A New Probability Distribution for Simultaneous Representation of Uncertain Position and Orientation

Igor Gilitschenski*, **Gerhard Kurz***, **Simon J. Julier**†, and **Uwe D. Hanebeck***

*Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
gilitschenski@kit.edu, gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

†Virtual Environments and Computer Graphics Group
Department of Computer Science
University College London (UCL), United Kingdom
s.julier@ucl.ac.uk

***Abstract*—This work proposes a novel way to represent uncertainty on the Lie group of rigid-body motions in the plane. This is achieved by using dual quaternions for representation of a planar rigid-body motion and proposing a probability distribution from the exponential family of distributions that inherently respects the underlying structure of the representation. This is particularly beneficial in scenarios involving strong measurement noise. A relationship between the newly proposed distributional model and the Bingham distribution is discussed. The presented results involve formulas for computation of the normalization constant, the mode, parameter estimation techniques, and a closed-form Bayesian measurement fusion.**

Keywords*—*Pose estimation, dual quaternions, Bingham distribution, directional statistics, SE(2), Lie groups, probability theory

I. Introduction

Estimation of a rigid-body motion, i.e., simultaneous estimation of translation and rotation presents a major problem in many applications involving robotic perception, processing of sensor data for mixed and augmented reality applications, and navigation. This problem is particularly challenging for two reasons. First, it has a nonlinear structure because the values are defined on an inherently nonlinear domain. The cases of particular practical interest are the manifolds of rigid-body motions in the plane $SE(2)$ and in three-dimensional space $SE(3)$. Second, there is no canonical way to describe dependencies between position and orientation.

These problems are negligible in scenarios involving strong prior knowledge or high-precision sensors. Here, one can easily approximate arising nonlinearities involved in the estimation problem by linear models making use of the locally linear structure of the considered manifolds. Thus, usually two strategies are used to address this estimation scenario. First, costly high-precision sensors are deployed to avoid strong measurement noise. Second, the classical Kalman filter is modified to consider nonlinear system models. The most famous adaptations involve the extended Kalman filter (EKF) and the unscented Kalman filter (UKF) [1].

These adaptions are usually based on a more intelligent propagation technique (such as the unscented transform) and a sound state representation. Most of these approaches maintain a Gaussian assumption because of its convenient properties and natural appearance as a limit distribution. However, as the noises become larger, the errors introduced by this assumption become non-negligible. The reasons for this involve the fact that the Gaussian distribution does not capture the nonlinear structure of the underlying manifold. Recently, a number of authors have developed recursive estimation algorithms which use directional probability distributions instead. The proposed algorithms address angular [2], [3] and orientation estimation [4], [5] problems. These approaches show benefits over estimation techniques which maintain a Gaussian assumption.

Our earlier works on estimating angular and orientational quantities build upon some of the numerous probability distributions defined on periodic manifolds such as the wrapped normal distribution, the von Mises distribution [6], or the Bingham distribution [7]. One of the approaches using dual quaternions for representing rigid body motions was presented in [8]. Dual quaternions are a convenient representation because they generalize the idea of rotations represented by quaternions to a representation of the entire rigid-body transformation. However, this approach is based on an iterated EKF in order to account for the fact that the Gaussian assumption is inherently wrong and might yield poor results.

Combination of angular and linear quantities is of particular interest in robotic perception and inertial navigation systems. However, only a very limited number of approaches considers these quantities simultaneously in the underlying probability distribution. Two approaches are particularly notable for estimation of rigid transformations. First, the projected Gaussian approach proposed by Feiten et al. [9], [10], [11]. It is based on projecting parts of a higher dimensional Gaussian random vector to the unit sphere in an intelligent way. Furthermore, dual quaternions are used for state space representation. Unfortunately, this approach has the drawback of using an approximation in the Bayesian update step for a typical dynamic state estimation scenario. In [12], an approach related to the Bingham distribution was used for estimating a "mean" rigid-body transformation. It maintains some properties of the Bingham distribution and uses a combination of orthogonal matrices and vectors for representing uncertain orientation and translation respectively. However, it is not clear whether this approach makes a closed-form Bayes update possible.

A. Main Contribution

In this work, we propose a novel probability distribution which is capable of capturing the underlying structure of the manifold of planar rigid transformations $SE(2)$. This is achieved through the use of dual quaternions (restricted to planar transformations). Thus, there is no need to maintain an entire matrix to describe the actual transformation. Furthermore, choosing dual quaternions for our formulation simplifies a future generalization to the $SE(3)$ case. Similar to the Bingham distribution and its natural use for representing uncertain orientations represented by quaternions, the proposed distribution involves antipodal symmetry and makes closed form Bayesian inference in classical measurement update and fusion scenarios possible. Furthermore, we show that the Bingham distribution appears as a marginal distribution for the spherical part of our proposed model and describe the relationship between the modes and normalization constant of both distributions.

The remainder of this paper is structured as follows. In the next section, we introduce dual quaternions and discuss their convenience for composing and representing rigid transformations. Our new distributional model is presented in Sec. III, where a parameter estimation technique is proposed and the mode of the distribution is discussed. In Sec. IV, we show how the distribution can be applied to representing uncertainty of positions and orientations in the plane. Furthermore, a closed-form Bayesian inference procedure is derived for a typical estimation and fusion scenario. The work is concluded in Sec. V.

II. Dual Quaternions for Representing Rigid Body Motions

Unit quaternions offer a convenient way of representing and combining orientations. The latter is done by quaternion multiplication. Other operations from the skew field of quaternions obtain in this context a natural geometrical interpretation, e.g., multiplicative inversion represents the inverse of a rotation and exponentiation is applied to derive an interpolation algorithm [13].

