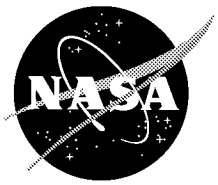
NASA/CP—2001–209986



2001 Flight Mechanics Symposium

Edited by
John P. Lynch
NASA Goddard Space Flight Center
Greenbelt, Maryland

Proceedings of a conference sponsored and held at
NASA Goddard Space Flight Center
Greenbelt, Maryland
June 19–21, 2001

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

June 2001

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076–1320
Price Code: A17

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Price Code: A10

FOREWORD

The papers included here have been presented in summary form at the Flight Mechanics Symposium held June 19-21, 2001, at the NASA Goddard Space Flight Center. For completeness, abstracts have been included for those papers that were presented but unavailable at the time of printing. The papers in this document are presented as received from the authors, with little or no editing.

Table of Contents

Session 1: Trajectory Design, Autonomous Navigation and Computational Techniques

Session 2: Constellation Design and Formation Flying

Session 3: Attitude Determination, Sensor Calibration and Computational Techniques

Session 4: Attitude Control System Design and Simulation

Session 5: In-Flight Orbit Experience

Session 6: In-Flight Attitude Experience

Session 1: Trajectory Design, Autonomous Navigation and Computational Techniques

SOLAR ACTIVITY FORECASTING FOR USE IN ORBIT PREDICTION

Kenneth Schatten*
a.i. solutions, Inc.
10001 Derekwood Lane, Suite 215
Lanham, MD 20706

ABSTRACT

Orbital prediction for satellites in low Earth orbit (LEO) or low planetary orbit depends strongly on exospheric densities. Solar activity forecasting is important in orbital prediction, as the solar UV and EUV inflate the upper atmospheric layers of the Earth and planets, forming the exosphere in which satellites orbit. Geomagnetic effects also relate to solar activity. Because of the complex and ephemeral nature of solar activity, with different cycles varying in strength by more than 100%, many different forecasting techniques have been utilized. The methods range from purely numerical techniques (essentially curve fitting) to numerous oddball schemes, as well as a small subset, called "Precursor techniques." The situation can be puzzling, owing to the numerous methodologies involved, somewhat akin to the numerous ether theories near the turn of the last century. Nevertheless, the Precursor techniques alone have a physical basis, namely dynamo theory, which provides a physical explanation for why this subset seems to work. I discuss this solar cycle's predictions, as well as the Sun's observed activity. I also discuss the SODA (SOlar Dynamo Amplitude) index, which provides the user with the ability to track the Sun's hidden, interior dynamo magnetic fields. As a result, one may then update solar activity predictions continuously, by monitoring the solar magnetic fields as they change throughout the solar cycle. This paper ends by providing a glimpse into what the next solar cycle (#24) portends.

* This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under GSA Schedule Contract GS-35F-4899G, Purchase Order S-37707-G.

INTRODUCTION

This paper will focus on long term solar activity predictions. I discuss why solar activity forecasting is useful for satellite orbit determination, and then concentrate on solar activity forecasting methods. I shall end this with a rough outline of how the "SOlar Dynamo Amplitude" method (SODA index) is used to predict solar activity, and also the physical basis of the solar dynamo method. This method has predicted 3 solar cycles quite well, having first been tested with 8 prior cycles of data. The paper begins with a discussion of how solar activity forecasting is useful for flight dynamics.

Satellites in low Earth orbit (LEO) or low planetary orbit have paths that depend strongly on exospheric densities, since these densities affect satellite drag. Of course, satellite and orbital properties (ballistic coefficient related to drag vs. mass, orbital properties, etc.) are also important. It is not through the satellite properties, but rather through the variations in the exospheric densities, that solar activity plays a strong role in orbital paths.

Interestingly, the Sun, whose nearly constant energy source provides a very stable harbor for life on Earth, provides a very unstable environment for satellites in orbit. The reason is that the solar UV and EUV irradiances vary dramatically with solar activity (showing changes at certain wavelengths of more than 100%), and this energy inflates the upper atmospheric layers of the Earth and planets, forming the exosphere in which satellites orbit. The solar UV radiation ionizes oxygen, forming ozone in the Earth's stratosphere; and above this, the more intense EUV forms the hotter thermosphere and exosphere, in which terrestrial satellites orbit. Solar activity energy, as opposed to the more constant solar luminosity, thus inflates the Earth's atmosphere into its upper layers. These upper layers vary exponentially with the exotic forms of solar radiation, making them more sensitive to the some of the most variable forms of solar radiation. Hence satellite drag is greatly magnified by solar activity. Consequently, solar activity forecasting is a valuable tool for orbit predictions. This exospheric behavior contrasts greatly with tropospheric behavior, where meteorologists traditionally, yet safely, ignore solar irradiance changes. The solar irradiance is called as a misnomer, the "solar constant," since its changes are about 0.1%, much smaller than the EUV variations which often exceed 100%! Let us now examine solar activity behavior and how this is predicted.

Figure 1 shows solar activity as measured by sunspots for the past few centuries. Firstly, the most prominent feature is the famous 11-year Schwabe solar cycle. Additionally, there is "power" at a variety of other periods (including days, months, 80 - 100 years, and even secular variations with longer periods not seen here), however, a closer examination of these periods shows that they are <u>not</u> strictly periodic. They appear "chaotic;" unable to be predicted on the basis of simple spectral techniques. Additionally, even the main 11-year periodicity is <u>not</u> a sharp period; some cycles have been as short at 8 years, and some as long as 17 years! Further, the amplitude of the cycles varies by more than 100%, in a rather chaotic manner. This is even more striking, if one realizes that there were long periods of time (50 - 100 years), most recently in the 1600's called the "Maunder minimum" when solar activity was near 0. Hence solar scientists have developed numerous methods in their attempts to forecast solar activity. Unfortunately, a number of these techniques, although successful for some cycles, have failed in the long run.

Methods of solar activity forecasting may be placed in general categories. A NASA-funded NOAA panel investigating these methods several years ago (see Joselyn et al.[1]) chose the following general solar forecasting categories: Even/Odd Behavior, Spectral, Recent Climatology, Climatology, Neural Networks, and Precursors. Briefly, these categories indicate the following:

- Even/Odd Behavior - uses the relationship that this century (see Figure 1), and for most of the last century as well, the Odd numbered cycles have been larger than the preceding Even Numbered cycles.
- Spectral - analyzing activity by spectral methods, such as Fourier analysis.
- Recent Climatology - simply averaging recent cycles, say the last 5.
- Climatology - using statistics of the longest duration of solar cycles (although usually leaving out the Maunder Minimum) to obtain a mean cycle and standard deviation to obtain a simple mean and uncertainty.
- Neural Networks - using AI - artificial intelligence methods on solar activity.
- Precursors - physical phenomena related to future solar activity levels. The Precursor category has been subdivided into geomagnetic and solar magnetic branches, since they gave slightly different prediction levels for this cycle[2].

The above methods suggest the values for solar cycle #23's smoothed peak sunspot numbers shown in Table 1. It now appears from the behavior of this cycle that it will peak near a smoothed sunspot number near 125 ± 15, or an F10.7 Radio Flux of 180 ± 15. The sunspot value is near the lower end of most estimates seen in Table 1. From an examination of Table 1, one sees the solar cycle behavior, for this cycle, supports the Climatology, Neural Networks, and Solar Precursor methods. Geomagnetic Precursors predicted too large a value, partially due to the Geomagnetic Precursors reaching their minimum values late. Additionally, the Even/Odd behavior predicted too large a value.

There are other methods that might be added to these categories, such as the McNish-Lincoln method[3], which although not originally thought to be applicable to predictive capability beyond a year in advance, has been used by MSFC for longer duration forecasts. It simply is a kind of "curve fitting" technique, with a regression towards a mean cycle, and thus is considered in the climatological category. In addition, some have advocated using planetary orbits, since Jupiter's 11-year period is close to a solar cycle. Hence gravitational changes in the solar system do correlate for short intervals with solar activity, though any agreement is simply a fortuitous short-lived one.

I shall make a few general comments about these techniques and then concentrate upon Precursor methods, which have been shown to be generally useful over several solar cycles and have a physical basis, as well as statistical support. As mentioned, the Precursor category may be subdivided into solar and geomagnetic varieties. This past cycle the solar Precursors suggested a peak value[2] near smoothed sunspot number of 138 ± 30 in 1996 and F10.7 values of 182 ± 30 (Schatten, Myers, and Sofia[2]). This value was lower than the geomagnetic predictions. Let us now go through the Table 1 list, and provide some comments about these "traditional" methods.

The first method, Even/Odd, can be seen in the 20th century of Figure 1, wherein if one numbers the cycles (the present one is #23), then for this past century all odd numbered cycles have been larger than the preceding even-numbered one. This Even/Odd effect seemed to us like a small statistical fluke. Consider that there have been only 4 previous such pairs this century - cycles 14 through 21. If one takes the first pair, one or the other must be larger, hence one can't count the first pair to support any effect. Hence, there are only 3 matching subsequent pairs, the chance of these three agreeing with the first is only one in eight, or a significance of 87%, not highly significant! If one includes the last century's cycles, one adds 4 more pairs, but only 3 agree with the pattern. Additionally, the previous pair falls half within the previous century, but this pair disagrees with the Even/Odd effect. This marginally changes the statistics. Additionally, using this method would suggest that cycle #23 should be the largest cycle ever, it being an odd cycle following the largest even numbered cycle! One is fortunate not to believe in the Even/Odd effect, using the reason is that there was no physical basis for it because this enormous activity does not appear to be happening!

The second method, Spectral, has seen the most attention throughout the history of solar activity predictions, since Fourier analyses have been readily available, and the methods generally useful for periodic phenomena. In recent times, however, with new knowledge of "chaotic" systems, it has become generally recognized that Fourier or generally spectral methods do not lend themselves to understandings of chaotic phenomena. The most obvious area where this is seen is in weather prediction. A weather forecaster would not be well received if s/he tried to forecast the weather at a location on Earth simply by taking parameters at that location and Fourier analyzing them. The reason, of course, is that knowledge in a chaotic system becomes "lost," whereas in a Fourier or spectral method, the coefficients are as dependent on what happened recently as on what happened in the distant past. Further, examination of long-term solar activity, using cosmogenic isotopes reveals that commonly accepted cycles (e.g. the Gleissberg cycle of 80-100 years), do not bear out over long periods, with significant power in the solar activity cycle over periods in the few hundred to thousand year range. This would be expected if one considers the "time constant" expected for changes, at the base of the convection zone, where magnetic fields are regenerated.

Next, we come to Recent Climatology, and Climatology. These basically, simply use the statistics of known solar cycles, with climatology using the last few hundred years, from after the Maunder Minimum to present. One obtains a mean and standard deviation to get the chances of any size cycle. Recent climatology recognizes the "chaotic nature" of solar activity, and thus eliminates the early data. One develops a "recent average" and "recent statistical behavior" from the last few cycles. This method relies strictly upon "persistence," which is at the heart of the Precursor methods too, although they have an added physical component. One of the earliest solar activity prediction methods in this category is the curve fitting techniques of McNish and Lincoln. In this method the recent

behavior gradually blends into the average behavior, however, the technique was to be utilized for no more than a year in advance.

The next set of methods is Neural Networks - using AI - artificial intelligence methods on solar activity. Unfortunately, these have been used primarily to "curve fit." Although AI can be a powerful technique, but since, at present, this method has only been supplied with past activity, an inadequate source to predict future activity from, the technique suffers from an inadequate data base to make long term predictions. Nevertheless, the AI technique seems to provide a curve fitting technique similar to that of the McNish-Lincoln method, which works for about 1 year in advance. Over a long time interval, this method reverts to the climatological method, and thus provides little more than an average. The paper now discusses Precursor methods more fully, distinguishing solar from geomagnetic Precursor methods. In general, for more information on solar activity prediction methods including Precursor methods, the NOAA panel discussions[1] provides an excellent source. Further, other views including more detail on climatological and statistical methods may be found in Holland and Vaughan[4], Kerridge, et al.[5] and Hathaway, et al.[6].

SOLAR AND GEOMAGNETIC PRECURSOR METHODS

Geomagnetic Precursors

As far as geomagnetic Precursors, one of the first to point out the significance of the geomagnetic Aa index in tracking long-term solar activity was Feynman[7]. Although she never used the information for directly making predictions, it seems clear that Ohl[8,9] and other Geomagnetic Precursor practitioners used the same or similar methods to predict activity. Feynman separated the geomagnetic Aa index into two components: one in phase with sunspot number, and one out of phase. This effectively led to "active" and "quiet" components. She found that this quiet signal tracked the sunspot numbers several years in advance, similar to the Ohl results. The maximum in this signal occurs at sunspot minimum and is proportional to the sunspot number during the following maximum. How this signal propagated or why it should be present, however, was not clear.

