

## ATTITUDE ANALYSIS IN FLATLAND: THE PLANE TRUTH

**Malcolm D. Shuster**

The Johns Hopkins University Applied Physics Laboratory  
Laurel, Maryland 20723-6099

### ABSTRACT

Many results in attitude analysis are still meaningful when the attitude is restricted to rotations about a single axis. Such a picture corresponds to attitude analysis in the Euclidean plane. The present report formalizes the representation of attitude in the plane and applies it to some well-known problems. In particular we study the connection of the “additive” and “multiplicative” formulations of the differential corrector for the quaternion in its two-dimensional setting.

### Introduction

*I call our world Flatland, not because we call it so, but to  
make its nature clearer to you, . . . who are privileged to live  
in Space.*

— *A. Square in Flatland*

The treatment of attitude, because of the non-linearity and non-commutivity of the composition rule, is much more difficult to treat than position, for which components may be combined by simple addition. The complexity of the attitude composition rule leads to virtually all attitude problems being intrinsically three-dimensional or, in the case of the quaternion, four-dimensional. There is, however, a class of attitude problems which are much simpler, namely, single-axis problems, and the study of these will in many cases illuminate the more complex problems. The present report attempts to formalize such a treatment.

## Attitude in Flatland

*Having amused myself til a late hour with my favourite recreation of Geometry, I had retired to rest with an unsolved problem in my mind.*

Let us imagine that the world, Flatland, has only two dimensions and a constant isotropic Euclidean metric. Such a world was imagined by Edwin Abbott [1], with the intent of satirizing the social and political foibles of his day as much as of clarifying the concepts related to the dimensionality of space. Our interest here is more limited than Abbott's. We develop the mathematical structure of Flatland somewhat further in order to better understand those aspects of attitude which do not depend on the dimensionality of space. The quotations which appear in this report are from [1]. Following Abbott we will refer to our three-dimensional world as *Space*.

In Flatland, vectors are, of course, two-dimensional

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \quad (1)$$

The “dot” product takes the usual form

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2, \quad (2)$$

while the “cross product” is now a scalar

$$\mathbf{u} \times \mathbf{v} = u_1 v_2 - u_2 v_1. \quad (3)$$

There is, therefore, no *vector* product, and as alternate names to scalar and vector products we might prefer symmetric and asymmetric products. The lack of a meaningful vector product in two dimensions was ultimately the cause of many years of grief for Hamilton [2–4].

The *attitude matrix* in two dimensions is a  $2 \times 2$  proper orthogonal matrix,  $A$ , which transforms column vectors in the usual way

$$\mathbf{W} = A \mathbf{V}, \quad (4)$$

with

$$A^T A = A A^T = \mathbf{I}, \quad (5)$$

$$\det A = +1, \quad (6)$$

where  $\mathbf{I}$  denotes the  $2 \times 2$  identity matrix,

$$\mathbf{I} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (7)$$

It is a simple matter to show that in two dimensions the attitude matrix may be represented as

$$A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}, \quad (8)$$

and  $\theta$  is the *angle of rotation*. If we define the matrix  $\mathbf{J}$  according to

$$\mathbf{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad (9)$$

which satisfies

$$\mathbf{J}^2 = -\mathbf{I}, \quad (10)$$

then *Euler's formula* becomes simply

$$A = \cos \theta \mathbf{I} + \sin \theta \mathbf{J}, \quad (11)$$

which is much simpler than the three-dimensional form [5–10]. Note that  $\mathbf{J}$  acting on a vector always generates a vector perpendicular to it. The matrices  $\mathbf{I}$  and  $\mathbf{J}$  in Flatland have an importance similar to that of the  $3 \times 3$  identity matrix and the Levi-Civita symbol in Space. They are, in fact, the representations of these objects in two dimensions.

If we define now

$$[[a]] \equiv a \mathbf{J}, \quad (12)$$

then trivially

$$[[a]] [[b]] = -ab \mathbf{I}, \quad (13)$$

which again is much simpler than the three-dimensional variant, and Euler's formula becomes

$$A = \exp [[\theta]], \quad (14)$$

as in Space.

Corresponding to the quaternion in Space, in Flatland we must be content with the *biernion* (pronounced “by-Ernie-on”). The biernion is defined as

$$\bar{q} \equiv \begin{bmatrix} \sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix}, \quad (15)$$

for which

$$\bar{q}^T \bar{q} = 1. \quad (16)$$

We continue to use the notation  $\bar{q}$  (rather than  $\bar{b}$ ) in order to retain a greater resemblance to the equations in Space.

In terms of the biernion Euler's formula becomes

$$A(\bar{q}) = (q_2^2 - q_1^2) \mathbf{I} + 2 q_2 q_1 \mathbf{J} \quad (17a)$$

$$= (q_2^2 - q_1^2) \mathbf{I} + 2 q_2 [[q_1]] \quad (17b)$$

$$= (q_1 \mathbf{I} + q_2 \mathbf{J})^2. \quad (17c)$$

The biernion may be extracted from the attitude matrix in a manner similar to the method for extracting the quaternion from the attitude matrix in Space,

$$q_2 = \frac{1}{2} \sqrt{2 + \text{tr} A}, \quad (18a)$$

$$q_1 = \frac{1}{4q_2} (A_{12} - A_{21}), \quad (18b)$$

where

$$\text{tr}A \equiv A_{11} + A_{22}. \quad (19)$$

Biernion composition follows directly from the trigonometric formula and reads

$$\bar{q}'' = \bar{q}' \otimes \bar{q} \quad (20a)$$

$$= \{\bar{q}'\} \bar{q} = \{\bar{q}\} \bar{q}', \quad (20b)$$

where

$$\{\bar{q}\} \equiv \begin{bmatrix} q_2 & q_1 \\ -q_1 & q_2 \end{bmatrix} = q_2 \mathbf{I} + q_1 \mathbf{J}. \quad (21)$$

Note that biernion composition is commutative, as is the multiplication of attitude matrices in two dimensions.

