The Journal of the Astronautical Sciences, Vol. 55, No 2, April–June 2007, pp. 245–257

The Optimization of TRIAD

Malcolm D. Shuster[1]

MATERIAM SUPERABAT OPUS.[2]

Publius Ovidius Naso (43 BCE–17 CE)
Metamorphoses Book II

Abstract

The optimized TRIAD algorithm of Bar-Itzhack and Harman, proposed in 1997 on the basis of insupportable premises, is shown to provide, nonetheless, the optimal solution of the Wahba problem to within terms on the order of the measurement variances. A careful rigorous derivation of the algorithm is presented as well as a comparison with the arguments of Bar-Itzhack and Harman. The algorithm is compared to a result of Markley's.

Introduction

The simple and elegant TRIAD algorithm, invented by Harold D. Black in 1964 [1, 2] offers a simple and elegant means for estimating spacecraft attitude. Unfortunately, Black's algorithm does not provide an optimal attitude estimate nor can it take account of more than two measurements, a situation that was remedied by the Wahba problem [3] and its solutions [4] published a year and two years after Black's algorithm. These early solutions to the Wahba problem, however, were not very practical for mission support, and Black's algorithm held pride of place for nearly two decades, when it was supplanted by Davenport's q-algorithm [2, 5, 6], a solution to the Wahba problem, particularly in its most popular implementation, QUEST [2]. The QUEST algorithm, published in 1981, being also a solution of the Wahba problem, had the advantage also of being able to make use of more than two measurements and of providing an optimal attitude estimate. Since then, numerous other algorithms have been proposed [6].[3]

In the present study, we examine the possibility of computing a more accurate attitude estimate than is provided by the TRIAD estimate as a linear combination

[1]Director of Research, Acme Spacecraft Company, 13017 Wisteria Drive, Box 328, Germantown, Maryland 20874. E-mail: mdshuster@comcast.net. Website: http://home.comcast.net/~mdshuster.
[2]The work surpassed the material.
[3]Reference [6] provides a masterful overview of the many solutions of the Wahba problem. On its numerical results and their interpretation see references [7] and [8].

of the TRIAD attitude estimates for different permutations of the two measured directions. The resulting algorithm turns out to be optimal, that is, a solution of the Wahba problem, to within terms on the order of the measurement error variances. For measurements with an accuracy of 20 arcsec, these terms are on the order of 0.002 arcsec, a suitably negligible quantity. It also turns out to be the optimized TRIAD algorithm of Bar-Itzhack and Harman [9].[4]

We will begin with a careful rigorous derivation of the optimized TRIAD algorithm and then compare our derivation with the arguments offered by reference [9].

The TRIAD Algorithm

In the TRIAD algorithm, one is given two unit-vector measurements (i.e., 3×1 arrays of numbers or, equivalently, 3×1 column vectors) $\hat{\mathbf{W}}_1'$ and $\hat{\mathbf{W}}_2'$, generally the observed directions of the Sun, a star, the magnetic field, or the nadir in the spacecraft body frame. These are the realizations of two random column vectors $\hat{\mathbf{W}}_1^{\text{r.v.}}$ and $\hat{\mathbf{W}}_2^{\text{r.v.}}$. Our notation follows reference [10]. Here, $\hat{\mathbf{W}}_1^{\text{r.v.}}$ and $\hat{\mathbf{W}}_2^{\text{r.v.}}$ are the representations with respect to the spacecraft body frame of two *physical* random vectors $\hat{\mathsf{W}}_1^{\text{r.v.}}$ and $\hat{\mathsf{W}}_2^{\text{r.v.}}$ with physical realizations $\hat{\mathsf{W}}_1'$ and $\hat{\mathsf{W}}_2'$. Note the differences in typeface. If $\hat{\mathsf{W}}_1^{\text{true}}$ and $\hat{\mathsf{W}}_2^{\text{true}}$ are the true (and, therefore, non-random) physical vectors, then their representations with respect to space (generally inertial) axes are written $\hat{\mathbf{V}}_1$ and $\hat{\mathbf{V}}_2$. The random body representations satisfy

$$\hat{\mathbf{W}}_1^{\text{r.v.}} = A^{\text{true}}\hat{\mathbf{V}}_1 + \Delta\hat{\mathbf{W}}_1^{\text{r.v.}} \qquad \text{and} \qquad \hat{\mathbf{W}}_2^{\text{r.v.}} = A^{\text{true}}\hat{\mathbf{V}}_2 + \Delta\hat{\mathbf{W}}_2^{\text{r.v.}} \tag{1ab}$$

with A^{true} the attitude matrix and $\Delta\hat{\mathbf{W}}_1^{\text{r.v.}}$ and $\Delta\hat{\mathbf{W}}_2^{\text{r.v.}}$ the random measurement noise in the body frame. Effectively, we assume that the errors in the ground-based measurements which led to the space representations of the measured physical vectors are negligible compared to those of the spacecraft sensors which lead to $\hat{\mathbf{W}}_1'$ and $\hat{\mathbf{W}}_2'$ and may be replaced by non-random $\hat{\mathbf{V}}_1$ and $\hat{\mathbf{V}}_2$, respectively. These two equations are generally not solvable in practice for the attitude matrix A^{true}, owing to the presence of the noise terms, whose realizations, generally, cannot be known.[5]

The TRIAD prescription for constructing the estimator of a (proper orthogonal) attitude matrix [1, 2] is as follows: Define two right-hand orthonormal triads of column vectors, $\{\hat{\mathbf{r}}_1, \hat{\mathbf{r}}_2, \hat{\mathbf{r}}_3\}$ and $\{\hat{\mathbf{s}}_1^{\text{r.v.}}, \hat{\mathbf{s}}_2^{\text{r.v.}}, \hat{\mathbf{s}}_3^{\text{r.v.}}\}$ according to

$$\hat{\mathbf{r}}_1 = \hat{\mathbf{V}}_1, \qquad \hat{\mathbf{r}}_2 = \frac{\hat{\mathbf{V}}_1 \times \hat{\mathbf{V}}_2}{|\hat{\mathbf{V}}_1 \times \hat{\mathbf{V}}_2|}, \qquad \hat{\mathbf{r}}_3 = \hat{\mathbf{r}}_1 \times \hat{\mathbf{r}}_2 \tag{2abc}$$

$$\hat{\mathbf{s}}_1^{\text{r.v.}} = \hat{\mathbf{W}}_1^{\text{r.v.}}, \qquad \hat{\mathbf{s}}_2^{\text{r.v.}} = \frac{\hat{\mathbf{W}}_1^{\text{r.v.}} \times \hat{\mathbf{W}}_2^{\text{r.v.}}}{|\hat{\mathbf{W}}_1^{\text{r.v.}} \times \hat{\mathbf{W}}_2^{\text{r.v.}}|}, \qquad \hat{\mathbf{s}}_3^{\text{r.v.}} = \hat{\mathbf{s}}_1^{\text{r.v.}} \times \hat{\mathbf{s}}_2^{\text{r.v.}} \tag{3abc}$$

and set

$$A^{\text{TRIAD}*} = [\hat{\mathbf{s}}_1^{\text{r.v.}} \quad \hat{\mathbf{s}}_2^{\text{r.v.}} \quad \hat{\mathbf{s}}_3^{\text{r.v.}}][\hat{\mathbf{r}}_1 \quad \hat{\mathbf{r}}_2 \quad \hat{\mathbf{r}}_3]^{\text{T}} \tag{4}$$

Voilà!

[4]The algorithm presented here differs slightly from that of Bar-Itzhack and Harman in that the latter is not exactly orthogonal.

