Advances in the Astronautical Sciences, Vol. 108, December 2001, pp. 383–396

AAS-01-125

MAN, LIKE THESE ATTITUDES ARE TOTALLY RANDOM! I. QUATERNIONS*

Malcolm D. Shuster**

We address the problem of modeling the probability density function of the attitude quaternion when there is no *a priori* knowledge of the attitude of any kind. We first define generally, on the basis of purely physical arguments, what is meant by such a probability density function and develop a general expression from which the probability density function for any three-dimensional attitude representation can be determined. We then calculate explicit expressions for this completely *a priori* probability density function for the attitude quaternion, both for the vectorial components alone and as a function of all four components treated independently in R^4.

INTRODUCTION

Generally, one needs to model a totally random attitude for which there is no *a priori* information in two situations:

1. In simulation when one wishes to test an algorithm for all attitudes in a "uniform" and unbiased manner; and
2. In Bayesian attitude estimation, when there is no *a priori* information on the attitude.

In the present work, which is in two parts, we address ourselves entirely to the first topic. The second topic will be treated at length in a succeeding paper devoted to Bayesian techniques in attitude estimation.

In order to carry out "uniform" simulation tests of an attitude-dependent algorithm, one must be able to generate random samples of the attitude which have a probability density function [1] (pdf) which is *uniform* in the sense that no set of attitudes will be more probable than another. Unfortunately, except in the most trivial cases, it is not obvious how to construct the pdf

*Presented at the 11th AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, California, February 11–14, 2001.
** Scientist, Space Systems Group, Orbital Sciences Corporation, Germantown, MD 20874. email: m.shuster@ieee.org. Current Address: Acme Space Company, 13017 Wisteria Drive, PMB 328, Germantown, MD 20874.

of a random variable which is "uniform" in this way. As it turns out, it is possible to generate an unambiguous expression for the pdf, if the random variable and a given operation forms a group.

As a first step to deriving the more general methods, consider the very simple case of attitude about a fixed axis of rotation, so that the only attitude degree of freedom is the angle of rotation. This case will serve as an example in miniature of our general approach.

We write the pdf of the angle as $p_\theta(\theta')$ where θ is the random variable, which serves here also as a label on the pdf, and θ' is a possible value of the sampled angle, the *realization* of the random variable. If there is no *a priori* knowledge of any kind on the angle of rotation, then the pdf of the angle of rotation should be the same independent of the fixed direction, perpendicular to the axis of rotation, from which the angle of rotation is measured. Thus, if the reference direction is rotated by an angle $-c$, the random angle of rotation becomes equivalently $\theta + c$, and we must have

$$p_\theta(\theta') = p_{\theta+c}(\theta')\,, \tag{1}$$

with the addition of angles being understood as modulo 2π. However, it is also true that

$$p_{\theta+c}(\theta') = p_\theta(\theta' - c)\,\left|\frac{d(\theta' - c)}{d\theta'}\right| = p_\theta(\theta' - c)\,. \tag{2}$$

Hence,

$$p_\theta(\theta') = p_\theta(\theta' - c)\,. \tag{3}$$

Since c is arbitrary it follows that $p_\theta(\theta')$ must be independent of θ', whence, from the constraint that the total probability be unity, it follows that

$$p_\theta(\theta') = \frac{1}{2\pi}\,, \tag{4}$$

the answer which we expected intuitively.

This derivation makes clear why the pdf for θ without any *a priori* knowledge is a constant function of θ', but not the pdf for θ^2 as a function of $(\theta')^2$ under these same conditions, since the composition rule for the later is not simple addition.

The key equation here is Eq. (1), which relates the "uniformity" of the pdf to a physical process, the selection of the reference point. The next step is to change the relation between the two random variables to an equation describing the dependence on the argument of the pdf. From there one calculates the explicit function for the pdf.

These same steps will be followed for three-axis attitude, with the following differences: (1) the change of variable will involve at least three coordinates, and will not be as simple as Eq. (2); the composition rule is not simple addition for any attitude representation, but is non-linear, as a result of which the "uniform" pdf of the attitude will never be a constant function in Euclidean space. Clearly, since we assume no prior information from measurements, the only knowledge of the attitude that can be used to construct the pdf for these "uniform" distributions is the nature of the composition rule for the representation.

It should be said at this point that not all of the results presented in this work can be said to be original, even though the author has derived them independently. This area was studied by mathematicians in the first half of the twentieth century, and belongs to the Theory of Representations of Compact Groups, whose study requires the knowledge of Measure Theory, Topology, Group Theory, and the Theory of Differentiable Manifolds, none of which are part of the usual education of aerospace engineers at any degree level.

Fortunately, the results needed by aerospace engineers can be derived solely on the basis of Riemann-Stieljes integration, which forms a part of every undergraduate Calculus course. The vast majority of the results in this work are new, although many of the new results are not of particular interest to pure mathematicians, whose focus is on rotations as an abstract compact

Lie group and not on the specific attitude representations and their application to problems of Engineering.

Because of the importance of the quaternion in all areas of attitude analysis a great deal of attention is devoted to the representation of the probability of the four-component quaternion. In fact, we study this problem twice, using approaches which offer very different insights. The first approach, more familiar to mathematicians, replaces the unit 3-sphere of the quaternion with a finite covering of overlapping open sets. On each open set the quaternion and its probability density function have a different parameterization, but simple relationships exist relating these where the open sets overlap, and these provide a rigorous though somewhat clumsy means for evaluating the probability density function for any value of the unit quaternion. In a roundabout way the constancy of the quaternion probability density on the unit 3-sphere is demonstrated, and this is used to construct the probability function in Euclidean 4-space. The second approach, more familiar to theoretical physicists, examines the probability density function from the outset in the full four-dimensional space (which thus includes quaternions of arbitrary norm) and uses the Dirac δ-function to construct a simple explicit form. In this way one avoids the segmentation of the unit hypersphere, and one arrives at a concise and global description of the quaternion probability density in which the geometrical properties of the probability density are manifest. These studies of the pdf for the four-component quaternion lead in Part II of this work to a very simple formula for the probability density function for any uniformly distributed three-component representation of the attitude.

