EFFICIENT ESTIMATION OF ATTITUDE SENSOR COALIGNMENTS

Malcolm D. Shuster*
University of Florida, Gainesville, FL 32611

Abstract

Simple and statistically correct algorithms are presented for computing the coalignments of attitude sensors from inflight data consisting of both vectors and complete three-axis attitudes. The effect of autocollimators are considered and specific algorithms are given.

Introduction

This work extends algorithms developed in two earlier works[1,2] for the estimation of alignments from sensors which measure either directions, entire vectors, or attitudes. The present report presents these algorithms as simply as possible without studies of alternate methods, special cases, or expected performance.

Consider a spacecraft with a suite of sensors. These may measure either a single vector (possibly only a unit vector), such is the case for Sun sensors, magnetometers, conical horizon scanners, and the older star trackers such as the Ball Brothers CT-401, or they may be able to determine a complete three-axis attitude. This is usually done by measuring two or more directions, such as in the newer star trackers provided with charge-coupled devices (CCDs). How the attitude is determined for the complete-attitude sensors (for a CCD star tracker one could, for example, use the QUEST algorithm[3] or the FOAM algorithm[4]), is not important for this report. What we assume is that we have two classes of sensors, sensors which measure a complete attitude and sensors which measure only a single vector.

Vector sensors, which measure a single vector, have measurements which may be described by

$$\hat{\boldsymbol{U}}_{i,k} = A_{i,k}\hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{U}}_{i,k} \,, \tag{1}$$

where $\hat{\boldsymbol{U}}_{i,k}$ is the measured unit vector in the sensor coordinates of sensor i at time t_k, $A_{i,k}$ is the attitude of that sensor with respect to the primary reference frame (typically Geocentric Inertial (GCI)) and $\Delta\hat{\boldsymbol{U}}_{i,k}$ is the (Gaussian) measurement noise, which is assumed to satisfy

$$E\{\hat{\boldsymbol{U}}_{i,k}\} = \mathbf{0} \,, \tag{2}$$

and

$$\hat{\boldsymbol{U}}_{i,k} \cdot \Delta\hat{\boldsymbol{U}}_{i,k} = 0 \,. \tag{3}$$

Here $E\{\cdot\}$ denotes the expectation. We denote the covariance matrix of $\hat{\boldsymbol{U}}_{i,k}$ by $R_{\hat{\boldsymbol{U}}_{i,k}}$. A convenient form which is usually justified for sensors with small fields of view is given by the QUEST model[3],

$$R^{\text{QUEST}}_{\hat{\boldsymbol{U}}_{i,k}} = \sigma^2_{i,k}\left[I_{3\times 3} - \hat{\boldsymbol{U}}_{i,k}\hat{\boldsymbol{U}}^T_{i,k}\right] \,, \tag{4}$$

where $\sigma^2_{i,k}$ is the variance characteristic of the sensor assuming the QUEST error model.

Complete-attitude sensors furnish a complete three-axis attitude. For these sensors the measurement model may be taken to be

$$A^*_{i,k} = (\delta A_{i,k})A_{i,k} \,, \tag{5}$$

where $A^*_{i,k}$ is the measured attitude, which we write as an estimate for a specific time, and $(\delta A_{i,k})$ is the attitude error, which we have written as a small rotation. If this rotation is characterized by a small rotation vector $\boldsymbol{\xi}_{i,k}$, then we can write

$$\delta A_{i,k} = e^{[[\boldsymbol{\xi}_{i,k}]]} \tag{6}$$

$$= I + \left(\frac{\sin|\boldsymbol{\xi}_{i,k}|}{|\boldsymbol{\xi}_{i,k}|}\right)[[\boldsymbol{\xi}_{i,k}]] + \left(\frac{1-\cos|\boldsymbol{\xi}_{i,k}|}{|\boldsymbol{\xi}_{i,k}|^2}\right)[[\boldsymbol{\xi}_{i,k}]]^2 \tag{7}$$

$$= I + [[\boldsymbol{\xi}_{i,k}]] + O(|\boldsymbol{\xi}_{i,k}|^2) \,, \tag{8}$$

and we assume that

$$E\{\boldsymbol{\xi}_{i,k}\} = \mathbf{0} \,, \tag{9}$$

$$E\{\boldsymbol{\xi}_{i,k}\boldsymbol{\xi}^T_{i,k}\} = R_{\boldsymbol{\xi}_{i,k}} \,. \tag{10}$$

*Professor, Department of Aerospace Engineering, Mechanics and Engineering Science, Associate Fellow AIAA.

Here $[[\boldsymbol{v}]]$ denotes the antisymmetric matrix

$$[[\boldsymbol{v}]] \equiv \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix} . \qquad (11)$$

If the actual sensor measurements consist of $n_{i,k}$ measured directions $\hat{\boldsymbol{U}}_{i,k,\ell}$, $\ell = 1, \ldots, n_{i,k}$, whose covariance matrix is represented by Eq.(4) and which are used as input to the QUEST algorithm, then the resulting attitude covariance matrix would be given by

$$R^{\text{QUEST}}_{\boldsymbol{\xi}_{i,k}} = \left[\sum_{\ell=1}^{n_{i,k}} \frac{1}{\sigma^2_{i,k,\ell}} \left(I_{3\times3} - \hat{\boldsymbol{U}}_{i,k,\ell} \hat{\boldsymbol{U}}^T_{i,k,\ell} \right) \right]^{-1} . \qquad (12)$$

The algorithms which we will present below do not assume that the QUEST algorithm has been used. We assume that the data has already been made simultaneous using a dynamical model or gyro information. Our purpose here is to create effective measurements $\boldsymbol{z}_k$ of the general form

$$\boldsymbol{z}_k = H_k \boldsymbol{\theta} + \boldsymbol{v}_k , \qquad (13)$$

where $\boldsymbol{v}_k$ is white noise, and $\boldsymbol{\theta}$ is the correction to the coalignment (i.e., relative alignment) of sensor j to sensor i, to be defined in more detail below. The dimension of $\boldsymbol{z}_k$ may be 1, 2, or 3 depending on whether neither, one or both sensors have sufficient data to determine an attitude.