Precursor methods were developed by the Soviet geophysicist Ohl[8,9] to make solar predictions and taken up by Brown and Williams[10], who later noticed an extremely high correlation (close to 1) between geomagnetic activity near solar minimum and the size of the next solar cycle. High correlations were found between the number of "geomagnetic abnormal quiet days" and the size of the next solar cycle. Although the abnormal quiet day geomagnetic index was an unusual one, later the correlations remained high when objective geomagnetic indices, such as Ap, and Aa were employed. Bartels[11] discusses these indices. Thompson[12] further improved upon the relationships between geomagnetically "disturbed" days and the amplitude of the next sunspot maximum.

Let us move on to Precursor methods by pointing out that the correlations found by the geophysicists were very puzzling because the Sun's activity might cause a terrestrial effect, but not vice versa! So the order of the causality seemed to be reversed. Trying to unravel the mystery of how the Sun could broadcast to the Earth, in advance, the level of its future activity, Schatten et al.[13] searched for a physical mechanism to understand the phenomena. To place these puzzling correlations in a physical context meant relating these geomagnetic effects somehow to solar dynamo theory. Let us examine how this is done.

Solar Dynamo Theory and Solar Precursors

In any MHD (magnetohydrodynamics) dynamo, there is a behavior similar to the simple "disk dynamo" shown in Figure 2. This property is that the amplified magnetic field is proportional to the initial magnetic field. The solar dynamo goes through more "gyrations" than the simple disk dynamo, but the previous behavior remains. Figure 3 shows the Babcock dynamo, generally agreed upon for the Sun, as there are many observed solar features explained by this model. Namely, in this model the Sun's polar fields near solar minimum are wrapped up by differential rotation to form the toroidal fields, which later float to the Sun's surface and erupt to form active regions. As these fields dissipate, they then regenerate the polar field allowing the solar cycle to recur. Modern helioseismological studies have shed new light on the Sun's dynamo. For example, the solar community now knows that the buried toroidal dynamo field is located just below the base of the convection zone. Nevertheless, the broad view outlined by Babcock still remains valid.

Let us now see how making key observations and processing them, based on this dynamo paradigm, will allow us to gain an understanding of the Sun's buried magnetic flux and to better predict solar activity. The dynamo process outlined is neither as simplified nor "perfect" as outlined, but rather subject to the irregularities of the individual active regions formed. Hence, over an 11- year solar cycle, the amplification sometimes regenerates more polar field and sometimes less, leading to a growth or decay in the solar cycle. If one assumes the dynamo is fairly linear, then one expects a direct correlation between the number of active regions formed in a cycle with the strength of the Sun's polar field near the prior solar minimum. Figure 4 shows a simple schematic picture of the Babcock dynamo mechanism for two different cycles: a weak cycle (left) and a strong cycle (right). The complex picture of the temporal dynamics of a solar cycle is simplified in this picture so that the polar field and toroidal field are both displayed. The main point, however, is that during a weak cycle, a weak polar field is amplified by dynamo physics below the Sun's surface into few sunspots, solar activity, etc. For a strong cycle, the reverse is true. Since the polar field of the Sun is later amplified into the sunspot fields, one can use it as a precursor or predictor of solar activity. Namely, by monitoring the observed magnetic fields of the Sun, one can use these observations to predict future levels of solar activity. This is similar to the way meteorologists monitor pressure regions to predict cloud formation. Hence it is the first "physics-based" forecasting technique.

To understand the Geomagnetic Precursor methods, we then only needed to see whether the Sun's polar field near solar minimum might be correlated with the amount of geomagnetic activity at that time. The extended solar field is called the Interplanetary magnetic field (IMF), and was found to be correlated to geomagnetic activity[14] supporting Dungey's theory of magnetic reconnection. It was found that the southward component of the IMF (and also the total field) correlated well with geomagnetic activity. Additionally, models of coronal structure near solar minimum, when the low latitude solar fields are weak, show that the polar fields arc towards the equator to fill the low latitude heliosphere. Hence, the correlation between geomagnetic activity near solar minimum and the size of the next solar cycle seemed physically reasonable.

To test this hypothesis Schatten et al.[13] used 8 solar cycles of historic data, and found reasonable correlations, although not as good as those found by the geomagneticians[8,9,10,12]. Until recently, solar magnetic measurements could not be used directly, and instead solar "proxy" fields were used (estimated from numerous solar indices, ranged from solar polar faculae, to the shape of the Sun's corona) which were not as well measured as the geomagnetic indices. Nevertheless, the correlations were reasonable. At present we can measure directly the Sun's polar fields; Schatten and colleagues have been basing their predictions primarily from solar magnetism (the so-called "Solar Precursor method"), although geomagnetic methods were also examined to augment and check the methods (Sofia et al.[15]). The Sun's polar fields represent an excellent Sun-Earth connection to explain the correlation, as these fields are the main apparent physical manifestation of the Sun's dynamo near solar minimum.

The SODA (SOlar Dynamo Amplitude) Index

When the solar dynamo method was first developed[13], it was only possible to assess the state of the Sun's dynamo near each solar minimum, when the Sun's buried magnetic fields poke through the Sun's surface at the poles and these fields may directly be observed. Schatten and Pesnell[16] developed a more sophisticated method for undertaking the analysis than was done in the early days of this field. This now allows an estimation of the "magnetic state" of the Sun to be ascertained during any phase of the solar cycle, rather than only at solar minimum. As the solar cycle progresses, there is an interchange between poloidal and toroidal magnetic field (see Figure 4). This interchange is similar to the interchange between the kinetic and potential energies of a pendulum. One can measure both, and obtain a measure of the total energy of the pendulum rather than measuring only one, when that one maximizes. Expanding this idea allowed Schatten and Pesnell[16] to capitalize on all the aspects of solar activity and magnetism to obtain a combined index, called the "SOlar Dynamo Amplitude" or SODA index. Just as with the total energy of a pendulum, use of this index can be updated during any phase of the Sun's solar cycle. Through a combined measure (the SODA index), the strength of the Sun's buried magnetic flux is obtained. Figure 5 shows the 11-year oscillations of the poloidal and toroidal field, plus their secular changes. By using both indices, the combined SODA index, shows less 11-year variation, but retains the Sun's secular changes, thereby capturing the slowly varying strength of the Sun's dynamo fields while allowing the state of the dynamo to be monitored continuously! Note that it is important that removing the 11-year variation is not done with spectral filtering, as this would require having current conditions dependant upon old temporal variations, and hence would completely mitigate the benefits gained by updating conditions with the latest information (it would smooth the data out)!

Let us discuss other properties of the SODA index. Firstly, it provides a continuous measure of the strength of the magnetic field buried within the Sun's interior. Additionally, since the magnetic field in the interior of the Sun is "buoyant" (as the magnetic field pressure excludes plasma), the field acts like a gas in a liquid (e.g. carbon dioxide inside a carbonated drink). Hence, the SODA index terminology is not only an acronym for the "solar dynamo amplitude" measure of the magnetic field, but also as a descriptor of the amount of magnetic "fizz" inside the Sun's interior! Figure 5 shows the SODA index in recent times. It has been somewhat down from cycle #22, suggesting (several years ago) that cycle #23 would be somewhat reduced (which has been born out). This incidentally goes against the Even/Odd behavior mentioned earlier about cycles this century! Using the SODA index we[2] predicted a value at the lower end of the Precursor methods shown in Table 1, namely a smoothed sunspot number of 138 ± 30 in 1996 and F10.7 values of 182 ± 30. This is somewhat less than the NOAA panel estimate[1]. Let us now see how cycle #23 has progressed.

SOLAR CYCLE #23 FORECASTS AND OBSERVED BEHAVIOR

It now appears that the solar cycle is certainly near, and likely just past, its peak activity for this cycle, #23. In any case, a good estimate of the size of the current cycle can now be ascertained from the observed data, rather than from the forecasting techniques outlined, barring catastrophic events. As the cycle was predicted to peak in the year 2000, if one takes the time period 1/2000 through 12/2000, one obtains a good estimate of the average maximum activity. For sunspot number the mean value is close to 125 and for F10.7, 180. Activity would need to shoot up markedly for these numbers to be far off from the eventual smoothed average for cycle #23. Figure 6 shows F10.7 radio flux data over the past 50 years, along with the past three predictions[13,2,16], including cycle #23. Examining Figure 6, one notes that timing of earlier cycles was off by ± 1 year roughly. We have, however, developed methods of improved timing this cycle seems to have been much closer to the predictions. Additionally, although perhaps fortuitously, the accuracy of the smoothed peak prediction also seems to have improved, namely cycle #23's prediction fits the observed data better than the earlier predictions.

An Early Prospective on Cycle #24

Let us end with an early prospective on solar activity for cycle #24 (years 2005 - 2016). Although the polar field of the Sun is just beginning its growth towards a new peak, it has already reversed sign and gone through zero from last cycle's polar field. This, incidentally, is another indication that cycle #23 may have already reached its peak amplitude. Let us examine the first year's polar field rise.

Wilcox Solar Observatory polar field strength measured in the polemost 3' aperture shows magnetograms averaged each 10 days. They provide the following behavior[17]: this past year, 2000, the smoothed mean polar field rose from zero to 0.22 Gauss, compared with only 0.12 Gauss over a similar period for the last cycle, a decade earlier. Thus this cycle's polar field is beginning its rise twice as fast as last cycle's. Although one year is a small fraction of a full solar cycle, the polar field often rises in only 2-3 years, and one year is a significant fraction of this. Additionally, as MHD (magnetohydrodynamic) magnetic fields are only amplified from pre-existing fields, the sharp rise may portend a large cycle #24. Nevertheless, one shall certainly have to wait for the next couple of years to see how the Sun's fields grow before one can make more definitive predictions of the next cycle.

CONCLUSIONS AND FUTURE PROGRESS

This paper discusses predicted levels of activity for this solar cycle, how the cycle has behaved, and where solar activity predictions may lead. Currently, solar cycle #23 seems to have reversed the trend this past century of odd numbered cycles having larger amounts of solar activity than even numbered cycles. At present, the peak smoothed sunspot number for this cycle is near 125 ± 15 and F10.7 Radio Flux near 180 ± 15, based upon the observed behavior. For this cycle, climatological and neural network prediction methods have given reasonably good values. Precursor methods, dominated by the geomagnetic variety, have given slightly high values. The solar Precursor method seems to have been very close to the observed cycle's behavior. Additionally, it has been reasonably successful in two prior predictions. Further, it gains support by having a physical basis for its workings - solar dynamo theory. Let us also mention that the SODA (solar dynamo amplitude) method went against common expectations in a couple of predictions: it predicted cycle #22 (an even-numbered cycle) would be exceptionally large (for an even numbered cycle), and also that cycle #23 would break the Even/Odd effect.

Thus, as far as the field of solar activity prediction, there is some support that just as in the early days of weather forecasting, we have hits and misses, but the field as a whole is making gains both in understanding and improvements in forecasting skills. Nevertheless, there are still areas where the field of solar activity prediction has little or minimal ability. Let us at least mention the areas where the solar community has little ability and hence where improvements would be helpful. Even if embarrassing, discussing these areas at least points the way where room exists for future progress.