We preface this work with a section on mathematical preliminaries. The section is intended largely to introduce notation and to refresh the reader's memory of multivariate calculus, probability distributions, and groups. This material is not presented with mathematical rigor, but the statements made there are, nonetheless, rigorously true.

MATHEMATICAL PRELIMINARIES

Univariate Probability Distributions

We define the *probability function*[1] of a one-dimensional random variable x as

$$\mathsf{P}_x(x') \equiv \text{probability that } x < x' \equiv \text{Prob}\big(\{x < x'\}\big)\,. \tag{5}$$

Here x denotes the random variable and x' a possible value that can be realized by this random variable (the realization of x). We assume that $\mathsf{P}_x(x')$ is defined on some continuous interval, either limit of which may be infinite. If we write the interval formally as $[x_i, x_f]$, then we must have $\mathsf{P}(x_i) = 0$ and, by convention, $\mathsf{P}_x(x_f) = 1$.

If now $y(x)$ is a continuous, strictly increasing function of x over the entire domain of $y(x)$, then, defining similarly

$$\mathsf{P}_y(y') \equiv \text{Prob}\big(\{y < y'\}\big)\,, \tag{6}$$

we must have

$$\mathsf{P}_y\big(y(x')\big) \equiv \text{Prob}\big(\{y < y(x')\}\big) = \text{Prob}\big(\{x < x'\}\big) = \mathsf{P}_x\big(x'\big)\,, \tag{7}$$

so that a probability function is a function in the usual sense.* Equation (7) requires that $y(x)$ be strictly increasing. Had we chosen $y(x) = -x$, for example, we would have had $\mathsf{P}_x(x') = 1 - \mathsf{P}_y\big(y(x')\big)$.

If the probability functions for x and y are differentiable as well, we can write

$$\mathsf{P}_x(x') = \int_{x_i}^{x'} p_x(x'')\,dx''\,, \quad \mathsf{P}_y(y') = \int_{y_i}^{y'} p_y(y'')\,dy''\,, \tag{8}$$

*Note that the symbol y is doing triple duty: (1) as a random variable, (2) as a function $y(x)$, and (3) as a label for the probability and probability density functions.

with $p_x(x')$ and $p_y(y')$ the respective *probability density functions* (pdf), which we take to be piecewise continuous functions of their respective arguments

$$p_x(x') = \frac{d\mathsf{P}_x(x')}{dx'}\,, \quad p_y(y') = \frac{d\mathsf{P}_y(y')}{dy'}\,, \tag{9}$$

Let us suppose further that $y(x)$ is, in addition, a differentiable function of x. Then, in obvious notation,

$$\int_{x_1}^{x_2} p_x(x')\,dx' = \int_{y_1}^{y_2} p_y(y')\,dy' = \int_{x_1}^{x_2} p_y\big((y(x')\big) \left(\frac{dy(x')}{dx'}\right) dx'\,, \tag{10}$$

from which it follows that

$$p_x(x') = p_y\big(y(x')\big) \left(\frac{dy(x')}{dx'}\right)\,. \tag{11}$$

We call such a function a *density*. The probability function is a function; the probability density function is a density.

Discontinuous Probabilities

What if $\mathsf{P}_x(x')$ is not differentiable almost everywhere and not even continuous? Suppose, for example, that $\mathsf{P}_x(x')$ is differentiable at every point of $[x_i,\,x_f]$ except at the point x_1 at which it has a finite discontinuity, i.e.,

$$\lim_{\substack{x'\to x_1\\ x'<x_1}} \mathsf{P}_x(x') = \mathsf{P}_x(x_1^-)\,, \qquad \lim_{\substack{x'\to x_1\\ x'>x_1}} \mathsf{P}_x(x') = \mathsf{P}_x(x_1^+)\,, \tag{12}$$

with $\mathsf{P}_x(x_1^-) < \mathsf{P}_x(x_1^+)$. Then the probability of x realizing exactly the value x_1 is

$$\mathrm{Prob}(\{x = x_1\}) = \lim_{\epsilon\to 0^+}\big(\mathsf{P}_x(x_1+\epsilon) - \mathsf{P}_x(x_1-\epsilon)\big) = \mathsf{P}_x(x_1^+) - \mathsf{P}_x(x_1^-) > 0\,. \tag{13}$$

The point x_1 is said to be a point of concentration, because there is a finite probability that $x' = x_1$ exactly.

If there is a point of concentration, a pdf cannot be written for the probability function $\mathsf{P}_x(x')$ in the usual way. We can, however, write the pdf for $\mathsf{P}_x(x')$ formally as

$$p_x(x') = \overline{p_x(x')} + \big(\mathsf{P}_x(x_1^+) - \mathsf{P}_x(x_1^-)\big)\;\delta(x - x_1)\,, \tag{14}$$

where $\overline{p_x(x')}$ is a finite piecewise-continuous function defined by

$$\overline{p_x(x')} \equiv \begin{cases} \dfrac{d\mathsf{P}_x(x')}{dx'} & \text{for} \quad x' \neq x_1 \\ 0 & \text{for} \quad x' = x_1 \end{cases}\,, \tag{15}$$

and $\delta(x)$ is the Dirac δ-function, which satisfies

$$\delta(x) = 0 \quad \text{for} \quad x \neq 0\,, \tag{16a}$$

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1\,, \tag{16b}$$

$$f(x)\,\delta(x) = f(0)\,\delta(x)\,, \tag{16c}$$

$$\int_{-\infty}^{\infty} f(x)\,\delta^{(n)}(x)\,dx = (-1)^n \int_{-\infty}^{\infty} \frac{d^n f(x)}{dx^n}\,\delta(x)\,dx\,, \tag{16d}$$

$$\int_{-\infty}^{\infty} f(x)\,\delta\big(g(x)\big)\,dx = \sum_i \int_{-\infty}^{\infty} f(x)\,\left|\frac{dg}{dx}(x_i)\right|^{-1} \delta(x - x_i)\,dx\,, \tag{16e}$$

where $\delta^{(n)}(x)$ denotes the n-th derivative of $\delta(x)$, and the x_i are the roots of $g(x)$.