Alignments and Coalignments

The sensor relative-alignment matrix (coalignment) $S_{i\leftarrow j}$ is the orthogonal matrix which transforms column vectors in the coordinate frame of sensor j into column vectors in the coordinate frame of sensor i. Thus,

$$\boldsymbol{V}_{\mathcal{E}_i} = S_{i\leftarrow j} \boldsymbol{V}_{\mathcal{E}_j} , \qquad (14)$$

where $\mathcal{E}_i$ denotes the coordinate frame of sensor i, and $\boldsymbol{V}_{\mathcal{E}_i}$ denotes the 3×1 matrix of the components of the abstract vector $\boldsymbol{V}$ with respect to the basis $\mathcal{E}_i$. The coalignment correction (comisalignment) matrix is the orthogonal matrix which corrects an *a priori* value of the coalignment matrix $S^o_{i\leftarrow j}$ according to

$$S_{i\leftarrow j} = M_{i\leftarrow j} S^o_{i\leftarrow j} . \qquad (15)$$

The absolute alignments[1,2] are defined in terms of some fiducial body frame. Thus

$$\boldsymbol{V}_{\mathcal{B}} = S_i \boldsymbol{V}_{\mathcal{E}_i} \qquad (16)$$

and the (absolute) misalignment M_i is defined by

$$S_i \equiv M_i S^o_i , \qquad (17)$$

where S^o_i is an *a priori* value of the absolute alignment. We may define the fiducial body frame in terms of the attitude of one of the sensors, say sensor 1. Then

$$\boldsymbol{V}_{\mathcal{B}} = S^o_1 \boldsymbol{V}_{\mathcal{E}_1} , \qquad (18)$$

and the alignment of sensor 1 is assumed not to change from its prelaunch value

$$M_1 \equiv I . \qquad (19)$$

Otherwise, one may define the fiducial body frame independent of any constellation of sensors. In that case, the post-launch absolute alignments can be determined with only poor accuracy, as shown in Ref. 2.

Clearly, the absolute alignments and coalignments are related by

$$S_{i\leftarrow j} = S^T_i S_j , \qquad (20)$$

and therefore coalignments with respect to sensor 1 are given by

$$S_{1\leftarrow i} = S^{oT}_1 S_i , \quad i = 2, \ldots, n , \qquad (21)$$

or

$$S_i = S^o_1 S_{1\leftarrow i} , \quad i = 2, \ldots, n . \qquad (22)$$

Since M_i and $M_{i\leftarrow j}$ are expected to be very small rotations, we may write

$$M_{i\leftarrow j} \equiv e^{[[\boldsymbol{\theta}_{i\leftarrow j}]]} \approx I + [[\boldsymbol{\theta}_{i\leftarrow j}]] , \qquad (23)$$

and

$$M_i \equiv e^{[[\boldsymbol{\theta}_i]]} \approx I + [[\boldsymbol{\theta}_i]] . \qquad (24)$$

θ_i and $\theta_{i\leftarrow j}$ are the (absolute) misalignment vectors and comisalignment vectors, respectively.

In the earlier work[1,2] the representation of the absolute alignments in the body frame led to measurements which depended only on the difference of the misalignment vectors, $\boldsymbol{\theta}_i - \boldsymbol{\theta}_j$. This same property is preserved by the coalignments. From Eqs.(15), (17), and (20) it follows that

$$M_{i\leftarrow j} = S^{oT}_i M^T_i M_j S^o_i , \qquad (25)$$

whence

$$\boldsymbol{\theta}_{i\leftarrow j} = S^{oT}_i (\boldsymbol{\theta}_j - \boldsymbol{\theta}_i) . \qquad (26)$$

Measurement Models for Sensor Coalignment

In the situation of most practical interest, sensor 1 is a complete attitude sensor, especially a CCD star tracker, but may on occasion measure only a single vector. We shall consider only frames of data in which some data from sensor 1 is always present, and we shall consider as effective measurements the those involving sensor 1 and one other sensor. There are, therefore, four cases: (1) sensor 1 and sensor 2 both measure a complete attitude; (2) sensor 1 measures a complete attitude sensor but sensor 2 measures only a single vector; (3) sensor (1) measures a single vector while sensor 2 measures a complete attitude; (4) both sensors 1 and 2 measure only a single vector.

In this work we will assume not only that sensor 1 is a complete-attitude sensor but also that it measures a complete attitude in every frame. Thus, we need consider only the first two cases. We shall also assume that only sensor 1 and one other sensor have data available at any one time, the case freqently when the other sensors are scientific instruments.

Attitude-Attitude Measurement Model

Let us suppose that sensor i, $i > 1$, measures a complete attitude $A_{i,k}$. Then,

$$S_{1\leftarrow i} = A_{1,k}\, A_{i,k}^T\,, \tag{27}$$

where $A_{1,k}$ is the attitude of sensor 1 at time t_k, and $A_{i,k}$ is the attitude of sensor i at time t_k. Recalling Eq.(5) this becomes

$$S_{1\leftarrow i} = (\delta A_{1,k})^T A_{1,k}^*\, A_{i,k}^{*T}(\delta A_{i,k})\,, \tag{28}$$

or

$$\begin{aligned} A_{1,k}^* A_{i,k}^{*T} &= (\delta A_{1,k})\, S_{1\leftarrow i}\, (\delta A_{i,k})^T & (29)\\ &= (\delta A_{1,k})\, M_{1\leftarrow i}\, S_{1\leftarrow i}^o\, (\delta A_{i,k})^T\,. & (30)\end{aligned}$$

We define now

$$\mathcal{Z}_{i,k} \equiv A_{1,k}^*\, A_{i,k}^{*T}\, S_{1\leftarrow i}^{oT}\,. \tag{31}$$

Then

$$\mathcal{Z}_{i,k} = (\delta A_{1,k})\, M_{1\leftarrow i}\, S_{1\leftarrow i}^o\, (\delta A_{i,k})^T S_{1\leftarrow i}^{oT}\,, \tag{32}$$

with

$$S_{1\leftarrow i}^o \equiv S_1^{oT} S_i^o\,. \tag{33}$$

Recalling Eq.(6), Eq.(32) becomes

$$\begin{aligned} \mathcal{Z}_{i,k} &= e^{[[\boldsymbol{\xi}_{1,k}]]}\, e^{[[\boldsymbol{\theta}_{1\leftarrow i}]]}\, e^{-[[S_{1\leftarrow i}^o \boldsymbol{\xi}_{i,k}]]} & (34)\\ &\approx e^{[[\boldsymbol{\theta}_{1\leftarrow i} + \boldsymbol{\xi}_{1,k} - S_{1\leftarrow i}^o \boldsymbol{\xi}_{i,k}]]}\,. & (35)\end{aligned}$$

If we define now Rtov(·) to be the function which computes the rotation vector corresponding to a particular direction-cosine matrix and

$$\boldsymbol{z}_{i,k}^{(2)} \equiv \mathrm{Rtov}(\mathcal{Z}_{i,k})\,, \tag{36}$$

then

$$\begin{aligned} \boldsymbol{z}_{i,k}^{(2)} &= \boldsymbol{\theta}_{1\leftarrow i} + \boldsymbol{\xi}_{1,k} - S_{1\leftarrow i}^o \boldsymbol{\xi}_{i,k} & (37)\\ &\equiv H_{i,k}^{(2)}\, \boldsymbol{\theta}_{1\leftarrow i} + \boldsymbol{v}_{i,k}^{(2)}\,, & (38)\end{aligned}$$

where now

$$\begin{aligned} H_{i,k}^{(2)} &= I\,, & (39)\\ \boldsymbol{v}_{i,k}^{(2)} &= \boldsymbol{\xi}_{1,k} - S_{1\leftarrow i}^o \boldsymbol{\xi}_{i,k}\,. & (40)\end{aligned}$$

It follows that

$$\boldsymbol{v}_{i,k}^{(2)} \sim \mathcal{N}\left(\mathbf{0}, R_{\boldsymbol{z}_{i,k}^{(2)}}\right)\,, \tag{41}$$

with

$$R_{\boldsymbol{z}_{i,k}^{(2)}} = R_{\boldsymbol{\xi}_{1,k}} + S_{1\leftarrow i}^o R_{\boldsymbol{\xi}_{i,k}} S_{1\leftarrow i}^{oT}\,. \tag{42}$$

This is the desired form for the effective coalignment measurement. We have denoted the measurement by $\boldsymbol{z}_{i,k}^{(2)}$ to distintinguish it from $\boldsymbol{z}_{i,k}$, the alignment measurement constructed from one vector in each sensor, hereafter denoted by $\boldsymbol{z}_{i,k}^{(1)}$. We will follow this practice throughout this work of distinguishing different measurement models by a numerical superscript.