The *Gibbs scalar* or *Rodrigues scalar* is given by

$$g = q_1/q_2 = \tan(\theta/2). \quad (22)$$

Thus,

$$\bar{q} = \frac{1}{\sqrt{1+g^2}} \begin{bmatrix} g \\ 1 \end{bmatrix}, \quad (23)$$

and Cayley's formula takes the familiar form

$$A = \frac{\mathbf{I} + [[g]]}{\mathbf{I} - [[g]]}. \quad (24)$$

The composition of Gibbs scalars is given by

$$g'' = \frac{g' + g}{1 - g'g}, \quad (25)$$

in complete analogy to the formula for the Gibbs vector in Space.

The *Cayley-Klein parameters* are

$$\alpha \equiv q_2 + i q_1 = e^{i\theta/2}, \quad \text{and} \quad \beta \equiv q_2 - i q_1 = e^{-i\theta/2} = \alpha^c, \quad (26)$$

and the superscript  $c$  denotes complex conjugation. These obviously satisfy

$$\alpha\beta = 1. \quad (27)$$

It follows that

$$A = \frac{1}{2}(\alpha^2 + \beta^2) \mathbf{I} + \frac{1}{2i}(\alpha^2 - \beta^2) \mathbf{J}. \quad (28)$$

### Attitude Kinematics in Flatland

*Restraining my impatience—for I was now under a strong temptation to rush blindly at my visitor and precipitate him into Space ...*

The kinematic equation for the attitude matrix is given as usual by

$$\frac{d}{dt} A(t) = [[\omega(t)]] A(t), \quad (29)$$

which, in fact, defines  $\omega(t)$ . If we define the biernion analogue,

$$\bar{\omega} \equiv \begin{bmatrix} \omega \\ 0 \end{bmatrix}, \quad (30)$$

then the kinematic equation for the biernion is simply

$$\frac{d}{dt} \bar{q}(t) = \frac{1}{2} \bar{\omega}(t) \otimes \bar{q}(t) = \frac{1}{2} \Omega(\omega(t)) \bar{q}(t), \quad (31)$$

where

$$\Omega(\omega) \equiv \omega \mathbf{J}. \quad (32)$$

Likewise, we can partition  $\{\bar{q}\}$  defined by equation (21) in terms of column matrices as

$$\{\bar{q}\} \equiv [\Xi(\bar{q}) \bar{q}], \quad (33)$$

which leads to

$$\frac{d}{dt} \bar{q}(t) = \frac{1}{2} \Xi(\bar{q}(t)) \omega, \quad (34a)$$

and

$$\Xi(\bar{q}) = \begin{bmatrix} q_2 \\ -q_1 \end{bmatrix} = \mathbf{J} \bar{q}. \quad (34b)$$

The kinematic equation for the Gibbs scalar becomes finally

$$\frac{d}{dt} g(t) = \frac{1}{2} [1 + g^2(t)] \omega(t), \quad (35)$$

while that for the angle of rotation is just

$$\frac{d\theta}{dt} = \omega. \quad (36)$$

Euler's equation for rigid-body dynamics is simply

$$I \frac{d\omega}{dt} = N, \quad (37)$$

where  $N$ , the *torque*, is a scalar and  $I$ , the *moment of inertia*, another scalar, is given by

$$I = \int r^2 dm. \quad (38)$$

### Attitude Errors in Flatland

*If Fog were non-existent, all lines would appear equally and indistinguishably clear.*

The representation of attitude errors in Flatland follows that in Space, with obvious simplifications. The error in the attitude matrix, since it has only a single degree of freedom, can be written as

$$A^* = (\delta A) A^{\text{true}}, \quad (39)$$

with  $A^*$  a random variable, usually an attitude estimate, and  $\delta A$  is the (multiplicative) attitude error,

$$\delta A = \exp\{[[\Delta\xi]]\} \approx \mathbf{I} + [[\Delta\xi]], \quad (40)$$

with  $\Delta\xi$ , the attitude error angle, generally an infinitesimal quantity assumed to have zero mean. The attitude variance is defined to be

$$P_{\xi\xi} = E\{(\Delta\xi)^2\}, \quad (41)$$

where  $E\{\cdot\}$  denotes the expectation.