Equations (16d) and (16e) state that we can treat the δ-function like an ordinary density function with respect to a change of variable or to integration by parts. That would follow if we wrote

$$\delta(x) = \lim_{\sigma\to 0^+} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-x^2/2\sigma^2} \equiv \lim_{\sigma\to 0^+} \delta_\sigma(x)\,. \tag{17}$$

Equations (16d) and (16e) will hold for $\delta_\sigma(x)$ for every positive value of σ. Equations (16c), (16d), and (16e) will then hold in the limit that $\sigma \to 0^+$ provided that: (*i*) $f(x)$ is continuous at $x = 0$ (for Eq. (16c)); (*ii*) $d^n f/dx^n$ is continuous at $x = 0$ (for Eq. (16d)); (*iii*) $g(x)$ is differentiable and $f(x)$ is continuous at each of the x_i (for Eq. (16e)). Note that because of Eq. (16e) the Dirac δ-function, is, in fact, a density. The Dirac δ-function is not a function in the usual sense, but only a formal shorthand for the presence of a discontinuity in the probability function and, accordingly, a point of concentration.

Multivariate Probability Distributions and the Implicit Function Theorem

For the multivariate case in n dimensions the probability density is defined by

$$\mathsf{P}_{\mathbf{x}}(\mathbf{x}') \equiv \text{Prob}\left(\{x_1 < x_1', \ldots, x_n < x_n'\}\right)\,, \tag{18}$$

with $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ and the multivariate pdf, assuming that $\mathsf{P}_{\mathbf{x}}(\mathbf{x}')$ is sufficiently differentiable, is defined by

$$p_{\mathbf{x}}(\mathbf{x}') \equiv \frac{\partial}{\partial x_1'} \frac{\partial}{\partial x_2'} \cdots \frac{\partial}{\partial x_n'} \mathsf{P}_{\mathbf{x}}(\mathbf{x}')\,, \tag{19}$$

whence

$$\mathsf{P}_{\mathbf{x}}(\mathbf{x}') = \int_{x_{i,1}}^{x_1'} \int_{x_{i,2}}^{x_2'} \cdots \int_{x_{i,n}}^{x_n'} p_{\mathbf{x}}(\mathbf{x}'')\, dx_1''\, dx_2'' \ldots dx_n''\,, \tag{20}$$

and

$$\int_{x_{i,1}}^{x_{f,1}} \int_{x_{i,2}}^{x_{f,2}} \cdots \int_{x_{i,n}}^{x_{f,n}} p_{\mathbf{x}}(\mathbf{x}'')\, dx_1''\, dx_2'' \ldots dx_n'' = 1\,, \tag{21}$$

assuming that the domain of $p_x(x')$ is rectangular. As in the one-dimensional case, in n-dimensions the probability function can have discontinuities. These can be at a point, on a curve, on a surface, or on a hypersurface. The situation is much more complicated than on the real line. We will encounter an example of a probability discontinuity on a hyperspherical surface in our treatment of the pdf for the four components of the quaternion.

In n dimensions the rule for a change of variables is provided by the implicit function theorem,[2,3] which may be stated for the present purposes as follows: Let the multivariate function $\mathbf{y}(\mathbf{x}) = [y_1(x_1, \ldots, x_n), \ldots, y_n(x_1, \ldots, x_n)]^T$ be such that for every $\mathbf{x}'$ in the domain of $\mathbf{y}(\mathbf{x})$ the Jacobian matrix

$$\left[\frac{\partial \mathbf{y}'}{\partial \mathbf{x}'}\right]_{ij} = \left[\frac{\partial(y_1', y_2', \ldots, y_n')}{\partial(x_1', x_2', \ldots, x_n')}\right]_{ij} \equiv \frac{\partial y_i'}{\partial x_j'} \tag{22}$$

exists and has non-zero determinant everywhere in the domain of $\mathbf{y}(\mathbf{x})$. If $f(\mathbf{y}')$ is a piecewise-continuous function, then the implicit function theorem tells us that the integral over $\mathbf{y}'$ can be transformed to an integral over $\mathbf{x}'$ according to

$$\int_{\mathbf{Y}} f(\mathbf{y}')\, d^n\mathbf{y}' = \int_{\mathbf{X}} f\big(\mathbf{y}(\mathbf{x}')\big) \left| \frac{\partial \mathbf{y}(\mathbf{x}')}{\partial \mathbf{x}'} \right| d^n\mathbf{x}'\,, \tag{23}$$

where $|\partial\mathbf{y}'(\mathbf{x}')/\partial\mathbf{x}'|$ denotes the absolute value of the determinant of the Jacobian matrix. The absolute value of the determinant of the Jacobian matrix (henceforth, for brevity: the Jacobian determinant) must be used lest the sign of the integral depend on the ordering of the elements of $\mathbf{x}$ and $\mathbf{y}$. Here $\mathbf{X}$ and $\mathbf{Y}$ are the domains of integration with $\mathbf{Y}$ the image of $\mathbf{X}$ or, equivalently, $\mathbf{X}$ the pre-image of $\mathbf{Y}$. Except in the case that every component of $\mathbf{y}(\mathbf{x})$ is a monotonically increasing function of every component of $\mathbf{x}$, the multivariate probability function will not satisfy a relation analogous to Eq. (11). Thus, in general,

$$\mathsf{P}_{\mathbf{x}}(\mathbf{x}') \neq \mathsf{P}_{\mathbf{y}}\big(\mathbf{y}(\mathbf{x}')\big)\,, \tag{24}$$