This concept is generalized by dual quaternions of unit magnitude making the representation of translations possible and offering a natural way for composition of several translation and rotation sequences by dual quaternion multiplication. Applications of this concept cover the field of robotics and computer graphics such as [14], [15]. In the following, dual quaternions are revisited and their application to representation of rigid translations of the $SE(3)$ is discussed. The consideration of planar motions results as a natural special case.

A. Dual Quaternions

Understanding of dual numbers and plain quaternions is helpful in order to understand the concept of dual quaternions. In this work, the skew-field of quaternions will be denoted by $\mathbb{H}$. First, consider the quaternion $\mathbf{q} \in \mathbb{H}$

$$\mathbf{q} = q_1 + q_2 i + q_3 j + q_4 k \ ,$$

which has the typical basis elements i, j, k. $\mathbf{q}$ will be described using the vector representation $(q_1, q_2, q_3, q_4)^\top$ whenever it is more convenient. As described above, we are particularly interested in quaternion multiplication, because it can be used to apply a sequence of rotation operations. For two quaternions $\mathbf{a}, \mathbf{b} \in \mathbb{H}$, it is defined as

$$\mathbf{a} \cdot \mathbf{b} = \begin{pmatrix} a_1 b_1 - a_2 b_2 - a_3 b_3 - a_4 b_4 \\ a_1 b_2 + a_2 b_1 + a_3 b_4 - a_4 b_3 \\ a_1 b_3 - a_2 b_4 + a_3 b_1 + a_4 b_2 \\ a_1 b_4 + a_2 b_3 + a_3 b_2 - a_4 b_1 \end{pmatrix} \ .$$

The inverse of a unit quaternion is obtained by quaternion conjugation, i.e., by changing the sign in front of each imaginary basis element. Thus $\mathbf{q}^{-1} = \mathbf{q}^* = (q_1, -q_2, -q_3, -q_4)^\top$.

The second concept necessary to define dual quaternions are dual numbers, which can be seen as an extension of real numbers. Similarly to complex numbers (which introduce the imaginary unit i), dual numbers are defined by introducing a new element. It is usually denoted by ε and characterized by its nilpotency property $\varepsilon^2 = 0$. Thus, a dual number is given by

$$a + \varepsilon b \tag{1}$$

for $a, b \in \mathbb{R}$. The multiplication of two dual numbers is straightforward

$$(a_1 + \varepsilon b_1)(a_2 + \varepsilon b_2) = a_1 a_2 + \varepsilon(b_1 a_2 + a_1 b_2) \ . \tag{2}$$

Dual quaternions (denoted by $\mathbb{H}_D$) are a quaternion equivalent to dual numbers. Thus, they are defined by replacing real numbers a, b in (1) with quaternions $\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{H}$. That is, a dual quaternion $\mathbf{dq}$ can be written as

$$\mathbf{dq} = \mathbf{q}_1 + \varepsilon \mathbf{q}_2 \ .$$

Multiplication of dual quaternions works in the same way as multiplication of dual numbers (2) except for the fact that multiplication of dual quaternions is not commutative because quaternion multiplication is not commutative. Thus, the product of two dual quaternions $\mathbf{da}, \mathbf{db} \in \mathbb{H}_D$ is given by

$$\begin{aligned} \mathbf{da} \cdot \mathbf{db} =& (\mathbf{a}_1 + \varepsilon \mathbf{a}_2) \cdot (\mathbf{b}_1 + \varepsilon \mathbf{b}_2) \\ =& \mathbf{a}_1 \cdot \mathbf{b}_1 + \varepsilon(\mathbf{a}_1 \cdot \mathbf{b}_2 + \mathbf{a}_2 \cdot \mathbf{b}_1) \ . \end{aligned}$$

Several different types of conjugation exists for dual quaternions. We are particularly interested in conjugating a dual quaternion by conjugating both its quaternions individually. Thus, we define the conjugated dual quaternion as $\mathbf{dq}^* := \mathbf{q}_1^* + \varepsilon \mathbf{q}_2^*$, where $\mathbf{q}_i^*$ means the quaternion conjugation.

The magnitude of a dual quaternion is given by

$$||\mathbf{dq}|| = \sqrt{\mathbf{dq} \cdot \mathbf{dq}^*} \ .$$

Dual quaternions with magnitude 1 are of our particular interest. They are also known as unit dual quaternions. The multiplicative inverse of such a unit dual quaternion $\mathbf{q}_1 + \varepsilon \mathbf{q}_2$ is given by $\mathbf{q}_1^* + \varepsilon \mathbf{q}_2^*$. It is important to note, that every unit quaternion (with a zero dual part) is itself a unit dual quaternion.

B. Representation of Rigid Body Motions

Quaternions can be used to represent orientation in $\mathbb{R}^3$. A rotation around the unit-length axis $(x_1, x_2, x_3)^\top$ with angle α is represented by the quaternion

$$\mathbf{r}_a = \cos\left(\frac{\alpha}{2}\right) + \sin\left(\frac{\alpha}{2}\right)(x_1 i + x_2 j + x_3 k) \ .$$

A more in-depth discussion of quaternions for orientation representation can be found in [16]. In a dual quaternion a pure rotation is represented by a quaternion where the non-dual part is the unit quaternion representing the rotation and the dual part is zero, i.e., $\mathbf{dr} = \mathbf{r}_a + \varepsilon\, 0$. Thus, classical unit quaternions are a special case of orientation representation using dual quaternions. A translation $\underline{t} = (t_x, t_y, t_z)^\top \in \mathbb{R}^3$ is represented by the dual quaternion

$$\begin{aligned}\mathbf{dt} =& \mathbf{t}_1 + \varepsilon\, \mathbf{t}_2 \\ =& 1 + \varepsilon \frac{1}{2}(t_x i + t_y j + t_z k) \ .\end{aligned}$$

It is important to note, that the quaternion in the dual part is not necessarily of unit length, and thus there is no restriction imposed on the translation. However, the resulting dual quaternion is still a unit dual quaternion with respect to the definition given above. A composition of rotation and orientation is represented by dual quaternion multiplication. The inverse of this transformation is simply obtained by conjugating the dual quaternion as described above. It is important to note that — similarly to the purely rotational case — the dual quaternions $\mathbf{dq}$ and $-\mathbf{dq}$ represent the same position and orientation.

