AIAA PAPER
NO. 78-1249

APPROXIMATE ALGORITHMS FOR FAST OPTIMAL ATTITUDE COMPUTATION

MALCOLM D. SHUSTER
Computer Sciences Corporation
System Sciences Division
Silver Spring, Maryland 20910

AIAA GUIDANCE AND CONTROL CONFERENCE

PALO ALTO, CALIFORNIA/AUGUST 7-9, 1978

APPROXIMATE ALGORITHMS FOR FAST OPTIMAL ATTITUDE COMPUTATION†

Malcolm D. Shuster*

Computer Sciences Corporation, System Sciences Division
Silver Spring, Maryland 20910

Abstract

Fast accurate algorithms are presented for computing an optimal attitude which minimizes a quadratic loss function. These algorithms compute an optimal rotation which carries a set of reference vectors into a set of corresponding observation vectors. Simplifications of these algorithms are obtained for the case of small rotation angles. Applications to the Magsat mission are discussed.

I. Introduction

Spacecraft typically carry a sufficient number of sensors that the attitude is overdetermined. This permits a more accurate determination of spacecraft attitude than could be obtained with a smaller number of sensors. A method must be devised, however, for combining these sensor measurements to obtain the best possible determination of the spacecraft attitude.

We present here a fast method for constructing the optimal estimate of spacecraft attitude given n reference unit vectors, n observation unit vectors, and a quadratic loss function in these vectors. This method is based on an algorithm of P. Davenport,[1] which determines an optimal estimate of the rotation which carries a set of reference unit vectors into a set of observation unit vectors.

In the application of this algorithm to spacecraft attitude determination, the observation unit vectors** $\hat{W}_1, \ldots, \hat{W}_n$ are the directions of the Earth, the Sun, a star, or some other object, in the body-fixed coordinate system of the spacecraft. The reference unit vectors, $\hat{V}_1, \ldots, \hat{V}_n$, are the directions of these same objects in some reference coordinate system (often a spacecraft-centered inertial coordinate system).

Ideally, the observation vectors may be obtained by applying a rotation to the reference vectors. The rotation matrix which accomplishes this is the attitude matrix of the spacecraft in the reference coordinate system. In practice, this attitude matrix cannot be constructed unambiguously from the reference and observation vectors due to errors in the measurement of the latter.

It is possible, however, to determine unambiguously the rotation which minimizes a quadratic loss function. The loss function studied here is

$$\mathcal{L}(R) = \frac{1}{2}\sum_{i=1}^{n} a_i |\hat{W}_i - R\,\hat{V}_i|^2 \qquad (1)$$

where the a_i, $i = 1, \ldots, n$, are a set of (positive) weights. Usually, the reference unit vectors are better known than the observation unit vectors. In this case, the a_i are inversely proportional to the mean square angular errors of observation.

If an exact simultaneous rotation of the reference unit vectors into the observation unit vectors is possible, $\mathcal{L}(R)$ can be made to vanish. In this case, the rotation matrix (the attitude) is determined exactly. Generally, this is not the case and $\mathcal{L}(R)$ is always larger than some minimum positive value. The rotation R_{opt} which minimizes the loss function of Equation (1) is the optimal estimate of the rotation which carries the set of reference vectors into the set of observation vectors.

The optimization procedure sketched above is not universally applicable to attitude determination since it is assumed that the sensor measurements can be related to the direction of some object. For the majority of attitude sensors (star sensors, two-axis Sun sensors, vector magnetometers, and horizon scanners), this is indeed the case. Gyros are an example of a sensor for which the methods presented here are not applicable.

The problem of finding the optimal rotation which minimizes the loss function above has been studied by Davenport,[1] who showed that the quaternion representing this optimal rotation (the optimal quaternion) was the eigenvector belonging to the largest eigenvalue of a 4 x 4 matrix. Davenport's algorithm has been examined by Keat,[2] who analyzed its use in software for the High Energy Astronomy Observatory-1 (HEAO-1) mission.

In the implementation of Davenport's q-algorithm in the HEAO-1 mission software, the eigenvalue problem was solved using general methods for determining the eigenvalues and eigenvectors of a square matrix of arbitrary rank. This makes the algorithm both rather slow and costly of computer core storage. It is thus unsuitable for processing on an on-board computer or when the algorithm must be executed with high frequency. This was not a consideration in the HEAO-1

†Work supported by the Magsat Project Office, Goddard Space Flight Center, National Aeronautics and Space Administration, under Contract No. NAS 5-24300.