The modeling of vector measurement errors follow a similar pattern. We write

$$\hat{\mathbf{W}} = e^{[[\epsilon]]} A \hat{\mathbf{V}}, \quad (42)$$

where  $\epsilon$  is a zero-mean random variable with variance  $\sigma_W^2$ . In linear approximation this may be written as

$$\hat{\mathbf{W}} = A \hat{\mathbf{V}} + \Delta \hat{\mathbf{W}}, \quad (43)$$

with

$$\Delta \hat{\mathbf{W}} = \epsilon \mathbf{J} A \hat{\mathbf{V}} = [[A \hat{\mathbf{V}}]] \epsilon, \quad (44)$$

and we have defined  $[[\mathbf{v}]]$  with *vector* argument to be

$$[[\mathbf{v}]] \equiv \mathbf{J} \mathbf{v} = \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix}. \quad (45)$$

Thus,

$$[[\mathbf{u}]]^T \mathbf{v} = -\mathbf{u} \times \mathbf{v} = -\mathbf{u}^T [[\mathbf{v}]], \quad (46a)$$

$$[[\mathbf{v}]]^T \mathbf{v} = \mathbf{0}, \quad (46b)$$

$$[[\mathbf{u}]]^T [[\mathbf{v}]] = \mathbf{u} \cdot \mathbf{v}, \quad (46c)$$

$$[[\mathbf{u}]] [[\mathbf{v}]]^T = (\mathbf{u} \cdot \mathbf{v}) \mathbf{I} - \mathbf{v} \mathbf{u}^T, \quad (46d)$$

$$[[a]] \mathbf{v} = [[\mathbf{v}]] a. \quad (46e)$$

### Batch Attitude Determination in Flatland

*I answer that though we cannot see angles, we can infer them, and this with great precision.*

We can now examine some well-known algorithms in their Flatland setting. These are the DYAD algorithm, the two-dimensional analogue of the TRIAD algorithm [11–12], and the BEST algorithm, the two-dimensional analogue of the QUEST algorithm [12]. The development of these algorithms in two dimensions is very similar to that of their forbears in Space. As can be expected, the results are much simpler in the smaller dimension.

*The DYAD Algorithm*

For the DYAD algorithm\* we seek an attitude matrix which satisfies

$$\mathbf{W} = A \mathbf{V}, \quad (47)$$

where  $\mathbf{V}$  and  $\mathbf{W}$  are arbitrary vectors. In a space of  $n$  dimensions,  $n - 1$  linearly independent vector measurements are required to uniquely determine the attitude matrix [13]. In two dimensions, therefore, a single measurement suffices. (In one dimension, zero measurements are sufficient.)

To construct the attitude matrix we first construct orthonormal dyads of reference and observation vectors as

$$\hat{\mathbf{r}}_1 = \frac{\hat{\mathbf{V}}}{|\hat{\mathbf{V}}|} \quad \text{and} \quad \hat{\mathbf{r}}_2 = \mathbf{J} \hat{\mathbf{r}}_1, \quad (48a)$$

and

$$\hat{\mathbf{s}}_1 = \frac{\hat{\mathbf{W}}}{|\hat{\mathbf{W}}|} \quad \text{and} \quad \hat{\mathbf{s}}_2 = \mathbf{J} \hat{\mathbf{s}}_1. \quad (48b)$$

From

$$\mathbf{J}^3 = -\mathbf{J}, \quad (49)$$

it follows that

$$\mathbf{J} A \mathbf{J}^T = A. \quad (50)$$

Hence,

$$\hat{\mathbf{s}}_i = A \hat{\mathbf{r}}_i, \quad i = 1, 2, \quad (51)$$

Defining now orthogonal matrices (labeled by their columns)

$$M_R = [\hat{\mathbf{r}}_1 \quad \hat{\mathbf{r}}_2], \quad \text{and} \quad M_S = [\hat{\mathbf{s}}_1 \quad \hat{\mathbf{s}}_2], \quad (52)$$

it follows that

$$M_S = A M_R, \quad (53)$$

whence

$$A = M_S M_R^T. \quad (54)$$

The development of the DYAD attitude variance follows almost identical steps as in the calculation of the TRIAD attitude covariance in Space [12] with the result

$$P_{\text{DYAD}} = \sigma_W^2. \quad (55)$$

*The BEST Algorithm*

The BEST (Biernion ESTimator) algorithm in Flatland is only slightly less complicated than the QUEST algorithm in Space. As usual, we seek an attitude matrix which minimizes [12, 14]

$$J(A) = \frac{1}{2} \sum_{i=1}^n a_i |\hat{\mathbf{W}}_i - A \hat{\mathbf{V}}_i|^2, \quad (56)$$

---

\*F. Landis Markley has suggested that the DYAD algorithm be renamed the BAD algorithm in contrast with BEST.

where the  $a_i$ ,  $i = 1, \dots, n > 2$ , are a set of positive weights, whose sum, we will assume, is unity. As in the Space we define a gain function,  $g(A)$ , such that

$$g(A) = 1 - J(A) = \text{tr}(B^T A), \quad (57)$$

which is a maximum when  $J(A)$  is a minimum, and, as before the attitude profile matrix is given by

$$B = \sum_{i=1}^n a_i \hat{\mathbf{W}}_i \hat{\mathbf{V}}_i^T. \quad (58)$$

Substituting equation (17a) in equation (57) leads straightforwardly to

$$g(\bar{q}) = (q_2^2 - q_1^2) s + 2 q_1 q_2 z, \quad (59)$$

where

$$s \equiv \text{tr}(B^T) = \text{tr} B = B_{11} + B_{22}, \quad (60a)$$

$$z \equiv \text{tr}(\mathbf{J}B^T) = -\text{tr}(\mathbf{J}B) = B_{12} - B_{21}. \quad (60b)$$