In order to apply a rigid body transform represented by a dual quaternion $\mathbf{dq}$ to a given vector $\underline{x} = (x_1, x_2, x_3)$, we represent the vector as a dual quaternion

$$\mathbf{dx} = 1 + \varepsilon(x_1 i + x_2 j + x_3 k) \ .$$

Now the transformed vector is represented by $\mathbf{dq} \cdot \mathbf{dx} \cdot \overline{\mathbf{dq}}^*$ (where $\overline{\mathbf{dq}}^*$ denotes conjugation of each involved quaternion and dual conjugation, i.e., changing the sign in front of the dual unit ε). A consequence of this operation is that there is no need for the factor $1/2$ in the dual part of $\mathbf{dx}$. Further discussion can be found in [15].

III. A New Distribution Model

An antipodally symmetric distribution is required to account for the fact that dual quaternions $\mathbf{dq}$ and $-\mathbf{dq}$ represent the same rigid transformation. In the case of quaternions, this can be simply done by using the Bingham distribution, which is obtained by restricting the random variables of a zero mean Gaussian distribution to unit length. A similar strategy can be used to obtain a probability distribution for unit dual quaternions. In the following, S_1 is used to denote the unit circle in $\mathbb{R}^2$. Now, we can define the proposed distribution.

Definition 1. *A random vector $\underline{x} \in S_1 \times \mathbb{R}^2$ is distributed according to the proposed distribution if its probability density function (p.d.f.) is given by*

$$f(\underline{x}) = \frac{1}{N(\mathbf{C})} \exp(\underline{x}^\top \mathbf{C} \underline{x}) \ ,$$

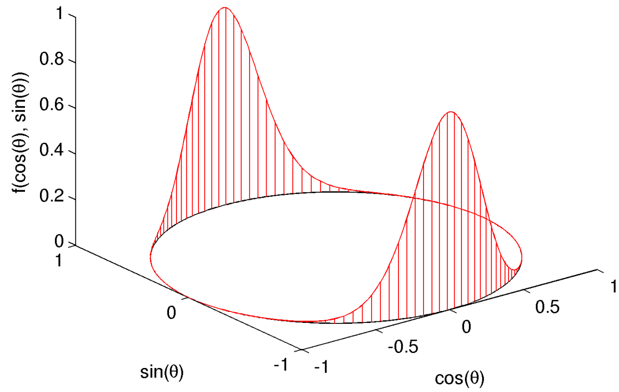


Figure 1. Probability density function of the Bingham distribution, which is the marginal distribution for the first two components of the proposed distribution model.

where $\mathbf{C}$ is a suitable symmetric parameter matrix and $N(\mathbf{C})$ a corresponding normalization constant.

We can decompose $\underline{x}$ into two components $\underline{x} = (\underline{x}_s^\top, \underline{x}_t^\top)^\top$ with $\underline{x}_s \in S^1 \subset \mathbb{R}^2$ and $\underline{x}_t \in \mathbb{R}^2$. Later, we will see that $\underline{x}_s$ can be interpreted as the non-dual part of a unit dual quaternion and $\underline{x}_t$ will be interpreted as its dual part. This is possible, when we consider dual quaternions restricted to representing a rigid-body motion in the plane (otherwise a higher dimensional probability distribution would be needed). Furthermore, the normalization constant is given by

$$N(\mathbf{C}) = \int_{S^1} \int_{\mathbb{R}^2} \exp\left(\begin{pmatrix} \underline{x}_s \\ \underline{x}_t \end{pmatrix}^\top \mathbf{C} \begin{pmatrix} \underline{x}_s \\ \underline{x}_t \end{pmatrix} \right) \, \mathrm{d}\underline{x}_t \, \mathrm{d}\underline{x}_s \ .$$

Due to the symmetry of $\mathbf{C}$, it can be represented as

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2^\top \\ \mathbf{C}_2 & \mathbf{C}_3 \end{pmatrix} \ ,$$

with $\mathbf{C}_i \in \mathbb{R}^{2\times 2}$. This notation can be used for deriving a more convenient representation of the density function, and to make clear which choices of the matrices $\mathbf{C}_i$ are feasible.

Lemma 2. *The proposed probability distribution can be represented by*

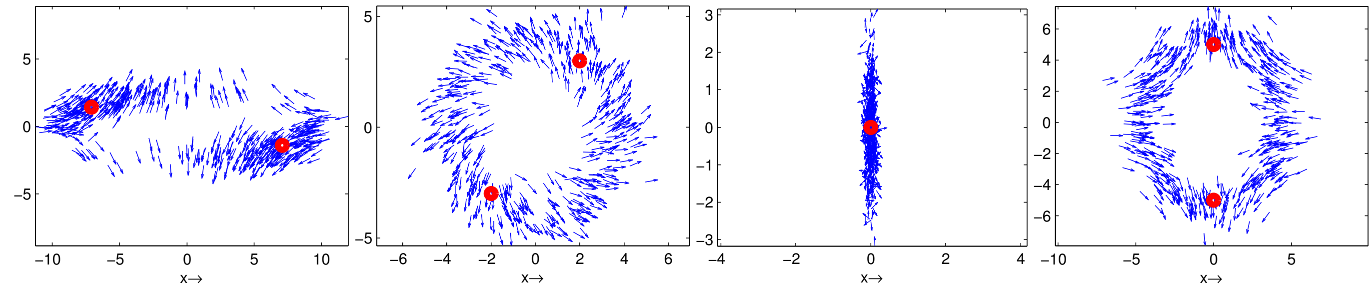
$$\begin{aligned} f(\underline{x}) =& \frac{1}{N(\mathbf{C})} \exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s \\ &\qquad + (\underline{x}_t - \mathbf{A}\underline{x}_s)^\top \mathbf{C}_3 (\underline{x}_t - \mathbf{A}\underline{x}_s)\big) \ , \end{aligned} \tag{3}$$

where $\mathbf{A} = -\mathbf{C}_3^{-1}\mathbf{C}_2$. This is a well-defined probability density function for symmetric $\mathbf{C}_1$, arbitrary $\mathbf{C}_2$ and symmetric negative definite $\mathbf{C}_3$.