*Member AIAA.

**Unit vectors will always be denoted by a caret.

attitude ground support software because the algorithm was invoked only infrequently. It is, however, a consideration for Magsat where the attitude must be computed every quarter second.

We develop here a more direct method for determining the optimal quaternion which exploits the special properties of the 4 × 4 matrix. A study[3] of this method for the sensor configuration of the Magsat mission showed it to be capable of yielding very accurate results with relatively little computation. This method, therefore, is very fast.

In this paper, we first derive Davenport's q-algorithm, (i.e., we obtain the eigenvalue equation for the quaternion representation of the optimal rotation). We then develop a fast method for constructing the exact solution to this eigenvalue equation. A number of approximations are presented both for the general case and when the attitude matrix is close to the identity matrix (null attitude). The accuracy of these algorithms is discussed for the sensor configuration of the Magsat mission.

II. Davenport's q-Algorithm*

Let $\hat{W}_1, \hat{W}_2, \ldots, \hat{W}_n$ be a set of n observation unit vectors and $\hat{V}_1, \hat{V}_2, \ldots, \hat{V}_n$ be a set of n reference unit vectors. Let $a_1, a_2, \ldots, a_n$ be a set of n (positive) weights. For any rotation matrix R, the loss $L(R)$ is defined as

$$L(R) = \frac{1}{2}\sum_{i=1}^{n} a_i |\hat{W}_i - R\hat{V}_i|^2 \qquad (1)$$

The rotation matrix R_{opt} which minimizes this loss function is said to be the optimal rotation (in a least squares sense) which carries the vectors $\hat{V}_i$, $i = 1, \ldots, n$, into the vectors $\hat{W}_i$, $i = 1, \ldots, n$.

Since the loss function may be scaled without affecting the determination of the optimal rotation, it is possible to set

$$\sum_{i=1}^{n} a_i = 1 \qquad (2)$$

Then

$$L(R) = 1 - g(R) \qquad (3)$$

with

$$g(R) = \sum_{i=1}^{n} a_i \hat{W}_i \cdot R\hat{V}_i \qquad (4)$$

Interpreting each $\hat{V}_i$, $i = 1, \ldots, n$, as a 3 × 1 matrix, Equation (4) may be rewritten as

$$g(R) = \sum_{i=1}^{n} a_i \hat{W}_i^T R \hat{V}_i \qquad (4a)$$

$\hat{W}_i^T$ denotes the transpose of $\hat{W}_i$. $L(R)$ will be a minimum if and only if the gain function, $g(R)$, is a maximum. Henceforth, all attention will be directed toward finding R_{opt} which maximizes $g(R)$.

Interpreting the individual terms of Equation (4a) as 1 × 1 matrices, it follows that

$$g(R) = \sum_{i=1}^{n} a_i \, Tr\left[\hat{W}_i^T R \hat{V}_i\right] = Tr[RB^T] \qquad (5)$$

where

$$B = \sum_{i=1}^{n} a_i \hat{W}_i \hat{V}_i^T \qquad (6)$$

as can be verified by explicitly evaluating both expressions. The result has been used that the trace of a product of matrices is unchanged by a cyclic permutation of the order of the matrices appearing in the product.

The maximization of $g(R)$ is complicated by the fact that the nine elements of R are subject to six nonlinear constraints. It is, therefore, more convenient to reexpress the gain function in terms of the quaternion.

The quaternion $\bar{q}$ representing a rotation is given by

$$\bar{q} = \begin{Bmatrix} \vec{Q} \\ q \end{Bmatrix} = \begin{Bmatrix} \hat{X} \sin(\theta/2) \\ \cos(\theta/2) \end{Bmatrix} \qquad (7)$$

where $\hat{X}$ is the axis of rotation and θ is the angle of rotation about $\hat{X}$. It should be noted that

$$\bar{q}^T \bar{q} = |\vec{Q}|^2 + q^2 = 1 \qquad (8)$$

The rotation matrix R is related to the quaternion $\bar{q}$ by

$$R(\bar{q}) = (q^2 - \vec{Q}\cdot\vec{Q})I + 2\vec{Q}\vec{Q}^T - 2q\underset{\sim}{Q} \qquad (9)$$

*The derivation here follows closely that of Reference 2.