[5]Equations (1) are deceptively simple. While they appear to be linear, they are not, because the nine elements of A satisfy six quadratic constraints, and the six constraint equations must be considered simultaneously with equations (1). The problem of constraint in the attitude representations is one of the central hurdles of attitude estimation and the improper treatment of constraint, whether for the direction-cosine matrix or for the quaternion, is not an infrequent occurrence, unfortunately, and can lead to catastrophic results [11, 12]. Black through his remarkable algorithm is the first person to successfully linearize the attitude estimation problem.

In equation (4) the brackets denote two matrices labeled by their column vectors, and the superscript "T" denotes the matrix transpose. Note that $A^{\text{TRIAD}*}$ in equation (4) is an *estimator,* a random matrix. The related *estimate,* its realization, is

$$A^{\text{TRIAD}*\prime} = [\hat{\mathbf{s}}_1' \quad \hat{\mathbf{s}}_2' \quad \hat{\mathbf{s}}_3'][\hat{\mathbf{r}}_1 \quad \hat{\mathbf{r}}_2 \quad \hat{\mathbf{r}}_3]^{\text{T}} \tag{5}$$

and the true value of the attitude matrix is given by

$$A^{\text{true}} = (A^{\text{TRIAD}})^{\text{true}} = [\hat{\mathbf{s}}_1^{\text{true}} \quad \hat{\mathbf{s}}_2^{\text{true}} \quad \hat{\mathbf{s}}_3^{\text{true}}][\hat{\mathbf{r}}_1 \quad \hat{\mathbf{r}}_2 \quad \hat{\mathbf{r}}_3]^{\text{T}} \tag{6}$$

The random proper-orthogonal matrix $A^{\text{TRIAD}*}$ perforce satisfies

$$A^{\text{TRIAD}*}\hat{\mathbf{r}}_k = \hat{\mathbf{s}}_k^{\text{r.v.}}, \qquad k = 1, 2, 3 \tag{7}$$

and satisfies $\hat{\mathbf{W}}_1^{\text{r.v.}} = A^{\text{TRIAD}*}\hat{\mathbf{V}}_1$ exactly. If there were no measurement noise, then $\hat{\mathbf{W}}_2^{\text{true}} = (A^{\text{TRIAD}})^{\text{true}}\hat{\mathbf{V}}_2 = A^{\text{true}}\hat{\mathbf{V}}_2$ would also be satisfied. Otherwise, the equation $\hat{\mathbf{s}}_3^{\text{r.v.}} = A^{\text{TRIAD}*}\hat{\mathbf{r}}_3$ shows plainly that $A^{\text{TRIAD}*}$ aligns the component of $\hat{\mathbf{V}}_2$ perpendicular to $\hat{\mathbf{V}}_1$ along the component of $\hat{\mathbf{W}}_2^{\text{r.v.}}$ perpendicular to $\hat{\mathbf{W}}_1^{\text{r.v.}}$.

Note that the TRIAD attitude estimator or estimate depends on the order of the two measurement random variables or realizations, respectively.

The information matrix (not necessarily the Fisher information matrix [13]) for the spacecraft attitude, asymptotically the inverse attitude covariance matrix, is given[6] for the TRIAD algorithm by reference [2],[7] [14], and [15]

$$P_{\xi\xi}^{\text{TRIAD}} \equiv E\{\Delta\boldsymbol{\xi}^{\text{TRIAD}}\Delta\boldsymbol{\xi}^{\text{TRIAD T}}\} = (F_{\xi\xi}^{\text{TRIAD}})^{-1} \tag{8a}$$

$$\begin{aligned} F_{\xi\xi}^{\text{TRIAD}} &= \frac{1}{\sigma_1^2}\hat{\mathbf{s}}_3^{\text{true}}\hat{\mathbf{s}}_3^{\text{true T}} + \frac{1}{\sigma_1^2}\hat{\mathbf{s}}_2^{\text{true}}\hat{\mathbf{s}}_2^{\text{true T}} + \frac{1}{\sigma_2^2}\hat{\mathbf{s}}_4^{\text{true}}\hat{\mathbf{s}}_4^{\text{true T}} \\ &= \frac{1}{\sigma_1^2}(I_{3\times3} - \hat{\mathbf{s}}_1^{\text{true}}\hat{\mathbf{s}}_1^{\text{true T}}) + \frac{1}{\sigma_2^2}\hat{\mathbf{s}}_4^{\text{true}}\hat{\mathbf{s}}_4^{\text{true T}} \end{aligned} \tag{8b}$$

where σ_1^2 and σ_2^2 are the variance parameters for $\hat{\mathbf{W}}_1$ and $\hat{\mathbf{W}}_2$, respectively, in the QUEST measurement model [2, 16] (see equation (12)).

$$\hat{\mathbf{s}}_4 \equiv \hat{\mathbf{W}}_2 \times \hat{\mathbf{s}}_2 \tag{9}$$

Here, $\Delta\boldsymbol{\xi}^{\text{TRIAD}*}$, of which $F_{\xi\xi}^{\text{TRIAD}}$ is the information matrix, is the random attitude error increment defined by

$$A^{\text{TRIAD}*} = e^{[[\Delta\xi^{\text{TRIAD}*}]]}A^{\text{true}} \tag{10}$$

where, for a 3×1 column vector $\mathbf{u}$,

$$[[\mathbf{u}]] \equiv \begin{bmatrix} 0 & u_3 & -u_2 \\ -u_3 & 0 & u_1 \\ u_2 & -u_1 & 0 \end{bmatrix} \tag{11}$$

Equation (8) has assumed the QUEST measurement model [2, 16], namely, that the $\Delta\hat{\mathbf{W}}_i^{\text{r.v.}}$, $i = 1, 2$, are zero-mean, mutually uncorrelated, and

[6]When the same formula holds for the random variable, the realization of the random variable, and the true value of the random variable, the common formula will be given without superscripts. An asterisk always indicates an estimator and, therefore, a random variable. Likewise, "Δ" will always distinguish a zero-mean random infinitesimal quantity.

[7]Reference [2] had defined the attitude covariance in an almost identical manner but included a factor 1/4, because it employed the quaternion rather than the rotation vector of an infinitesimal rotation.

$$\Delta \hat{\mathbf{W}}_i^{\text{r.v.}} \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 (I_{3\times 3} - \hat{\mathbf{W}}_i^{\text{true}} \hat{\mathbf{W}}_i^{\text{true T}})), \qquad i = 1, 2 \tag{12}$$

The TRIAD-II Algorithm

The result of the TRIAD algorithm presented above depends on which observed direction is chosen to be first. Let us consider the case where the order of the two observed directions is reversed. We shall call this the TRIAD-II algorithm and the original TRIAD algorithm we will call TRIAD-I, although we shall retain the name "TRIAD" for TRIAD-I in some contexts.[8] In obvious notation

$$A^{\text{TRIAD-II}}(\hat{\mathbf{W}}_1, \hat{\mathbf{W}}_2) \equiv A^{\text{TRIAD}}(\hat{\mathbf{W}}_2, \hat{\mathbf{W}}_1) = A^{\text{TRIAD-I}}(\hat{\mathbf{W}}_2, \hat{\mathbf{W}}_1) \tag{13}$$