Thus,

$$g(\bar{q}) = \bar{q}^T K \bar{q}, \quad (61)$$

with

$$K = \begin{bmatrix} -s & z \\ z & s \end{bmatrix}. \quad (62)$$

The maximization of  $g(\bar{q})$  leads to the familiar eigenvalue problem

$$K \bar{q}^* = \lambda_{\max} \bar{q}^*, \quad (63)$$

but in Flatland  $\lambda_{\max}$  can be calculated in closed form as

$$\lambda_{\max} = \sqrt{s^2 + z^2}, \quad (64)$$

and

$$\bar{q}^* = \frac{1}{\sqrt{z^2 + (\lambda_{\max} + s)^2}} \begin{bmatrix} z \\ \lambda_{\max} + s \end{bmatrix}. \quad (65)$$

The attitude variance of the BEST algorithm is calculated most easily from the Fisher information matrix using the fact that the BEST algorithm is a maximum-likelihood estimator [15]. Assuming the errors to have a Gaussian distribution, the calculation is straightforward and leads to

$$P_{\text{BEST}}^{-2} = \sum_{i=1}^n \sigma_{W_i}^{-2}. \quad (66)$$

The optimal angle of rotation can also be computed directly by noting that the gain function can be written in the form

$$g(\theta) = s \cos \theta + z \sin \theta, \quad (67)$$

which is obviously a maximum when  $\theta = \theta^*$ , with

$$\cos \theta^* = \frac{s}{\sqrt{s^2 + z^2}} \quad \text{and} \quad \sin \theta^* = \frac{z}{\sqrt{s^2 + z^2}}. \quad (68)$$

We write the solution of equation (68) more conveniently as

$$\theta^* = \arctan_2(z, s), \quad (69)$$

where  $\arctan_2$  is the same function as ATAN2 in FORTRAN. Equation (68) leads directly to a solution for the optimal attitude matrix, namely,

$$A^* = \frac{1}{\sqrt{s^2 + z^2}} \begin{bmatrix} s & z \\ -z & s \end{bmatrix}. \quad (70)$$

Substitution of equation (65) into equation (17) leads somewhat less directly to the same result, which should be compared with the construction of the optimal attitude matrix in Space developed by Markley [16]. Markley's FOAM algorithm [17] carries over with little change into Flatland and yields necessarily the same result as equation (70).

### General Comments on Attitude Estimation in Flatland (and Space)

*I am about to appear very inconsistent.*

There seems to be some confusion concerning the use of representations in attitude estimation, which we will now attempt to muddy further. Typically in attitude determination, one is given a set of measurements,  $\{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ , from which one wishes to infer the attitude, which we will denote without reference to a representation by  $\mathcal{A}$ . The space of  $\mathcal{A}$  we know from long experience is an  $m$ -dimensional manifold, where  $m = 1$  in Flatland,  $m = 3$  in Space, and  $m = 6$  in worlds so unfortunate as to be four-dimensional. An important milestone in every probabilistic estimate of the attitude is the construction of the probability density function (pdf) of the measurements as a function of the attitude,  $p_{\mathbf{z}_1, \dots, \mathbf{z}_N}(\mathbf{z}'_1, \dots, \mathbf{z}'_N; \mathcal{A})$ , where the primed variables denote the values taken by the (unprimed) random variables, and the attitude manifold is assumed to be simply a parameter space rather than a space of random variables. If  $\mathcal{A}$  is also a random variable then the pdf of interest is  $p_{\mathbf{z}_1, \dots, \mathbf{z}_N, \mathcal{A}}(\mathbf{z}'_1, \dots, \mathbf{z}'_N, \mathcal{A}')$ . When one constructs a square loss function, one is, in fact, constructing part of the appropriate pdf assuming Gaussian random noise.

The maximum-likelihood estimate is simply the value of  $\mathcal{A}$  (or  $\mathcal{A}'$ ) which maximizes the appropriate pdf [18]. In mathematical notation we can write\*

$$\mathcal{A}_{ML}^* \equiv \arg \max_{\mathcal{A}} p_{\mathbf{z}_1, \dots, \mathbf{z}_N}(\mathbf{z}'_1, \dots, \mathbf{z}'_N, \mathcal{A}), \quad (71a)$$

or

$$\mathcal{A}'_{ML} \equiv \arg \max_{\mathcal{A}'} p_{\mathbf{z}_1, \dots, \mathbf{z}_N, \mathcal{A}}(\mathbf{z}'_1, \dots, \mathbf{z}'_N, \mathcal{A}'), \quad (71b)$$

---

\*In the particular case where  $\mathcal{A}$  is a random variable one usually speaks of a maximum *a posteriori* estimate.

according to whether or not  $\mathcal{A}$  is a random variable, and the maximum is taken over the manifold of  $\mathcal{A}$  (or  $\mathcal{A}'$ ). Likewise, in the special case that  $\mathcal{A}$  is a random variable, we can define a minimum variance estimate of the attitude as

$$\mathcal{A}'_{MV} \equiv E\{\mathcal{A} \mid \mathbf{z}'_1, \dots, \mathbf{z}'_N\}, \quad (72)$$

that is, as the conditional expectation of  $\mathcal{A}$ . This form of the minimum variance as given by equation (72) is not meaningful, generally, unless the representation of  $\mathcal{A}$  has minimum dimension. Otherwise, the conditional expectation will usually lead to a value which is not on the manifold, and, therefore, unacceptable as a solution. It is difficult, in general, to calculate the minimum variance estimate except in the case where the measurement depends linearly on the quantities to be estimated and on the measurement noise, which is assumed to be Gaussian, in which case the minimum-variance estimate is identical to the maximum-likelihood estimate.