where I is the identity matrix and $\underset{\sim}{\overline{Q}}$ is the skew-symmetric matrix given by

$$\underset{\sim}{\overline{Q}} = \begin{bmatrix} 0 & -Q_3 & Q_2 \\ Q_3 & 0 & -Q_1 \\ -Q_2 & Q_1 & 0 \end{bmatrix} \quad (10)$$

The gain function may now be written

$$\begin{aligned} g(\bar{q}) &= g(R(\bar{q})) \\ &= (q^2 - \overline{Q}\cdot\overline{Q})\,\mathrm{Tr}\,B^T + 2\,\mathrm{Tr}[\overline{Q}\overline{Q}^T B^T] \quad (11) \\ &\quad - 2q\,\mathrm{Tr}[\underset{\sim}{\overline{Q}} B^T] \end{aligned}$$

It is convenient to define scalar, matrix, and vector quantities σ, S, and $\overline{Z}$ according to

$$\sigma = \mathrm{Tr}\,B = \sum_{i=1}^{n} a_i\,\hat{W}_i \cdot \hat{V}_i \quad (12)$$

$$S = B + B^T = \sum_{i=1}^{n} a_i\left[\hat{W}_i\,\hat{V}_i^T + \hat{V}_i\,\hat{W}_i^T\right] \quad (13)$$

$$\overline{Z} = \sum_{i=1}^{n} a_i\,\hat{W}_i \times \hat{V}_i \quad (14)$$

In terms of these quantities

$$g(\bar{q}) = \sigma(q^2 - \overline{Q}\cdot\overline{Q}) + \overline{Q}^T S\,\overline{Q} + 2q\,\overline{Q}\cdot\overline{Z} \quad (15)$$

To obtain Equation (15), we have used the various invariance properties of the trace operation and the identities

$$[\underset{\sim}{\overline{Q}}]_{ij} = -\sum_k \epsilon_{ijk}\,Q_k \quad (10a)$$

and

$$(\bar{u} \times \bar{v})_i = \sum_{jk} \epsilon_{ijk}\,u_j\,v_k$$

where ϵ_{ijk} is the Levi-Civita symbol which is defined to be totally antisymmetric with $\epsilon_{123} = 1$.

Equation (15) may be written equivalently as

$$g(\bar{q}) = \bar{q}^T K \bar{q} \quad (16)$$

where

$$K = \begin{bmatrix} S - \sigma I & \overline{Z} \\ \overline{Z}^T & \sigma \end{bmatrix} \quad (17)$$

The problem of obtaining the optimal estimate of the rotation has been reduced to finding the quaternion $\bar{q}_{opt}$ which maximizes the bilinear form of Equation (16) subject to the constraint of Equation (8). R_{opt} is then obtained by substituting $\bar{q}_{opt}$ in Equation (9).

The constraint of Equation (8) may be taken into account by the method of Lagrange multipliers.[5] A new gain function $g'(\bar{q})$ is defined as

$$g'(\bar{q}) = \bar{q}^T K \bar{q} - \lambda \bar{q}^T \bar{q} \quad (18)$$

which is maximized without constraint. λ is then chosen to satisfy the constraint. It may be verified by straighforward differentiation that $g'(\bar{q})$ attains a stationary value provided

$$K\bar{q} = \lambda \bar{q} \quad (19)$$

Thus, $\bar{q}_{opt}$ must be an eigenvector of K. Equation (19) is independent of the normalization of $\bar{q}$. Hence, Equation (8) does not determine λ, although, clearly, λ must be one of the four eigenvalues of K. However, for each of the four eigenvectors of K

$$g(\bar{q}) = \bar{q}^T K \bar{q} = \lambda \bar{q}^T \bar{q} = \lambda \quad (20)$$

Thus, $g(\bar{q})$ will be maximized if $\bar{q}_{opt}$ is chosen to be the eigenvector belonging to the largest eigenvalue of K. Therefore, we write

$$K\bar{q}_{opt} = \lambda_{max}\,\bar{q}_{opt} \quad (21)$$

This is Davenport's result.

A unique solution to Equation (21) will be possible as long as there exist two independent observations. This may be expressed algebraically by the condition that the vector

$$\sum_{ij} a_i\,a_j\,\hat{W}_i \times \hat{W}_j$$

not vanish. The larger the magnitude of this vector, the greater will be the accuracy with which the attitude is determined.