For the TRIAD-II algorithm we define[9]

$$\hat{\mathbf{r}}_5 = \hat{\mathbf{V}}_2, \qquad \hat{\mathbf{r}}_2 = \frac{\hat{\mathbf{V}}_1 \times \hat{\mathbf{V}}_2}{|\hat{\mathbf{V}}_1 \times \hat{\mathbf{V}}_2|}, \qquad \hat{\mathbf{r}}_4 = \hat{\mathbf{r}}_5 \times \hat{\mathbf{r}}_2 \tag{14abc}$$

$$\hat{\mathbf{s}}_5^{\text{r.v.}} = \hat{\mathbf{W}}_2^{\text{r.v.}}, \qquad \hat{\mathbf{s}}_2^{\text{r.v.}} = \frac{\hat{\mathbf{W}}_1^{\text{r.v.}} \times \hat{\mathbf{W}}_2^{\text{r.v.}}}{|\hat{\mathbf{W}}_1^{\text{r.v.}} \times \hat{\mathbf{W}}_2^{\text{r.v.}}|}, \qquad \hat{\mathbf{s}}_4^{\text{r.v.}} = \hat{\mathbf{s}}_5^{\text{r.v.}} \times \hat{\mathbf{s}}_2^{\text{r.v.}} \tag{14def}$$

and then set

$$A^{\text{TRIAD-II}*} = [\hat{\mathbf{s}}_5^{\text{r.v.}} \quad \hat{\mathbf{s}}_2^{\text{r.v.}} \quad \hat{\mathbf{s}}_4^{\text{r.v.}}][\hat{\mathbf{r}}_5 \quad \hat{\mathbf{r}}_2 \quad \hat{\mathbf{r}}_4]^{\text{T}} \tag{15}$$

The new indices follow the conventions of reference [14]. The expressions for the TRIAD-II estimate and the true value of the TRIAD-II estimate follow similarly. The TRIAD-II attitude information matrix is given analogously by

$$\begin{aligned} F_{\xi\xi}^{\text{TRIAD-II}} = (P_{\xi\xi}^{\text{TRIAD-II}})^{-1} &= \frac{1}{\sigma_2^2} \hat{\mathbf{s}}_4^{\text{true}} \hat{\mathbf{s}}_4^{\text{true T}} + \frac{1}{\sigma_2^2} \hat{\mathbf{s}}_2^{\text{true}} \hat{\mathbf{s}}_2^{\text{true T}} + \frac{1}{\sigma_1^2} \hat{\mathbf{s}}_3^{\text{true}} \hat{\mathbf{s}}_3^{\text{true T}} \\ &= \frac{1}{\sigma_2^2} (I_{3\times 3} - \hat{\mathbf{s}}_5^{\text{true}} \hat{\mathbf{s}}_5^{\text{true T}}) + \frac{1}{\sigma_1^2} \hat{\mathbf{s}}_3^{\text{true}} \hat{\mathbf{s}}_3^{\text{true T}} \end{aligned} \tag{16}$$

The TRIAD-I and TRIAD-II Estimation Errors

In the present section we rely on the methods of references [2], [14], and [15], although it should not be necessary to refer to those works in order to follow the derivations here. Because $\{\hat{\mathbf{W}}_1^{\text{r.v.}}, \hat{\mathbf{s}}_2^{\text{r.v.}}, \hat{\mathbf{s}}_3^{\text{r.v.}}\}$ and $\{\hat{\mathbf{W}}_2^{\text{r.v.}}, \hat{\mathbf{s}}_2^{\text{r.v.}}, \hat{\mathbf{s}}_4^{\text{r.v.}}\}$ are each right-hand orthonormal triads, we may write, assuming the QUEST measurement model [2, 16],

$$\hat{\mathbf{W}}_1^{\text{r.v.}} = A\hat{\mathbf{V}}_1^{\text{true}} + v_1\, \hat{\mathbf{s}}_2^{\text{true}} + v_2\, \hat{\mathbf{s}}_3^{\text{true}} + O(\sigma_1^2) \tag{17a}$$

$$\hat{\mathbf{W}}_2^{\text{r.v.}} = A\hat{\mathbf{V}}_2^{\text{true}} + v_4\, \hat{\mathbf{s}}_2^{\text{true}} + v_5\, \hat{\mathbf{s}}_4^{\text{true}} + O(\sigma_2^2) \tag{17b}$$

where

$$v_1, v_2 \sim \mathcal{N}(0, \sigma_1^2) \qquad \text{and} \qquad v_4, v_5 \sim \mathcal{N}(0, \sigma_2^2) \tag{18ab}$$

are statistically-independent noise terms. The choice of indices for the noise terms in equations (17) and (18) follows reference [15] but is not significant here. Generally, we have not written "r.v." on v_k or on z_k (below), $k = 1, 2, 4, 5$. These are always random variables in the present work.

[8]It is Mortari [17] who first introduced the nomenclature TRIAD-I and TRIAD-II.
[9]Note that $\hat{\mathbf{W}}_1$ and $\hat{\mathbf{W}}_2$ are the same unit column vectors for both TRIAD-I and TRIAD-II.

Writing[10]

$$A = e^{[[\boldsymbol{\xi}]]} A^{\text{true}} \qquad \text{and} \qquad \hat{\mathbf{W}}_k^{\text{true}} = A^{\text{true}} \hat{\mathbf{V}}_k, \qquad k = 1, 2 \tag{19ab}$$

then from equations (17) we can write the four effective scalar measurements, the components of $\hat{\mathbf{W}}_1^{\text{r.v.}}$ and $\hat{\mathbf{W}}_2^{\text{r.v.}}$ normal to $\hat{\mathbf{W}}_1^{\text{true}}$ and $\hat{\mathbf{W}}_2^{\text{true}}$, respectively, to $O(\sigma_1^2, \sigma_2^2)$ as [15]

$$z_1 \equiv \hat{\mathbf{S}}_2^{\text{true}\,T} \hat{\mathbf{W}}_1^{\text{r.v.}} = -\hat{\mathbf{S}}_3^{\text{true}\,T} \boldsymbol{\xi} + v_1 \tag{20a}$$

$$z_2 \equiv \hat{\mathbf{S}}_3^{\text{true}\,T} \hat{\mathbf{W}}_1^{\text{r.v.}} = \hat{\mathbf{S}}_2^{\text{true}\,T} \boldsymbol{\xi} + v_2 \tag{20b}$$

$$z_4 \equiv \hat{\mathbf{S}}_2^{\text{true}\,T} \hat{\mathbf{W}}_2^{\text{r.v.}} = -\hat{\mathbf{S}}_4^{\text{true}\,T} \boldsymbol{\xi} + v_4 \tag{20c}$$

$$z_5 \equiv \hat{\mathbf{S}}_4^{\text{true}\,T} \hat{\mathbf{W}}_2^{\text{r.v.}} = \hat{\mathbf{S}}_2^{\text{true}\,T} \boldsymbol{\xi} + v_5 \tag{20d}$$

In the TRIAD-I construction, $A^{\text{TRIAD*}}$ is determined effectively from $\hat{\mathbf{W}}_1^{\text{r.v.}}$ and $\hat{\mathbf{S}}_3^{\text{r.v.}}$, since the remaining unit vector of the triad is superfluous. But by explicit construction $\hat{\mathbf{S}}_3^{\text{r.v.}}$ has no component along $\hat{\mathbf{W}}_1^{\text{r.v.}}$, hence not, to order σ_2^2, along $\hat{\mathbf{S}}_1^{\text{true}}$. Likewise, the random part of $\hat{\mathbf{S}}_3^{\text{r.v.}}$ has no component along $\hat{\mathbf{S}}_3^{\text{true}}$ because of the norm constraint. Thus, only the component of $\hat{\mathbf{S}}_3^{\text{r.v.}}$ along $\hat{\mathbf{S}}_2^{\text{true}}$ can contribute to the TRIAD attitude estimator, and this is also the component of $\hat{\mathbf{W}}_2^{\text{r.v.}}$ along $\hat{\mathbf{S}}_2^{\text{true}}$, i.e., z_4. This, microscopically, is the nature of the data truncation of the TRIAD-I algorithm which leads to an unambiguous answer, i.e., what makes the TRIAD algorithm deterministic.[11] This result is derived more explicitly and at somewhat greater length in reference [15]. Thus, we can write the cost function for the TRIAD-I estimate as[12]

$$J^{\text{TRIAD-I}}(A | \hat{\mathbf{W}}_1', \hat{\mathbf{W}}_2') = J^{\text{TRIAD}}(A | \hat{\mathbf{W}}_1', \hat{\mathbf{S}}_3') = J(A | z_1', z_2', z_4') \tag{21a}$$