The general method of solution by maximum-likelihood estimation to an attitude estimation problem given a set of measurements and a probabilistic measurement model is to write the negative-log-likelihood function

$$J(\mathcal{A}') \equiv -\log p_{\mathbf{z}'_1, \dots, \mathbf{z}'_N, \mathcal{A}}(\mathbf{z}'_1, \dots, \mathbf{z}'_N, \mathcal{A}'), \quad (73)$$

where for definiteness we consider the case that  $\mathcal{A}$  is a random variable. The negative-log-likelihood function is a minimum at the maximum-likelihood estimate. The procedure is thus to minimize the expression in equation (73) by an iterative method, such as the Newton-Raphson method. Thus, if  $\mathcal{A}'_i$  is the  $i$ -th iteration we write

$$\mathcal{A}' = \mathcal{A}(\boldsymbol{\beta}) \otimes \mathcal{A}'_i, \quad (74)$$

where  $\mathcal{A}(\boldsymbol{\beta})$  denotes the general attitude as a function of  $\boldsymbol{\beta}$ , which is one of the many minimum-dimensional representations of the attitude which is Euclidean at the origin and for which  $\mathcal{A}(\mathbf{0})$  is the identity rotation. Expanding  $\mathcal{A}$  as a function of  $\boldsymbol{\beta}$  leads to

$$\begin{aligned} J(\mathcal{A}') &= J(\mathcal{A}'_i) + \left[ \frac{\partial}{\partial \boldsymbol{\beta}} J(\mathcal{A}(\boldsymbol{\beta}) \otimes \mathcal{A}'_i) \right]_{\boldsymbol{\beta}=\mathbf{0}} \boldsymbol{\beta} \\ &\quad + \frac{1}{2} \boldsymbol{\beta}^T \left[ \frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} J(\mathcal{A}(\boldsymbol{\beta}) \otimes \mathcal{A}'_i) \right]_{\boldsymbol{\beta}=\mathbf{0}} \boldsymbol{\beta} + O(|\boldsymbol{\beta}|^3), \end{aligned} \quad (75)$$

and minimizing this expression keeping terms only up to second order in  $\boldsymbol{\beta}$  leads to the next iteration

$$\boldsymbol{\beta}_{i+1} = - \left[ \frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} J(\mathcal{A}(\boldsymbol{\beta}) \otimes \mathcal{A}'_i) \right]_{\boldsymbol{\beta}=\mathbf{0}}^{-1} \left[ \frac{\partial}{\partial \boldsymbol{\beta}} J(\mathcal{A}(\boldsymbol{\beta}) \otimes \mathcal{A}'_i) \right]_{\boldsymbol{\beta}=\mathbf{0}}, \quad (76)$$

$$\mathcal{A}'_{i+1} = \mathcal{A}(\boldsymbol{\beta}_{i+1}) \otimes \mathcal{A}'_i. \quad (77)$$

This procedure will generally converge to a minimum of the negative-log-likelihood function. In well-defined attitude problems this minimum is usually unique and hence,

$$\mathcal{A}'_{ML} = \lim_{i \rightarrow \infty} \mathcal{A}'_i. \quad (78)$$

In the limit that the amount of data is infinite the attitude covariance matrix can be written as

$$P_{\beta\beta}^{-1} = E \left\{ \left[ \frac{\partial^2}{\partial\beta\partial\beta^T} J(\mathcal{A}(\beta) \otimes \mathcal{A}') \right] \right\}_{\beta=\mathbf{0}, \mathcal{A}'=\mathcal{A}'_{ML}} \quad (79)$$

If the measurement errors are Gaussian, then within the linearization approximation, equation (79) will be true even for small samples.

### Treatment of the Biernion/Quaternion in Attitude Estimation

*It is high time that I should pass from these brief and discursive notes about things in Flatland to the central event*

...

Several schemes have been proposed [19] for mechanizing the Kalman filter update for the quaternion. The effect of these and other schemes has been studied via numerical examples by Bar-Itzhack, Deutschman and Markley [20, 21]. These latter authors make a distinction between the update step of the Kalman filter using what they call the additive as opposed to the multiplicative update. This distinction is artificial and misleading, as we shall now show.

Let us write the relation between the updated and predicted quaternions/biernions as

$$\bar{q}_k(+) = \bar{q}_k(-) + \Delta\bar{q}_k(+), \quad (80)$$

which Bar-Itzhack *et al.* call the additive approach. The components are all resolved with respect to inertial axes. Let us examine the same equation expressed with respect to the predicted spacecraft body frame, i.e., we express all rotations as rotations from the predicted spacecraft body frame. Denoting the quaternions/biernions of rotation with respect to this frame by  $\bar{q}'$  where

$$\bar{q}'_k = \bar{q}_k \otimes \bar{q}_k^{-1}(-), \quad (81)$$

it follows that

$$\bar{q}'_k(+) = \bar{1} + \Delta\bar{q}'_k(+), \quad (82)$$

where  $\bar{1} = [0 \ 0 \ 0 \ 1]^T$  for quaternions, and  $\bar{1} = [0 \ 1]^T$  for biernions. If we write now

$$\delta\bar{q}_k(+) \equiv \bar{q}'_k(+), \quad (83)$$

then it follows that

$$\bar{q}_k(+) = \delta\bar{q}_k(+) \otimes \bar{q}_k(-), \quad (84)$$

which is the so-called multiplicative correction. Thus, the distinction between the additive and the multiplicative formulations of the Kalman filter is not one of the fundamental mechanization of the filter but simply the frame in which it is desired to compute the update. These two formulations are present in Reference [19], where they are given the names “truncated covariance representation” and “body-fixed covariance representation.” Admittedly, the presentation by those authors gave the appearance of there being one more distinct formulation of the Kalman filter than was actually the case. This has even led one careful study to test both formulations, as if they were distinct [22].