If the individual sensor attitudes are given as quaternions, then the above relations become equivalently

$$\bar{z}_{i,k} \equiv \bar{q}_{1,k}^* \otimes (\bar{q}_{i,k}^*)^{-1} \otimes (\bar{s}^o)_{1\leftarrow i}^{-1}\,, \tag{43}$$

where $\bar{s}_{1\leftarrow i}^o$ is the *a priori* coalignment quaternion given by

$$(\bar{s}^o)_{1\leftarrow i} \equiv (\bar{s}_1^o)^{-1} \otimes \bar{s}_i^o\,, \tag{44}$$

with $\bar{s}_i^o$ the *a priori* alignment quaternion of sensor i. We may write in analogy to Eq.(36)

$$\bar{s}_i = \mathrm{Rtoq}\,(S_i) \quad \text{and} \quad S_i = \mathrm{qtoR}\,(\bar{s}_i)\,. \tag{45}$$

The effective measurement is now given equivalently by

$$\boldsymbol{z}_{i,k}^{(2)} = \mathrm{qtov}\,(\bar{z}_{i,k})\,. \tag{46}$$

Attitude-Vector Measurement Model

Suppose now that sensor i measures only a single vector. Then we can write successively

$$\begin{aligned} \hat{\boldsymbol{U}}_{i,k} &= A_{i,k}\, \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{U}}_{i,k} & (47)\\ &= S_{i\leftarrow 1}\, A_{1,k}\, \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{U}}_{i,k} & (48)\\ &= S_{1\leftarrow i}^T\, A_{1,k}\, \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{U}}_{i,k} & (49)\\ &= S_{1\leftarrow i}^{oT}\, M_{1\leftarrow i}^T\, A_{1,k}\, \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{U}}_{i,k}\,. & (50)\end{aligned}$$

Let us define now

$$\hat{\boldsymbol{W}}_{i,k}^o = S_{1\leftarrow i}^o\, \hat{\boldsymbol{U}}_{i,k}\,. \tag{51}$$

Then

$$\hat{\boldsymbol{W}}_{i,k}^o = M_{1\leftarrow i}^T\, A_{1,k}\, \hat{\boldsymbol{V}}_{i,k} + S_{1\leftarrow i}^o\, \Delta\hat{\boldsymbol{U}}_{i,k}\,. \tag{52}$$

This quantity is similar to the one defined in Ref. 1 except that the final frame here is the frame of sensor 1 and not the fiducial body frame.

Recalling Eq.(5), we have that

$$\begin{aligned}
\hat{W}^o_{i,k} &= M^T_{1\leftarrow i}\,(\delta A_{1,k})^T A^*_{1,k}\,\hat{V}_{i,k} \\
&\quad + S^o_{1\leftarrow i}\,\Delta\hat{U}_{i,k} && (53)\\
&\approx M^T_{1\leftarrow i}\,A^*_{1,k}\,\hat{V}_{i,k} \\
&\quad -[[\boldsymbol{\xi}_1]]A^*_{1,k}\,\hat{V}_{i,k} + S^o_{1\leftarrow i}\,\Delta\hat{U}_{i,k} && (54)\\
&= M^T_{1\leftarrow i}\,A^*_{1,k}\,\hat{V}_{i,k} \\
&\quad +[[A^*_{1,k}\,\hat{V}_{i,k}]]\,\boldsymbol{\xi}_1 + S^o_{1\leftarrow i}\,\Delta\hat{U}_{i,k} && (55)\\
&= M^T_{1\leftarrow i}\,\hat{W}^{o*}_{i,k} + \Delta\hat{W}^o_{i,k}\,, && (56)
\end{aligned}$$

and we have defined

$$\hat{W}^{o*}_{i,k} \equiv A^*_{1,k}\,\hat{V}_{i,k}\,, \qquad (57)$$

$$\Delta\hat{W}^o_{i,k} \equiv [[\hat{W}^{o*}_{i,k}]]\,\boldsymbol{\xi}_1 + S^o_{1\leftarrow i}\,\Delta\hat{U}_{i,k}\,. \qquad (58)$$

It is now a simple matter to show that

$$\begin{aligned}
\hat{W}^{o*}_{i,k} \times \hat{W}^o_{i,k} &= -[[\hat{W}^{o*}_{i,k}]]\Big\{[[\boldsymbol{\theta}_{1\leftarrow i}]]^T\hat{W}^{o*}_{i,k} \\
&\quad + \Delta\hat{W}^o_{i,k}\Big\} && (59)\\
&= -[[\hat{W}^{o*}_{i,k}]]^2\,\boldsymbol{\theta}_{1\leftarrow i} \\
&\quad -[[\hat{W}^{o*}_{i,k}]]\,\Delta\hat{W}^o_{i,k}\,, && (60)
\end{aligned}$$

and

$$\begin{aligned}
[[\hat{W}^{o*}_{i,k}]]\,\Delta\hat{W}^o_{i,k} &= [[\hat{W}^{o*}_{i,k}]]^2\,\boldsymbol{\xi}_1 \\
&\quad +[[\hat{W}^{o*}_{i,k}]]\,S^o_{1\leftarrow i}\,\Delta\hat{U}_{i,k}\,. && (61)
\end{aligned}$$

The quantity $\hat{W}^{o*}_{i,k} \times \hat{W}^o_{i,k}$ has the disadvantage of being three-dimensional but having only two degrees of freedom. Hence, its covariance matrix will be singular and noninvertible. It will be to our advantage, therefore, to project from it a two-dimensional quantity which will have a non-singular covariance matrix.