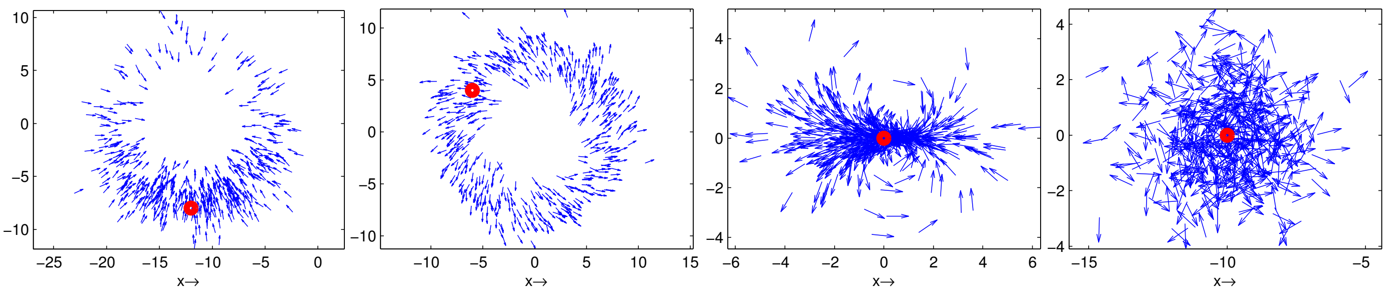
A proof is given in Appendix A.

A. Parameter Estimation

For parameter estimation, we adapt the methodology outlined in [12] to our distribution. This is made clear by formulating a relationship of the proposed distribution model with



(a) Original samples before transformation.



(b) Translation and rotation represented by these samples.

Figure 2. Illustration of several different random sample sets and the rotation / translation combination represented by these random samples as obtained using Algorithm 1. Each sample is represented by an arrow. The position of the arrow represents the translational part and its direction represents the angular part. Modes of the underlying distributions are shown as red dots. Due to antipodal symmetry, both modes of a considered distribution represent the same rotation. Thus, there is no such symmetry in the transformed samples.

the Bingham distribution, which is an antipodally symmetric distribution on the hypersphere. Its p.d.f. is given by

$$f(\underline{x}) = \frac{1}{F(\mathbf{Z})} \exp\left(\underline{x}^\top \mathbf{M}\,\mathbf{Z}\,\mathbf{M}^\top \underline{x}\right) \ ,$$

where $\mathbf{Z}$ is a diagonal matrix, $\mathbf{M}$ is an orthogonal matrix, and $\underline{x} \in S_n$. This parametrization is the same as in our previous work [4]. The matrix $\mathbf{M}$ is left out of the normalization constant, because it can be shown that $F(\mathbf{M}\,\mathbf{Z}\,\mathbf{M}^\top) = F(\mathbf{Z})$. A typical density of a Bingham distribution on S_1 is shown in Fig. 1. Its role as a marginal distribution of our proposed model is made precise by the following lemma.

Lemma 3. *Let $\underline{x} = (\underline{x}_s, \underline{x}_t)$ be a random variable following the distribution defined above. Then, the Bingham distribution is the marginal distribution of $\underline{x}_s$ with parameter matrices $\mathbf{M}$,$\mathbf{Z}$ being the eigendecomposition of the Schur complement of $\mathbf{C}_3$ in $\mathbf{C}$, i.e., $\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2$.*

A proof is given in Appendix B.

This result gives us the first part of the parameter estimation procedure. Using the approach discussed in [3] (or any other approach for estimating parameters of the Bingham distribution), we can estimate $\mathbf{M}$, $\mathbf{Z}$ and thus we obtain an estimate of $\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2$.

For the further derivation of the estimation procedure, it is important to note that $\mathcal{N}(-\mathbf{A}\underline{x}_s, -\frac{1}{2}\mathbf{C}_3^{-1})$ is the conditional distribution of $\underline{x}_t$ given a fixed $\underline{x}_s$. This is a simple consequence of the representation (3). That is, the unconstrained part follows a Gaussian distribution, with an uncertain transformed Bingham-distributed mean. Estimation of $\mathbf{A}$ and $-\frac{1}{2}\mathbf{C}_3^{-1}$ is done using multivariate linear regression as described in [17]. These estimates can be used to reconstruct estimates of the original parameter matrices directly. Furthermore, we can derive a mode of the distribution, which is given by $\underline{m} = (\underline{m}_r^\top, \underline{m}_t^\top)^\top$, where $\underline{m}_r \in S_1$ is the normalized eigenvector corresponding to the largest eigenvalue of $\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2$, and $\underline{m}_t = \mathbf{A}\,\underline{m}_s$. Because of antipodal symmetry $-\underline{m}$ is also a mode of the distribution.

B. Normalization Constant

Computing the normalization constant $N(\mathbf{C})$ is not straightforward, as it involves an integration over $S^1 \times \mathbb{R}^2$. Fortunately, we can reduce this computation to computing a Bingham normalization constant.