One avoids this problem by defining the multivariate probability function not as a function of coordinates as in Eq. (18) but as a function of sets in general. If $\mathbf{X}$ is a subset of the domain of $\mathbf{y}(\mathbf{x})$, say a multivariate interval, and $\mathbf{y}(\mathbf{X})$ is its image in the range of $\mathbf{y}(\mathbf{x})$, then

$$\mathsf{P}_{\mathbf{x}}(\mathbf{X}) = \mathsf{P}_{\mathbf{y}}\big(\mathbf{y}(\mathbf{X})\big)\,. \tag{25}$$

This idea was already implicit in Eqs. (5) and (18). The multivariate probability density function is a density in the previous sense, since it is defined at a point

$$p_{\mathbf{x}}(\mathbf{x}') = p_{\mathbf{y}}\big(\mathbf{y}(\mathbf{x}')\big)\,\left|\frac{\partial \mathbf{y}(\mathbf{x}')}{\partial \mathbf{x}'}\right|\,. \tag{26}$$

The Rotation Group

Every attitude representation forms a group.[4] A *group* $\boldsymbol{G} = \{\,G,\,\circ\,\}$ consists of a set of objects $G = \{\,\alpha,\,\beta,\,\ldots\,\}$ and an operation $\circ$ which satisfy the following conditions:

(1) For any two α and β in G, $\alpha\circ\beta$ is also in G.

(2) The operation is associative

$$\alpha\circ(\beta\circ\gamma) = (\alpha\circ\beta)\circ\gamma\,. \tag{27}$$

(3) There exists an identity element ι in G which satisfies

$$\iota\circ\alpha = \alpha\circ\iota = \alpha \tag{28}$$

for every α in G.

(4) For every α in G there exits an inverse element α^{-1} such that

$$\alpha\circ\alpha^{-1} = \alpha^{-1}\circ\alpha = \iota\,. \tag{29}$$

These properties allow us to write and solve algebraic equations for the attitude representations in a general manner without having to reduce them to numerical form.

For the rotation matrix or direction-cosine matrix (DCM), G is the set of all proper orthogonal 3×3 matrices; the group operation is matrix multiplication; the identity element is the 3×3 identity matrix; and the inverse of a DCM is just its transpose. For the attitude quaternion, G is the set of all 4×1 arrays with unit norm; the group operation is quaternion composition;[5] the identity element is the quaternion $[0,\,0,\,0,\,1]^T$; and the inverse of $\bar{\eta}$ is $\bar{\eta}^*$ with

$$\bar{\eta}^* \equiv \begin{bmatrix} -\boldsymbol{\eta} \\ \eta_4 \end{bmatrix}\,. \tag{30}$$

Properties like these hold true for every representation of the attitude, although in some cases, as for the Euler angles, the nature of the group property may not be very transparent.[6] What enters into our calculation of the pdf's for uniformly distributed representations of random attitude is the analytic form of the group operation.

UNIFORM ATTITUDE PROBABILITY DENSITIES

A uniform attitude probability density function[1] is one which makes all attitudes "equally" probable. To understand what this means, let the random attitude representation be denoted by ξ. As before, we write the pdf as $p_{\xi}(\xi')$ where ξ' denotes the value of a possible realization of ξ. Note first that a uniformly distributed random attitude cannot mean that in general

$$p_{\xi}(\xi') = p_{\xi}(\xi'') \tag{31}$$

for arbitrary values ξ' and ξ'' of the attitude representation, which would say that the pdf was a constant function of the attitude representation. Equation (26) shows that this cannot be true in general, because the pdf is a density. Hence, if the pdf were a constant function for one choice of the attitude representation, on performing a change of variable to a different choice of the attitude representation, it would cease to be a constant function because of the factor of the Jacobian determinant.

An attitude representation by definition is the representation of a rotation. Thus, if ξ is the representation of the uniformly distributed random attitude from the space coordinate frame to the body coordinate frame, then $\xi \circ \chi$ is the representation of the rotation from a new space frame to the body frame, with the representation of the (constant) rotation from the old to the new space frame being χ^{-1}. Likewise, if ζ^{-1} is the the representation of the (constant) rotation carrying the old body axes to new body axes, then the random representation of the attitude from the old space axes to the new body axes becomes $\zeta \circ \xi$. Thus, for the pdf of the attitude to provide no information we must have

$$p_{\xi}(\xi') = p_{\xi\circ\chi}(\xi') = p_{\zeta\circ\xi}(\xi') = p_{\zeta\circ\xi\circ\chi}(\xi'), \tag{32}$$

for all values of ξ' no matter what the choice of χ and ζ, so long as they have fixed values. The rightmost member of Eq. (32) is the most general statement of the invariance but will not be of practical use in the development of this work.

Practitioners of maximum-likelihood estimation[7] will immediately recognize the significance of Eq. (32). The representations χ and ζ play the role of parameters of the pdf. Eq. (32) thus says that the likelihood function is independent of the parameters to be estimated, and hence the parameters cannot be estimated from the likelihood.

The group properties of the attitude have thus determined the nature of the invariance which characterizes a uniformly distributed random attitude. To exploit the implicit function theorem further, however, requires that the discussion be specialized to three-parameter representations of the attitude. This is because the Jacobian determinant will be indeterminate or vanish (both outcomes extremely unpleasant) unless all of the variables in its definition are independent. Since there are only three independent attitude parameters, we must set aside for the moment, any application of our results thus far to the four-component quaternion or the DCM, which have four and nine parameters, respectively.