III. Construction of the Optimal Quaternion

Nature of the Solutions

We now develop a method for constructing the optimal quaternion, whence, by Equation (9), the optimal rotation. Equation (19) may be rewritten

$$(S - \sigma I)\,\overline{Q} + \overline{Z}\,q = \lambda \overline{Q} \quad (22)$$

$$\overline{Z}\cdot\overline{Q} + \sigma q = \lambda q \quad (23)$$

which may be rearranged to read

$$\overline{Y} = [(\lambda + \sigma) I - S]^{-1}\,\overline{Z} \quad (24)$$

$$\lambda = \sigma + \vec{Z} \cdot \vec{Y} \tag{25}$$

where $\vec{Y}$ is the Gibbs vector, defined as

$$\vec{Y} = \vec{Q}/q = \hat{X} \tan(\theta/2) \tag{26}$$

Again, $\hat{X}$ is the axis of rotation and θ is the angle of rotation about $\hat{X}$. In terms of the Gibbs vector,

$$\bar{q} = \frac{1}{\sqrt{1 + |\vec{Y}|^2}} \begin{Bmatrix} \vec{Y} \\ 1 \end{Bmatrix} \tag{27}$$

When λ is equal to λ_{max}, $\vec{Y}$ and $\bar{q}$ are representations of the optimal rotation. Once λ_{max} is known, $\bar{q}_{opt}$ is immediately determined by Equations (24) and (27). An implicit equation may be obtained for λ_{max} by substituting Equation (24) into Equation (25) to yield

$$\lambda = \sigma + \vec{Z}^T \frac{1}{(\lambda + \sigma) I - S} \vec{Z} \tag{28}$$

Equation (28) is equivalent to the characteristic equation for the four eigenvalus of K, the explicit solution of which we wish to avoid. However, combining Equations (1), (3), (20), and (21), we obtain

$$\lambda_{max} = 1 - \frac{1}{2} \sum_{i=1}^{n} a_i |\hat{W}_i - R_{opt} \hat{V}_i|^2 \tag{29}$$

It follows immediately that λ_{max} is unity if an exact rotation of the reference vectors into the observation vectors is possible. It is to be expected that λ_{max} will deviate from unity by an amount on the order of half the mean square angular accuracy of the sensors.

Thus, the deviation of λ_{max} from unity will be very small. For sensor accuracies on the order of 1 degree ($\sim .017$ rad) λ_{max} will differ from unity by an amount on the order of 3×10^{-4} ($\sim (.017)^2$). From Equation (24) we note that if the matrix

$$[(\lambda_{max} + \sigma) I - S]$$

is nonsingular, then $\vec{Y}$ may be expanded in a Taylor series in $(\lambda_{max} - 1)$. Therefore, substituting

$$\lambda_{max} \approx 1 \tag{30}$$

in Equation (24) yields $\vec{Y}$ to this same accuracy, which corresponds to a computational error of 1 arc-minute ($\sim 3 \times 10^{-4}$ rad). For Magsat, where sensor accuracies are better than 20 arc-seconds, the anticipated computational error using Equations (24) and (30) is .002 arc-seconds.

The premise that the matrix

$$[(\lambda_{max} + \sigma) I - S]$$

is nonsingular does not hold when the angle of rotation is π.* In that case, the Gibbs vector is infinite and, therefore, the matrix must be singular. Equation (30) is not a useful approximation in this case.

The difficulties encountered when the angle of rotation is close to π may be reduced in two ways. In the first, the Gibbs vector is eliminated as an intermediate variable and a more accurate expression is obtained for λ_{max}. This reduces the bad cases to a sufficiently small interval that they may be ignored. The second method eliminates the possibility of a rotation through π by expressing the rotation as a sequence of two rotations. Both of these methods, which may be used conjointly, will be developed in the subsections which follow.

More Accurate Expressions

We present here a method whereby the optimal quaternion can be constructed without first computing the Gibbs vector. This method relies on the Cayley-Hamilton Theorem,[6] which states that a square matrix satisfies its own characteristic equation. The material in this section relies on a useful representation of the matrix

$$[(\omega + \sigma) I - S]^{-1}$$

for arbitrary ω, which is due to Markley.[7]