Lemma 4. *The normalization constant of the proposed distribution can be rewritten as*

$$N(\mathbf{C}) = \frac{2\pi\sqrt{\det\left(-\frac{1}{2}\mathbf{C}_3^{-1}\right)}}{F(\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2)} \ ,$$

where $F(\cdot)$ is the normalization constant of a Bingham distribution.

A proof is given in Appendix C. Unfortunately, the computation of a Bingham normalization constant is not straight forward either. However, there are several approaches addressing this problem, which are based on precomputed lookup tables [18], saddlepoint approximations [19], and holonomic gradient descent [20].

IV. Uncertain Rigid-Body Motions in the Plane

A naïve approach would be to use the distribution presented in the previous section to represent an uncertain motion in the

Algorithm 1: Translation and rotation from unit dual quaternion

Input: Dual quaternion $\underline{w}_{\mathbf{dq}} = (w_1, w_2, w_3, w_4)^T$
Output: Rotation angle φ and translation vector $\underline{t} = (t_1, t_2)$

```
/* Compute rotation angle */
```
$\varphi \leftarrow 2 \cdot \text{atan2}(w_2, w_1)$

```
/* Compute translation */
```
$t_1 \leftarrow 2 \cdot (w_1\, w_3 - w_2\, w_4)$
$t_2 \leftarrow 2 \cdot (w_2\, w_3 + w_1\, w_4)$

plane by using the first two entries to represent the rotation and the second two entries to represent a subsequent translation. Unfortunately, this naïve approach has several drawbacks. First, $\underline{x}$ and $-\underline{x}$ would always have the same probability mass (even if it is not desired that the positions $\underline{x}_t$ and $-\underline{x}_t$ have the same probability mass). Second, it would not be possible to represent a mean for the position. Thus, we use dual quaternions (restricted to rotation and translation in the plane) for a natural representation of uncertainty.

A. Representation of Uncertain Motions in the Plane

Rotation in the (x, y)-plane can be represented by the dual quaternion

$$\mathbf{dr} = \left(\cos\left(\frac{\alpha}{2}\right) + 0 \cdot i + 0 \cdot j + \sin\left(\frac{\alpha}{2}\right) \cdot k\right) + \varepsilon \cdot 0 \ .$$

A translation (t_1, t_2) in the (x, y)-plane can be represented by the dual quaternion

$$\mathbf{dt} = 1 + \frac{\varepsilon}{2}\left(0 + t_1 \cdot i + t_2 \cdot j + 0 \cdot k\right) \ .$$

Thus, a combination of rotation and translation (where the rotation is performed first) is given by

$$\begin{aligned}\mathbf{dt} \cdot \mathbf{dr} = &\left(\cos\left(\frac{\alpha}{2}\right) + \sin\left(\frac{\alpha}{2}\right) \cdot k\right) \\ &+ \frac{\varepsilon}{2}\left[\left(\cos\left(\frac{\alpha}{2}\right) t_1 + \sin\left(\frac{\alpha}{2}\right) t_2\right) \cdot i \right. \\ &\left. + \left(-\sin\left(\frac{\alpha}{2}\right) t_1 + \cos\left(\frac{\alpha}{2}\right) t_2\right) \cdot j\right] \\ =: &d_1 + d_2\, k + \varepsilon(d_3\, i + d_4\, j)\end{aligned}$$

It is important to note that, in this planar setting, the dual part of the dual quaternion $\mathbf{dq} := \mathbf{q}_a + \varepsilon \mathbf{q}_b$ has a zero real and a zero k component. Furthermore, the non-dual part always keeps a zero i and a zero j component. Thus, we require only four values to represent our rigid body motion on the plane. This gives us a new way of interpretation of a random vector following the proposed distribution. Its first two entries are interpreted as the real and the k component of the non-dual part of $\mathbf{dq}$ and its last two entries are interpreted as the i and j component of the dual part of $\mathbf{dq}$. That is

$$\underline{d}_{\mathbf{dq}} = (d_1, d_2, d_3, d_4)^\top \ .$$

A way to recover the original rotation angle and translation is presented in Algorithm 1. This algorithm assumes that the translation is performed after the rotation. Thus, its returned values can be used in a straight forward manner for constructing a homogeneous matrix of an affine transformation. In Fig. 2, we apply this algorithm to several sets of random samples from the proposed distribution. Consequently, we obtain the rigid transformation represented by these samples.

It is also possible to define a matrix representation of the dual quaternions involved that obeys the corresponding operations. The derivation is based on combining matrix representations of dual numbers with matrix representations of quaternions, and then leaving out columns and rows unnecessary for the planar case. This results in

$$\mathbf{D}_{\mathbf{dq}} = \begin{pmatrix} w_1 & w_2 & 0 & 0 \\ -w_2 & w_1 & 0 & 0 \\ -w_3 & w_4 & w_1 & -w_2 \\ -w_4 & -w_3 & w_2 & w_1 \end{pmatrix} \ .$$

This matrix representation has several beautiful properties such as $\mathbf{D}_{\mathbf{da} \cdot \mathbf{db}} = \mathbf{D}_{\mathbf{da}} \cdot \mathbf{D}_{\mathbf{db}}$ and thus $\mathbf{D}_{\mathbf{dq}^{-1}} = \mathbf{D}_{\mathbf{dq}}^{-1}$. Furthermore, we have that $\det(\mathbf{D}_{\mathbf{dq}}) = 1$, because $\mathbf{D}_{\mathbf{dq}}$ is a Block diagonal matrix and the diagonal 2×2 blocks are rotation matrices and thus have determinant 1. It is also possible to combine both representations for performing dual quaternion multiplication

$$\begin{aligned}\underline{d}_{\mathbf{da} \cdot \mathbf{db}} =& \text{diag}(1, -1, -1, -1) \cdot \mathbf{D}_{\mathbf{da}} \\ &\cdot \text{diag}(1, -1, -1, -1) \cdot \underline{d}_{\mathbf{db}} \ .\end{aligned}$$

Finally, the inverse of a unit dual quaternion can be obtained by $\underline{d}_{\mathbf{da}^{-1}} = \text{diag}(1, -1, -1, -1) \cdot \underline{d}_{\mathbf{da}}$.