Where the important distinctions do lie is in how  $\Delta\bar{q}_k$  or  $\Delta\bar{q}'_k$  is calculated, and consequently, whether  $\delta\bar{q}(+)$  has unit norm. From the earlier discussion it is clear that a correct

approach is obtained by expressing this quantity in terms of some representation of the attitude of minimal degree. In this case it is clearly advantageous to work from the spacecraft body frame so that this minimal-dimensional representation will be far from a singularity, and it will be most revealing to compare the results of References [19] and [20, 21] in that frame. The results of [20, 21], however, are not directly comparable to [19] because the former rests on the attitude Kalman filter of Bar-Itzhack and Oshman [23]. However, many points of commonality will be apparent.

Consider now the estimation of a constant biernion from scalar measurements of the form

$$z_k = \hat{\mathbf{u}}_k^T \mathbf{W}_k, \quad k = 1, \dots, N, \quad (85)$$

where  $\hat{\mathbf{u}}_k$  is a known direction in the spacecraft body and  $\mathbf{W}_k$  is some vector measured in the body frame. We assume that  $\mathbf{W}_k$  is related to a representation of the same vector in the inertial frame according to

$$\mathbf{W}_k = A \mathbf{V}_k + \mathbf{v}_k, \quad (86)$$

where  $A$  is the attitude matrix and  $\mathbf{v}_k$  is white Gaussian noise. We wish to compute the batch attitude estimate from these measurements, using an good approximate estimate of the attitude as a point of departure.

If we write now

$$A = (\delta A) A_o, \quad (87)$$

where  $A_o$  is the approximate value of the optimal attitude estimate, then the measurement equation becomes

$$z_k = \hat{\mathbf{u}}_k^T (\delta A) \mathbf{W}_{o,k} + \hat{\mathbf{u}}_k^T \mathbf{v}_k, \quad (88)$$

where  $\hat{\mathbf{W}}_{o,k} = A_o \mathbf{V}_k$ , the expected value of the measurement in the body frame.  $\delta A$  is now an infinitesimal rotation, which we shall parameterize in terms of the additive biernion error as in equation (82). Recalling equation (17c) it is a simple manner to expand  $z_k$  to lowest order in  $\Delta\bar{q}$  with the result

$$\nu_k \equiv z_k - z_{o,k} = H_k \Delta\bar{q} + v_k, \quad (89)$$

where  $z_{o,k}$  is the value of the measurement with  $\Delta\bar{q} = \bar{0}$ ,  $v_k$  is the scalar white Gaussian noise term appearing in equation (88) and the  $1 \times 2$  sensitivity matrix  $H_k$  is given by

$$H_k = [H_{1,k} \quad H_{2,k}] = [(\hat{\mathbf{u}}_k \times \mathbf{W}_{o,k}), (\hat{\mathbf{u}}_k \cdot \mathbf{W}_{o,k})]. \quad (90)$$

The maximum likelihood estimate of  $\Delta\bar{q}$  (for the additive biernion correction, which is not constrained to preserve the norm) is given by

$$\Delta\bar{q}_{\text{add}}^* = P_{qq} \mathbf{p}, \quad (91)$$

where the covariance matrix,  $P_{qq}$ , and the information vector,  $\mathbf{p}$ , are given by

$$P_{qq} = \left[ \sum_{k=1}^N H_k^T R_k^{-1} H_k \right]^{-1}, \quad \mathbf{p} = \sum_{k=1}^N H_k^T R_k^{-1} \nu_k. \quad (92)$$

For the multiplicative correction (which is norm-preserving) the estimate for the same data is

$$\Delta\bar{q}_{1,\text{mult}}^* = P_{\text{mult}} p_1,$$

with

$$P_{\text{mult}} = \left[ \sum_{k=1}^N H_{1,k}^T R_k^{-1} H_{1,k} \right]^{-1} = (P_{qq}^{-1})_{11}^{-1}, \quad (93a)$$

$$p_1 = \sum_{k=1}^N H_{1,k}^T R_k^{-1} \nu_k, \quad (93b)$$

so that  $p_1$  is just the first component of  $\mathbf{p}$ . Note that we have written unnecessary (but not incorrect) transpose signs and not commuted symbols, even between scalars, to preserve the resemblance with the equations of Space. We can find a relation between the additive and the multiplicative corrections to the biernion by solving equation (91) for  $\mathbf{p}$  in terms of  $\Delta \bar{q}_{\text{add}}^*$  and using the value of  $p_1$  from this expression in equation (93).

This leads to

$$\Delta \bar{q}_{1,\text{mult}}^* = \Delta \bar{q}_{1,\text{add}}^* + (P_{qq}^{-1})_{11}^{-1} (P_{qq}^{-1})_{12} \Delta \bar{q}_{2,\text{add}}^*. \quad (94)$$

We will return to this equation soon.