Solar activity prediction/forecasting is lacking, or could make improvements in the following areas:

1. Short Time Scales - Solar predictions on short time scales, less than decadal, have not been highly successful. For each time scale, one needs to consider the data involved, and must utilize sufficient observations consistent with the time scale one wishes to predict. Since solar variations occur both on very short time scales (days) and moderate time scales (monthly to yearly), there certainly is room for improved prediction on these time scales. High quality data sources with appropriate physical methods would need to be employed, considering Sun-Earth geometry, the location of active regions on the Sun's disk, etc.
2. Solar Cycle Timing - Although progress in this area[18] has been made by utilizing the equatorward march of active regions as the cycle progresses, a number of solar physicists have pointed out the inadequacies in how one labels and number solar cycles. Namely, solar physicists do this numbering from solar minimum, which is an ill-defined quantity, subject to minor fluctuations in a number near zero! This just points out one of many inadequacies in our methodologies.
3. Solar Activity "Size" - By this is meant the indices, levels, strength, or generically "size" of solar activity as it is quantified. From sunspot number to the present day, commonly used F10.7 radio flux, the field uses either ill-defined indices or indices which only approximately measure what really is needed. Namely, for flight dynamics as an example, one often uses F10.7 Radio Flux, and geomagnetic activity levels to calculate exospheric densities. The real exospheric densities in which satellites orbit, are affected by upper atmospheric processes and their interaction with solar UV and EUV fluxes, as well as geomagnetic behavior. Thus if one were to search for improvements in satellite orbital prediction on short time scales, one might find that our overall indices scheme (e.g. F10.7) is inadequate to quantify the solar flux. Thus, one would need better ways to quantify solar activity. Tobiska[19] and colleagues are making progress in this area.
4. Solar Activity "Shape" - This may be folded into time scales. However, one may consider "shape" a separate topic here. The point is, that even though the curve one uses for predictions may capture "average" behavior reasonably well, one could also use updated solar observations to make shape changes, not just to improve the overall level. For example, just as the SODA index updates the state of buried solar magnetic flux, one could, in principal, use updated solar field information for details of future activity "shape." If solar magnetism increases on the Sun at a particular latitude, one could ascertain where and when this magnetism would be amplified and affect future solar activity levels. Although this may seem a tall order, it should be mentioned that there is some basis for this; for the past few cycles I have examined North vs. South solar activity levels and find some correlation with polar field levels in those hemispheres! So, even though this is at a global level, it has the hope of being directed more locally in the future.

Thus, the field of solar activity predictions is interesting scientifically and practically, both for the valuable information it provides in dynamo physics and for its usefulness to NASA and other agencies interested in solar activity related phenomena, ranging from power grid spikes, to communication blackouts, to satellite orbital dynamics.

REFERENCES

1. Joselyn, J.A., J.B. Anderson, H. Coffey, K. Harvey, D. Hathaway, G. Heckman, E. Hildner, W. Mende, K. Schatten, R. Thompson, A.W.P. Thomson, and O.R. White, "Panel achieves consensus prediction of Solar Cycle 23", *EOS, Trans. Amer. Geophys. Union, 78,* pp. 205, 211-212, 1997.

2. Schatten, K. H., Myers, D. J., and Sofia, S., "Solar Activity Forecast for Solar Cycle 23", *Geophys. Res. Lett., 6,* 605-608, 1996.

3. McNish, A.G., and J.V. Lincoln, Prediction of Sunspot Numbers, *Eos, Trans. Amer. Geophys. Union, 30,* p. 673, 1949.

4. Holland, R.L. and W.W. Vaughan, Lagrangian Least-Squares Prediction of Solar Flux (F10.7), *J. Geophys. Res., 89,* 11-16, 1984.

5. Kerridge, D.J., V. Carlaw, and D. Beamish, "A Review of Methods for Solar and Geomagnetic Activity Forecasting for Application in Space Missions Planning", BGS Technical Report, WM/89/14C (62 pp.)1989.

6. Hathaway, D. Wilson, R., and Reichmann, *J. Geophys. Res. 104,* 22,375-22, 388, 1999.

7. Feynman, J., and X.Y. Gu, "Prediction of geomagnetic activity on time scales of one to ten years", *Rev. Geophys.,24,* 650, 1986.

8. Ohl, A. I. *Soln. Dann.* No. 12, 84, 1966.

9. Ohl, A. I. and Ohl, G. I. in R. Donnelly, ed. Solar - Terr. Pred. Proc. Boulder, CO, USA. Vol 2, p 246, 1979.

10. Brown, G. M. and Williams, W. R., "Some Properties of the Day-to-Day Variability of Sq(H)", *Planet. Sp. Sci., 17,* 455-469, 1969.

11. Bartels, J., "Discussion of time variations of geomagnetic activity indices Kp and Ap, 1932-1961", *Ann. Geophys., 19,* 1, 1963.

12. Thompson, R. J., "A Technique for Predicting the Amplitude of the Solar Cycle", *Solar Phys. 148,* 383, 1993.

13. Schatten, K. H., Scherrer, P. H. , Svalgaard L., and Wilcox, J. M., "Using Dynamo Theory to Predict the Sunspot Number During Solar Cycle 21", *Geophys. Res. Lett., 5,* 411, 1978.

14. Schatten, K. H. and Wilcox, J. M., "Response of the Geomagnetic Activity Index Kp to the Interplanetary Magnetic Field," *J. Geophys. Res., 72,* 5185, 1967.

15. Sofia, S., Fox, P., Schatten, K., "Forecast Update for Activity Cycle 23 from a Dynamo-Based Method", *Geophys. Res. Lett., 25,* # 22, pp. 4149-4152, 1998.

16. Schatten, K.H., and W.D. Pesnell, "An early solar dynamo prediction: Cycle 23 ~ Cycle 22," *Geophys. Res. Lett, 20,* 2275-2278, 1993.

17. Wilcox Solar Observatory Polar Field strengths, 3' aperture, 20 nhz low pass filter, 2001.

18. Schatten, K. H. and J. A. Orosz "A Solar Cycle Timing Predictor--The Latitude of Active Regions," *Solar Physics, 125,* 185-189, 1990.

19. Tobiska, W.K, et al., "The SOLAR2000 empirical solar irradiance model and forecast tool," *J. Atm. Solar Terr. Phys.,* 62, 1233-1250, 2000.

TABLE

TABLE 1 Forecasting techniques and predicted levels (in smoothed International Sunspot Units) for solar cycle #23, based upon the NOAA panel report by Joselyn, et al.[1]. The Precursor levels have been divided into the Geomagnetic and Solar Precursor methods.

PREDICTED SUNSPOT NUMBERS FOR SOLAR CYCLE #23

Technique	Low End of Range	Smoothed Max R_z	High End of Range
Even/Odd Behavior	165	200	235
Spectral	135	155	185
Recent Climatology	125	155	185
Climatology	75	115	155
Neural Networks	110	140	170
Precursor:			
Geomagnetic Precursor	140	160	180
Solar Precursor	108	138	168

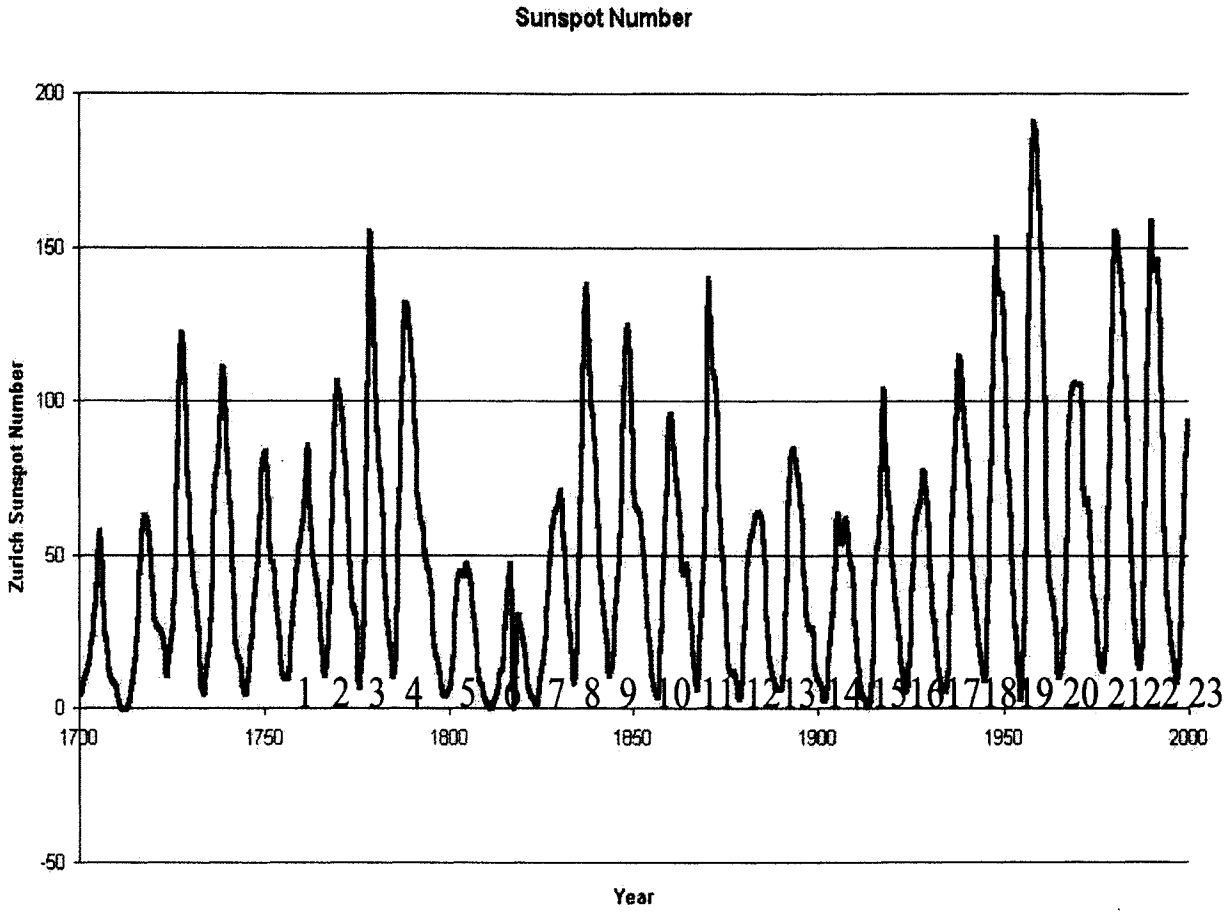


Figure 1. Sunspot Number vs. time for the past few centuries. The figure shows "power" in a wide variety of periods beyond the famous 11-year Schwabe periodicity. Additionally, major variations exist both on longer and shorter timescales. Further, amplitudes of the cycles vary by more than 100%, in a rather chaotic manner. Epochs occur, such as during the "Maunder minimum," when solar activity dropped precipitously to near 0. The numbering on the chart shows the "Even/Odd" effect, where this century odd numbered cycles have always been larger than the previous even numbered cycle (e.g. cycle #19 > cycle #18).

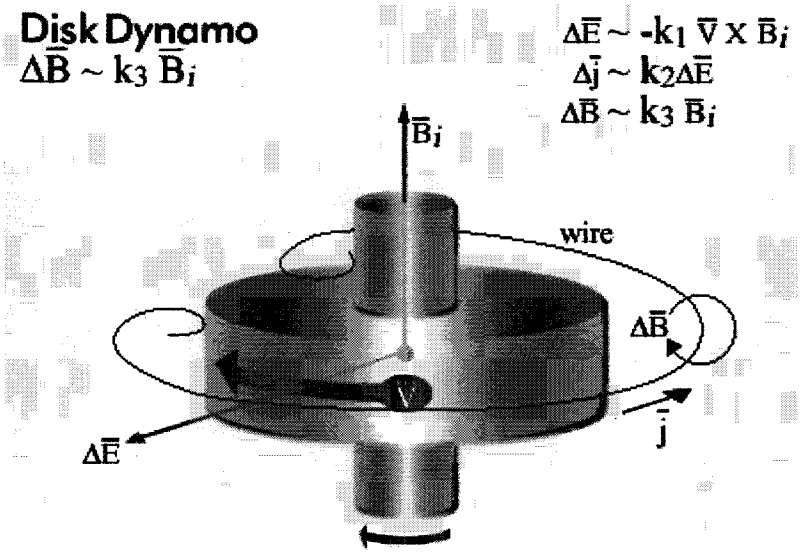


Figure 2. Model of a simple Disk Dynamo is shown. The rotation of a conductor inside a magnetic field generates an electric field. The electric field drives a current, amplifying the initial magnetic field. The resultant amplified magnetic field is then proportional to the initial magnetic field.