We now exploit the implicit function theorem once more. Changing the random variable from $\xi \circ \chi$ in the second member of Eq. (32) to ξ is equivalent to changing the argument of the pdf from ξ' to $\xi' \circ \chi^{-1}$, and similarly for the third member. Applying the implicit function theorem to Eq. (32) then leads to

$$p_{\xi}(\xi') = p_{\xi\circ\chi}(\xi') = p_{\xi}(\xi' \circ \chi^{-1}) \left| \frac{\partial(\xi' \circ \chi^{-1})}{\partial \xi'} \right| , \tag{33a}$$

and

$$p_\xi(\xi') = p_{\zeta\circ\xi}(\xi') = p_\xi(\zeta^{-1}\circ\xi') \left|\frac{\partial(\zeta^{-1}\circ\xi')}{\partial\xi'}\right| . \tag{33b}$$

Equations (33) are remarkable in that they tell us that for a uniformly distributed random attitude, the value of the pdf for one value of the attitude representation is related to the value of the pdf for a different value of the representation by the Jacobian determinant. Thus, if the value of the pdf is known for one value of the attitude representation, the value of the pdf for any other value of the attitude representations can be obtained by choosing χ or ζ appropriately. The group properties of the attitude representations assure that an appropriate χ or ζ can always be found. For simplicity, let us chose that initial value of the attitude representation to be ι, the identity element of representation (corresponding to $\mathbf{0}$ for the vector components of the quaternion, the rotation vector, the Rodrigues parameters, and the modified Rodrigues parameters, to $I_{3\times3}$ for the direction-cosine matrix, or to $[0, 0, 0, 1]^T$ for the quaternion as discussed in Ref. 5). Then Eqs. (32) can be transformed to

$$p_\xi(\xi') = p_\xi(\iota) \left|\frac{\partial\boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha}\circ\xi')}\right|_{\alpha=\iota} = p_\xi(\iota) \left|\frac{\partial\boldsymbol{\alpha}}{\partial(\xi'\circ\boldsymbol{\alpha})}\right|_{\alpha=\iota} . \tag{34}$$

To obtain Eq. (34) we make the substitution in Eq. (33a) of $\boldsymbol{\alpha} = \xi'\circ\chi^{-1}$, which implies $\xi' = \boldsymbol{\alpha}\circ\chi$. Equation (33a) then becomes

$$p_\xi(\boldsymbol{\alpha}\circ\chi) = p_\xi(\boldsymbol{\alpha}) \left|\frac{\partial\boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha}\circ\chi)}\right| . \tag{35}$$

Note that we are not carrying out a change of variable, so there is no additional factor of a Jacobian determinant. We are simply expressing the same numerical value in a different manner. The numerical values of the respective arguments of Eq. (33a) and Eq. (35) are identical.

Now let χ, whose value is arbitrary, have the value ξ' yielding

$$p_\xi(\boldsymbol{\alpha}\circ\xi') = p_\xi(\boldsymbol{\alpha}) \left|\frac{\partial\boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha}\circ\xi')}\right| . \tag{36}$$

Taking the limit that $\boldsymbol{\alpha}\to\iota$ then yields the first relation in Eq. (34) provided that the limits of the two factors in the right member exist separately and $p_\xi(\iota)$ not vanish. When this condition is violated one must either rely on Eq. (36) for values of $\boldsymbol{\alpha}$ different from ι, or else apply L'Hôpital's rule as $\boldsymbol{\alpha}\to\iota$.

Equation (34) implies that the pdf can be uniform in the strict sense, i.e., a constant function, only if the Jacobian determinant is a constant. This can happen only if the composition rule for the representation is simple componentwise addition. This is true for no attitude representation globally. Hence, it is not expected that the pdf will be a constant over all space for any attitude representation. Note that translations add, so that one expect the uniform pdf for position to be a constant. Also, the components of the rotation vector for a small rotation add, so that we should expect the pdf of that representation to be constant to first order near the origin.

Equation (34) shows clearly that the entire pdf is determined from the implicit function theorem (which provides the factor of the Jacobian determinant) and the group property of attitude representation, which enters into the computation of the Jacobian determinant. The value of the pdf at ι need not be known *ab initio*, since it is determined completely from the condition that the total probability be unity. Thus

$$1 = \int p_\xi(\xi')\, d^3\xi' = \left[\int_\Xi p_\xi(\iota) \left|\frac{\partial\boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha}\circ\xi')}\right|_{\alpha=\iota} d^3\xi'\right] = 1 , \tag{37}$$

and, therefore,

$$p_{\xi}(\iota) = \left[\int\!\!\int_{\Xi} \left| \frac{\partial \boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha} \circ \boldsymbol{\xi}')} \right|_{\boldsymbol{\alpha}=\iota} d^3\xi' \right]^{-1} . \tag{38}$$

and Ξ is the entire domain of the pdf.

Equations (26), (32), and their consequential results, Eqs. (34), and (38) are the cornerstones of the present work and its sequel.

THE UNIFORM PDF FOR VECTOR COMPONENTS OF THE QUATERNION

The pdf for the vector components of the quaternion are obtained directly from Eqs. (34) and (38). The quaternion, we note, is related to the axis and angle of rotation by[5]

$$\bar{\eta} = \begin{bmatrix} \boldsymbol{\eta} \\ \eta_4 \end{bmatrix} = \begin{bmatrix} \sin(\theta/2)\,\hat{\mathbf{n}} \\ \cos(\theta/2) \end{bmatrix} , \tag{39}$$

with θ the angle of rotation and $\hat{\mathbf{n}}$ the axis of rotation, a unit vector. The quaternion composition rule for the vector components alone under the constraint that the scalar component be positive is[5]

$$\boldsymbol{\alpha} \circ \boldsymbol{\eta}' = \mathrm{sgn}(\boldsymbol{\alpha}, \boldsymbol{\eta}') \left(\alpha_4 \boldsymbol{\eta}' + \eta'_4 \boldsymbol{\alpha} - \boldsymbol{\alpha} \times \boldsymbol{\eta}' \right) , \tag{40}$$

where α_4 and $\boldsymbol{\eta}'_4$ are here shorthand for $+\sqrt{1 - |\boldsymbol{\alpha}|^2}$ and $+\sqrt{1 - |\boldsymbol{\eta}'|^2}$, respectively, and

$$\mathrm{sgn}(\boldsymbol{\alpha}, \boldsymbol{\eta}') = \mathrm{sign}\left(\alpha_4 \eta'_4 - \boldsymbol{\alpha} \cdot \boldsymbol{\eta}' \right) . \tag{41}$$