By similar arguments

$$J^{\text{TRIAD-II}}(A | \hat{\mathbf{W}}_1', \hat{\mathbf{W}}_2') = J^{\text{TRIAD}}(A | \hat{\mathbf{W}}_2', \hat{\mathbf{S}}_4') = J(A | z_1', z_4', z_5') \tag{21b}$$

The optimization of the two cost functions of equations (21) leads to the TRIAD-I and TRIAD-II algorithms. The reformulation of the TRIAD algorithm as an optimal algorithm is the subject of reference [15]. Since the four random noise terms of equations (20) are zero-mean, Gaussian, and independent, it is a trivial matter to construct the two maximum-likelihood cost functions[13] of equations (21). Thus, in obvious notation,

$$J^{\text{TRIAD-I}}(A) = \frac{1}{2}\left[\frac{1}{\sigma_1^2}|z_1' + \hat{\mathbf{S}}_3^{\text{true}}\,\boldsymbol{\xi}|^2 + \frac{1}{\sigma_1^2}|z_2' - \hat{\mathbf{S}}_2^{\text{true}}\boldsymbol{\xi}|^2 + \frac{1}{\sigma_2^2}|z_4' + \hat{\mathbf{S}}_4^{\text{true}}\boldsymbol{\xi}|^2\right] \tag{22a}$$

$$J^{\text{TRIAD-II}}(A) = \frac{1}{2}\left[\frac{1}{\sigma_1^2}|z_1' + \hat{\mathbf{S}}_3^{\text{true}}\,\boldsymbol{\xi}|^2 + \frac{1}{\sigma_2^2}|z_4' + \hat{\mathbf{S}}_4^{\text{true}}\boldsymbol{\xi}|^2 + \frac{1}{\sigma_2^2}|z_5' + \hat{\mathbf{S}}_2^{\text{true}}\boldsymbol{\xi}|^2\right] \tag{22b}$$

[10]Note that we assume that A^{true} is known. This is not a defect of our derivation, since our goal is only to construct a model for the attitude estimation errors $\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}}$ and $\Delta\boldsymbol{\xi}^{\text{TRIAD-II*}}$. An implementable algorithm could always be constructed by using $A^{\text{TRIAD-I*}\prime}$ or $A^{\text{TRIAD-II*}\prime}$ as the reference attitude, but this is an avenue we do not need or wish to explore.

[11]Note that while the three triad vectors are interchangeable as geometric objects, as statistical objects, $\hat{\mathbf{S}}_1^{\text{r.v.}}$ is different from the other two.

[12]To be rigorous we should have included also the contribution of the two "constraint" measurements z_3 and z_6 (see reference [15]) to avoid discrete degeneracies in the attitude solutions. The constraint measurements, however, do not contribute to the attitude information matrix.

[13]That is, the appropriate portions of the respective negative-log-likelihood functions [13].

for which the minimizing values of the arguments lead to random (zero-mean) errors in the *estimators* given by[14]

$$\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}} = (F_{\xi\xi}^{\text{TRIAD-I}})^{-1}\Delta\mathbf{g}^{\text{TRIAD-I}} \quad \text{and} \quad \Delta\boldsymbol{\xi}^{\text{TRIAD-II*}} = (F_{\xi\xi}^{\text{TRIAD-II}})^{-1}\Delta\mathbf{g}^{\text{TRIAD-II}} \tag{23ab}$$

with $F_{\xi\xi}^{\text{TRIAD-I}}$ given by equation (8b), $F_{\xi\xi}^{\text{TRIAD-II}}$ given by equation (16), and the corresponding gradient error vectors given by

$$\Delta\mathbf{g}^{\text{TRIAD-I}} = -\frac{1}{\sigma_1^2}\,v_1\,\hat{\mathbf{S}}_3^{\text{true}} - \frac{1}{\sigma_2^2}\,v_4\,\hat{\mathbf{S}}_4^{\text{true}} + \frac{1}{\sigma_1^2}\,v_2\,\hat{\mathbf{S}}_2^{\text{true}} \tag{24a}$$

$$\Delta\mathbf{g}^{\text{TRIAD-II}} = -\frac{1}{\sigma_1^2}\,v_1\,\hat{\mathbf{S}}_3^{\text{true}} - \frac{1}{\sigma_2^2}\,v_4\,\hat{\mathbf{S}}_4^{\text{true}} + \frac{1}{\sigma_2^2}\,v_5\,\hat{\mathbf{S}}_2^{\text{true}} \tag{24b}$$

Note from equations (24) that the TRIAD-I and TRIAD-II attitude errors are correlated strongly, as averred earlier. The attitude errors for the two estimators along any axis in the plane of the true measurements, $\hat{\mathbf{W}}_1^{\text{true}}$ and $\hat{\mathbf{W}}_2^{\text{true}}$, have unit correlation. Only the attitude errors about the $\hat{\mathbf{S}}_2^{\text{true}}$ axis are uncorrelated for the two estimators.

In an optimization scheme it is the expectation of the sum of the squares of the random errors which we wish to minimize, so we write only the attitude error increments and not the estimators themselves (which would have identical formulas but with $\Delta\mathbf{g}$ replaced with $\mathbf{g}$ and v_k replaced by z_k, $k = 1, 2, 4, 5$).[15]

We now note an interesting fact. If we decompose the three-dimensional space into a one-dimensional subspace spanned by $\hat{\mathbf{S}}_2^{\text{true}}$ and a two-dimensional subspace spanned by $\hat{\mathbf{S}}_3^{\text{true}}$ and $\hat{\mathbf{S}}_4^{\text{true}}$ (equivalently, by $\hat{\mathbf{W}}_1^{\text{true}}$ and $\hat{\mathbf{W}}_2^{\text{true}}$, so that the two-dimensional subspace is the true measurement plane), then we can achieve a very convenient direct decomposition of the Fisher information matrices and gradient error vectors.[16] Also from equations (8) and (16) we see that the 2×2 submatrices of $F_{\xi\xi}^{\text{TRIAD-I}}$ and $F_{\xi\xi}^{\text{TRIAD-II}}$ in this direct decomposition are identical, as are the corresponding two-dimensional components of $\Delta\mathbf{g}^{\text{TRIAD-I}}$ and $\Delta\mathbf{g}^{\text{TRIAD-II}}$, which we can see from equations (24). It follows that we can write

$$\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}} = \Delta\boldsymbol{\xi}_{\text{o}} - v_2\,\hat{\mathbf{S}}_2^{\text{true}} \quad \text{and} \quad \Delta\boldsymbol{\xi}^{\text{TRIAD-II*}} = \Delta\boldsymbol{\xi}_{\text{o}} - v_5\,\hat{\mathbf{S}}_2^{\text{true}} \tag{25ab}$$

with $\hat{\mathbf{S}}_2^{\text{true}} \cdot \Delta\boldsymbol{\xi}_{\text{o}} = 0$. It follows that

$$\begin{aligned}\Delta\boldsymbol{\xi}_{\text{o}} &= -\mathcal{P}^{\text{T}}\left[\mathcal{P}\left(\frac{1}{\sigma_1^2}\,\hat{\mathbf{S}}_3^{\text{true}}\hat{\mathbf{S}}_3^{\text{true T}} + \frac{1}{\sigma_2^2}\,\hat{\mathbf{S}}_4^{\text{true}}\hat{\mathbf{S}}_4^{\text{true T}}\right)\mathcal{P}^{\text{T}}\right]^{-1}\mathcal{P}\left[\frac{1}{\sigma_1^2}\,v_1\hat{\mathbf{S}}_3^{\text{true}} + \frac{1}{\sigma_2^2}\,v_4\hat{\mathbf{S}}_4^{\text{true}}\right]\\ &= v_1\,\hat{\mathbf{S}}_3^{\text{true}} + v_4\,\hat{\mathbf{S}}_4^{\text{true}}\end{aligned} \tag{26}$$

where $\mathcal{P}$ is a projection operator onto the plane spanned by $\hat{\mathbf{W}}_1^{\text{true}}$ and $\hat{\mathbf{W}}_2^{\text{true}}$, and

$$E\{\Delta\boldsymbol{\xi}_{\text{o}}\Delta\boldsymbol{\xi}_{\text{o}}^{\text{T}}\} = \sigma_1^2\hat{\mathbf{S}}_3^{\text{true}}\hat{\mathbf{S}}_3^{\text{true T}} + \sigma_2^2\hat{\mathbf{S}}_4^{\text{true}}\hat{\mathbf{S}}_4^{\text{true T}} \equiv P_{\text{o}} \tag{27}$$

[14]See footnote 6.