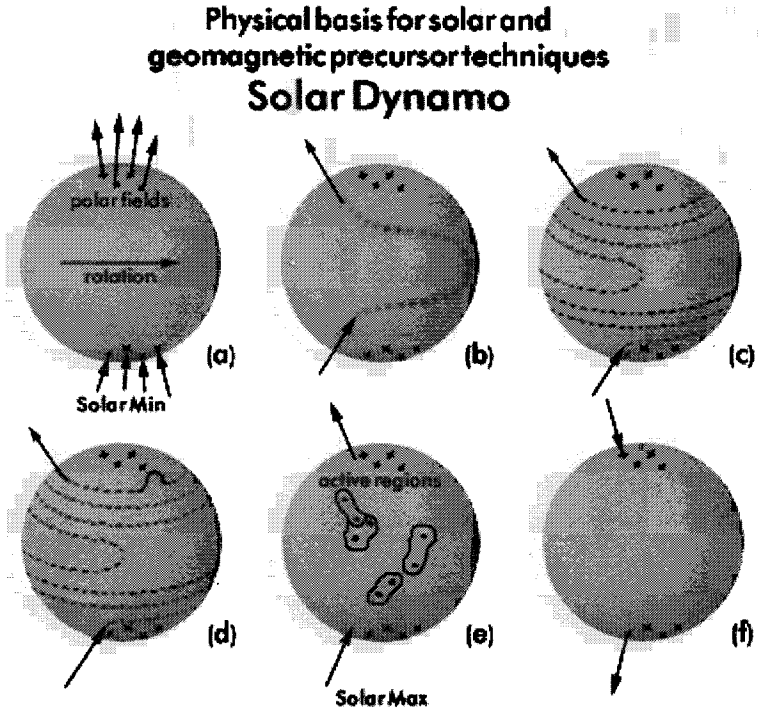
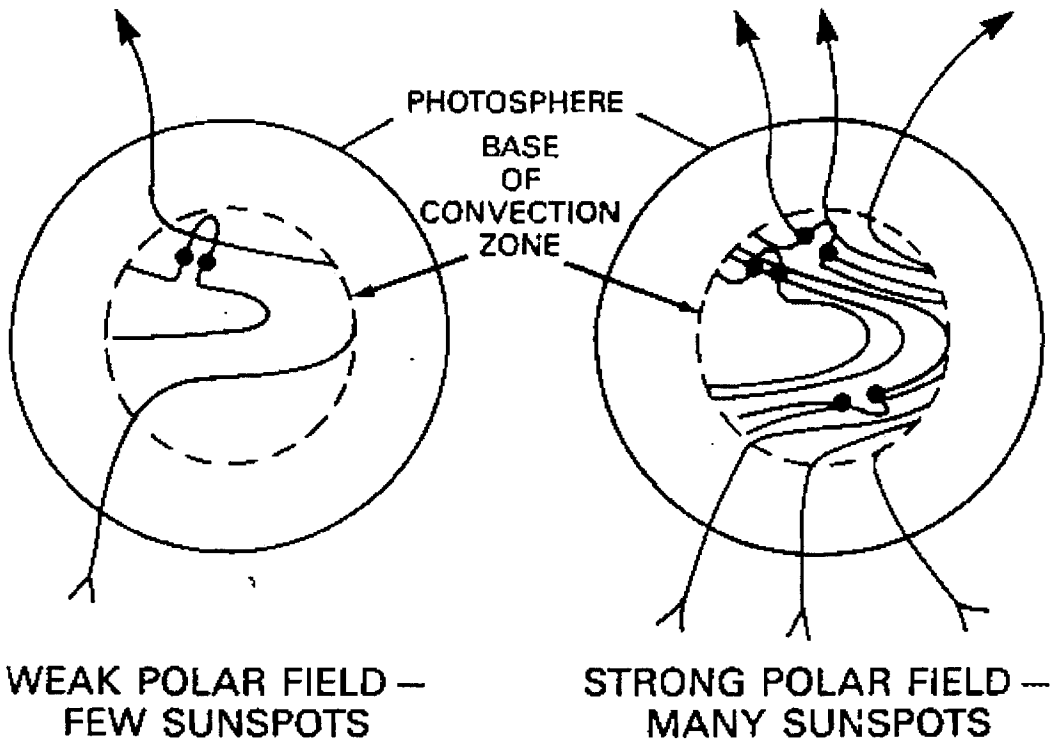


Figure 3. In the Babcock dynamo, the Sun's polar fields near solar minimum (a) are wrapped up by differential rotation (b) to form toroidal fields (c). These fields, later in the cycle, float to the Sun's surface and erupt (d) to form active regions containing sunspots (e). The breakup of these active region fields regenerate the Sun's polar field with a reverse sign (f), allowing the process to repeat anti-symmetrically.



$$\Delta B \propto B_{POLAR}$$

Figure 4. Shown is a simple schematic picture of the Babcock dynamo mechanism for two different cycles: a weak cycle (left) and a strong cycle (right). The complex picture of the temporal dynamics of a solar cycle is simplified in this picture so that the polar fields and toroidal fields of each cycle are both displayed.

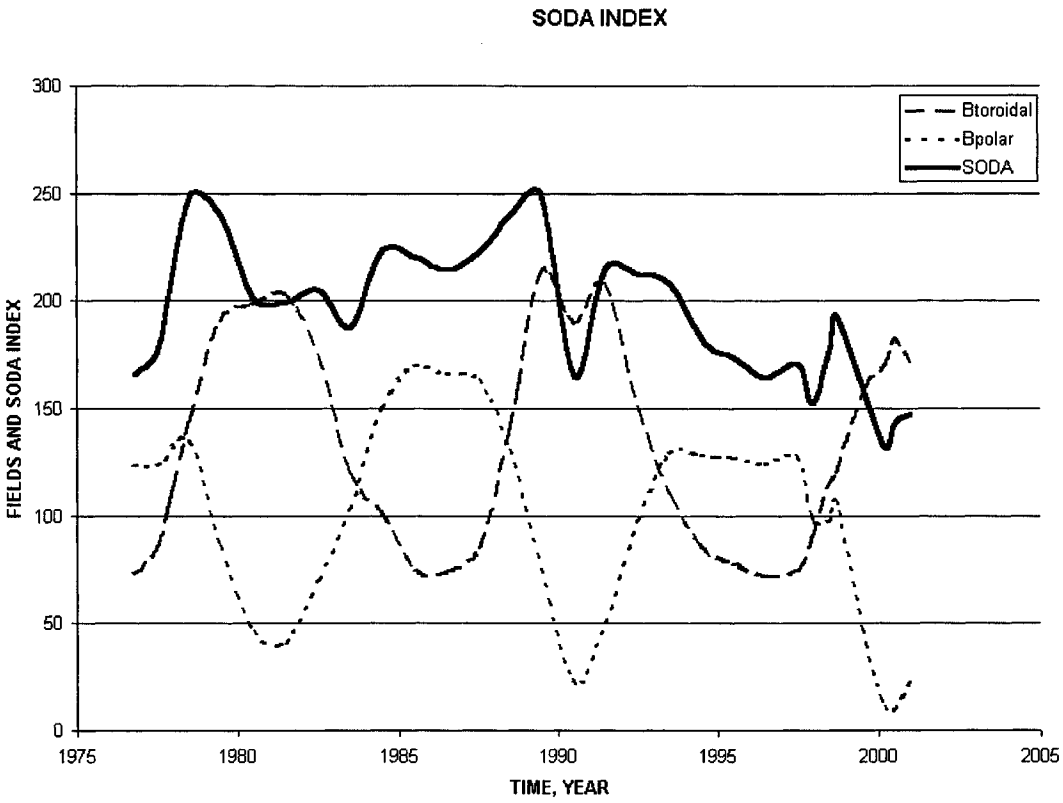


Figure 5. The "SOlar Dynamo Amplitude" or SODA index is a composite index attempting to combine the changing toroidal and poloidal fields of the Sun. As these fields vary with time, the combined SODA index allows us to monitor the "buried magnetic flux" present in the Sun's ever-changing dynamo.

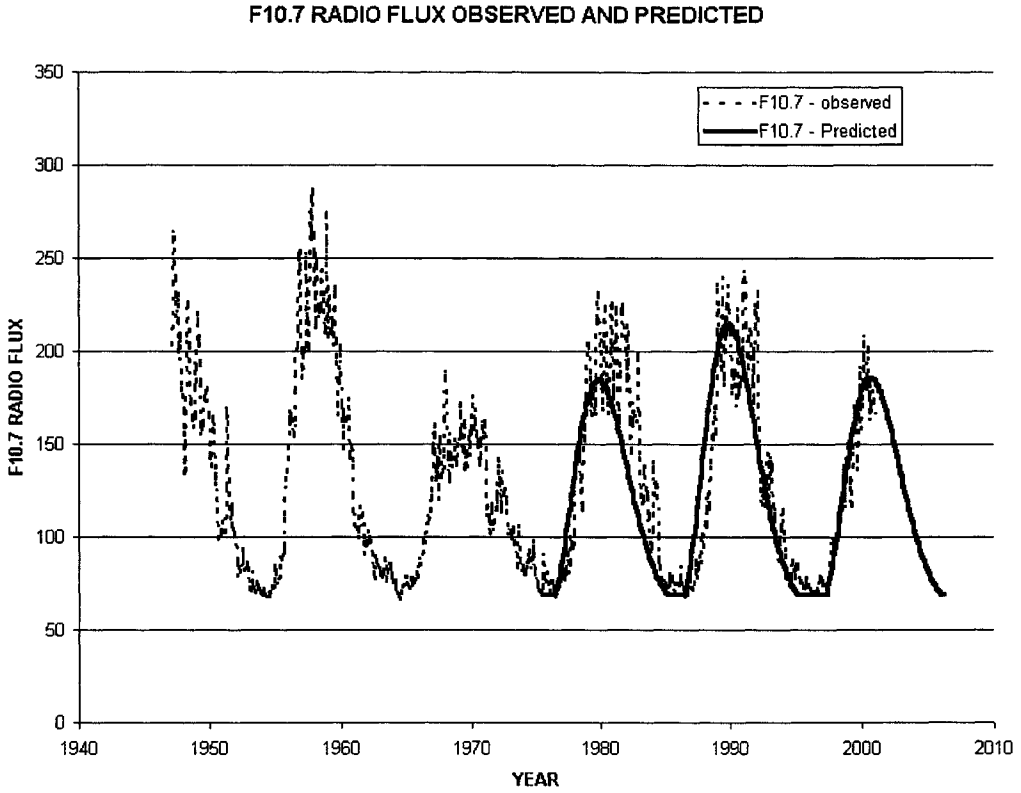


Figure 6. Shown is F10.7 Radio Flux for the past 50 years, and Schatten et al. predictions for the last 3 cycles, published in advance. Note that cycle #23, the present cycle, seems to be a better fit to the predicted values, both in timing and amplitude than previous cycles. Although this may be fortuitous, it may also be a sign that our skill level is increasing.

A Non-Linear Approach to Spacecraft Trajectory Control in the Vicinity of a Libration Point

Richard J. Luquette
NASA Goddard Space Flight Center

Robert M. Sanner
University of Maryland, Dept. of Aerospace Engineering

ABSTRACT

An expanding interest in mission design strategies that exploit libration point regions demands the continued development of enhanced, efficient, control algorithms for station-keeping and formation maintenance. This paper discusses the development of a non-linear, station-keeping, control algorithm for trajectories in the vicinity of a libration point. The control law guarantees exponential convergence, based on a Lyaponov analysis. Controller performance is evaluated using FreeFlyer® and MATLAB® for a spacecraft stationed near the L2 libration point in the Earth-Moon system, tracking a pre-defined reference trajectory. Evaluation metrics are fuel usage and tracking accuracy. Simulation results are compared with a linear-based controller for a spacecraft tracking the same reference trajectory. Although the analysis is framed in the context of station-keeping, the control algorithm is equally applicable to a formation flying problem with an appropriate definition of the reference trajectory.

INTRODUCTION

The restricted-three body problem examines the behavior of an infinitesimal mass in the combined gravitational field of two finite masses rotating in an orbit about their common center of mass. Research on this problem began prior to 1772, the year Lagrange published a set of particular solutions known as the Lagrange or libration points. Libration points, defined within a rotating two body system, represent locations within the rotating frame at which the dynamical forces due to gravity and rotation are neutralized. The equilibrium points are grouped in a set of three collinear points, referred to as L1, L2 and L3; and a set of two triangular points, L4 and L5. With our emerging capability to implement space-based missions, a growing research interest is focused on methods for exploiting the dynamics in the vicinity of these points. In particular, this paper builds on the current research associated with control strategies for station-keeping in the vicinity of a Libration point. Earlier works (refs. 1, 2, 3) have focused on linear control strategies, based on linearized dynamical equations. This requires assumptions about the motion of the two primary masses, i.e. either circular or elliptical motion about their common center of mass. Further, the validity of the control design is confined to some local region about the point of linearization, typically a libration point.

This study examines the design and performance of a control strategy based on the full non-linear dynamics associated with the restricted three-body problem. The performance of the control strategy is evaluated using the example of a halo orbit about the L2 point in the Earth-Moon system. Interestingly, the orbits of the Earth and Moon about their barycenter are non-planar and non-elliptic. The controller design is based on the Euler-Lagrange (Hamiltonian) form of the dynamics. This design technique, presented by Slotine and Li (ref. 4), is also explored by de Queiroz, et.al. (ref.5), for an Earth orbiting spacecraft. However, their results cannot be simply extended to the restricted three-body problem, since the governing dynamics are significantly different.