We begin by noting that an eigenvalue ξ of any matrix S satisfies the characteristic equation

$$\det |S - \xi I| = 0 \tag{31}$$

For a 3 × 3 matrix, this equation takes the form

$$-\xi^3 + 2\sigma\xi^2 - \kappa\xi + \Delta = 0 \tag{32}$$

with

$$\sigma = \frac{1}{2} \operatorname{Tr} S \tag{33a}$$

$$\kappa = \operatorname{Tr}(\operatorname{adj} S) \tag{33b}$$

$$\Delta = \det S \tag{33c}$$

Tr, adj, and det denote the trace, adjoint matrix, and determinant, respectively. By the Cayley-Hamilton theorem, S satisfies this same equation so that

$$S^3 = 2\sigma S^2 - \kappa S + \Delta I \tag{34}$$

In general, any analytic or rational matrix function of S may be expanded, at least formally, as

$$f(S) = b_0 I + \sum_{k=1}^{\infty} b_k S^k \tag{35}$$

*When no units of angular measure are given these are radians.

By repeated application of Equation (34), this same series can always be reduced to a quadratic in S . In particular, it must be possible to write

$$[(\omega + \sigma) I - S]^{-1} = \gamma^{-1}(\alpha I + \beta S + S^2) \tag{36}$$

α , β , and γ may be obtained by multiplying both members of this equation by

$$[(\omega + \sigma) I - S]$$

and noting that I , S , and S^2 are linearly independent. The result is

$$\alpha = \omega^2 - \sigma^2 + \kappa \tag{37}$$

$$\beta = \omega - \sigma \tag{38}$$

$$\gamma = (\omega + \sigma)(\omega^2 - \sigma^2 + \kappa) - \Delta \tag{39}$$

Letting ω take on the value λ_{max} and substituting Equation (36) into Equation (24) leads to

$$\overline{Y} = \overline{X}/\gamma \tag{40}$$

where

$$\overline{X} = (\alpha I + \beta S + S^2)\overline{Z} \tag{41}$$

is the unnormalized optimal axis of rotation. It follows from Equation (40) that

$$\bar{q}_{opt} = \frac{1}{\sqrt{\gamma^2 + |\overline{X}|^2}} \begin{Bmatrix} \overline{X} \\ \gamma \end{Bmatrix} \tag{42}$$

Since $|\overline{X}|$ and γ can never be infinite, Equation (42) avoids the difficulties associated with the divergence of the Gibbs vector as the angle of rotation tends toward π . From Equation (40), we see that γ must vanish when the angle of rotation is π . However, it is possible in some instances that $\gamma^2 + |\overline{X}|^2$ vanishes making Equation (42) indeterminant. We now determine the condition for the vanishing of $\gamma^2 + |\overline{X}|^2$.

We note first that

$$\gamma = \det |(\lambda_{max} + \sigma) I - S| \tag{43}$$

which may be verified by direct computation. Therefore, γ vanishes if and only if the angle of rotation is π . From Equations (36) and (43), it follows that

$$\alpha I + \beta S + S^2 = \frac{\det |(\lambda_{max} + \sigma) I - S|}{(\lambda_{max} + \sigma) I - S} \tag{44}$$

whence

$$\det |\alpha I + \beta S + S^2| = \gamma^2 \tag{45}$$

Hence, the matrix $\alpha I + \beta S + S^2$ is singular if and only if the angle of rotation is π . However, unless this matrix vanishes identically, which is not possible, $\overline{Z}$ cannot be an eigenvector of this matrix with vanishing eigenvalue. Otherwise, $|\overline{Y}|$ would not diverge as the angle of rotation approached π . Thus, $\overline{X}$ vanishes if and only if $\overline{Z}$ vanishes and, therefore, $\gamma^2 + |\overline{X}|^2$ vanishes if and only if the angle of rotation is π and $\overline{Z}$ vanishes.

It is now necessary to determine the conditions under which $\overline{Z}$ vanishes when the angle of rotation is π . If the errors of observation are neglected, then

$$\overline{Z} = \sum_{i=1}^{n} a_i (R_{opt} \hat{V}_i) \times \hat{V}_i \tag{46}$$