To accomplish this let $\hat{W}$ be any unit vector and let $\hat{a}(\hat{W})$ and $\hat{b}(\hat{W})$ be any two other unit vectors such that $\{\hat{W},\, \hat{a}(\hat{W}),\, \hat{b}(\hat{W})\}$ is a right-hand orthonormal triad. Since the component of $\hat{W}^{o*}_{i,k} \times \hat{W}^o_{i,k}$ along $\hat{W}^{o*}_{i,k}$ vanishes identically, it is clear that the desired projection operator is the 2×3 matrix

$$\mathcal{P}(\hat{W}^{o*}_{i,k}) \equiv \begin{bmatrix} \hat{a}^T(\hat{W}^{o*}_{i,k}) \\ \hat{b}^T(\hat{W}^{o*}_{i,k}) \end{bmatrix}, \qquad (62)$$

and we define the effective coalignment measurement as

$$z^{(3)}_{i,k} \equiv \mathcal{P}(\hat{W}^{o*}_{i,k})\,(\hat{W}^{o*}_{i,k} \times \hat{W}^o_{i,k})\,. \qquad (63)$$

We note immediately that

$$\mathcal{P}(\hat{W})\,[[\hat{W}]]^2 = \mathcal{P}(\hat{W})[\hat{W}\hat{W}^T - I] = -\mathcal{P}(\hat{W})\,, \qquad (64)$$

from which it follows that

$$\begin{aligned}
z^{(3)}_{i,k} &= \mathcal{P}(\hat{W}^{o*}_{i,k})\,\boldsymbol{\theta}_{1\leftarrow i} \\
&\quad -\mathcal{P}(\hat{W}^{o*}_{i,k})\,[[\hat{W}^{o*}_{i,k}]]\,\Delta\hat{W}^o_{i,k}) && (65)\\
&\equiv H^{(3)}_{i,k}\,\boldsymbol{\theta}_{1\leftarrow i} + v^{(3)}_{i,k}\,. && (66)
\end{aligned}$$

It follows after much manipulation that

$$v^{(3)}_{i,k} \sim \mathcal{N}\left(\mathbf{0},\, R_{z^{(3)}_{i,k}}\right), \qquad (67)$$

with

$$\begin{aligned}
R_{z^{(3)}_{i,k}} &= \mathcal{Q}(\hat{W}^{o*}_{i,k})\,S^o_{1\leftarrow i}R_{\hat{U}_{i,k}}S^{oT}_{1\leftarrow i}\,\mathcal{Q}^T(\hat{W}^{o*}_{i,k}) \\
&\quad +\mathcal{P}(\hat{W}^{o*}_{i,k})R_{\boldsymbol{\xi}_{i,k}}\mathcal{P}^T(\hat{W}^{o*}_{i,k})\,, && (68)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{Q}(\hat{W}^{o*}_{i,k}) &\equiv -\mathcal{P}(\hat{W}^{o*}_{i,k})\,[[\hat{W}^{o*}_{i,k}]] && (69)\\
&= \begin{bmatrix} -\hat{b}^T(\hat{W}^{o*}_{i,k}) \\ \hat{a}^T(\hat{W}^{o*}_{i,k}) \end{bmatrix}. && (70)
\end{aligned}$$

Note that we could also have written

$$z^{(3)}_{i,k} = \mathcal{Q}\left(\hat{W}^{o*}_{i,k}\right)\hat{W}^o_{i,k}\,. \qquad (71)$$

The Alignment Estimator

Given a set of measurements for a sensor, which may consist of single vectors or complete attitudes, we construct the negative-log-likelihood function as

$$\begin{aligned}
J(\boldsymbol{\theta}_{1\leftarrow i}) &= \frac{1}{2}\sum_{k=1}^{n_i} (z_{i,k} - H_{i,k}\boldsymbol{\theta}_{1\leftarrow i})^T\, R^{-1}_{z_{i,k}} \\
&\quad \times (z_{i,k} - H_{i,k}\boldsymbol{\theta}_{1\leftarrow i})\,. && (72)
\end{aligned}$$

The effective measurement is $z^{(2)}_{i,k}$ or $z^{(3)}_{i,k}$ depending on whether the sensor furnishes on a single vector observation in the frame k or sufficient observations to construct the entire attitude, respectively. The optimal estimate of $\boldsymbol{\theta}_{1\leftarrow i}$ minimizes this quantity. This has the value

$$\boldsymbol{\theta}^*_{1\leftarrow i} = P_{\theta\theta}\sum_{k=1}^{n_i} H^T_{i,k}\,R^{-1}_{z_{i,k}}\,z_{i,k}\,, \qquad (73)$$

where

$$P_{\theta\theta} = \left[\sum_{k=1}^{n_i} H^T_{i,k}\,R^{-1}_{z_{i,k}}\,H_{i,k}\right]^{-1}, \qquad (74)$$

is the estimate error covariance matrix.

Sensors with Autocollimators

Sometimes the alignment or coalignment of a sensor is instrumented through an *autocollimator* or *attitude transfer system*, whose output gives the alignment of the sensor relative to some reference frame fixed in the spacecraft. This autocollimator may be misaligned as well. Thus, in modelling the alignment of sensors equipped with autocollimators one must consider two interfaces. The interface between the sensor and the autocollimator and the interface between the autocollimator and the spacecraft body frame.

Ideally in the absence of misalignment, the representation of a vector in the body frame given its representation in sensor coordinates is given by

$$\left(\hat{\boldsymbol{W}}_{i,k}\right)_{\text{body}} = S_i\, C_{i,k}\, \hat{\boldsymbol{U}}_{i,k}\,, \qquad (75)$$

where $C_{i,k}$ is the direction-cosine matrix of the transformation supplied by the autocollimator. This is generally a vendor-supplied function of two or three scalar outputs. When there are only two scalar outputs, the missing output generally parameters the rotation about the axis of the autocollimator, which is usually poorly known in the best of cases. If we consider the case where sensor i is provided with an autocollimator, then Eq.(75) becomes

$$\left(\hat{\boldsymbol{W}}_{i,k}\right)_{\text{body}} = S_i\, C_{i,k}\, T_i\, \hat{\boldsymbol{U}}_{i,k}\,. \qquad (76)$$

There are now two misalignment transformations, which are both proper orthogonal transformations. In general, $C_{i,k}$ will be a small rotation.

$A_{i,k}$, the attitude of sensor i, is now given by

$$A_{i,k} = T_i^T\, C_{i,k}^T\, S_i^T\, S_1\, A_{1,k}\,, \qquad (77)$$

and the coalignment of sensor i relative to sensor 1 is

$$S_{1\leftarrow i} = S_1^T\, S_i\, C_{i,k}\, T_i\,. \qquad (78)$$

We now define misalignments by

$$S_i \equiv S_i^o\, N_i \quad \text{and} \quad T_i \equiv O_i\, T_i^o\,, \qquad (79)$$

where N_i and O_i are both proper orthogonal matrices corresponding to very small rotations. Note the order of the matrices in the definition of N_i. We could equally well have defined M_i according to

$$S_i \equiv M_i\, S_i^o\,. \qquad (80)$$

However, N_i will prove more convenient, and M_i can always be recovered using

$$M_i = S_i^o\, N_i\, S_i^{o\,T}\,. \qquad (81)$$

The estimation of the two alignment matrices associated with sensors that are equipped with autocollimators will not be possible unless $C_{i,k}$ shows sufficient variation. There are two important cases to consider: (1) the autocollimator instruments all three axes of $C_{i,k}$; and (2) the autocollimator instruments only two axes. In the second case, there weill be one axis along which the two misalignments associated with N_i and O_i cannot be determined uniquely. We will treat the first case, initially, since it will provide the background for treating the second case.