B. Bayesian Inference

A typical scenario in which Bayesian inference is applied considers a system state (which is typically unknown) $\underline{x}$ and a noisy measurement of this state $\underline{z}$. The noise is represented by some random vector $\underline{v}$. This can be formulated in our dual quaternion setting resulting in

$$\mathbf{dz} = \mathbf{dv} \cdot \mathbf{dx} \ . \tag{4}$$

Bayesian inference can now be used to perform classical measurement updates, e.g., for measurement fusion. We will show that the combination of dual quaternions and the proposed probability distribution yields a closed form measurement update for simultaneous consideration of position and orientation.

Lemma 5. *Consider* (4)*, where* $\mathbf{dv}$ *and* $\mathbf{dx}$ *are distributed according to our proposed distribution with respective parameter matrices* $\mathbf{C}_v$ *and* $\mathbf{C}_x$*. Then,* $\mathbf{dx}$ *given a fixed* $\mathbf{dz}$ *is also distributed according to our proposed distribution.*

A proof is given in Appendix D. It also describes a procedure to obtain the parameter matrix of the posterior distribution. In the following example we show the application of this result to fusing measurements.

Example: Measurement Fusion
Consider the proposed distribution with parameter matrix

$$\mathbf{C}_{true} = \mathbf{B}_1 \cdot \text{diag}(-1, -100, -50, -50) \cdot \mathbf{B}_1^T \ ,$$

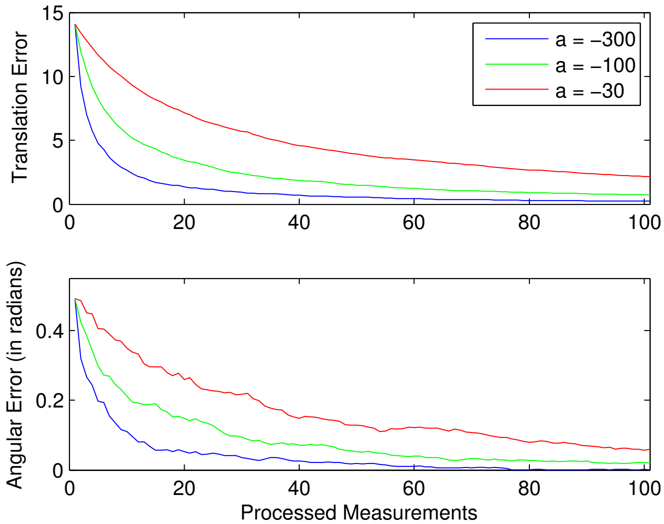


Figure 3. Measurement fusion example.

where $\mathbf{B}_1$ is a suitable repositioning matrix ensuring the mode of the resulting distribution is at $(\cos(340°), \sin(340°), 10, 10)^T$ after converting the dual quaternion to coordinates directly representing the rotation and translation as two vectors in $\mathbb{R}^2$. It is created by multiplying $\mathrm{diag}(1, -1, -1, -1)$ with a matrix representation of the quaternion representing the desired mode. In this example, we sample one value from this distribution and try to estimate this value by performing sequential measurement updates. The prior is given by the parameter matrix $\mathbf{C}_{prior} = \mathrm{diag}(-1, -500, -500, -500)$. This prior has its mode at $(1, 0, 0, 0)$, i.e., the mode of the prior presents a wrong estimate of the true hidden value. Now, we perform 100 sequential measurement updates, where the measurements are sampled from our distribution with parameter matrix

$$\mathbf{C}_{obs} = \mathbf{B}_2 \cdot \mathrm{diag}(-1, -a, -a, -a) \cdot \mathbf{B}_2^T \ .$$

Here $\mathbf{B}_2$ is chosen in a way relocating the mode of the distribution to being the true hidden value. For a, we chose three different scenarios with $a = -30$, $a = -100$ and $a = -300$. In this setting, larger a denotes stronger noise. Measurement updates are performed using the formulas from the proof of Lemma 5. The resulting angular error and translational error are computed by extracting the mode of the posterior after each measurement update step and comparing them to the true sampled value. This is shown in Fig. 3.

V. Conclusions

In this work, we proposed a new distribution capable of describing uncertainty on the group of rigid transformations in the plane $SE(2)$. This is of particular interest, because it considers the inherently nonlinear structure of $SE(2)$, and thus, errors based on a wrong linearity assumption can be avoided.

Uncertain rigid transformations appear in a broad scope of technical applications, involving mixed and augmented reality systems, robotics, and tracking. Consequently, our future work will focus on presenting dynamic state estimation algorithms based on this distributional model. Furthermore, we also want to focus on investigating theoretical aspects of distributions inherently defined on the $SE(2)$ and $SE(3)$.

Acknowledgment

This work was partially supported by grants from the German Research Foundation (DFG) within the Research Training Groups RTG 1194 "Self-organizing Sensor-Actuator-Networks", the RTG 1126 "Soft-tissue Surgery: New Computer-based Methods for the Future Workplace", and EU grant FP7-ICT-2011-318787 (IQmulus).