THEORY

The dynamics (1) and kinematics (2) for most physical systems can be expressed in the Hamiltonian (or Euler-Lagrange) form.

$$\mathbf{H(q)} * \dot{\mathbf{v}} + \mathbf{C(q,v)} * \mathbf{v} + \mathbf{E(q,v)} = \mathbf{u} \tag{1}$$

$$\dot{\mathbf{q}} = \mathbf{J(q)} * \mathbf{v} \tag{2}$$

Where:

$\mathbf{q}$ – Configuration Variables
$\boldsymbol{\nu}$ – Velocity Variables
$\mathbf{H(q)} = \mathbf{H(q)}^{\mathrm{T}} > \mathbf{0}, \forall \mathbf{q}, \mathbf{v}$
$\mathbf{C(q,v)}$ is defined such that $\dot{\mathbf{H}}(\mathbf{q}) - \mathbf{2} * \mathbf{C(q,v)}$ is skew symmetric for all $\mathbf{q}, \boldsymbol{\nu}$.

Assuming $\dot{\mathbf{q}} = \mathbf{v}$, i.e. $\mathbf{J(q)} = \mathbf{I}$, then equations (1) and (2) are combined as

$$\mathbf{H(q)} * \ddot{\mathbf{q}} + \mathbf{C(q,\dot{q})} * \dot{\mathbf{q}} + \mathbf{E(q,\dot{q})} = \mathbf{u} \tag{3}$$

The goal is to compute $\mathbf{u}(t)$, such that the system tracks a desired trajectory, $\mathbf{q_d}(t)$. Define the tracking error as $\mathbf{e(t)} = \mathbf{q(t)} - \mathbf{q_d(t)}$, an auxiliary error metric $\mathbf{s(t)} = \dot{\mathbf{e}}(\mathbf{t}) + \boldsymbol{\Lambda} * \mathbf{e(t)}$, with $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}^{\mathbf{T}} > \mathbf{0}$, and a reference velocity $\dot{\mathbf{q}}_{\mathbf{r}}(\mathbf{t}) = \dot{\mathbf{q}}_{\mathbf{d}}(\mathbf{t}) - \boldsymbol{\Lambda} * \mathbf{e(t)}$.

$$\mathbf{s(t)} = \dot{\mathbf{q}}(\mathbf{t}) - \dot{\mathbf{q}}_{\mathbf{r}}(\mathbf{t}) \tag{4}$$

Reference (4) proposes the following control law with $\mathbf{K_d} = \mathbf{K_d} > \mathbf{0}$:

$$\mathbf{u(t)} = \mathbf{H(q)} * \ddot{\mathbf{q}}_{\mathbf{r}} + \mathbf{C(q,\dot{q})} * \dot{\mathbf{q}}_{\mathbf{r}} + \mathbf{E(q,\dot{q})} - \mathbf{K_d} * \mathbf{s(t)} \tag{5}$$

Combining equations (3), (4) and (5) yields:

$$\mathbf{H(q)} * \dot{\mathbf{s}}(\mathbf{t}) + \mathbf{C(q,\dot{q})} * \mathbf{s(t)} + \mathbf{K_d} * \mathbf{s(t)} = \mathbf{0} \tag{6}$$

The above closed-loop dynamics drive the tracking error, $\mathbf{e}(t)$, to zero. Before proceeding with the proof, it is beneficial to recall the following corollary to Barbalat's Lemma (ref. 4).

If a scalar function $V(\mathbf{x},t)$ satisfies the following conditions, then $\dot{V}(\mathbf{x},t) \rightarrow 0$, ***as*** $t \rightarrow \infty$.

- $V(\mathbf{x},t)$ *is lower bounded*
- $\dot{V}(\mathbf{x},t)$ *is negative semi-definite*
- $\dot{V}(\mathbf{x},t)$ *is uniformly continuous in time*

Note: The third condition is met if $\ddot{V}(x,t)$ *is bounded.*

The proof that $\mathbf{e}(t) \rightarrow 0$, is based on a Lyapunov analysis. Consider the following candidate Lyapunov function, and it's derivative (ref. 4).

$$V(\mathbf{s},t) = \tfrac{1}{2}\mathbf{s(t)} * \mathbf{H(t)} * \mathbf{s(t)} \geq 0 \tag{7}$$

$$\begin{aligned} \dot{V}(\mathbf{s},t) &= \mathbf{s(t)} * \mathbf{H(t)} * \dot{\mathbf{s}}(\mathbf{t}) + \tfrac{1}{2}\mathbf{s(t)} * \dot{\mathbf{H}}(\mathbf{t}) * \mathbf{s(t)} \\ &= \mathbf{s(t)} * [-\mathbf{C(q,\dot{q})} - \mathbf{K_d}] * \mathbf{s(t)} + \tfrac{1}{2}\mathbf{s(t)} * \dot{\mathbf{H}}(\mathbf{t}) * \mathbf{s(t)} \\ &= -\mathbf{s(t)} * \mathbf{K_d} * \mathbf{s(t)} + \tfrac{1}{2}\mathbf{s(t)} * [\dot{\mathbf{H}}(\mathbf{t}) - \mathbf{2} * \mathbf{C(q,\dot{q})}] * \mathbf{s(t)} \\ &= -\mathbf{s(t)} * \mathbf{K_d} * \mathbf{s(t)} \leq 0 \end{aligned} \tag{8}$$

(Recall $[\dot{\mathbf{H}}(\mathbf{t}) - 2 * \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})]$ is skew symmetric, $\mathbf{K}_d = \mathbf{K}_d^T > \mathbf{0}$)

Since $\mathbf{H(t)}$ is positive definite, *V(s,t)* is non-negative (lower bounded). From equation (8), $\dot{V}(\mathbf{s},t)$ is negative semi-definite. It is assumed $\mathbf{q(t)}$ and $\mathbf{q_d(t)}$ are at least twice differentiable. Therefore, $\dot{V}(\mathbf{s},t)$ is uniformly continuous in time. Thus, Barbalat's Lemma guarantees, $\dot{V}(\mathbf{s},t) \rightarrow 0$, (and $\mathbf{s}(t) \rightarrow 0$) as $t \rightarrow \infty$. Consider $\mathbf{e(t)}$ as the output of a stable linear system, $\mathbf{s(t)} = \dot{\mathbf{e}}(\mathbf{t}) + \mathbf{\Lambda} * \mathbf{e(t)}$, with $\mathbf{s(t)}$ as the input. It follows that $\mathbf{s}(t) \rightarrow 0$, implies $\mathbf{e}(t) \rightarrow 0$. In fact, the system is globally stable, and the tracking error converges to zero exponentially.

This analysis ignores disturbance forces, $\mathbf{F_d}$. Consequently, the control strategy will not yield perfect tracking under a disturbance. However, an adaptive control strategy guarantees the desired tracking, provided the disturbed dynamics assume the linear form:

$$\mathbf{H(q)} * \ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) * \dot{\mathbf{q}} + \mathbf{E}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{q}_d, \dot{\mathbf{q}}_d) * \mathbf{a}$$

where $\mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{q}_d, \dot{\mathbf{q}}_d)$ is known, and here, $\mathbf{a}$ is a constant, unknown, vector.
Global stability and convergence of the tracking error to zero are guaranteed by the following control law and adaptive rule. The proof is provided in reference [4].

$$\begin{aligned} &\mathbf{u} = \mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{q}_r, \dot{\mathbf{q}}) * \hat{\mathbf{a}} - \mathbf{K}_d * \mathbf{s}, \\ &\dot{\hat{\mathbf{a}}} = -\mathbf{\Gamma} * \mathbf{Y}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{q}_r, \dot{\mathbf{q}}) * \mathbf{s}, \quad \mathbf{\Gamma} = \mathbf{\Gamma}^T > \mathbf{0} \end{aligned} \tag{9}$$

THE RESTRICTED-THREE BODY PROBLEM

Under the assumptions of the restricted-three body problem, the gravitational influence of two primary masses (rotating about their common, inertially-fixed, center of mass) govern the dynamics of a small mass (spacecraft). The spacecraft (S/C) dynamics (per unit mass) in inertial coordinates are given by:

$$\ddot{\mathbf{R}}_{sc} = -GM_1 * \mathbf{R}_{1s} / \left\| \mathbf{R}_{1s} \right\|_2^3 - GM_2 * \mathbf{R}_{2s} / \left\| \mathbf{R}_{2s} \right\|_2^3 + \mathbf{F}_d \tag{10}$$

where:

GM_i	= Gravitational Parameter of Mass *i*
$\mathbf{R}_{is}$	= Position of S/C with respect to Mass *i*
$\mathbf{R}_{sc}$	= Position of S/C
$\mathbf{F}_d$	= Disturbance Force

Although (10) fits the Hamiltonian form, it is not expressed in a convenient coordinate frame for defining our desired trajectory. Recalling the stated goal is to follow a trajectory in the vicinity of a Libration point, it is more convenient to express the dynamics in the rotating coordinate frame defined by the motion of the smaller primary.

$$\hat{\mathbf{i}} = \mathbf{R}_2 / \left\| \mathbf{R}_2 \right\|_2, \hat{\mathbf{k}} = (\mathbf{R}_2 \times \mathbf{V}_2) / \left\| (\mathbf{R}_2 \times \mathbf{V}_2) \right\|_2, \hat{\mathbf{j}} = \hat{\mathbf{k}} \times \hat{\mathbf{i}}$$

Hence:

$$\mathbf{R} = \mathbf{A}_{ir} * \mathbf{r} \; ; \; \mathbf{A}_{ir} = [\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}] \tag{11}$$

Where $\mathbf{r}$ represents the coordinates in the rotating frame.

Differentiating (11) and combining the result with (10), yields:

$$\mathbf{u} = \mathbf{A}_{ir}^T * \mathbf{U} = \mathbf{H} * \ddot{\mathbf{r}} + \mathbf{C} * \dot{\mathbf{r}} + \mathbf{E}(\mathbf{r}) + \mathbf{A}_{ir}^T * \mathbf{F}_d \tag{12}$$

where:

$$\mathbf{H} = \mathbf{I}_3$$

$$\mathbf{C} = 2 * \mathbf{A}_{ir}^{T} * \dot{\mathbf{A}}_{ir} = 2 * \mathbf{A}_{ir}^{T} * \mathbf{A}_{ir} * \boldsymbol{\Omega} = 2 * \boldsymbol{\Omega}$$

$$\mathbf{E}(\mathbf{r}) = \mathbf{A}_{ir}^{T} * (\ddot{\mathbf{A}}_{ir} * \mathbf{r} + GM_1 * \mathbf{R}_{1s} / \|\mathbf{R}_{1s}\|_2^3 + GM_2 * \mathbf{R}_{2s} / \|\mathbf{R}_{2s}\|_2^3)$$

$$\boldsymbol{\Omega} = Skew\{(\mathbf{r} \times \mathbf{v}) / \|\mathbf{r}\|_2\}$$

Note: $[\dot{\mathbf{H}}(\mathbf{t}) - 2 * \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})] = 2 * \boldsymbol{\Omega}$, which is skew symmetric.

It is important to note that $\mathbf{A_{ir}}$, and it's derivatives, are defined by the motion of the two primary masses. Hence, the behavior of $\mathbf{A_{ir}}$ is generally well known for trajectories within our solar system.

CONTROL LAW DESIGN FOR THE THREE-BODY PROBLEM

The results of the two preceding sections combine to generate the control law. Here $\mathbf{F_d}$ is considered a constant, unknown, force in inertial coordinates. This is reasonable, since the significant perturbations (the gravitational influence of other bodies and solar pressure) will be reasonably constant in inertial space over short periods of time. In fact, $\mathbf{F_d}$ is expressed as a combination of known and unknown components. For example, the gravitational forces due to select bodies are directly computed as known disturbances. The remaining disturbances are estimated with the adaptive rule given in (9).