Treatment of Autocollimators with Three Instrumented Axes

Attitude-Attitude Measurement Model

We have now in analogy to Eq.(30) that

$$\begin{aligned} A_{1,k}^*\, A_{i,k}^{*\,T} &= (\delta A_{1,k})\, S_1^{o\,T}\, S_i\, C_{i,k}\, T_i\, (\delta A_{i,k})^T \quad (82)\\ &= (\delta A_{1,k})\, S_1^{o\,T}\, S_i^o\, N_i\, C_{i,k}\\ &\qquad \times O_i\, T_i^o\, (\delta A_{i,k})^T\,. \quad (83)\end{aligned}$$

It might seem that we have abandoned coalignments in favor of alignments in our definition of N_i and O_i. However, it would clearly be difficult to define two coalignments from these two quantities, since there is only one coalignment but two corrections. From Eq.(83) we see that N_i can be interpreted equally well either as a correction to the alignment S_i or the coalignment $S_1^T\, S_i$.

We define now

$$\mathcal{Z}_{i,k} \equiv S_i^{o\,T} S_1^o A_{1,k}^*\; A_{i,k}^{*\,T}\, T_i^{o\,T} C_{i,k}^T\,, \qquad (84)$$

which should be compared with Eq.(31) above. This expression differs from that equation in that the contributions of the two alignments are now separated and the appearance of $C_{i,k}$. We may now transform $\mathcal{Z}_{i,k}$ as

$$\begin{aligned} \mathcal{Z}_{i,k} &= S_i^{o\,T} S_1^o\, (\delta A_{1,k})\, S_1^{o\,T}\, S_i^o\, N_i\, C_{i,k}\\ &\qquad \times O_i\, T_i^o\, (\delta A_{i,k})^T\, T_i^{o\,T} C_{i,k}^T \quad (85)\\ &= \left\{S_i^{o\,T}\, S_1^o\, (\delta A_{1,k})\, S_1^{o\,T}\, S_i^o\right\}\\ &\qquad \times N_i\, \left\{C_{i,k}\, O_i\, C_{i,k}^T\right\}\\ &\qquad \times \left\{C_{i,k}\, T_i^o\, (\delta A_{i,k})^T\, T_i^{o\,T} C_{i,k}^T\right\} \quad (86)\\ &= e^{[[\, S_i^{o\,T}\, S_1^o\, \boldsymbol{\xi}_{1,k}\,]]}\, e^{[[\, \boldsymbol{\phi}_i\,]]}\\ &\qquad \times e^{[[\, C_{i,k}\, \boldsymbol{\psi}_i\,]]}\, e^{-[[\, C_{i,k}\, T_i^o\, \boldsymbol{\xi}_{i,k}\,]]}\,, \quad (87)\end{aligned}$$

where we have defined

$$\begin{aligned} \boldsymbol{\phi}_i &= \text{Rtov}(N_i)\,, \quad (88)\\ \boldsymbol{\psi}_i &= \text{Rtov}(O_i)\,. \quad (89)\end{aligned}$$

Since all of the rotations appearing in Eq.(87) are very small, we can write

$$\mathcal{Z}_{i,k} \approx e^{[[\,\boldsymbol{\phi}_i + C_{i,k}\,\boldsymbol{\psi}_i + S_i^{o\,T}\,S_1^o\,\boldsymbol{\xi}_{1,k} - C_{i,k}\,T_i^o\,\boldsymbol{\xi}_{i,k}\,]]} . \tag{90}$$

Defining now

$$\boldsymbol{z}_{i,k}^{(4)} \equiv \mathrm{Rtov}(\mathcal{Z}_{i,k}) , \tag{91}$$

we obtain

$$\begin{aligned} \boldsymbol{z}_{i,k}^{(4)} &= \boldsymbol{\phi}_i + C_{i,k}\,\boldsymbol{\psi}_i + S_i^{o\,T}\,S_1^o\,\boldsymbol{\xi}_{1,k} \\ &\quad - C_{i,k}\,T_i^o\,\boldsymbol{\xi}_{i,k} \qquad (92) \\ &= H_{i,k}^{(4)} \begin{bmatrix} \boldsymbol{\phi}_i \\ \boldsymbol{\psi}_i \end{bmatrix} + \boldsymbol{v}_{i,k}^{(4)} , \qquad (93) \end{aligned}$$

where

$$H_{i,k}^{(4)} \equiv [\ I \quad C_{i,k}\ ] , \tag{94}$$

and

$$\boldsymbol{v}_{i,k}^{(4)} \sim \mathcal{N}\left(\mathbf{0},\, R_{\boldsymbol{z}_{i,k}^{(4)}}\right) , \tag{95}$$

with

$$\begin{aligned} R_{\boldsymbol{z}_{i,k}^{(4)}} &= S_i^{o\,T}\,S_1^o\,R_{\boldsymbol{\xi}_{1,k}}\,S_1^{o\,T}\,S_i^o \\ &\quad + C_{i,k}\,T_i^o\,R_{\boldsymbol{\xi}_{i,k}}\,T_i^{o\,T}\,C_{i,k}^T . \qquad (96) \end{aligned}$$

Attitude-Vector Measurement Model

To construct an effective measurement when sensor i measures only a single vector we note that the representation of the measurement in the body frame is given by

$$\left(\hat{\boldsymbol{W}}_{i,k}\right)_{\mathrm{body}} = S_i\,C_{i,k}\,T_i\,\hat{\boldsymbol{U}}_{i,k} , \tag{97}$$

and

$$\left(\hat{\boldsymbol{W}}_{i,k}\right)_{\mathrm{body}} = S_1\,A_{1,k}\,\hat{\boldsymbol{V}}_{i,k} + S_i\,C_{i,k}\,T_i\,\Delta\hat{\boldsymbol{U}}_{i,k} . \tag{98}$$

In analogy to Eqs.(51) and (57) we define

$$\begin{aligned} \hat{\boldsymbol{W}}_{i,k}^{o} &\equiv C_{i,k}\,T_i^o\,\hat{\boldsymbol{U}}_{i,k} , \qquad (99) \\ \hat{\boldsymbol{W}}_{i,k}^{o\,*} &\equiv S_{1\leftarrow i}^{o\,T}\,A_{1,k}^*\,\hat{\boldsymbol{V}}_{i,k} . \qquad (100) \end{aligned}$$

Equations (97) and (98) can then be equated and solved using Eq.(99) to obtain

$$\begin{aligned} \hat{\boldsymbol{W}}_{i,k}^{o} &= \left(C_{i,k}\,O_i^T\,C_{i,k}\right)\,N_i^T\,\left(S_{1\leftarrow i}^{o\,T}\,(\delta A_{1,k}^*\,S_{1\leftarrow i}^o\right)\,\hat{\boldsymbol{V}}_{i,k} \\ &\quad + \left(C_{i,k}\,O_i^T\,C_{i,k}\right)\,C_{i,k}\,T_i\,\Delta\hat{\boldsymbol{V}}_{i,k} . \qquad (101) \end{aligned}$$