If $\hat{n}$ is the axis of rotation,* then $\hat{V}_i$ may be decomposed as

$$\hat{V}_i = \overline{V}_i'' + \overline{V}_i' \tag{47}$$

with

$$\overline{V}_i'' = \hat{n} (\hat{n} \cdot \hat{V}_i) \tag{48a}$$

$$\overline{V}_i' = -\hat{n} \times (\hat{n} \times \hat{V}_i) \tag{48b}$$

For a rotation through π

$$R_{opt} \hat{V}_i = \overline{V}_i'' - \overline{V}_i' \tag{49}$$

whence,

$$\overline{Z} = 2 \sum_{i=1}^{n} a_i \overline{V}_i'' \times \overline{V}_i' \tag{50}$$

$$\overline{Z} = 2 \sum_{i=1}^{n} a_i (\hat{n} \cdot \hat{V}_i) (\hat{n} \times \hat{V}_i) \tag{51}$$

or

$$\overline{Z} = \hat{n} \times (S_o \hat{n}) \tag{52}$$

with

$$S_o = 2 \sum_{i=1}^{n} a_i \hat{V}_i \hat{V}_i^T \tag{53}$$

*We do not write $\hat{X}$ as before to avoid confusion with $\overline{X}$ which, when it vanishes, does not have a well-defined direction associated with it.

Thus, $\gamma^2 + |\overline{X}|^2$ vanishes if and only if the angle of rotation is π and the axis of rotation $\hat{n}$ is an eigenvector of S_o. For a spacecraft with three identical mutually perpendicular sensors, S_o will be equal to (2/3) 1 and $\gamma^2 + |\overline{X}|^2$ will vanish whenever the angle of rotation is π for any axis of rotation. Thus, there is some small advantage in not making the sensors mutually orthogonal when there are three sensors of equal accuracy.

It may be pointed out that while $\gamma^2 + |\overline{X}|^2$ may not vanish, there may still be large computational errors from truncation errors in the computation of γ and $\overline{X}$. These occur only for angles of rotation near π and will be discussed in more detail in Section V.

Markley's formula also leads to a convenient expression for the characteristic equation for λ. Substituting Equation (36) with ω equal to λ into Equation (28) and multiplying both members by γ leads to[7]

$$\lambda^4 - (a + b)\lambda^2 - c\lambda + (ab + c\sigma - d) = 0 \tag{54}$$

with

$$a = \sigma^2 - \kappa \tag{55}$$

$$b = \sigma^2 + \overline{Z}^T \overline{Z} \tag{56}$$

$$c = \Delta + \overline{Z}^T S \overline{Z} \tag{57}$$

$$d = \overline{Z}^T S^2 \overline{Z} \tag{58}$$

Since λ_{max} is known to be very close to unity, the Newton-Raphson method[5] applied to Equation (54) with unity as a starting value provides a very fast means of computing λ_{max}. For Magsat, where the sensor accuracies are better than 20 arc-seconds ($\sim 10^{-4}$ rad), the deviation of λ_{max} from unity is on the order of 10^{-8}. After one application of the Newton-Raphson method, λ_{max} is obtained with an accuracy of 10^{-16}. This is comparable to the accuracy of a double-precision word (64 bits) in IBM FORTRAN, namely, $10^{-16.8}$. For sensor accuracies on the order of 1 degree, this same accuracy is obtained with only two iterations.

The Method of Sequential Rotations

The accuracy of the computation of the optimal quaternion when the angle of rotation is close to π may be further improved as follows.

When the angle of rotation is greater than $\pi/2$, the rotation can be expressed as a rotation through π about one of the coordinate axes followed by a rotation about some new axis through an angle less than $\pi/2$. An initial rotation through π about one of the coordinate axes is equivalent to changing the signs of two of the components of each reference vector. The quaternion $\overline{p} = (p_1, p_2, p_3, p_4)^T$ of the optimal rotation transforming the new reference vectors $\hat{V}_i'$ to the observation vectors $\hat{W}_i$ is related very simply to the desired optimal quaternion $\overline{q} = (q_1, q_2, q_3, q_4)^T$ for the total optimal rotation. The results are as follows:

Rotation through π about the x-axis

$$\hat{V}_i' = \left(\hat{V}_{ix}, -\hat{V}_{iy}, -\hat{V}_{iz}\right)^T \tag{59}$$

$$q_1 = p_4 \qquad q_2 = -p_3$$
$$q_3 = p_2 \qquad q_4 = -p_1 \tag{60}$$

Rotation through π about the y-axis

$$\hat{V}_i' = \left(-\hat{V}_{ix}, \hat{V}_{iy}, -\hat{V}_{iz}\right)^T \tag{61}$$

$$q_1 = p_3 \qquad q_2 = p_4$$
$$q_3 = -p_1 \qquad q_4 = -p_2 \tag{62}$$

Rotation through π about the z-axis

$$\hat{V}_i' = \left(-\hat{V}_{ix}, -\hat{V}_{iy}, \hat{V}_{iz}\right)^T \tag{63}$$