Equation (34) for $\eta'_4 \neq 0$ leads to

$$\left| \frac{\partial \boldsymbol{\alpha}}{\partial(\boldsymbol{\alpha} \circ \boldsymbol{\eta}')} \right|_{\boldsymbol{\alpha}=\mathbf{0}} = \left| \frac{\partial(\boldsymbol{\alpha} \circ \boldsymbol{\eta}')}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\mathbf{0}}^{-1} = \frac{1}{\eta'_4} = \frac{1}{\sqrt{1 - |\boldsymbol{\eta}'|^2}} , \tag{42}$$

and from Eq. (38), $p_{\boldsymbol{\eta}}(\mathbf{0}) = 1/\pi^2$. Hence,

$$p_{\boldsymbol{\eta}}(\boldsymbol{\eta}') = \frac{1}{\pi^2 \sqrt{1 - |\boldsymbol{\eta}'|^2}} . \tag{43}$$

The author derived this result for the first time in 1993, when the problem was first posed to him by Markley. Note that the pdf is constant to first order in $|\boldsymbol{\eta}|$ near $\boldsymbol{\eta} = 0$ as predicted by the discussion following Eq. (34).

We emphasize that it is only through the nature of the composition rule that any information about the particular attitude representation enters the calculation of the pdf. Equation (39), although a helpful reminder, plays no role in the derivation of Eq. (43).

The pdf for the quaternion as a four-component object is of particular interest, but its derivation will be delayed until later in this work because of the complexities arising from the four-dimensionality.

THE UNIFORM PDF FOR THE FOUR-COMPONENT QUATERNION: A PARTITIONED APPROACH

We present here a mathematically rigorous derivation of the pdf of the quaternion on the whole of S^3, the unit 3-sphere.* Previously, we determined $p_{\boldsymbol{\eta}}(\boldsymbol{\eta}')$ restricted to the open hemihypersphere $\eta'_4 > 0$. Let us define now eight open hemihyperspheres $H(i, \kappa)$ according to

$$H(i, \kappa) = \begin{cases} \{ \bar{\eta}' \,|\, \eta'_i > 0 \} , & \text{for } i = 1, 2, 3, 4, \quad \kappa = 1 \\ \{ \bar{\eta}' \,|\, \eta'_i < 0 \} , & \text{for } i = 1, 2, 3, 4, \quad \kappa = 2 \end{cases} \tag{44}$$

*S^2, the unit 2-sphere, is the familiar spherical surface in three-dimensional space (the surface of the solid unit 3-ball), S^1 is the unit circle.

Almost every point in S^3 belongs to four open hemihyperspheres, and every point of S^3 belongs to at least one open hemihypersphere. (All but eight points in S^3 belong to at least two open hemihyperspheres.) Thus, these eight open hemihyperspheres constitute a finite open covering of S^3.

Let us define now

$$\boldsymbol{\eta}(1) \equiv \begin{bmatrix} \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}, \qquad \boldsymbol{\eta}(2) \equiv \begin{bmatrix} \eta_1 \\ \eta_3 \\ \eta_4 \end{bmatrix}, \qquad \boldsymbol{\eta}(3) \equiv \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_4 \end{bmatrix} \tag{45}$$

(hence, $\boldsymbol{\eta} = \boldsymbol{\eta}(4)$). It follows trivially that for every $\bar{\eta} \in H(i, \kappa)$

$$p_{\boldsymbol{\eta}(i),\kappa}\big(\boldsymbol{\eta}'(i)\big) = \frac{C(i, \kappa)}{|\eta_i'|}, \qquad i = 1, 2, 3, 4, \quad \kappa = 1, 2, \tag{46}$$

for some constant $C(i, \kappa)$. Since the probability density function at every point of S^3 is finite when written in terms of the appropriate coordinates, it follows that the uniform distribution of $\bar{\eta}$ on S^3 cannot have a point of concentration.

We note further that wherever it is finite

$$\left| \frac{\partial \boldsymbol{\eta}'(i)}{\partial \boldsymbol{\eta}'(j)} \right| = \frac{|\eta_i'|}{|\eta_j'|}, \qquad i = 1, 2, 3, 4. \tag{47}$$

Hence, if $\bar{\alpha}$ belongs to both $H(i, \kappa)$ and $H(j, \lambda)$, then

$$\frac{C(i, \kappa)}{|\alpha_i|} = p_{\boldsymbol{\eta}(i),\kappa}\big(\boldsymbol{\alpha}(i)\big) = p_{\boldsymbol{\eta}(j),\lambda}\big(\boldsymbol{\alpha}(j)\big) \left| \frac{\partial \boldsymbol{\alpha}(j)}{\partial \boldsymbol{\alpha}(i)} \right| = \frac{C(j, \lambda)}{|\alpha_j|} \frac{|\alpha_j|}{|\alpha_i|} = \frac{C(j, \lambda)}{|\alpha_i|}. \tag{48}$$

It follows that $C(i, \kappa) = C(j, \lambda)$ for any two hemihyperspheres. Therefore, for any point of S^3 for which $\eta_i' \neq 0$,

$$p_{\boldsymbol{\eta}(i),\kappa}\big(\boldsymbol{\eta}'(i)\big) = \frac{C}{|\eta_i'|}, \qquad i = 1, 2, 3, 4, \quad \kappa = 1, 2, \tag{49}$$

for a common constant C. Since there are no points of concentration on S^3 and $S^3 = H(4, 1) \cup H(4, 2) \cup \partial H(4, 1)$, with $\partial H(4, 1)$, the boundary of $H(4, 1)$, a set of measure zero, we can determine C from

$$1 = \int\limits_{H(4,1)\cup H(4,2)} p_{\boldsymbol{\eta}}(\boldsymbol{\eta}')\, d^3\boldsymbol{\eta}' = 2C \int\limits_{|\boldsymbol{\eta}'|<1} \frac{d^3\boldsymbol{\eta}'}{\sqrt{1 - |\boldsymbol{\eta}'|^2}} = 2\pi^2 C, \tag{50}$$

whence

$$p_{\boldsymbol{\eta}(i),\kappa}\big(\boldsymbol{\eta}'(i)\big) = \frac{1}{2\pi^2 |\eta_i'|}, \qquad i = 1, 2, 3, 4, \quad \kappa = 1, 2, \tag{51}$$

everywhere in S^3 that $n_i' \neq 0$.