Keeping only first order terms in the noise component and expanding very small rotations, Eq.(100) becomes

$$\begin{aligned} \hat{\boldsymbol{W}}_{i,k}^{o} &= (I - [[C_{i,k}\,\boldsymbol{\psi}_i]])\,(I - [[\,\boldsymbol{\phi}_i\,]]) \\ &\quad \times (I - [[S_{1\leftarrow i}^{o\,T}\boldsymbol{\xi}_{1,k}]])\ \hat{\boldsymbol{W}}_{i,k}^{o\,*} \\ &\quad + C_{i,k}\,T_i^o\,\Delta\hat{\boldsymbol{U}}_{i,k} . \qquad (102) \end{aligned}$$

Defining

$$\boldsymbol{z}_{i,k}^{(5)} \equiv \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\left(\hat{\boldsymbol{W}}_{i,k}^{o\,*} \times \hat{\boldsymbol{W}}_{i,k}^{o}\right) = \mathcal{Q}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\hat{\boldsymbol{W}}_{i,k}^{o} , \tag{103}$$

it follows that

$$\begin{aligned} \boldsymbol{z}_{i,k}^{(5)} &= \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,(\boldsymbol{\phi}_i + C_{i,k}\,\boldsymbol{\psi}_i) \\ &\quad + \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,\big(S_{1\leftarrow i}^{o\,T}\boldsymbol{\xi}_{1,k} \\ &\qquad - [[\,\hat{\boldsymbol{W}}_{i,k}^{o\,*}]]\,C_{i,k}\,T_i^o\,\Delta\hat{\boldsymbol{U}}_{i,k}\big) \qquad (104) \\ &= H_{i,k}^{(5)} \begin{bmatrix} \boldsymbol{\phi}_i \\ \boldsymbol{\psi}_i \end{bmatrix} + \boldsymbol{v}_{i,k}^{(5)} , \qquad (105) \end{aligned}$$

with

$$\begin{aligned} H_{i,k}^{(5)} &= \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,[\ I \quad C_{i,k}\ ] , \qquad (106) \\ \boldsymbol{v}_{i,k}^{(5)} &= \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,S_{1\leftarrow i}^{o\,T}\,\boldsymbol{\xi}_{1,k} \\ &\quad + \mathcal{Q}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,C_{i,k}\,T_i^o\,\Delta\hat{\boldsymbol{U}}_{i,k} . \qquad (107) \end{aligned}$$

Thus,

$$\boldsymbol{v}_{i,k}^{(5)} \sim \mathcal{N}\left(\mathbf{0},\, R_{\boldsymbol{z}_{i,k}^{(5)}}\right) , \tag{108}$$

with

$$\begin{aligned} R_{\boldsymbol{z}_{i,k}^{(5)}} &= \mathcal{P}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,S_{1\leftarrow i}^{o\,T}\,R_{\boldsymbol{\xi}_{1,k}}\,S_{1\leftarrow i}^{o}\,\mathcal{P}^T(\hat{\boldsymbol{W}}_{i,k}^{o\,*}) \\ &\quad + \mathcal{Q}(\hat{\boldsymbol{W}}_{i,k}^{o\,*})\,C_{i,k}\,T_i^o\,R_{\hat{\boldsymbol{U}}_{i,k}} \\ &\quad \times T_i^{o\,T}\,C_{i,k}^T\,\mathcal{Q}^T(\hat{\boldsymbol{W}}_{i,k}^{o\,*}) . \qquad (109) \end{aligned}$$

The estimation of $\boldsymbol{\phi}_i$ and $\boldsymbol{\psi}_i$ proceeds as in Eqs.(72) through (74) except now six parameters ($= 3 + 3$) are estimated and $P_{\theta\theta}$ is replaced by a 6×6 matrix

$$\begin{bmatrix} P_{\phi\phi} & P_{\phi\psi} \\ P_{\psi\phi} & P_{\psi\psi} \end{bmatrix} . \tag{110}$$

Treatment of Poor Observability

We see from Eqs.(92) and (104) that for sensors equipped with autocollimators, the sensed quantity is $\boldsymbol{\phi}_i - C_{i,k}\boldsymbol{\psi}_i$. The two "misalignment" vectors can be separated from a sequence of measurements only if $C_{i,k}$ varies over the data set. However, for short segments of data (say one orbit or less), this may not be the case. Hence, on a single alignment calibration exercise, it may not be possible to separate the two "misalignment" vectors.

Over a much longer time span, a separation of these two quantities may be possible. It is clearly not possible to conduct an alignment calibration exercise over a period of several days or longer. Hence, a modification of the procedure will be necessary.

Suppose that $C_{i,k}$ has an essentially constant value over a data span, which we denote by $\overline{C}_i(m)$, where m is the index of the data set. Since ϕ_i and ψ_i are not separately observable, we will make the following model reduction for data set m

$$T_i(m) = T_i^o , \qquad (111)$$

and estimate only $S_i(m)$. Recalling

$$\hat{\boldsymbol{U}}_{i,k} = T_i^T C_{i,k}^T S_i^T S_1 A_{1,k} \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{V}}_{i,k} , \qquad (112)$$

we define effective measurements for the m-th data set

$$\hat{\boldsymbol{X}}_{i,k}(m) \equiv \overline{C}_i(m) T_i^o \hat{\boldsymbol{U}}_{i,k} , \qquad (113)$$
$$\Delta\hat{\boldsymbol{X}}_{i,k}(m) \equiv \overline{C}_i(m) T_i^o \Delta\hat{\boldsymbol{U}}_{i,k} , \qquad (114)$$

whence

$$\hat{\boldsymbol{X}}_{i,k}(m) = S_i^T S_1 A_{1,k} \hat{\boldsymbol{V}}_{i,k} + \Delta\hat{\boldsymbol{X}}_{i,k}(m) . \qquad (115)$$

This is just the measurement equation for a sensor without an autocollimator. We may therefore use the earlier results provided we make the substitution $\hat{\boldsymbol{U}}_{i,k} \to \hat{\boldsymbol{X}}_{i,k}$.

After treating several data sets this way we are left with a sequence of reduced alignment estimates $S_i^*(m)$, $m = 1, \ldots, M$, from which we wish to estimate N_i and O_i, assuming that there is sufficient variation in $\overline{C}_i(m)$ for this to be possible. The individual alignment estimates may be written as

$$S_i^*(m) = (\delta S_i(m))\, S_i(m) = e^{[[\boldsymbol{\chi}_i(m)]]} S_i(m) , \qquad (116)$$

with $S_i(m)$ the true value. As usual, we assume that $\boldsymbol{\chi}_i(m)$ is zero-mean and Gaussian with covariance matrix $R_{\boldsymbol{\chi}_i(m)}$.

