at: <https://people.eecs.berkeley.edu/wkahan/Math128/NearestQ.pdf>, 2011.
- [21] G.H. Golub and W. Kahan, "Least squares, singular values and matrix approximations," *Aplikace Matematiky*, Vol. 13, pp. 44-51, 1968.
- [22] L. Dorst, "First order error propagation of the Procrustes method for 3-D attitude estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 2, pp. 221-229, 2005.
- [23] M.D. Shuster, "The generalized Wahba problem," *Journal of the Astronautical Sciences*, Vol. 54, No. 2, pp. 245-259, 2006.
- [24] G. Marsaglia, "Choosing a point from the surface of a sphere," *Annals of Mathematical Statistics*, Vol. 43, pp. 645-646, 1972.

S. Sarabandi

Institut de Robòtica i Informàtica Industrial (CSIC-UPC)

Llorens Artigas 4-6, 08028 Barcelona, Spain

e-mail: ssarabandi@iri.upc.edu

A. Perez-Gracia

Department of Mechanical Engineering

Idaho State University, Pocatello, Idaho, 83209, USA

e-mail: perealba@isu.edu

F. Thomas

Institut de Robòtica i Informàtica Industrial (CSIC-UPC)

Llorens Artigas 4-6, 08028 Barcelona, Spain

e-mail: fthomas@iri.upc.edu