We can describe the probability density function for the quaternion as a single function of the quaternion defined on all of S^3 and avoid the changes of variables associated with Eq. (51). To accomplish this task we write the general quaternion as a function of the vector components of the Euler-Rodrigues symmetric parameters (the unit quaternion) and the quaternion length as

$$\bar{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} q\,\eta_1 \\ q\,\eta_2 \\ q\,\eta_3 \\ q\sqrt{1 - |\boldsymbol{\eta}|^2} \end{bmatrix}. \tag{52}$$

The Jacobian determinant of the transformation from Cartesian quaternion coordinates to $(q, \eta_1 . \eta_2, \eta_3)$ is given by

$$\left| \frac{\partial(q_1, q_2, q_3, q_4)}{\partial(q, \eta_1, \eta_2, \eta_3)} \right| = \frac{q^3}{|\eta_4|} , \tag{53}$$

and, therefore,

$$d^4\bar{q} \equiv dq_1\, dq_2\, dq_3\, dq_4 = \frac{q^3}{|\eta_4|}\, dq\, d^3\boldsymbol{\eta} \equiv q^3\, dq\, d^3\boldsymbol{\sigma} , \tag{54}$$

which implies that

$$d^3\boldsymbol{\sigma} = \frac{1}{|\eta_4|}\, d^3\boldsymbol{\eta} \tag{55}$$

is the invariant "hyper-area" element on S^3, which should be called more properly the invariant "volume" element.

From this it follows that

$$d^3\mathrm{P}_{\bar{\eta}}(\bar{\eta}') = p_{\boldsymbol{\eta}(i),\,\kappa}\big(\boldsymbol{\eta}'(i)\big)\, d^3\boldsymbol{\eta}'(i) = \frac{1}{2\pi^2}\, d^3\boldsymbol{\sigma}' , \quad i = 1, 2, 3, 4 , \quad \kappa = 1, 2 . \tag{56}$$

Equation (56) shows that $1/2\pi^2$ is the invariant hypersurface probability density on S^3. Note again that the value of the probability density is $1/2\pi^2$ and not $1/\pi^2$, because we consider the entire hypersphere and not just a hemi-hypersphere on which each attitude corresponds to a single quaternion. The uniform pdf of the quaternion in four dimensions has a three-sphere of concentration.

THE UNIFORM PDF FOR THE FOUR-COMPONENT QUATERNION: A GLOBAL APPROACH

There is a very different way to develop and understand the results of the previous section. Instead of a pdf defined on S^3, let us consider a scalar function $f(\bar{q})$ defined in four-dimensional Euclidean space which has the same invariance property *mutatis mutandis* as the pdf, namely

$$f(\bar{q}') = f(\bar{p} \otimes \bar{q}') \left| \frac{\partial(\bar{p} \otimes \bar{q}')}{\partial \bar{q}'} \right| , \tag{57}$$

for all unit quaternions $\bar{p}$. Thus, $\bar{p}$ is a quaternion of rotation, while $\bar{q}'$ can have arbitrary norm. We may now treat $\bar{q}'$ and $\bar{p}$ as four-component objects, for which the composition rule is much simpler than for a subset of three components of $\bar{q}'$. Note that the transformation induced by $\bar{p}$ on $\bar{q}'$ can change only its position on a 3-sphere of constant radius, so that we are examining equivalently the same transformation as in Eqs. (32) but simultaneously on concentric hyperspheres rather than only on the unit hypersphere S^3.

Following the conventions of Ref. 5,

$$\bar{p} \otimes \bar{q}' = \{\, \bar{p} \,\}_L\, \bar{q}' , \tag{58}$$

with

$$\{\, \bar{p} \,\}_L \equiv \begin{bmatrix} p_4 & p_3 & -p_2 & p_1 \\ -p_3 & p_4 & p_1 & p_2 \\ p_2 & -p_1 & p_4 & p_3 \\ -p_1 & -p_2 & -p_3 & p_4 \end{bmatrix} . \tag{59}$$

It follows, then, that

$$\left| \frac{\partial(\bar{p} \otimes \bar{q}')}{\partial \bar{q}'} \right| = \left| \det \{\, \bar{p} \,\}_L \right| = 1 , \tag{60}$$

because $\{\bar{p}\}_L$ is a proper orthogonal matrix for any unit quaternion $\bar{p}$. Thus, we are led to search for functions of a general quaternion $\bar{q}$ which satisfy

$$f(\bar{q}') = f(\bar{p} \otimes \bar{q}') \tag{61}$$

for every unit quaternion $\bar{p}$.

We now restrict $\bar{q}'$ to have unit length and let $f(\bar{q}') = p_{\bar{q}}(\bar{q}')$. Because the unit quaternions form a group under quaternion multiplication, we are assured that for every value $\bar{q}''$ of the unit quaternion there exists a value of $\bar{p}$ such that $\bar{p} \otimes \bar{q}' = \bar{q}''$. It follows, therefore, that

$$p_{\bar{q}}(\bar{q}') = p_{\bar{q}}(\bar{q}'') . \tag{62}$$

The uniform pdf of the unit quaternion is a constant function of $\bar{q}'$.

Since the quaternion of rotation must have unit length, it follows that the quaternion pdf generalized to a function over all of R^4 must have the form

$$p_{\bar{q}}(\bar{q}') = c\,\delta(\bar{q}'^T \bar{q}' - 1) , \tag{63}$$

where c is a constant and $\delta(x)$ is the Dirac delta-function. The delta-function in Eq. (63) restricts the support of $p_{\bar{q}}(\bar{q}')$ defined on R^4 to S^3.