The additive correction,  $\Delta \bar{q}_{\text{add}}^*$ , allows us to construct an optimal biernion,  $\bar{q}_{\text{add}}^*$ ,

$$\bar{q}_{\text{add}}^* = \bar{\mathbf{I}} + \Delta \bar{q}_{\text{add}}^*. \quad (95)$$

Because it does not necessarily have unit norm,  $\bar{q}_{\text{add}}^*$  does not without further effort have an unambiguous connection to the attitude. However, we note that although  $\bar{q}_{\text{add}}^*$  is not a “biernion of rotation,” it is a sufficient statistic [18] for the attitude, certainly within the linear approximation of equation (89). It is, in fact, an estimate of the biernion of rotation, and we know also that were the measurement noise covariance to vanish (perfect measurements),  $\bar{q}_{\text{add}}^*$  would have unit norm and be the desired biernion. Thus, denoting the desired biernion of rotation by  $\bar{\eta}$ , we have that

$$\bar{q}_{\text{add}}^* = \bar{\eta} + \Delta \bar{\eta}_{\text{add}}, \quad (95)$$

and

$$\Delta \bar{\eta}_{\text{add}} \sim \mathcal{N}(\bar{\mathbf{0}}, P_{qq}). \quad (96)$$

Hence, the negative log-likelihood function of  $\bar{q}_{\text{add}}^*$  given  $\bar{\eta}$  is

$$J(\bar{q}_{\text{add}}^* | \bar{\eta}) = \frac{1}{2} [(\bar{q}_{\text{add}}^* - \bar{\eta})^T P_{qq} (\bar{q}_{\text{add}}^* - \bar{\eta}) + \log \det P_{qq} + 4 \log 2\pi]. \quad (97)$$

and the maximum-likelihood estimate of  $\bar{\eta}$  is simply

$$\bar{\eta}^* = \arg \max_{\bar{\eta}: \bar{\eta}^T \bar{\eta} = 1} J(\bar{q}_{\text{add}}^* | \bar{\eta}), \quad (98)$$

where, since we know that the true biernion must lie on the manifold of unit four-vectors, we must maximize the negative log-likelihood subject to the norm constraint.

We handle the constraint in the usual way, using Lagrange’s method of multipliers, and optimize

$$J(\bar{q}_{\text{add}}^* | \bar{\eta}) + \frac{1}{2} \lambda \bar{\eta}^T \bar{\eta}$$

without constraint and then choosing the value of the Lagrange multiplier,  $\lambda$ , for which the constraint is satisfied. Differentiating the above expression with respect to  $\bar{\eta}$  and setting the derivative equal to zero leads to

$$\bar{\eta}^* = (I + \lambda P_{qq})^{-1} \bar{q}_{\text{add}}^*, \quad (99)$$

and  $\lambda$  is a solution of

$$f(\lambda) \equiv \bar{\eta}^{*T}(\lambda) \bar{\eta}(\lambda) = \bar{q}_{\text{add}}^{*T} (I + \lambda P_{qq})^{-2} \bar{q}_{\text{add}}^* = 1. \quad (100)$$

We expect  $\lambda P_{qq}$  to be small. Therefore, it will usually be sufficient to calculate  $\lambda$  using one iteration of the Newton Raphson method with vanishing initial value. Thus,

$$\lambda \approx \frac{1 - f(0)}{f'(0)} = \frac{1}{2} (\bar{q}_{\text{add}}^{*T} P_{qq} \bar{q}_{\text{add}}^*)^{-1} (\bar{q}_{\text{add}}^{*T} \bar{q}_{\text{add}}^* - 1). \quad (101)$$

To first order

$$\bar{q}_{\text{add}}^{*T} \bar{q}_{\text{add}}^* = 2 \Delta q_2. \quad (102)$$

Hence,

$$\lambda = (P_{qq})_{22}^{-1} \Delta q_2. \quad (103)$$

Substituting this in equation (99) leads to lowest order in  $\Delta \bar{q}_{\text{add}}^*$

$$\bar{\eta}^* = (I + \lambda P_{qq})^{-1} \bar{q}_{\text{add}}^* \quad (104a)$$

$$\approx (I - \lambda P_{qq}) \bar{q}_{\text{add}}^* \quad (104b)$$

$$= \bar{q}_{\text{add}}^* - \Delta q_{2,\text{add}}^* (P_{qq})_{22}^{-1} P_{qq} \bar{q}_{\text{add}}^*. \quad (104c)$$

The first component of the desired optimal biernion is simply (to this same order)

$$\eta_1^* = \Delta q_{1,\text{add}}^* - (P_{qq})_{22}^{-1} (P_{qq})_{12} \Delta q_{2,\text{add}}^*. \quad (105)$$

But

$$-(P_{qq})_{12} (P_{qq})_{22}^{-1} = (P_{qq}^{-1})_{11}^{-1} (P_{qq}^{-1})_{12}, \quad (106)$$

so that, in fact, comparing equation (106) with equation (94) we have

$$\bar{\eta}_1^* = \Delta q_{1,\text{mult}}^*. \quad (107)$$

Since the other component must also agree to linear order in  $\Delta \bar{q}_{\text{mult}}^*$ , it follows that

$$\bar{\eta}^* = \delta \bar{q}_{\text{mult}}^*. \quad (108)$$

Thus, the additive correction to the biernion, followed by the normalization correction *dictated unambiguously by the maximum likelihood criterion*, is identical (at least up to linear terms in  $\Delta \bar{q}^*$ ) to the so-called multiplicative correction. It is hard to imagine that any other answer could be possible. It is obviously less burdensome to calculate the

multiplicative correction directly. Identical arguments hold for sequential correction of the biernion as in the Kalman filter.