Combining (5) and (9), let

$$\mathbf{u}(t) = \mathbf{A}_{ir}^{T} * \mathbf{U} = \mathbf{H} * \ddot{\mathbf{r}}_r + \mathbf{C}(t) * \dot{\mathbf{r}}_r + \mathbf{E}(\mathbf{r}) - \mathbf{K}_d * \mathbf{s} + \mathbf{A}_{ir}^{T} * \hat{\mathbf{F}}_d$$
$$\text{with } \dot{\hat{\mathbf{F}}}_d = -\boldsymbol{\Gamma} * \mathbf{A}_{ir} * \mathbf{s},\ \dot{\mathbf{r}}_r = \dot{\mathbf{r}}_d - \boldsymbol{\Lambda} * (\mathbf{r} - \mathbf{r}_d) \tag{13}$$

Expressed in inertial coordinates, the control is:

$$\mathbf{U}(t) = \mathbf{A}_{ir} * [\ddot{\mathbf{r}}_r + \mathbf{C}(t) * \dot{\mathbf{r}}_r + \mathbf{E}(\mathbf{r}) - \mathbf{K}_d * \mathbf{s}] + \hat{\mathbf{F}}_d \tag{14}$$

As previously shown, this control achieves perfect tracking. Note that with zero tracking error, $\mathbf{r_r} = \mathbf{r_d}$. In this case, equation (14) reduces to

$$\mathbf{U}_d(t) = \mathbf{A}_{ir} * [\ddot{\mathbf{r}}_d + \mathbf{C}(t) * \dot{\mathbf{r}}_d + \mathbf{E}(\mathbf{r}_d)] + \hat{\mathbf{F}}_d \tag{15}$$

So, as the tracking error goes to zero, the control effort converges to the exact value required to maintain the desired trajectory. Hence, under this control strategy (and the stated assumptions), optimal control is achieved by optimal trajectory design.

This development assumes the barycenter of the two masses remains inertially "fixed", and the spacecraft state is known. Additional assumptions regarding the relative motion of the primaries are not required. The dynamics model for traditional linear control designs constrain the primaries to planar circular (or elliptic) motion, and require linearization of the dynamics equation about a point (typically a libration point). Therefore, modeling errors are significantly reduced in the nonlinear design. Also, the linear design is locally optimal. In contrast, the nonlinear design is globally optimal, although constrained to the region governed by the restricted three-body dynamics.

SIMULATION

Implementation of the proposed adaptive control law, equation (13), is simulated using FreeFlyer® interfaced with MATLAB®. FreeFlyer® supplied the dynamics model and propagation tool. MATLAB® provides the computational tools to determine the control effort at each time step. Additionally, MATLAB® serves to capture and analyze the simulation data.

The simulation scenario considers a spacecraft stationed in the vicinity of the L2 point in the Earth-Moon system. This choice allows comparison with the linear control design presented by Hoffman (ref. 1). Also, the motion of the Earth/Moon about their barycenter is dynamically complex. It is eccentric, but non-elliptic and non-planar, and thus provides a challenging environment for any tracking problem. The control is successfully simulated as both a continuous and impulsive thrust. The results are similar, therefore all presented cases are based on a continuous thrust model. The linear control law is based on an LQR design of a classic PD controller. The gain matrices are provided in Reference 1, Table 5-II. Initial conditions are set with zero position and velocity error to limit the time required for the system to reach "steady state".

Five separate cases are studied. For each case the location of L2, and the orientation of the rotating coordinates are considered dynamic, determined by the instantaneous position of the Earth and Moon. The goal for cases 1 through 3, is to track the position of L2. For cases 4 and 5, the desired trajectory is a 3600 km halo orbit about the L2 point. The trajectory, based on Reference 1, Equation 3.20, is shown in Figure 1, as it appears in the rotating coordinate frame. For cases 1 through 4, the nonlinear control gains, $\mathbf{K_d}$ and $\mathbf{\Lambda}$, were chosen to be equivalent to the linear control gains. Case 5 is the same as case 4, except the nonlinear control gains are adjusted for improved tracking performance. Perturbations include the gravitational influence of the Sun and Jupiter, and solar pressure. Additionally, the Earth is modeled as a point mass for the cases without perturbations. Under perturbations, the Earth's gravitational field model includes zonal and tesseral terms up to J5. The simulation period is 100 days, starting on January 1, 2000. The integration step size is 600 seconds.

The distinguishing features of the five cases are summarized below:

Case 1	Track L2	Without Adaptation	No Perturbations	
Case 2	Track L2	With Adaptive Law	No Perturbations	
Case 3	Track L2	With Adaptive Law	With Perturbations	
Case 4	Track Halo Orbit	With Adaptive Law	With Perturbations	
Case 5	Track Halo Orbit	With Adaptive Law	With Perturbations	Modified Gains

Simulation results are exhibited in Figures 2 through 4.

Figure 2	Cases 1 through 4	Position Error
Figure 3	Cases 1 through 4	Control Effort
Figure 4	Case 5	Position Error/Velocity Error/Control Effort

Examination of the data leads to several observations. In general, the nonlinear control provides superior tracking with or without adaptation for all cases. Comparing cases 1 and 2, adaptation significantly improves the tracking performance of the nonlinear control. Case 3 shows improved tracking over cases 1 and 2. Although intuition suggests adding the perturbations would degrade performance, Figure 3 provides the necessary insight into the observed behavior. Case 3 requires less control to maintain the desired trajectory, suggesting the perturbations complement the control. This is true since the perturbations affect the motion of both the spacecraft and the position of L2 (indirectly) in a similar fashion. In this case compensation for the unmodeled perturbations are effectively built into the control law through the definition of the desired trajectory. This simplifies the demand on the adaptive rule, yielding improved tracking performance for both the nonlinear and linear control strategies.

As noted above, cases 4 and 5 explore the performance for a 3600 km halo orbit based on a design presented in reference 1. Although this trajectory is not considered optimal, it does provide a basis for comparing the control designs. For case 4, the nonlinear control still yields better tracking performance, but less significant than cases 1 through 3. As previously noted, the control gains are modified in case 5 to improve performance. Gain adjustments, established through a trial process, are limited to increasing the magnitude of $\mathbf{\Lambda}$, by a factor of 5. The modification increased the controller sensitivity to position error. The modification does not optimize the design, rather it demonstrates the controller can achieve improved position tracking without the constraint of matching the gains of the linear control law. Following control gain adjustment, the position tracking error is dramatically reduced, as expected (Figure 4).

Perfect tracking is not observed in the simulated results as the theory predicts. This is attributed to the limitations in the simulation environment, coupled with the large values of the spacecraft position vectors. The position vectors,

inertially referenced to the barycenter of the Earth-Moon system, are on the order of $4x10^8$ meters. Hence, a 40 meter tracking error equates to an accuracy of 10^{-5} percent. A control law based on relative position of two spacecraft eliminates these numerical problems, providing much tighter position tracking, a requirement for tight formation flying. This presents a topic for further research.

As noted, the simulation environment imposes additional limitations which contribute to the non-zero tracking error. The most significant limitation relates to computing the instantaneous acceleration of the Moon relative to the barycenter, a parameter not available as an output from FreeFlyer®. Therefore, the Moon's acceleration was estimated as $(\Delta V/\Delta t)$ for an interval centered at any given instance in time. This introduces numerical error in the computed control, since the desired trajectory is ultimately referenced to the Moon's position, velocity and acceleration. $\ddot{A}_{ir}$ is estimated in a similar fashion, since sufficient data is not available for direct computation. The combined influence of these and other numerical errors result in the non-zero tracking error observed in Case 1. The improved tracking for Case 2 implies the adaptive control is reacting to the computational errors as if the system is subjected to a disturbance force. Therefore, the adaptive control compensates for both unmodeled disturbances and numerical disturbances.

Finally, in each case the linear and nonlinear designs require similar control effort. This is expected since the major component of the control effort is required to maintain the desired trajectory.

Conclusion

An adaptive non-linear control law can be effectively implemented to perfectly track a pre-defined trajectory within the dynamic environment of the restricted, three body problem. These trajectories include orbits about either of the primary masses. Further, as the tracking error tends to zero, the resultant control will converge to its optimal value required for the desired trajectory. Thus, the control will perform optimally with an optimal trajectory.

This control is applicable to maintenance of a "leader/follower" formation. For this case the control is implemented on a "follower" spacecraft with the desired trajectory referenced to a "leader" spacecraft. However, for formation maintenance a control based on relative states between spacecraft is preferred. As future work, this design will be reformulated to base the control on relative states between spacecraft with superior performance expected, due to reduced sources of computational errors.

Finally, noise sources are not considered in this analysis, although the simulation environment introduces computational noise in the results. Therefore, additional work is required to evaluate the performance and stability of this control law when coupled with an appropriate estimator.

References:

1. Hoffman, D.A., "Station-keeping at the Collinear Equilibrium Points of the Earth-Moon System", Lyndon B. Johnson Space Center, Report: JSC-26189, September 1993.

2. Cielaszyk, D., and Wie, B., "New Approach to Halo Orbit Determination and Control", Journal of Guidance Navagation and Control, Vol. 19, No.2, March-April 1996.

3. Cielaszyk, D., and Wie, B., "Halo Orbit Determination and Control for the Elliptic Restricted Three-Body System", AIAA/AAS Astrodynamics Conference, August 1994, AIAA-94-3729-CP.

4. Slotine, J.E. and Li, W., "Applied Nonlinear Control", Prentice Hall, Inc., New Jersey, 1991

5. de Queiroz, M.S., Kapila, V., and Yan, Q., "Adaptive Nonlinear Control of Multiple Spacecraft Formation Flying", Journal of Guidance, Navigation, and Control, Vol. 23, No. 3, May-June 2000.

Figure 1: Sample Desired Trajectory, 3600 km Halo Orbit about Dynamic L2 (Earth-Moon)