We may write therefore for the correspondence between the reduced model and the complete model

$$S_i(m) \overline{C}_i(m) T_i^o = S_i \overline{C}_i(m) T_i , \qquad (117)$$

for $m = 1, \ldots, M$, or

$$(\delta S_i(m))^T S_i^*(m) \overline{C}_i(m) T_i^o = S_i \overline{C}_i(m) T_i , \qquad (118)$$

for $m = 1, \ldots, M$, which we may solve at each value of m as

$$S_i^*(m) = (\delta S_i(m))\, S_i \overline{C}_i(m) T_i T_i^{oT} \overline{C}_i^T(m) \qquad (119)$$
$$= (\delta S_i(m))\, S_i^o N_i \overline{C}_i(m) O_i \overline{C}_i^T(m) \qquad (120)$$

Let us define now

$$\boldsymbol{z}_i^{(6)}(m) \equiv \mathrm{Rtov}\left(S_i^{oT} S_i^*(m)\right) . \qquad (121)$$

Then for each value of m

$$\boldsymbol{z}_i^{(5)}(m) = \boldsymbol{\phi}_i + \overline{C}_i(m) \boldsymbol{\psi}_i + S_i^{oT} \boldsymbol{\chi}_i(m) \qquad (122)$$
$$= H_i^{(6)}(m) \begin{bmatrix} \boldsymbol{\phi}_i \\ \boldsymbol{\psi}_i \end{bmatrix} + \boldsymbol{v}_i^{(6)}(m) , \qquad (123)$$

with

$$H_i^{(6)}(m) = \begin{bmatrix} I & \overline{C}_i(m) \end{bmatrix} , \qquad (124)$$
$$R_{\boldsymbol{z}_i^{(6)}(m)} = S_i^{oT} R_{\boldsymbol{\chi}_i(m)} S_i^o . \qquad (125)$$

Realistic treatment of $\overline{C}_i(m)$

In general the values of $C_{i,k}$ may show some small variation over a data set, which, however is not sufficient to separate $\boldsymbol{\phi}_i$ and $\boldsymbol{\psi}_i$ convincingly. We must have some method, therefore, of determining the best value of $C_{i,k}$ from the data. Since all autocollimator readings are supposedly equally accurate, the simplest choice is to define

$$B_i(m) \equiv \frac{1}{n_i(m)} \sum_{k=1}^{n_i(m)} C_{i,k}(m) , \qquad (126)$$

which in general is no longer proper orthogonal, and define $\overline{C}_i(m)$ as the value which maximizes

$$g(C) = \mathrm{tr}\left[B_i^T(m) C\right] , \qquad (127)$$

where C is proper orthogonal, and $\mathrm{tr}(\cdot)$ is the trace function. The matrix $\overline{C}_i^*(m)$ so defined will minimize the cost function

$$J(C) = \frac{1}{2} \sum_{k=1}^{n_i(m)} \sum_{j=1}^{3} \sum_{\ell=1}^{3} |(C_{i,k}(m))_{j,\ell} - C_{j,\ell}|^2 , \qquad (128)$$

where j and ℓ are the row and column indices of the individual matrix elements. Equation (127) is simply the gain function of the QUEST algorithm[3], which can now be used to calculate $\overline{C}_i^*(m)$. The QUEST algorithm also gives us a measure of the covariance of $\overline{C}_i^*(m)$, which we can use as an additional term in the error associated with $S_i^*(m)$.

To take account of this additional error, we should modify Eq.(120) to read

$$S_i^*(m) = (\delta S_i(m))\, S_i^o N_i (\delta\overline{C}_i(m))^T \overline{C}_i^*(m) O_i \overline{C}_i^T(m) , \qquad (129)$$

with

$$\delta\overline{C}_i(m) = e^{[[\boldsymbol{\varepsilon}_i(m)]]} \approx I + [[\boldsymbol{\varepsilon}_i(m)]] , \qquad (130)$$

and the covariance of $\boldsymbol{\varepsilon}_i(m)$ is denoted by $R_{\boldsymbol{\varepsilon}_i(m)}$. Thus, we have

$$\boldsymbol{z}_i^{(6)}(m) = \boldsymbol{\phi}_i + \overline{C}_i^*(m) \boldsymbol{\psi}_i + S_i^{oT} \boldsymbol{\chi}_i(m) - \boldsymbol{\varepsilon}_i(m) , \qquad (131)$$
$$= H_i^{(6)}(m) \begin{bmatrix} \boldsymbol{\phi}_i \\ \boldsymbol{\psi}_i \end{bmatrix} + \boldsymbol{v}_i^{(6)}(m) , \qquad (132)$$

so that now

$$H_i^{(6)}(m) = [\ I \quad \overline{C}_i^*(m)\ ]\,, \tag{133}$$

$$v_i^{(6)}(m) = S_i^{o\,T}\,\chi_i(m) - \varepsilon_i(m)\,, \tag{134}$$

and

$$R_{z_i^{(6)}(m)} = S_I^{o\,T}\,R_{\chi_i(m)}\,S_i^o + R_{\varepsilon_i(m)}\,. \tag{135}$$

Normally, $R_{\varepsilon_i(m)}$ is a product of the QUEST[3] or some other attitude estimatin algorithm algorithm. However, if the autocollimator supplies only two independent angles characterizing $C_{i,k}$, then $R_{\varepsilon_i(m)}$ will be singular. In this case the QUEST algorithm cannot be expected to give the correct value for $R_{\varepsilon_i(m)}$, especially since this quantity will be singular. It will be sufficient for our purposes to calculate $R_{\varepsilon_i(m)}$ as a sampled covariance. To this end we define

$$\varepsilon_{i,k}^{\text{sampled}} = \text{Rtov}\left(C_{i,k}\,\overline{C}_i^{*(m)\,T}\right)\,, \tag{136}$$

and compute the sampled covariance as

$$R_{\varepsilon_i(m)} \approx \frac{1}{n_i(m)}\sum_{k=1}^{n_i(m)} \varepsilon_{i,k}^{\text{sampled}}\,\varepsilon_{i,k}^{\text{sampled}\,T}\,. \tag{137}$$

Treatment of Autocollimators with Two Instrumented Axes

If only two axes of the autocollimator are instrumented, we can write

$$C_{i,k} = e^{[[\,\alpha_{i,1}\hat{\mathbf{u}}_{i,1}+\alpha_{i,2}\hat{\mathbf{u}}_{i,2}\,]]} \tag{138}$$

Since $\alpha_{i,1}$ and $\alpha_{i,2}$ will be small, it follows that the quantity $\hat{\mathbf{u}}_{i,3}\cdot(\phi_i - \psi)$, where

$$\hat{\mathbf{u}}_{i,3} \equiv \hat{\mathbf{u}}_{i,1}\times\hat{\mathbf{u}}_{i,2}\,, \tag{139}$$

is unobservable. We have no choice then, but to estimate the parameters of a reduced model in which ψ_i is replaced by ψ_i' with

$$\psi_i' \equiv \psi_{i,1}'\,\hat{\mathbf{u}}_{i,1} + \psi_{i,2}'\,\hat{\mathbf{u}}_{i,2} \tag{140}$$

$$= \mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\,\psi'\,, \tag{141}$$

where $\mathcal{P}^T(\hat{\mathbf{u}})$ is given by Eq.(62). Equations (131) and (132) now become

$$z_i^{(6)}(m) = \phi_i + \overline{C}_i^*(m)\,\mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\,\psi_i' + S_i^{o\,T}\,\chi_i(m) - \varepsilon_i(m)\,, \tag{142}$$

$$= H_i^{(6)}(m)\begin{bmatrix}\phi_i\\ \psi_i\end{bmatrix} + v_i^{(6)}(m)\,, \tag{143}$$

with now

$$H_i^{(6)}(m) = [\ I \quad \overline{C}_i^*(m)\mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\ ]\,. \tag{144}$$