## Discussion

*... my Lord has shewn me the intestines of all my countrymen in the Land of Two Dimensions ...*

The representation of attitude in two dimensions has been described in detail. Two-dimensional analogues have been presented for the well known TRIAD and QUEST algorithms. General properties of attitude estimation in two and three dimensions have been discussed. The question of whether the multiplicative or additive correction to the quaternion is preferred has been answered in Flatland.

In Space the answer is no different. We have been careful to arrange the expression dealing with the differential correction of the biernion in such a way that they have the same form as the corresponding expression for the quaternion, in which many of the factors are matrices, which do not commute under multiplication. The only difference in the derivations for the quaternion and the biernion are the specific expressions for  $H_{1,k}$  and  $H_{2,k}$ . However, these differences do not effect the final result, which requires no specific expression for these quantities. Thus, the derivation of the result has been proved for the quaternion as well.

The additive correction, if done correctly, is identical to the multiplicative correction but is much more burdensome. The first commandment of biernion correction (and, one can show, also for quaternion correction in Space), therefore, is to multiply. We emphasize that this result is not the product of some heuristic argument or arbitrary procedure to be justified by experiment but the unavoidable conclusion to which one is led unambiguously and rigorously by the maximum likelihood criterion.

## Acknowledgements

*"You see ... how little your words have done."*

The author is grateful to F. Landis Markley and Itzhack Bar-Itzhack for many passionate discussions on this topic over the years.

## References

- [1] ABBOTT, E. A., "Flatland, a Romance of Many Dimensions," New York, Dover Publications, 1884.
- [2] ALTMANN, S. L., *Rotations, Quaternions, and Double Groups*, Oxford, Oxford University Press, 1986.
- [3] ALTMANN, S. L., "Hamilton, Rodrigues, and the Quaternion Scandal," *Mathematics Magazine*, Vol. 62, pp. 291–307, 1989.
- [4] VAN DER WAERDEN, B. L., *A History of Algebra*, New York and Heidelberg, Springer-Verlag, 1985.
- [5] GOLDSTEIN, H., *Classical Mechanics*, Reading, Mass., Addison-Wesley, 1980.

- [ 6 ] HUGHES, P. C., *Spacecraft Attitude Dynamics*, New York, John Wiley & Sons, 1986.
- [ 7 ] JUNKINS, J. L., and TURNER, J. D., *Optimal Spacecraft Rotational Maneuvers*, Amsterdam, Elsevier, 1986.
- [ 8 ] MARKLEY, F. L., "Parameterization of the Attitude," in WERTZ, J. R. (ed.) [ 9 ], *op. cit.*, pp. 410–420, 1978.
- [ 9 ] WERTZ, J. R. (ed.), *Spacecraft Attitude Determination and Control*, Dordrecht, the Netherlands, Kluwer Academic Publishers, 1978.
- [ 10 ] SHUSTER, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, (to appear).
- [ 11 ] BLACK, H. D., "A Passive System for Determining the Attitude of a Satellite," *AIAA Journal*, Vol. 2, pp. 1350–1351, 1964.
- [ 12 ] SHUSTER, M. D., and OH, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance, Control and Dynamics*, Vol. 4, No. 1, pp. 70–77, 1981.
- [ 13 ] SHUSTER, M. D., "Attitude Determination in Higher Dimensions," *Journal of Guidance, Control and Dynamics*, (to appear).
- [ 14 ] J. KEAT, "Analysis of Least Squares Attitude Determination Routine, DOAOP," Computer Sciences Corporation, CSC/TM-77/6034, February 1977.
- [ 15 ] SHUSTER, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, 1989, pp. 79–88.
- [ 16 ] MARKLEY, F. L., "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, Vol. 36, pp. 245–258, 1988.
- [ 17 ] MARKLEY, F. L., "Attitude Determination Using Vector Observations: A Fast Optimal Matrix Algorithm," Flight Mechanics/Estimation Theory Symposium, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 5–7, 1992.
- [ 18 ] SORENSON, H. W., *Parameter Estimation*, New York, Marcel Dekker, 1980.
- [ 19 ] LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics*, Vol. 5, No. 5, pp. 417–429, 1982.
- [ 20 ] BAR-ITZHACK, I. Y., MARKLEY, F. L., and DEUTSCHMAN, J., "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, May 1991.
- [ 21 ] DEUTSCHMAN, J., MARKLEY, F. L., and BAR-ITZHACK, I. Y., "Quaternion Normalization in Spacecraft Attitude Determination," *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, May 1992, and works cited therein.
- [ 22 ] FERRARESI, V. A., *Utilização conjunta de sensores inerciais e não-inerciais em determinação de atitude de satélites via filtro de Kalman*, Instituto Nacional das Pesquisas Espaciais, Publicação No. INPE-4313-TDL/280, August 1987.
- [ 23 ] BAR-ITZHACK, I. Y., and OSHMAN, Y., "Recursive Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronics Systems*, Vol. 21, 1985, pp. 128–135.