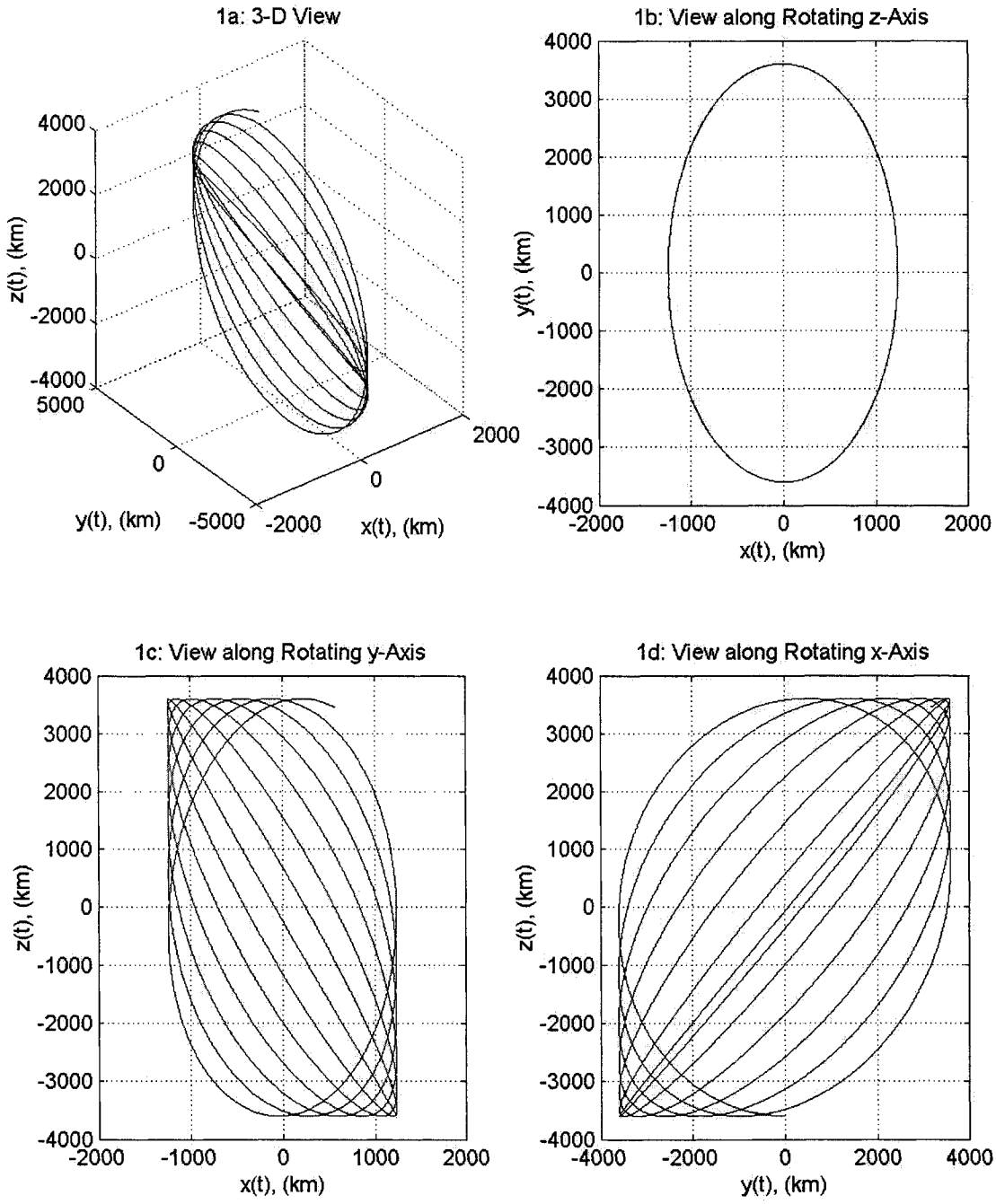


Figure 2: Magnitude of Position Error vs. time

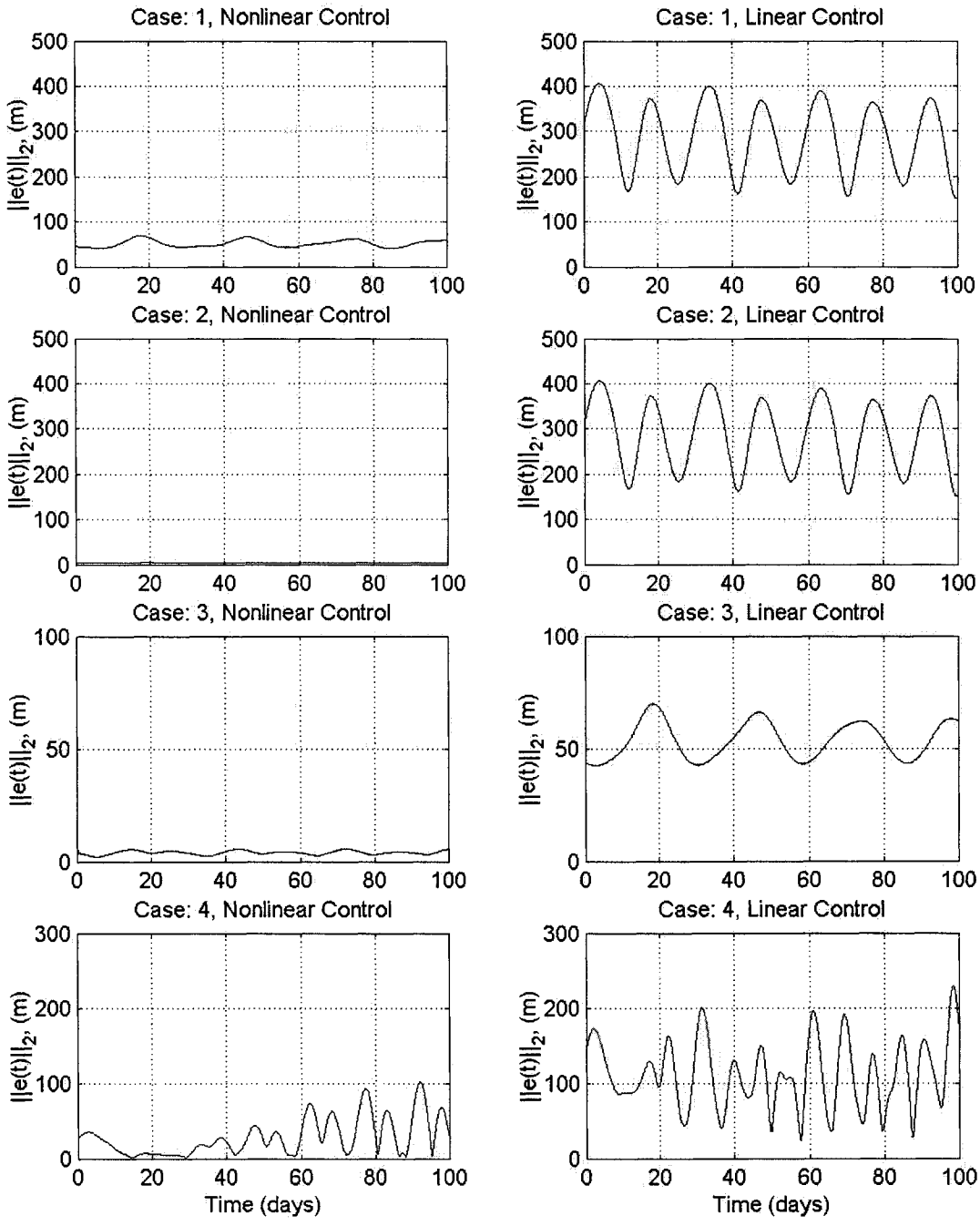


Figure 3: Magnitude of Control Thrust

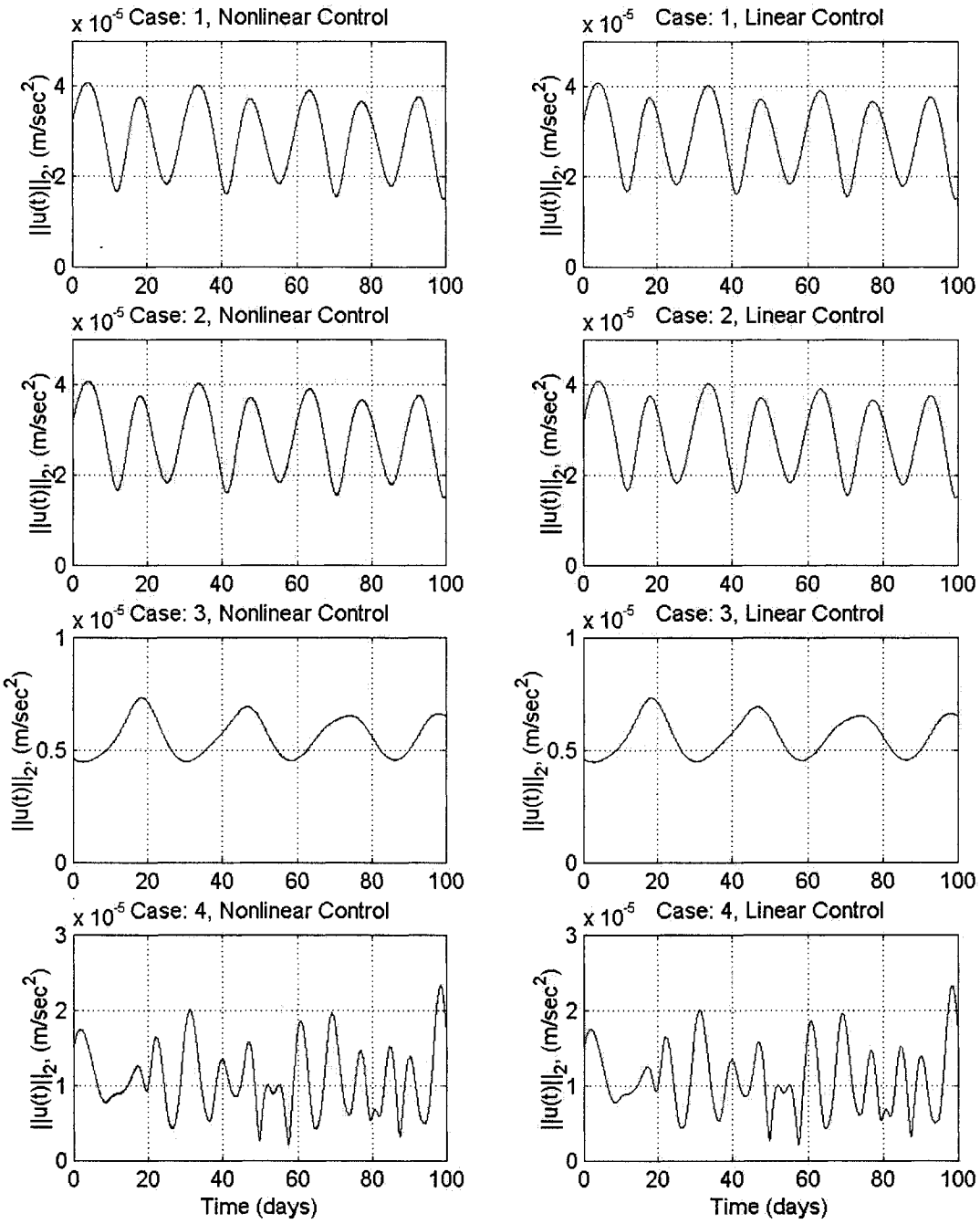
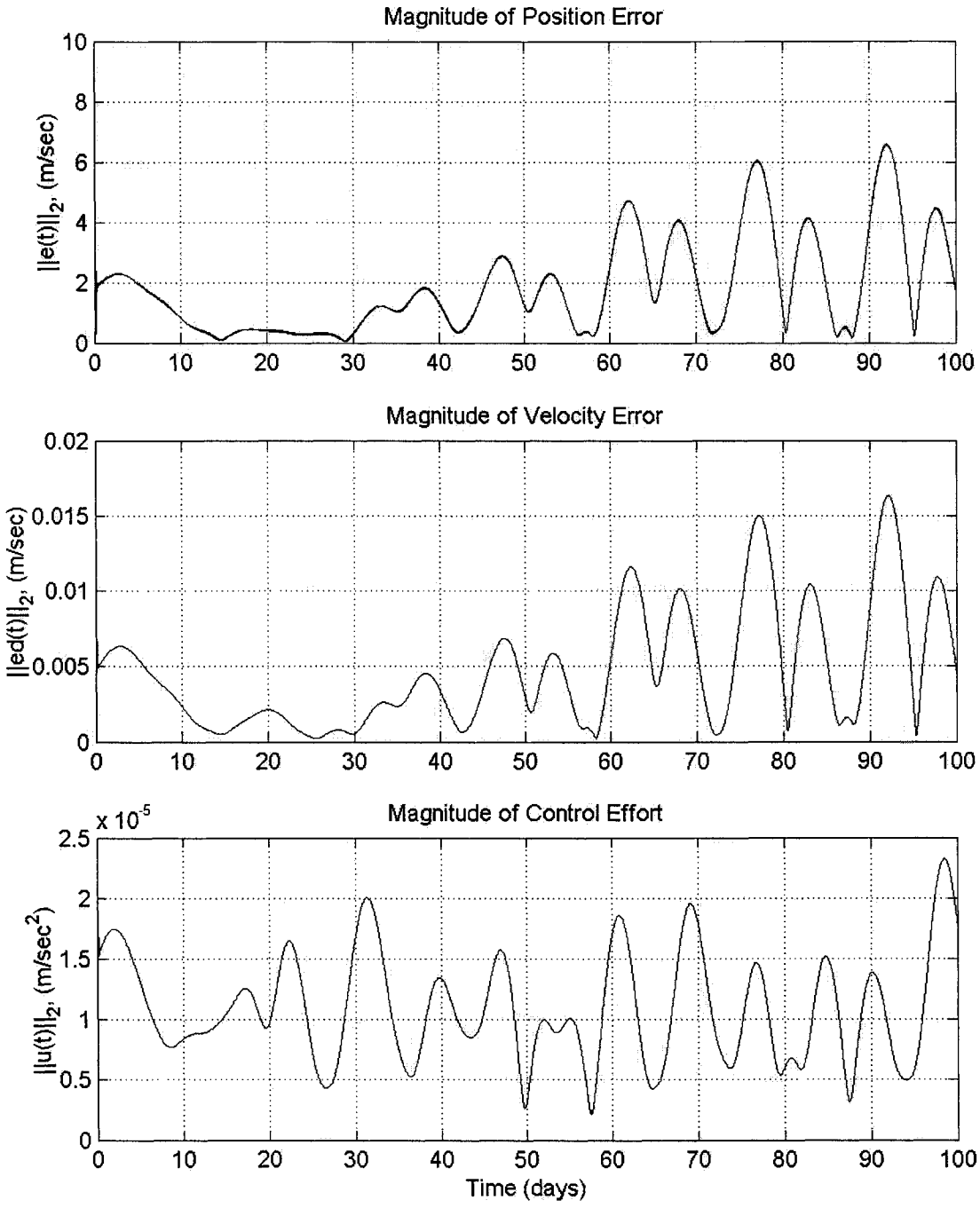


Figure 4: Results for Nonlinear Adaptive Control of a 3600 km Halo Orbit about L2 (Earth-Moon) with Improved Gains



PERIODIC ORBITS IN THE VICINITY OF L_1 AND L_2 IN THE COHERENT RESTRICTED FOUR-BODY PROBLEM

Jose J. Guzmán* and Kathleen C. Howell†

School of Aeronautics and Astronautics
Purdue University, West Lafayette, IN 47907-1282, USA

ABSTRACT

Consider the design of nominal orbits in the vicinity of the Sun-Earth collinear libration points. First, note that the equilibrium points actually exist only in the circular restricted three-body problem (**CR3BP**). As a result, for libration point missions, most of the analysis must be accomplished from the perspective of a model with, at least, three bodies. Unfortunately, for some cases in the Sun-Earth system, the qualitative behavior that is apparent in the CR3BP might be sometimes completely lost due to lunar perturbations. On the other hand, lunar gravity may be a key element in creating new types of solutions. Therefore, a formulation of the four-body problem is warranted, and the coherent restricted four-body problem is incorporated. In this model, the primaries satisfy the equations of motion, as derived in terms of a relative formulation, to some specified tolerance. To maximize the four-body model as a useful component, special solutions — such as periodic and quasi-periodic trajectories — can be initially isolated and identified. In this paper, the focus is on the computation of periodic orbits in the vicinity of the Moon perturbed Sun-Earth L_1 and L_2 libration points.