Once having estimated ψ_i', we can compute ψ_i^* from

$$\psi_i^* = \mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\,\psi_i'^{\,*}\,, \tag{145}$$

and the (singular) 6×6 estimate-error covariance matrix can be determined from the (nonsingular) 5×5 covariance matrix according to

$$\begin{bmatrix} P_{\phi\phi} & P_{\phi\psi}\\ P_{\psi\phi} & P_{\psi\psi}\end{bmatrix} = F\begin{bmatrix} P_{\phi\phi} & P_{\phi\psi'}\\ P_{\psi'\phi} & P_{\psi'\psi'}\end{bmatrix}F^T\,, \tag{146}$$

with

$$F = \begin{bmatrix} I_{3\times3} & 0_{3\times2}\\ 0_{3\times3} & \mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\end{bmatrix}\,. \tag{147}$$

In the case where $\hat{\mathbf{u}}_{i,3}$ is not known *a priori*, it can be determined by finding the singular axis of R_{ε_i},

$$R_{\varepsilon_i}\,\hat{\mathbf{u}}_{i,3} = \mathbf{0}\,, \tag{148}$$

In a similar manner, in the unlikely case that ϕ_i and ψ_i' are observable while ϕ_i and ψ_i are unobservable, we can replace Eq.(93) by

$$z_{i,k}^{(4)} = H_{i,k}'^{(4)}\begin{bmatrix}\phi_i\\ \psi_i'\end{bmatrix} + v_{i,k}^{(4)}\,, \tag{149}$$

where

$$H_{i,k}^{(4)} \equiv [\ I \quad C_{i,k}\,\mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\ ]\,. \tag{150}$$

Similarly, Eq.(105) becomes

$$z_{i,k}^{(5)} = H_{i,k}^{(5)}\begin{bmatrix}\phi_i\\ \psi_i\end{bmatrix} + v_{i,k}^{(5)}\,, \tag{151}$$

with

$$H_{i,k}'^{(5)} = \mathcal{P}(\hat{W}_{i,k}^{o\,*})\,[\ I \quad C_{i,k}\,\mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\ ]\,, \tag{152}$$

and Eq.(123) is likewise tranformed by the substitutions

$$\begin{bmatrix}\phi_i\\ \psi_i\end{bmatrix} \rightarrow \begin{bmatrix}\phi_i\\ \psi_i'\end{bmatrix} \tag{153}$$

and

$$C_{i,k} \rightarrow C_{i,k}\,\mathcal{P}^T(\hat{\mathbf{u}}_{i,3})\,. \tag{154}$$

Observability Criteria for Autocollimator Parameters

We require a criterion for when it will be desirable to estimate the two alignment matrices associated with sensors equipped with autocollimators and when one should attempt to estimate them in one step or two. To develop this criteria let us consider a simple one-dimensional case described by the measurement equation

$$z_k = x + h_k\, y + v_k\,, \quad k = 1, \ldots, N\,, \tag{155}$$

where v_k is a zero-mean Gaussian random variable with standard deviation σ and h_k has a zero-mean Gaussian distribution (as a function of k) with standard deviation τ, $\tau << 1$. Note that the distribution of h_k, although Gaussian, need not be random. In Eq.(155), x plays the role of $\phi + \psi$, and y plays the role of $\phi - \psi$. The estimate-error covariance matrix for x and y is given straightforwardly by

$$P = \frac{\sigma^2}{N}\begin{bmatrix} 1 & 0 \\ 0 & 1/\tau^2 \end{bmatrix}. \tag{156}$$

If the physical quantity of ultimate interest is of the form

$$u_k = x + h_k\, y\,, \tag{157}$$

then the error covariance of u will be

$$\sigma_u^2 = \frac{\sigma^2}{N}\,(1 + h_k^2/\tau^2)\,, \tag{158}$$

so that the contribution of the estimate errors for x and y contribute equally to the estimate errors of u.

If, on the other hand, we neglect y in our model, then we will make a modeling error of $h_k\, y$, which is on the order of $\tau\sigma_y$, where σ_y^2 is the second moment of y. Thus, the criterion for including y in the model is

$$\tau^2\,\sigma_y^2 > \sigma^2/N\,, \tag{159}$$

or

$$\sigma_y > \frac{\sigma}{\tau\sqrt{N}}\,. \tag{160}$$

If τ is on the order of 10 arc minutes, and σ is on the order of 5 arc seconds, then Eq.(160) becomes

$$\sigma_y > \frac{1}{120\sqrt{N}}\ \text{rad} \approx \frac{0.5}{\sqrt{M}}\ \text{deg}\,. \tag{161}$$

Thus, if we expect that the *a priori* uncertainly in the orientation of the autocollimator axes is on the order of 0.5 deg$/\sqrt{N}$, then it will be necessary to model both alignments associated with the autocollimator. If, however, the *a priori* uncertainty is expected to be much less than this, then there will be no need to model both alignments. A good practical measure of this uncertainly is simply to estimate ϕ_i assuming that $\psi_i = 0$. The magnitude of ϕ_i is then a good gauge of σ_y.

It is possible that τ may be small in the sense above for a given data interval of interest but large when the entire data set is considered. It is this case in which we will wish to first estimate only $phiv_i(m)$ for each data interval and then estimate ϕ_i and ψ_i from the cumulative sequence of estimates of $\phi_i(m)$.

Note also that if only two axes of $C_{i,k}$ are instrumented but that the angles associated with these two axes can assume large values (so that their product (in radians) is not negligible compared to the individual angles, then there will be sufficient variability along all three axes of $C_{i,k}$ that it will be possible to estimate ψ_i rather than just ψ_i'.

Acknowledgements

The author is grateful to H. Landis Fisher, Jr., for pointing out several corrections in this work.

References

[1]Shuster, M. D., Pitone, D. S., and Bierman, G. J., "Batch Estimation of Spacecraft Sensor Alignments, I. Relative Alignment Estimation," *Journal of the Astronautical Sciences*, Vol. 39, No. 4, pp. 519–546, October–December, 1991.

[2]Shuster, M. D., and Pitone, D. S., "Batch Estimation of Spacecraft Sensor Alignments, II. Absolute Alignment Estimation," *Journal of the Astronautical Sciences*, Vol. 39, No. 4, pp. 547–571, October–December, 1991.

[3]Shuster, M. D., and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, pp. 70–77, January-February 1981.

[4]Markley, F. L., "Attitude Determination Using Vector Observations : A Fast Opmtimal Matrix Algorithm," *Proceedings, Flight Mechanics/Estimation Theory Symposium Journal of Guidance and Control*, May 1992.