[15]Note that because $\boldsymbol{\xi}$ is defined with respect to A^{true}, it follows that $\Delta\boldsymbol{\xi} = \boldsymbol{\xi}$ and $\Delta\mathbf{g} = \mathbf{g}$. Had we defined $\boldsymbol{\xi}$ with respect to, say, $A^{\text{TRIAD-I*}\prime}$, then $\mathbf{g}^{\text{r.v.}}$ and $\boldsymbol{\xi}^{\text{r.v.}}$ would have been biased by the amounts $E\{\mathbf{g}^{\text{r.v.}}\}$ and $E\{\boldsymbol{\xi}^{\text{r.v.}}\}$, respectively, where $E\{\,\cdot\,\}$ denotes the expectation. The gradient error vector (or gradient-vector error) is defined as $\Delta\mathbf{g}^{\text{r.v.}} \equiv \mathbf{g}^{\text{r.v.}} - E\{\mathbf{g}^{\text{r.v.}}\}$ and given still by equations (24), and similarly for $\boldsymbol{\xi}$ and $\Delta\boldsymbol{\xi}$. The companion matrix [13] will still be equal to the Fisher information matrix.

[16]The information matrix is the Fisher information matrix over the minimal three-dimensional measurement space.

Thus,

$$E\{\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}}\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}\,\text{T}}\} = P_{\text{o}} + \sigma_1^2\,\hat{\mathbf{s}}_2^{\text{true}}\,\hat{\mathbf{s}}_2^{\text{true}\,\text{T}} \tag{28a}$$

$$E\{\Delta\boldsymbol{\xi}^{\text{TRIAD-II*}}\Delta\boldsymbol{\xi}^{\text{TRIAD-II*}\,\text{T}}\} = P_{\text{o}} + \sigma_2^2\,\hat{\mathbf{s}}_2^{\text{true}}\,\hat{\mathbf{s}}_2^{\text{true}\,\text{T}} \tag{28b}$$

$$E\{\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}}\Delta\boldsymbol{\xi}^{\text{TRIAD-II*}\,\text{T}}\} = P_{\text{o}} \tag{28c}$$

$$E\{\Delta\boldsymbol{\xi}^{\text{QUEST*}}\Delta\boldsymbol{\xi}^{\text{QUEST*}\,\text{T}}\} = P_{\text{o}} + \sigma_{\text{tot}}^2\,\hat{\mathbf{s}}_2^{\text{true}}\,\hat{\mathbf{s}}_2^{\text{true}\,\text{T}} \tag{28d}$$

The TRIAD-I and TRIAD-II estimation errors are very strongly correlated, as one should expect, since they each rely on the same directional data.

A More Optimal Attitude Estimation Algorithm

Let us consider $\Delta\boldsymbol{\xi}^{\text{trial}}$, a linear combination of the TRIAD-I and TRIAD-II estimation errors (which are three-dimensional not scalars). We find directly because of the specific nature of the direct decomposition that

$$\Delta\boldsymbol{\xi}^{\text{trial}} \equiv a_1\Delta\boldsymbol{\xi}^{\text{TRIAD-I*}} + a_2\Delta\boldsymbol{\xi}^{\text{TRIAD-II*}} = \Delta\boldsymbol{\xi}_{\text{o}} + (a_1v_2 + a_2v_5)\hat{\mathbf{s}}_2^{\text{true}} \tag{29}$$

where $\Delta\boldsymbol{\xi}_{\text{o}}$ is independent of the yet unspecified a_1 and a_2 and uncorrelated with v_2 and v_5. It is obvious from equations (26) and (29) that $\Delta\boldsymbol{\xi}_{\text{o}}$ is independent of the remaining portion of $\Delta\boldsymbol{\xi}^{\text{trial}}$.

Because $\Delta\boldsymbol{\xi}_{\text{o}}$ is independent of the remaining terms of $\Delta\boldsymbol{\xi}^{\text{trial}}$, we have that the attitude covariance matrix corresponding to $\Delta\boldsymbol{\xi}^{\text{trial}}$ must be

$$P_{\xi\xi}^{\text{trial}} = P_{\text{o}} + (a_1^2\sigma_1^2 + a_2^2\sigma_2^2)\,\hat{\mathbf{s}}_2^{\text{true}}\hat{\mathbf{s}}_2^{\text{true}\,\text{T}} \tag{30}$$

with $E\{\cdot\}$ again the expectation. Thus, in the subspace perpendicular to the true measurement plane, we have an isolated scalar linear Gaussian estimation problem with two uncorrelated noise sources, one with variance σ_1^2 and the other with variance σ_2^2. The covariance matrix is minimized, therefore, if we choose

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \frac{\sigma_{\text{tot}}^2}{\sigma_1^2} \qquad \text{and} \qquad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} = \frac{\sigma_{\text{tot}}^2}{\sigma_2^2} \tag{31ab}$$

with

$$\frac{1}{\sigma_{\text{tot}}^2} \equiv \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \tag{32}$$

which are the optimal weights for the Wahba problem for two measurements given the QUEST Measurement Model [2, 16] (see equation (12) above). The Wahba problem for two direction measurements seeks an attitude estimator which minimizes the cost function

$$J(A^{\text{r.v.}}) = \frac{1}{2}\left[a_1|\hat{\mathbf{W}}_1^{\text{r.v.}} - A^{\text{r.v.}}\hat{\mathbf{V}}_1|^2 + a_2|\hat{\mathbf{W}}_2^{\text{r.v.}} - A^{\text{r.v.}}\hat{\mathbf{V}}_2|^2\right] \tag{33}$$

When the two weights are chosen in accordance with equations (31) the QUEST estimator becomes the maximum-likelihood estimator for the attitude assuming the QUEST measurement model [2, 16].