$$q_1 = -p_2 \qquad q_2 = p_1$$
$$q_3 = p_4 \qquad q_4 = -p_3 \tag{64}$$

Although $\gamma^2 + |\overline{X}|^2$ may not be small, the computation of $\overline{q}_{opt}$ from Equation (42) will not be accurate if $|\gamma|$ is very small. For small angles of rotation, γ is of the order of unity. When $|\gamma|$ is much less than unity, one of the initial rotations given by Equations (59), (61), or (63) will lead to a second rotation for which γ is sufficiently large to ensure high accuracy. The optimal quaternion for the full rotation is obtained by applying Equations (60), (62), or (64).

IV. Approximations Near Null Attitude

When the attitude matrix is known to be very close to the identity matrix (null attitude) or, equivalently, when the angle of rotation is known to be small, simple approximations may be obtained. In that case, $\overline{Z}$ will be a small quantity of the order of the angle of rotation and S will differ from S_o of Equation (53) by terms of the same order. Thus, if δ is a quantity of the order the error of observation or the angle of rotation, whichever is larger, then

$$\overline{Z} = O(\delta) \tag{65}$$

$$S = S_o + O(\delta) \tag{66}$$

$$\sigma = 1 + O(\delta^2) \tag{67}$$

$$\lambda_{max} = 1 + O(\delta^2) \tag{68}$$

With these approximations Equation (24) becomes

$$\overline{Y} = [2I - S_o]^{-1} \overline{Z} + O(\delta^2) \tag{69}$$

In the special case that the spacecraft carries three mutually orthogonal sensors of the same accuracy, Equation (69) simplifies to

$$\overline{Y} = (3/4)\, \overline{Z} + O(\delta^2) \tag{70}$$

In the general case, near null attitude

$$\overline{Y} = \frac{1}{2\kappa_o - \Delta_o}\left[\kappa_o I + S_o^2\right] \overline{Z} + O(\delta^2) \tag{71}$$

where

$$\kappa_o = \mathrm{Tr}\,(\mathrm{adj}\; S_o) \tag{72}$$

$$\Delta_o = \det S_o \tag{73}$$

Equation (71) follows directly from Equation (40).

When there are only two observations, Equation (71) reduces to

$$\begin{aligned}\overline{Y} = \frac{1}{2}\Big\{ & a_1 \hat{W}_1 \times \hat{V}_1 + a_2 \hat{W}_2 \times \hat{V}_2 \\ & + a_1 \left[(a \hat{W}_2 - b \hat{W}_1) \cdot (\hat{V}_1 \times \hat{V}_2)\right] \hat{V}_1 \\ & + a_2 \left[(b \hat{W}_2 - a \hat{W}_1) \cdot (\hat{V}_1 \times \hat{V}_2)\right] \hat{V}_2 \Big\} \\ & + O(\delta^2)\end{aligned} \tag{74}$$

with a and b given by

$$a = \frac{1}{|\hat{V}_1 \times \hat{V}_2|^2} \tag{75}$$

$$b = \frac{(\hat{V}_1 \cdot \hat{V}_2)}{|\hat{V}_1 \times \hat{V}_2|^2} \tag{76}$$

More accurate though also more cumbersome expressions[3] can be obtained by setting

$$K = \begin{bmatrix} S - \sigma I & \overline{0} \\ \overline{0}^T & \sigma \end{bmatrix} + \begin{bmatrix} 0 & \overline{Z} \\ \overline{Z}^T & 0 \end{bmatrix} \tag{77}$$

and applying Rayleigh-Schroedinger perturbation theory. Such expressions, however, have no computational advantage over Equation (42) above.

Equations (69) through (76) are useful when the spacecraft has a high pointing accuracy requirement in an inertial reference coordinate system. In that case, the vectors $\hat{V}_i$, $i = 1, \ldots, n$, and the matrix S_o need be determined only once and the connection between the observation vectors and the Gibbs vector is simple.

Similarly, these small angle algorithms could be used in conjunction with an initial attitude estimate, obtained from some fast nonoptimal algorithm, to provide an optimal attitude solution, since the difference between the optimal and nonoptimal estimates of the attitude is usually small.

V. Applications to Magsat

Equations (42) and (54) have been tested extensively for the sensor configuration of the Magsat mission. The Magsat spacecraft, scheduled for launch in September 1979, will be placed in a Sun-synchronous orbit. An on-board semi-autonomous control system will keep the spacecraft Earth-pointing.