The constant c is determined again from the condition that the total probability be unity. Thus,

$$\begin{aligned} 1 &= c \int \delta(|\mathbf{q}'|^2 + q_4'^2 - 1)\, dq_4'\, d^3\mathbf{q}' \\ &= c \int_{|\mathbf{q}'|^2 \le 1} \Bigg[\int_{-\infty}^{\infty} \Bigg\{ \frac{1}{|2q_4'|}\, \delta(q_4' - \sqrt{1 - |\mathbf{q}'|^2}) \\ &\qquad\qquad + \frac{1}{|2q_4'|}\, \delta(q_4' + \sqrt{1 - |\mathbf{q}'|^2}) \Bigg\}\, dq_4' \Bigg]\, d^3\mathbf{q}' \\ &= c \int_{|\mathbf{q}'|^2 \le 1} \frac{1}{\sqrt{1 - |\mathbf{q}'|^2}}\, d^3\mathbf{q}' = \pi^2 c \quad . \end{aligned} \tag{64}$$

Therefore, $c = 1/\pi^2$ and

$$p_{\bar{q}}(\bar{q}') = \frac{1}{\pi^2}\, \delta(\bar{q}'^T \bar{q} - 1) , \tag{65}$$

and, consequently, from the penultimate line of Eq. (64)

$$p_{\boldsymbol{\eta}}(\boldsymbol{\eta}') = \int_{-\infty}^{\infty} p_{\bar{q}}(\bar{q}')\, dq_4 \bigg|_{\mathbf{q}=\boldsymbol{\eta}} = \frac{1}{\pi^2 \sqrt{1 - |\boldsymbol{\eta}'|^2}} , \tag{66}$$

which is the same as Eq. (43) above. Note again that Eq. (65) takes the domain of the pdf to be all of S^3 and not only a single hemihypersphere. Equation (63) could also have been written equivalently as

$$p_{\bar{q}}(\bar{q}') = \frac{1}{2\pi^2}\, \delta(\sqrt{\bar{q}'^T \bar{q}} - 1) = \frac{1}{2\pi^2}\, \delta(q - 1) . \tag{67}$$

Equations (63) and (65) are equivalent to Eq. (56).

THE UNIFORM PDF FOR A RIGID BODY

For a uniform rigid body we must consider the pdf not only of the attitude but also of the angular velocity. In terms of the attitude matrix and the angular velocity, the composition rule is

$$(A_3, \boldsymbol{\omega}_3) = (A_2, \boldsymbol{\omega}_2) \circ (A_1, \boldsymbol{\omega}_1)(A_2 A_1, \boldsymbol{\omega}_2 + A_2 \boldsymbol{\omega}_1) . \tag{68}$$

Thus, if we write the 6×1 state vector ξ for the rigid body as

$$\xi = \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\omega} \end{bmatrix} , \tag{69}$$

we obtain the Jacobian determinant

$$\left| \frac{\partial(\alpha \circ \xi')}{\partial \alpha} \right|_{\alpha=\iota} = \begin{vmatrix} \eta'_4 I_{3\times3} + [[\boldsymbol{\eta}']] & 0_{3\times3} \\ -2[[\boldsymbol{\omega}']] & I_{3\times3} \end{vmatrix} = \left| \eta'_4 I_{3\times3} + [[\boldsymbol{\eta}']] \right| = \eta'_4 . \tag{70}$$

Here $[[\mathbf{v}]]$ denotes the matrix

$$[[\mathbf{v}]] \equiv \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix} . \tag{71}$$

The pdf is uniform in $\boldsymbol{\omega}$ in the strict sense. Thus,

$$p_{\eta,\boldsymbol{\omega}}(\boldsymbol{\eta}', \boldsymbol{\omega}') = \frac{p_{\eta,\boldsymbol{\omega}}(\mathbf{0}, \mathbf{0})}{\sqrt{1 - |\boldsymbol{\eta}'|^2}} = p_\eta(\boldsymbol{\eta}') p_{\boldsymbol{\omega}}(\mathbf{0}) . \tag{72}$$

Clearly, $p_{\boldsymbol{\omega}}(\mathbf{0}) = 0$, since the range of $\boldsymbol{\omega}$ is infinite, so that we cannot write a pdf for $\boldsymbol{\omega}$ as easily as we could for $\boldsymbol{\eta}$, the problem with all functions which are constant on infinite intervals. This problem occurs because of topological differences between the rotation and translation groups. The former is compact, while the latter is not.

DISCUSSION

We have developed the theory of the uniform probability density function for the quaternion fairly completely. In part II of the paper,[8] we shall use these results to determine in a very simple manner, the uniform pdf's of all of the other common representations.

ACKNOWLEDGMENT

The author is grateful to Dr. F. Landis Markley for making him aware of the ball-in-the-box algorithm and for posing seven years ago the casual question which started the author on this work. The author is very grateful to Dr. Hanspeter Schaub, Mr. Reid Reynolds and Dr. Panagiotis Tsiotras for critical readings of the manuscript and for numerous helpful suggestions for its improvement. Mr. Robert Bauer has made numerous helpful comments. The author is grateful to Dr. Mark Pittelkau for bringing up the question of the uniform pdf for the attitude and the angular velocity of a rigid body.

REFERENCES

1. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Company, New York, 2nd Edition, 1984.

2. W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill Book Company, New York, 3rd Edition, 1973.

3. W. Kaplan, *Advanced Calculus*, Addison-Wesley Publishing Company, Reading, 1952.

4. G. D. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, A. K. Peters, London, 5th Edition, 1997.

5. M. D. Shuster, "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, October–December 1993, pp. 439–517.

6. J. L. Junkins and M. D. Shuster, "The Geometry of the Euler Angles," *Journal of the Astronautical Sciences*, Vol. 41, No 4, October–December, 1993, pp. 531–543.

7. H. W. Sorenson, *Parameter Estimation*, Marcel Dekker, New York, 1980.

8. M. D. Shuster, *Man, Like These Attitudes Are Totally Random!, II. Other Representations*, Paper No. AAS-01-126, AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, California, February 11-15, 2001.