This leads to the result for the optimal $\Delta\boldsymbol{\xi}^{\text{trial*}}$, which we now denote by $\Delta\boldsymbol{\xi}^{\text{opt*}}$

$$P_{\xi\xi}^{\text{opt}} = \sigma_1^2\hat{\mathbf{s}}_3^{\text{true}}\hat{\mathbf{s}}_3^{\text{true}\,\text{T}} + \sigma_2^2\hat{\mathbf{s}}_4^{\text{true}}\hat{\mathbf{s}}_4^{\text{true}\,\text{T}} + \sigma_{\text{tot}}^2\hat{\mathbf{s}}_2^{\text{true}}\hat{\mathbf{s}}_2^{\text{true}\,\text{T}} \tag{34}$$

The inverse attitude covariance matrix for the Wahba problem for two measurements given the QUEST measurement model is [2]

$$
\begin{aligned}
(P_{\xi\xi}^{\text{QUEST}})^{-1} &= \frac{1}{\sigma_1^2}(I_{3\times 3} - \hat{\mathbf{W}}_1^{\text{true}}\hat{\mathbf{W}}_1^{\text{true T}}) + \frac{1}{\sigma_2^2}(I_{3\times 3} - \hat{\mathbf{W}}_2^{\text{true}}\hat{\mathbf{W}}_2^{\text{true T}}) \\
&= \frac{1}{\sigma_1^2}\hat{\mathbf{S}}_3^{\text{true}}\hat{\mathbf{S}}_3^{\text{true T}} + \frac{1}{\sigma_2^2}\hat{\mathbf{S}}_4^{\text{true}}\hat{\mathbf{S}}_4^{\text{true T}} + \frac{1}{\sigma_{\text{tot}}^2}\hat{\mathbf{S}}_2^{\text{true}}\hat{\mathbf{S}}_2^{\text{true T}}
\end{aligned} \tag{35}
$$

It is a simple matter to verify that the expression in equation (35) is the inverse of that in equation (34). It follows that[17]

$$
P_{\xi\xi}^{\text{opt}} = P_{\xi\xi}^{\text{QUEST}} + O(\sigma_1^4, \sigma_2^4) \qquad \text{and} \qquad A_{\xi\xi}^{\text{opt*}} = A_{\xi\xi}^{\text{QUEST*}} + O(\sigma_1^2, \sigma_2^2) \tag{36ab}
$$

$P_{\xi\xi}^{\text{QUEST}}$ is the Cramér-Rao lower bound [13] for the attitude covariance matrix. Since the estimator which satisfies the Cramér-Rao lower bound is unique,[18] it follows that the optimal trial estimator is the QUEST estimator, i.e., it is the solution of the Wahba problem to $O(\sigma_1^2, \sigma_2^2)$, and we write it as A^{opt}, the value of A^{trial} for the optimal values of a_1 and a_2. We note that the expression for the covariance matrix, owing to the smallness of the measurement error levels, is generally computed only to lowest order in σ_1 and σ_2, that is to relative order $O(\sigma_1^2, \sigma_2^2)$. The QUEST measurement model, on which such calculations are always based, is, in fact, valid to only that order.

Note also that

$$
a_1(P_{\xi\xi}^{\text{TRIAD-I}})^{-1} + a_2(P_{\xi\xi}^{\text{TRIAD-II}})^{-1} = (P_{\xi\xi}^{\text{QUEST}})^{-1} \tag{37}
$$

which follows directly from equations (8), (16), and (35).

Given $\boldsymbol{\xi}^{\text{opt*}}$, the optimal value of $\boldsymbol{\xi}^{\text{trial}}$, it follows that

$$
\begin{aligned}
A^{\text{opt*}} &= e^{[[\boldsymbol{\xi}^{\text{opt*}}]]}A^{\text{true}} \\
&= e^{a_1[[a_1\boldsymbol{\xi}^{\text{TRIAD-I*}}]]+a_2[[a_1\boldsymbol{\xi}^{\text{TRIAD-II*}}]]}A^{\text{true}} \\
&= [I_{3\times 3} + a_1[[a_1\boldsymbol{\xi}^{\text{TRIAD-I*}}]] + a_2[[a_1\boldsymbol{\xi}^{\text{TRIAD-II*}}]]]A^{\text{true}} + O(\sigma_1^2, \sigma_2^2) \\
&= a_1A^{\text{TRIAD-I*}} + a_2A^{\text{TRIAD-II*}} + O(\sigma_1^2, \sigma_2^2) \\
&= \text{ortho}(a_1A^{\text{TRIAD-I*}} + a_2A^{\text{TRIAD-II*}}) + O(\sigma_1^2, \sigma_2^2)
\end{aligned} \tag{38}
$$

where ortho(M) gives the proper orthogonal matrix $M^{\text{prop-orth}}$ which is closest to M in the Frobenius norm, that is,

$$
\text{ortho}(M) \equiv \arg \min_{A\in SO(3)} \left\|M - A\right\|_{\text{Frobenius}} \equiv \arg \min_{A\in SO(3)} \sum_{i=1}^{3}\sum_{j=1}^{3}|M_{ij} - A_{ij}|^2 \tag{39}
$$

Such a proper orthogonal matrix can be found obviously by letting the realization of the attitude profile matrix B be given by

$$
B' = a_1A_1^{\text{TRIAD*}\prime} + a_2A^{\text{TRIAD-II*}\prime} \tag{40}
$$

in the Wahba problem or, equivalently, by extracting the quaternion $\bar{q}$ from M (assuming M is nearly proper orthogonal), unitizing it, and recomputing the attitude matrix from the unitized $\bar{q}$. Markley [18, 19] has presented a modification of

[17]We write $P_{\xi\xi}^{\text{QUEST}}$ not because of its association with the QUEST attitude estimation algorithm but because of its association with the QUEST measurement model. The Wahba problem is independent of any stochastic measurement model. However, for the QUEST measurement model the solution to the Wahba problem becomes the maximum-likelihood estimate of the attitude to $O(\sigma_1^2, \sigma_2^2)$ [16].

[18]Note that the linearization of the attitude estimation problem has discarded terms of order σ_1^2 and σ_2^2. Therefore, uniqueness is also true only to this order. This is all we desire.

Shepperd's algorithm [20] which performs this task efficiently without the possibility of a singularity in an intermediate step.[19]

We call this optimal attitude the optimized TRIAD attitude and write

$$A^{\text{opt-TRIAD}*} \equiv \text{ortho}(a_1 A^{\text{TRIAD-I}*} + a_2 A^{\text{TRIAD-II}*}) \tag{41}$$

We note that

$$A^{\text{opt-TRIAD}*} = A^{\text{QUEST}*} + O(\sigma_1^2, \sigma_2^2) \tag{42}$$

Of particular interest is a result of Markley [18], which we may write as

$$A^{\text{QUEST}*} = \frac{1}{\lambda_{\max}}(a_1 A^{\text{TRIAD-I}*} + a_2 A^{\text{TRIAD-II}*}) + \frac{\lambda_{\max} - a_1 - a_2}{\lambda_{\max}} \hat{\mathbf{s}}_2 \hat{\mathbf{r}}_2^{\mathrm{T}} \tag{43}$$

where $\lambda_{\max}$ is the maximum value of the Wahba gain-function [2], which is also the maximum overlap eigenvalue of the equivalent Davenport q-algorithm [2, 6].[20]

For two measurements there is an exact closed-form expression [2] for $\lambda_{\max}$, namely,

$$\lambda_{\max} = \sqrt{a_1^2 + 2a_1 a_2 \cos \Delta\theta + a_2^2} \tag{44}$$

with a_1 and a_2 as before and

$$\cos \Delta\theta = (\hat{\mathbf{W}}_1 \cdot \hat{\mathbf{W}}_2)(\hat{\mathbf{V}}_1 \cdot \hat{\mathbf{V}}_2) + |\hat{\mathbf{W}}_1 \times \hat{\mathbf{W}}_2||\hat{\mathbf{V}}_1 \times \hat{\mathbf{V}}_2| \tag{45}$$

As a result, Markley's formula, equation (43), is the most efficient way to compute the exactly orthogonal optimal attitude estimate from the weighted average of the two TRIAD estimates.[21]

Equation (43) holds independently of the condition on the sum $a_1 + a_2$ provided that $a_i = c/\sigma_i^2$, $i = 1, 2$, for some common c. Equation (43) may be expanded as

$$A^{\text{QUEST}*} = a_1 A^{\text{TRIAD-I}*} + a_2 A^{\text{TRIAD-II}*} + O(\sigma_{\text{tot}}^2) \tag{46}$$

because [22]

$$\lambda_{\max} = (a_1 + a_2)\left(1 - \frac{3}{2}\sigma_{\text{tot}}^2 \chi^2(3)\right) \tag{47}$$

where $\chi^2(n)$ denotes a chi-square random variable with n degrees of freedom.