The Magsat spacecraft will measure the geomagnetic field with an accuracy per component of 7 gamma ($= 7 \times 10^{-5}$ Gauss). This requires that the spacecraft attitude be known to an accuracy of at least 20 arc-seconds. To accomplish this, Magsat will carry two fixed-head star trackers and a fine Sun sensor.

The angle between any two sensor axes lies between 60 degrees and 75 degrees. Thus, the sensors provide relatively independent information and also the matrix S_o will be far from $(2/3)\, I$.

To test the algorithm, a set of three "exact" reference vectors $\hat{V}_i$ was constructed which corresponded to the directions of the three sensor axes. A set of "noised" reference vectors $\hat{U}_i$ was then constructed which differed from the $\hat{V}_i$ by errors of about 15 arc-seconds and chosen so that the optimal rotation carrying the $\hat{V}_i$ into the $\hat{U}_i$ was the null rotation. For given angle and axis of rotation, the observation vectors $\hat{W}_i$ were constructed by rotating the $\hat{U}_i$. This complicated procedure was necessary so that the angle of rotation could be given an exact meaning, even though the correspondence between the $\hat{V}_i$ and the $\hat{W}_i$ was not exact.

For many choices of the axis and angle of rotation, the input and output quaternions were computed. The input (true) quaternion was constructed from the known angle and axis of rotation according to Equation (7). The output quaternion was computed using Equations (42) and (54) with λ_{max} computed using one iteration of the Newton-Raphson method. Computations were performed using the IBM S/360-95 computer at NASA Goddard Space Flight Center.

It was found for all choices of the axis of rotation that the computational error was on the order of 10^{-15} radians ($\approx 2 \times 10^{-10}$ arc-seconds) until θ, the angle of rotation, became as large as 179.5 degrees ($|\pi - \theta| \geq 0.01$). Thereafter, for each decade decrease in $|\pi - \theta|$, the computational error of the solution increased by one decade until $|\pi - \theta|$ became smaller than 10^{-15}. This is the expected behavior for a purely truncational error. Thus, the error becomes greater than 1 arc-second for $|\pi - \theta| \leq 10^{-11}$.

To appreciate this magnitude, it should be noted that the circumference of the Magsat orbit is approximately 44,000 kilometers. Computational errors larger than 1 arc-second in the attitude estimation are restricted to a segment of this orbit no longer than 100 microns.

The Magsat spacecraft attitude will be computed once every quarter second. At this rate, it is anticipated that attitude computation errors will exceed 10^{-4} arc-seconds only once every 2 years and exceed 1 arc-second perhaps once in 20,000 years. Indeterminancy in the attitude due to the vanishing or near vanishing of $\gamma^2 + |\overline{X}|^2$ will occur perhaps once in 10^{20} years. These times are long compared to the expected Magsat mission lifetime of from 4 to 8 months. For this reason, there is no plan to implement the method of sequential rotations (developed at the end of Section III) in the Magsat mission software.

It should be noted that the computational error in determining the optimal attitude is not the same as the attitude determination accuracy. The computational error is the difference between the computed optimal attitude and the exact optimal attitude. Although the computational error can be made arbitrarily small, the difference of the optimal attitude from the true attitude will still be on the order of the statistical sum of the sensor errors.

Acknowledgements

The author is grateful to P. Davenport for interesting discussions and helpful criticisms and to F. L. Markley, whose contributions are cited in the text.

References

1. P. Davenport (private communication)

2. J. Keat, Analysis of Least Squares Attitude Determination Routine, DOAOP, Computer Sciences Corporation, CSC/TM-77/6034, February 1977

3. M. D. Shuster, Algorithms for Determining Optimal Attitude Solutions, Computer Sciences Corporation, CSC/TM-78/6056, April 1978

4. F. L. Markley, "Attitude Dynamics," in Attitude Determination and Control, edited by J. R. Wertz. Doordrecht, the Netherlands: D. Reidel Publishing Company (in press)

5. F. B. Hildebrand, Advanced Calculus for Applications, Englewood Cliffs, N. J., Prentice-Hall, Inc., 1962

6. K. Hoffman and R. Kunze, Linear Algebra, Englewood Cliffs, N. J., Prentice-Hall, Inc., 1961

7. F. L. Markley (private communication)