Appendix

A. Proof of Lemma 2

The proof is carried out by rewriting the exponent in the density function. First, we note that

$$\begin{aligned}
&\begin{pmatrix} \underline{x}_s \\ \underline{x}_t \end{pmatrix}^\top \cdot \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2^\top \\ \mathbf{C}_2 & \mathbf{C}_3 \end{pmatrix} \cdot \begin{pmatrix} \underline{x}_s \\ \underline{x}_t \end{pmatrix} \\
&\quad = \underline{x}_s^T \mathbf{C}_1 \underline{x}_s + 2 \cdot \underline{x}_t^\top \mathbf{C}_2 \underline{x}_s + \underline{x}_t^\top \mathbf{C}_3 \underline{x}_t \\
&\quad = \underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s \\
&\qquad + (\underline{x}_t + \mathbf{C}_3^{-1} \mathbf{C}_2 \underline{x}_s)^\top \mathbf{C}_3 (\underline{x}_t + \mathbf{C}_3^{-1} \mathbf{C}_2 \underline{x}_s) \ .
\end{aligned}$$

Thus, the original density can be rewritten as (3). Symmetry of $\mathbf{C}_1$ and $\mathbf{C}_3$ is necessary to ensure antipodal symmetry. In order to show that $f(x)$ is a proper p.d.f., we have to show that $N(\mathbf{C}) < \infty$. This is done by integrating the unrestricted part first and then integrating over the spherical part. This results in

$$\begin{aligned}
N(\mathbf{C}) = &\int_{S^1} \int_{R^2} \exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s \\
&\qquad + (\underline{x}_t - \mathbf{A}\underline{x}_s)^\top \mathbf{C}_3 (\underline{x}_t - \mathbf{A}\underline{x}_s)\big) \, \mathrm{d}\underline{x}_t \, \mathrm{d}\underline{x}_s \\
\propto &\int_{S^1} \exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s\big) \, \mathrm{d}\underline{x}_s \ .
\end{aligned}$$

We observe that the inner integral converges for arbitrary $\mathbf{C}_2$ and negative definite $\mathbf{C}_3$. This integration corresponds to integration over an unnormalized gaussian p.d.f. Thus, $\mathbf{A}\underline{x}_s$ is interpreted as a mean which disappears after integration. The remaining integral converges for arbitrary $\mathbf{C}_1$. □

B. Proof of Lemma 3

The lemma is shown by simply integrating out $\underline{x}_t$ from the density (3).

$$\begin{aligned}
f(\underline{x}_s) \propto &\int_{\mathbb{R}^2} \exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s \\
&\qquad + (\underline{x}_t - \mathbf{A}\underline{x}_s)^\top \mathbf{C}_3 (\underline{x}_t - \mathbf{A}\underline{x}_s)\big) \, \mathrm{d}\underline{x}_s \\
\propto &\exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s\big) \\
&\qquad \cdot \int_{\mathbb{R}^2} (\underline{x}_t - \mathbf{A}\underline{x}_s)^\top \mathbf{C}_3 (\underline{x}_t - \mathbf{A}\underline{x}_s)\big) \, \mathrm{d}\underline{x}_s \\
\propto &\exp\big(\underline{x}_s^\top (\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1} \mathbf{C}_2) \underline{x}_s\big) \ .
\end{aligned}$$

where once again $\mathbf{A} = -\mathbf{C}_3^{-1} \mathbf{C}_2$.

C. Proof of Lemma 4

We consider a random variable $\underline{x}^\top = (\underline{x}_s^\top, \underline{x}_t^\top)^\top$ distributed according to our proposed distribution. From Lemma 3, we know that the density of $\underline{x}_s$ is given by

$$\begin{aligned}f(\underline{x}_s) =& \frac{1}{N(\mathbf{C})} \exp\left(\underline{x}_s^\top(\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1}\mathbf{C}_2)\underline{x}_s\right) \\ &\cdot \int_{\mathbb{R}^2} \exp\left((\underline{x}_t - \mathbf{A}\underline{x}_s)^\top \mathbf{C}_3(\underline{x}_t - \mathbf{A}\underline{x}_s)\right) \\ =& \frac{2\pi\sqrt{\det(-\frac{1}{2}\mathbf{C}_3^{-1})}}{N(\mathbf{C})} \exp\left(\underline{x}_s^\top(\mathbf{C}_1 - \mathbf{C}_2^\top \mathbf{C}_3^{-1}\mathbf{C}_2)\underline{x}_s\right) ,\end{aligned}$$

where once again $\mathbf{A}$ is defined as in (3). The resulting normalization constant of the marginal distribution is simply the Bingham normalization constant $F(\cdot)$. Thus, we have

$$N(\mathbf{C}) = \frac{2\pi\sqrt{\det(-\frac{1}{2}\mathbf{C}_3^{-1})}}{F(\mathbf{C}_1 - \mathbf{C}_2^T \mathbf{C}_3^{-1}\mathbf{C}_2)} .$$

□

D. Proof of Lemma 5

Using Bayes theorem, we obtain

$$f_x(\underline{d}_{\mathbf{dx}}|\underline{d}_{\mathbf{dz}}) \propto f_z(\underline{d}_{\mathbf{dz}}|\underline{d}_{\mathbf{dx}}) \cdot f_x(\underline{d}_{\mathbf{dz}}) ,$$

where $\underline{d}_{\mathbf{dx}}$ and $\underline{d}_{\mathbf{dz}}$ are vector representations of our considered unit dual quaternions. Here, $f_x(\cdot)$ is the prior knowledge. Thus, it remains to derive $f_z(\underline{d}_{\mathbf{dz}}|\underline{d}_{\mathbf{dx}})$, which is directly related to the density of $\mathbf{dv}$. We now show how to derive f_z from the density of $\mathbf{dv}$. First, we note that

$$\mathbf{dv} = \mathbf{dz} \cdot \mathbf{dx}^{-1} .$$

After reformulating $\mathbf{dz}$ to its matrix form and $\mathbf{dx}^{-1}$ to its vector representation, we obtain