Thus, finally,

$$A^{\text{opt-TRIAD}*} = A^{\text{QUEST}*} + O(\sigma_{\text{tot}}^2) \qquad \text{and} \qquad P_{\xi\xi}^{\text{opt-TRIAD}} = P_{\xi\xi}^{\text{QUEST}} \tag{48ab}$$

The Optimized TRIAD Algorithm of Bar-Itzhack and Harman

The optimized TRIAD algorithm is virtually identical to the prescription published by Bar-Itzhack and Harman in 1997 as the optimized TRIAD algorithm [9], the only difference *in implementation* being that reference [9] orthogonalizes the weighted average of the TRIAD-I and TRIAD-II estimates only to lowest order using the first iteration of an infinite process [23]. Such a process will not be

[19]These orthogonalization procedures are not necessarily equivalent. However, since $a_1 A^{\text{TRIAD-I}*} + a_2 A^{\text{TRIAD-II}*}$ is already proper orthogonal to first order in the measurement variances, it hardly matters.

[20]Markley obtained equation (43) (derivation not published) not from an error analysis but from an examination of the FOAM expression [21] for the estimated attitude matrix with only two measurements. Note that $\lambda_{\max}$, because it depends on the measurements, is also a random variable.

[21]Equation (43) is equivalent to equation (40). Equation (40) does not provide a special case for the formula for ortho(M) when $M = a_1 A^{\text{TRIAD-I}*\prime} + a_2 A^{\text{TRIAD-II}*\prime}$, but is an acceptable substitute.

adequate for coarse sensors if one wishes to achieve a numerical result consistent with IEEE double precision. More important, however, is that the proposed procedure of reference [9] is based on explicit and implicit assumptions which are contradicted by the careful analysis presented here.

The development of the optimized TRIAD prescription begins with two *ad hoc* assumptions about the TRIAD-I and TRIAD-II estimates. The first assumption (implicit) is that these two estimates are independent. This is certainly not true as demonstrated by equation (25). The TRIAD-I and TRIAD-II estimates are very strongly correlated. The second assumption (explicit) is that the error in the TRIAD-I estimate is characterized entirely by σ_1 and that for the TRIAD-II estimate by σ_2. This cannot be true, obviously, because each estimate makes use of *both* $\hat{\mathbf{W}}'_1$ and $\hat{\mathbf{W}}'_2$, and, as demonstrated unequivocally by equations (8) and (17), the covariance matrices of the two attitude estimates depend to equal order on both variances. A third assumption is that a more optimal attitude estimation should be given, apart from the question or proper orthogonality, by the form

$$\alpha\, A^{\mathrm{TRIAD}} + \beta\, A^{\mathrm{TRIAD\text{-}II}}$$

which assumes (explicitly) that one can combine matrices in the same way as one can combine scalars. Reference [9] offers no derivation, it merely states its assertions and proceeds to something like equation (41). We have seen that this fact depends on very fine details of the structure of the TRIAD-I and TRIAD-II estimate errors.

Discussion and Conclusions

We have shown rigorously that the linear combination $a_1\, A^{\mathrm{TRIAD\text{-}I}*} + a_2\, A^{\mathrm{TRIAD\text{-}II}*}$, with a_1 and a_2 the unit-sum weights of the correspondingly terms of the Wahba cost function, is equal within errors of order σ_{tot}^2 to the optimal QUEST estimator of the attitude. This result followed from a detailed examination of the estimation errors of the TRIAD-I and TRIAD-II algorithms, and from the rigorous derivation of the optimized TRIAD attitude estimation algorithm. The result follows also from an earlier result of Markley [18]. The present work illustrates the importance of a rigorous mathematical foundation and a thorough mathematical analysis of an attitude estimation algorithm, especially of the attitude covariance matrix.

The optimality of the optimized TRIAD algorithm does not follow, certainly, from the assertions of reference [9]. That work, in fact, claimed only to lead to a more accurate estimator, not *the* optimal estimator. Rather, the optimality of the optimized TRIAD algorithm is due to fine details of the structure of the TRIAD attitude estimator and attitude covariance matrix.

Given the basis for the optimized TRIAD prescription assumed in reference [9], it is surprising that the prescription turned out to perform so well. Reference [9] is, in fact, unaware that the optimized TRIAD prescription is optimal to order σ_{tot}^2. Strangely, while the simulation tests of reference [9] seem to be correct, that work is distraught that the sampled mean of the attitude errors does not seem to tend toward zero as the number of samples tends toward infinity. However, the attitude error, as defined in reference [9], is intrinsically non-negative and, therefore, since it is not identically zero, intrinsically of non-zero mean. Despite the deficiencies in its development, the optimized TRIAD algorithm performs very well, as we know now from the present work.

Of greater importance for this work are our results for the structure of the TRIAD-I and TRIAD-II estimators, namely, that

$$\Delta \boldsymbol{\xi}^{\text{TRIAD-I*}} = \Delta \boldsymbol{\xi}_{\text{o}} + \Delta \theta^{\text{TRIAD-I*}} \hat{\mathbf{s}}_2^{\text{true}} \tag{49a}$$

$$\Delta \boldsymbol{\xi}^{\text{TRIAD-II*}} = \Delta \boldsymbol{\xi}_{\text{o}} + \Delta \theta^{\text{TRIAD-II*}} \hat{\mathbf{s}}_2^{\text{true}} \tag{49b}$$

with $\Delta\theta^{\text{TRIAD-I}}$ and $\Delta\theta^{\text{TRIAD-II}}$ error angles uncorrelated with $\Delta\boldsymbol{\xi}_{\text{o}}$ and with each other, and $\hat{\mathbf{s}}_2$ orthogonal to $\Delta\boldsymbol{\xi}_{\text{o}}$. The statistical arguments advanced wrongly for $A^{\text{TRIAD-I}}$ and $A^{\text{TRIAD-II}}$ in reference [9], in fact, are true for $\Delta\theta^{\text{TRIAD-I}}$ and $\Delta\theta^{\text{TRIAD-II}}$. It is this fact which is responsible for the good performance of the optimized TRIAD algorithm, not the baseless statistical arguments advanced in reference [9]. Equations (49ab) should be compared with a similar result for the QUEST estimate

$$\Delta \boldsymbol{\xi}^{\text{QUEST*}} = \Delta \boldsymbol{\xi}_{\text{o}} + \Delta \theta^{\text{QUEST*}} \hat{\mathbf{s}}_2^{\text{true}} \tag{49c}$$

That a linear combination of the two TRIAD estimates should be equal to order σ_{tot}^2 should not be a surprise. We know that $\hat{\mathbf{W}}_1'$, $\hat{\mathbf{W}}_2'$, $A^{\text{TRIAD-I*}\prime}\,\hat{\mathbf{V}}_1$, $A^{\text{TRIAD-I*}\prime}\,\hat{\mathbf{V}}_2$, $A^{\text{TRIAD-II*}\prime}\,\hat{\mathbf{V}}_1$, and $A^{\text{TRIAD-II*}\prime}\,\hat{\mathbf{V}}_2$ are all coplanar and all differ only by a single random rotation about the $\hat{\mathbf{s}}_2^{\text{true}}$ axis. Thus, some linear combination of the two TRIAD estimates should yield something very close to the optimal result. That this combination should be simply the weighted average of the two TRIAD estimates with coefficients a_1 and a_2 requires a careful analysis, however.