$$\begin{aligned}\underline{d}_{\mathbf{dz}\cdot\mathbf{dx}^{-1}} =& \mathbf{T} \cdot \mathbf{D}_{\mathbf{dz}} \cdot \mathbf{T} \cdot \underline{d}_{\mathbf{dx}^{-1}} \\ =& \mathbf{T} \cdot \mathbf{D}_{\mathbf{dz}} \cdot \mathbf{T}^2 \cdot \underline{d}_{\mathbf{dx}} \\ =& \underbrace{\mathbf{T} \cdot \mathbf{D}_{\mathbf{dz}}}_{=:\mathbf{H}(\mathbf{dz})} \cdot \underline{d}_{\mathbf{dx}} ,\end{aligned}$$

where $\mathbf{T} = \operatorname{diag}(1, -1, -1, -1)$. Thus, we have $\underline{d}_{\mathbf{dv}} = \mathbf{H}(\mathbf{dz}) \cdot \underline{d}_{\mathbf{dx}}$. Furthermore, we note that $|\det(\mathbf{H}(\mathbf{dz}))| = 1$. Applying the transformation theorem for densities yields

$$f_z(\underline{v}_{\mathbf{dz}}|\underline{v}_{\mathbf{dx}}) = f_v(\mathbf{H}(\mathbf{dz}) \cdot \underline{v}_{\mathbf{dx}}) .$$

This gives us

$$\begin{aligned}f_x(\underline{d}_{\mathbf{dx}}|\underline{d}_{\mathbf{dz}}) \propto& f_z(\underline{d}_{\mathbf{dz}}|\underline{d}_{\mathbf{dx}}) \cdot f_x(\underline{d}_{\mathbf{dz}}) \\ \propto& \exp\left(\underline{d}_{\mathbf{dx}}^\top \left(\mathbf{H}(\mathbf{dz})^\top \cdot \mathbf{C}_v \cdot \mathbf{H}(\mathbf{dz}) + \mathbf{C}_x\right) \underline{d}_{\mathbf{dx}}\right) ,\end{aligned}$$

which is just the shape of the desired pdf. □

References

[1] S. Julier and J. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401 – 422, 2004.

[2] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, "Recursive Nonlinear Filtering for Angular Data Based on Circular Distributions," in *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D. C., USA, Jun. 2013.

[3] G. Kurz, I. Gilitschenski, S. J. Julier, and U. D. Hanebeck, "Recursive Estimation of Orientation Based on the Bingham Distribution," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[4] I. Gilitschenski, G. Kurz, S. Julier, and U. D. Hanebeck, "Unscented Orientation Estimation Based on the Bingham Distribution," *arXiv preprint: Systems and Control (cs.SY)*, 2013.

[5] J. Glover and L. P. Kaelbling, "Tracking 3-D Rotations with the Quaternion Bingham Filter," MIT CSAIL, Tech. Rep., Mar. 2013.

[6] K. V. Mardia and P. E. Jupp, *Directional Statistics*, 1st ed. Wiley, 1999.

[7] C. Bingham, "An Antipodally Symmetric Distribution on the Sphere," *The Annals of Statistics*, vol. 2, no. 6, pp. 1201–1225, Nov. 1974.

[8] J. S. J. Goddard, "Pose and Motion Estimation from Vision Using Dual Quaternion-based Extended Kalman Filtering," Ph.D. dissertation, University of Tennessee, Jan. 1997.

[9] W. Feiten, P. Atwal, R. Eidenberger, and T. Grundmann, "6D Pose Uncertainty in Robotic Perception," in *Advances in Robotics Research*, T. Kröger and F. M. Wahl, Eds. Springer Berlin Heidelberg, 2009, pp. 89–98.

[10] M. Lang and W. Feiten, "MPG - Fast Forward Reasoning on 6 DOF Pose Uncertainty," in *7th German Conference on Robotics (ROBOTIK 2012)*, Munich, 2012, pp. 1–6.

[11] W. Feiten, M. Lang, and S. Hirche, "Rigid Motion Estimation Using Mixtures of Projected Gaussians," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, 2013, pp. 1465–1472.

[12] A. Kume, S. P. Preston, and A. T. A. Wood, "Saddlepoint Approximations for the Normalizing Constant of Fisher-Bingham Distributions on Products of Spheres and Stiefel Manifolds," *Biometrika*, vol. 100, no. 4, pp. 971–984, Aug. 2013.

[13] K. Shoemake, "Animating Rotation with Quaternion Curves," *ACM SIGGRAPH Computer Graphics*, vol. 19, no. 3, pp. 245–254, Jul. 1985.

[14] L. Kavan, S. Collins, J. Žára, and C. O'Sullivan, "Skinning with Dual Quaternions," in *Proceedings of the 2007 symposium on Interactive 3D graphics and games - I3D '07*. New York, New York, USA: ACM Press, Apr. 2007, p. 39.

[15] J. M. McCarthy, *Introduction to Theoretical Kinematics*. The MIT Press, 1990.

[16] J. B. Kuipers, *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press, 2002.

[17] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 2003.

[18] J. Glover, R. Rusu, and G. Bradski, "Monte Carlo Pose Estimation with Quaternion Kernels and the Bingham Distribution," in *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, Jun. 2011.

[19] A. Kume and A. T. A. Wood, "Saddlepoint Approximations for the Bingham and Fisher-Bingham Normalising Constants," *Biometrika*, vol. 92, no. 2, pp. 465–476, Jun. 2005.

[20] H. Nakayama, K. Nishiyama, M. Noro, K. Ohara, T. Sei, N. Takayama, and A. Takemura, "Holonomic Gradient Descent and its Application to the Fisher–Bingham Integral," *Advances in Applied Mathematics*, vol. 47, no. 3, pp. 639–658, Sep. 2011.