In the above work we have placed considerable emphasis on corrections being of order σ_1^2 and σ_2^2. The definition of the QUEST attitude is exact, so it is possible to speak of exact solutions to the Wahba problem and of solutions which are only to order σ_{tot}^2. However, if we wish to consider the Wahba problem as a maximum-likelihood estimation problem, so that the coefficients of the Wahba problem will be meaningful statistically, then we must be aware that this is possible only to order σ_{tot}^2. The QUEST measurement model, in fact, is mathematically consistent only to that order.

The optimized TRIAD algorithm is certainly not a candidate algorithm for mission support, because it imposes a significantly greater computational burden than the QUEST algorithm and is limited to only two measurements. In this sense also *materiam superabat opus.* The interest of the optimized TRIAD algorithm lies in the new insights it provides on the TRIAD algorithm, which has surely still not revealed all its secrets.

Markley's construction (equation (43)) may provide the most efficient way to calculate the QUEST attitude matrix for two measurements, namely,

$$A^{\text{QUEST*}} = \frac{a_1}{\lambda_{\max}}(\hat{\mathbf{W}}_1\hat{\mathbf{V}}_1^{\text{T}} + \hat{\mathbf{s}}_3\hat{\mathbf{r}}_3^{\text{T}}) + \frac{a_2}{\lambda_{\max}}(\hat{\mathbf{W}}_2\hat{\mathbf{V}}_2^{\text{T}} + \hat{\mathbf{s}}_4\hat{\mathbf{r}}_4^{\text{T}}) + \hat{\mathbf{s}}_2\hat{\mathbf{r}}_2^{\text{T}} \tag{50}$$

which appears in reference [18].

In a sense, the optimized TRIAD algorithm is not a TRIAD algorithm at all, because it does not produce an attitude solution which is the product of two proper orthogonal matrices, each constructed from a right-hand orthonormal triad, one from the observation vectors and one from the reference vectors. There are, in fact, many TRIAD algorithms, at least seven that have been investigated [24], of which TRIAD-I and TRIAD-II are two examples. One of these TRIAD algorithms turns out to be the QUEST attitude solution to $O(\sigma_1^2, \sigma_2^2)$ [24].

Acknowledgments

The author is grateful to John L. Crassidis, F. Landis Markley, Daniele Mortari, Yang Cheng, and an anonymous reviewer for interesting discussions and helpful criticisms. Sadly, Professor Bar-Itzhack passed away on May 9, 2007.

References

[1] BLACK, H. D. "A Passive System for Determining the Attitude of a Satellite," *AIAA Journal,* Vol. 2, July 1964, pp. 1350–1351.

[2] SHUSTER, M. D. and OH, S. D. "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control,* Vol. 4, No. 1, Jan.-Feb. 1981, pp. 70–77.

[3] WAHBA, G. "Problem 65-1: A Least Squares Estimate of Spacecraft Attitude," *SIAM Review,* Vol. 7, No. 3, July 1965, p. 409.

[4] FARRELL, J. L., STUELPNAGEL, J. C., WESSNER, R. H., and VELMAN, J. R. "Solution to Problem 65-1," *SIAM Review,* Vol. 8, No. 3, July 1966, pp. 384–386.

[5] LERNER, G. M. "Three-Axis Attitude Determination," in WERTZ, J. R. (ed.), *Spacecraft Attitude Determination and Control,* Springer, Berlin and New York, 1978, pp. 420–428.

[6] MARKLEY, F. L. and MORTARI, M. "Quaternion Attitude Estimation Using Vector Measurements, *The Journal of the Astronautical Sciences,* Vol. 48, Nos. 2 and 3, April–September 2000, pp. 359–380.

[7] CHENG, Y. and SHUSTER, M. D. "Robustness and Accuracy of the QUEST Algorithm," Paper AAS-07-102, *17th Space Flight Mechanics Meeting,* Sedona Arizona, January 28–February 2, 2007; Proceedings: *Advances in the Astronautical Sciences,* Vol. 127, 2007, pp. 41–61.

[8] CHENG, Y. and SHUSTER, M. D. "The Speed of Attitude Estimation," Paper AAS-07-105, *17th Space Flight Mechanics Meeting,* Sedona Arizona, January 28–February 2, 2007; Proceedings: *Advances in the Astronautical Sciences,* Vol. 127, 2007, pp. 101–116.

[9] BAR-ITZHACK, I. Y. and HARMAN, R. R. "Optimized TRIAD Algorithm for Attitude Determination," *Journal of Guidance and Control,* Vol. 20, No. 1, January–February 1997, pp. 208–211.

[10] SHUSTER, M. D. "A Survey of Attitude Representations," *The Journal of the Astronautical Sciences,* Vol. 41, No. 4, October–December 1993, pp. 439–517.

[11] SHUSTER, M. D. "Constraint in Attitude Estimation Part I: Constrained Estimation," M. D. Shuster, *The Journal of the Astronautical Sciences,* Vol. 51, No. 1, January–March 2003, pp. 51–74.

[12] SHUSTER, M. D. "Constraint in Attitude Estimation Part II: Unconstrained Estimation," *The Journal of the Astronautical Sciences,* Vol. 51, No. 1, January–March 2003, pp. 75–101.

[13] SORENSEN, H. W. *Parameter Estimation,* Marcel Dekker, New York 1980.

[14] CHENG, Y. and SHUSTER, M. D. "QUEST and the Anti-QUEST: Good and Evil Attitude Estimation," *The Journal of the Astronautical Sciences,* Vol. 53, No. 3, July–September 2005, pp. 337–351.

[15] SHUSTER, M. D. "The TRIAD Algorithm as Maximum-Likelihood Estimation," *The Journal of the Astronautical Sciences,* Vol. 54, No. 1, January–March 2006, pp. 113–123.

[16] SHUSTER, M. D. "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences,* Vol. 37, No. 1, January–March 1989, pp. 79–88.

[17] MORTARI, D. "EULER-2 and EULER-n Algorithms for Attitude Determination from Vector Observations," *Space Technology,* Vol. 16, Nos. 5–6, 1996, pp. 317–321.

[18] MARKLEY, F. L. "Attitude Determination Using Two Vector Measurements," *Proceedings, Flight Mechanics Symposium,* NASA Goddard Space Flight Center, Greenbelt, MD, May 1999, NASA Conference Publication NASA/CP-19989-209235, pp. 39–52.

[19] MARKLEY, F. L. "Unit Quarternion from Rotation Matrix," submitted to the *Journal of Guidance, Control and Dynamics.*

[20] SHEPPERD, S. W. "Quaternion from Rotation Matrix," *Journal of Guidance and Control,* Vol. 1, 1978, pp. 223–224.

[21] MARKLEY, F. L. "Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm," *The Journal of the Astronautical Sciences,* Vol. 41, No. 2, April–June 1993, pp. 261–280.

[22] SHUSTER, M. D. and FREESLAND, D. C. "The Statistics of TASTE and the Inflight Estimation of Sensor Precision," *Proceedings, NASA Flight Mechanics Symposium,* NASA Goddard Space Flight Center, Greenbelt, Maryland, October 18–20, 2005.

[23] BAR-ITZHACK, I. Y. and MEYER, J. "On the Convergence of Iterative Orthogonalization Processes," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-12, March 1976, pp. 146–151.

[24] TANYGIN, S. and SHUSTER, M. D. "The Many TRIAD Algorithms," Paper AAS-07-104, AAS/AIAA Space Flight Mechanics, Meeting, Sedona, Arizona, January 28–February 22, 2007; Proceedings: *Advances in the Astronautical Sciences,* Vol. 127, 2007, pp. 81–99.

Note added in proof:

The author has learned that F. Landis Markley has prepared an engineering note "Optimal Attitude Matrix from Two Vector Measurements" for eventual submission to the *Journal of Guidance, Control and Dynamics.* This note will contain a detailed derivation of equations (43) and (50) above.