INTRODUCTION

As libration point missions increase in complexity, a better understanding of the range of trajectory options is necessary. In this effort, the CR3BP formulation has been instrumental in developing a qualitative and quantitative understanding of the natural dynamics of the three-body problem. Not surprisingly, some difficulties arise. For example, consider a CR3B model composed of the Sun-Earth system plus a spacecraft (**S/C**). Suppose that a family of periodic halo orbits near L_1 is computed. When "shifted" to the four-body ephemeris model, the periodicity is lost due to lunar perturbations. Furthermore, the libration points no longer exist. As a result, to better isolate and incorporate the lunar perturbation, a baseline orbit constructed directly in the four-body problem is suggested. Analysis in this model will help bridge the gap between the CR3BP and the ephemeris model. Furthermore, to expand the solution space and (ultimately) automate the design process, a more sophisticated four-body model is a necessity. To maximize the four-body model as a useful component, special solutions must be isolated and identified. Specifically, periodic solutions, as well as quasi-periodic solutions, are sought. With these solutions, dynamical systems theory can then be applied to examine the phase space. New trajectory options can then be explored.

PREVIOUS CONTRIBUTIONS

In terms of four bodies, Su-Shu Huang published a development of the *very restricted four-body problem* in 1960 (ref. 1). Denote this model as the **very-R4BP**; the primaries move in circular orbits and, thus, do not precisely satisfy their relative equations of motion. In particular, Huang analyzed the motion of satellites in the Earth-Moon-Sun system. To accomplish this, Huang utilized two R3BPs: Sun-Earth-S/C and Earth-Moon-S/C (ref. 2). Using the zero velocity surfaces associated with L_1 for both systems, Huang identified some regions where the satellite would transition from the Earth-Moon system to the Sun-Earth system. Nevertheless, Huang utilized the R4BP model as a justification for the consideration of the two separate R3BPs. Danby (1965) argued that the very-R4BP was not adequate to completely capture the natural

*Graduate Student, Current Address: a.i. solutions, Inc., 10001 Derekwood Lane, Suite 215 Lanham, MD 20706
†Professor

dynamics and demonstrated the importance of including the secular perturbations of the Sun (ref. 3). Prior to Danby's publication, Musen and Carpenter (1963) had initiated a study of planetary perturbations and developed methods to incorporate those perturbations in a form convenient for computer programming (ref. 4). In particular, the perturbations are decomposed into position vectors possessing components along the undisturbed instantaneous position and velocity vectors, as well as along the normal vector (to the orbital plane) of the undisturbed motion. With this methodology, as well as Chebyshev series and Huang's very-R4BP, Kolenkiewicz and Carpenter (1967) studied periodic motion relative to the Earth-Moon triangular points (L_4 and L_5) — see ref. 5. Then, they utilized the perturbation methodology and trigonometric series, and developed a R4BP where the primaries do indeed satisfy their relative equations of motion (ref. 6). *This approach is incorporated in the current investigation and it is denoted the* **coherent-R4BP**. With this model in hand and, again, using the perturbation methodology and trigonometric series, Kolenkiewicz and Carpenter (1967) numerically located stable periodic solutions near the Sun perturbed Earth-Moon L_4 and L_5 points. Meanwhile, in preparation for mission design involving the collinear points, Farquhar (1969) further demonstrated that, when the gravitational impact of the Moon is included, it is more accurate to place the Sun-Earth L_2 point along the line connecting the Sun and the Earth-Moon barycenter (ref. 7). Carpenter (1970) subsequently applied the perturbation approach with trigonometric series in the planar R3BP to compute families of periodic orbits in the Sun-Jupiter system (ref. 8).

In 1985, Howell and Spencer utilized the R4BP to produce periodic orbits and study their stability (ref. 9). In 1991, the very-R4BP, also denoted as the **bi-circular model**, was utilized by Gómez, Jorba, Masdemont and C. Simó (ref. 10) to study the transfer problem (from the Earth to the vicinity of the Sun-Earth equilibrium points). Continuing the search for periodic solutions, Scheeres (1998) developed a restricted Hill four-body model and studied periodic motion in the vicinity of the Sun perturbed Earth-Moon L_4 and L_5 points (ref. 11). The invariant manifolds associated with these solutions were also examined. More recently, analysis in a coherent-R4BP model, denoted as the **quasi bi-circular model** was accomplished by Andreu (ref. 12). Andreu's Hamiltonian analysis is a detailed investigation into the phase space in the vicinity of the Sun perturbed, Earth-Moon collinear equilibrium points.

It is important to remark that in a number of the above analytical studies (with exception of those by Farquhar and Gómez et al.), the system of interest was specified such that the largest primary perturbed the three-body system composed of the smaller primaries and the infinitesimal "spacecraft". Of course, this is consistent with motion in a system such as the Sun perturbed, Earth-Moon system and a similar type of analysis was pursued in the 1960's in preparation for the lunar missions. Furthermore, astronomers long considered the possibility that matter might remain, for short or long intervals, in the vicinity of the Earth-Moon equilateral points.

APPROACH

As an extension of this CR3BP, consider the restricted four-body problem (**R4BP**). The motion of the primaries becomes a solution to the three-body problem. Unfortunately, there is no analytical solution for the three-body problem! Thus, if an analytical formulation is desired, the motion of the primaries must be approximated. Once a suitable approximation is available (*one that represents a solution to the equations of motion as derived in terms of a relative formulation, and accurate to some specified tolerance*), the motion of the particle in the primary system can be investigated. Different approaches have been employed to approximate the motion of the primaries in the R4BP. As noted above, these efforts include the very-R4BP (bi-circular model) and the coherent-R4BP. Although, both models are based on planar motion for the primaries, only in the coherent-R4BP do the primaries satisfy their relative equations of motion. For example, in the context of the Sun-Earth-Moon primary system, the very-R4BP prescribes planar circular two-body motion for the Sun-Earth and Earth-Moon systems. In the case of the coherent-R4BP, the planar motion of the primaries (e.g. Sun, Earth, and Moon) is approximated via trigonometric series in terms of the synodic frequency of the Sun-Earth-Moon system. The approximations then satisfy the relative equations of motion to a specified tolerance. The coefficients of these trigonometric approximations are easily obtained using the method of harmonic balance.

Once the motion of the primaries has been established, the S/C time evolution can be examined. Specifically, two methodologies are exploited to compute periodic motion for the S/C: (1) differential corrections,

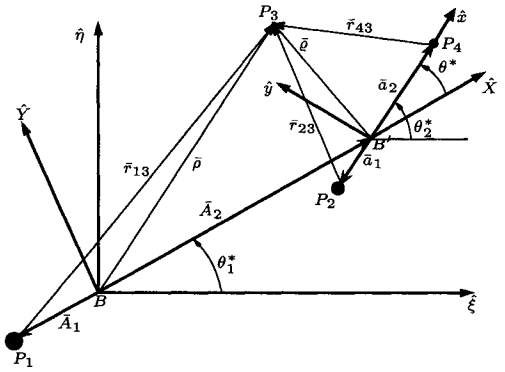


Figure 1: Restricted Four-Body Problem

and, (2) harmonic balance. For the differential corrections approach, the equations of motion as formulated in the rotating frame are used. On the other hand, for applying the method of harmonic balance, the relative formulation of the equations of motion is utilized.

ROTATING FRAME FORMULATION

To mathematically model the R4BP, begin by defining the bodies as P_i, for $i = 1, \ldots, 4$. Let P_3 represent the S/C and assume that its motion does not affect the motion of the primaries $P_i, i \neq 3$. Denote the barycenter of the three primaries as B and the barycenter corresponding only to P_2 and P_4 as B'. Define an inertial frame in terms of the unit vectors $(\hat{\xi}, \hat{\eta}, \hat{\zeta})$ with origin at B. Assume that the motion of the primaries is planar. See Figure 1. The unit vectors $\hat{\xi}$ and $\hat{\eta}$ are fixed in the plane of motion of the primaries, and $\hat{\zeta}$ completes the right handed triad. Define a rotating frame, also with origin at B, and unit vectors $(\hat{X}, \hat{Y}, \hat{Z})$. Let $\hat{X}$ be directed from P_1 towards B', $\hat{Z}$ along the normal to the plane of primary motion, and obtain $\hat{Y}$ by completing the orthogonal triad. This frame will be denoted as the B-B' rotating frame. Let the angle between the inertial frame and the rotating frame be denoted as θ_1^*. Define another rotating frame with origin at B', and with unit vectors $(\hat{x}, \hat{y}, \hat{z})$. Let $\hat{x}$ be parallel to a directed line segment from P_2 towards P_4, let $\hat{z} = \hat{Z}$, and then $\hat{y} = \hat{z} \times \hat{x}$. This second rotating frame is useful for visualization; it is labeled the B'-P_4 frame. Let the angle between the inertial frame and B'-P_4 frame be denoted as θ_2^*. Then, let $\theta^* = \theta_2^* - \theta_1^*$.

In Figure 1, the position vectors $\bar{A}_1$ and $\bar{A}_2$ are, respectively, the locations of P_1 and B' relative to B. Furthermore, the vectors $\bar{a}_1$ and $\bar{a}_2$ are, respectively, vectors representing the positions of P_2 and P_4 relative to B'. The vectors $\bar{r}_{13}$, $\bar{r}_{23}$, and $\bar{r}_{43}$ locate the S/C relative to P_1, P_2, and P_4, respectively. (Norms are indicated by omitting the bar.) The focus of any analysis is generally the motion of the S/C, so let $\bar{\rho}$ and $\bar{\varrho}$ denote the position vectors of the S/C relative to B and B', respectively. The mean distance between P_2 and P_4, the mean angular velocity of the B'-P_4 frame, and the sum of the masses of P_2 and P_4 are defined as the characteristic distance (l^*), angular velocity (n^*), and mass (m^*), respectively.[1] Furthermore, the characteristic time is computed as $t^* = 1/n$. Let the position vector of the S/C relative to B be expressed in component form as

$$\bar{\rho} = X\hat{X} + Y\hat{Y} + Z\hat{Z}. \tag{1}$$

The non-dimensional scalar equations of motion (**EOMs**) for P_3 relative to B are then obtained in the form

$$\ddot{X} - \dot{\Omega}_1 Y - 2\Omega_1 \dot{Y} = U_X, \tag{2a}$$

$$\ddot{Y} + \dot{\Omega}_1 X + 2\Omega_1 \dot{X} = U_Y, \tag{2b}$$

$$\ddot{Z} = U_Z, \tag{2c}$$

where Ω_1 and $\dot{\Omega}_1$ are the angular velocity and angular acceleration associated with θ_1^*, and the right side of the equations includes partial derivatives of the pseudo-potential U. The EOMs in Equation (2) appear

[1]The choice of characteristic quantities is unorthodox since the focus is motion in the vicinity of the equilibrium points for the P_1-P_2 system; nevertheless, it is effective for a scheme for small differential corrections with non-dimensional quantities.

in a form that is an exact replica of the equations that model the R3BP, but the angular velocity Ω_1, angular rate $\dot{\Omega}_1$, and the expressions for $\mathcal{U}_x$, $\mathcal{U}_y$, and $\mathcal{U}_z$ are not as readily available. The expression for the pseudo-potential that serves to model the four-body system is written as

$$\mathcal{U}(X,Y,Z) = \frac{1}{2}\Omega_1^2\left[X^2+Y^2\right] + \frac{M_1}{r_{13}} + \frac{M_2}{r_{23}} + \frac{M_4}{r_{43}}, \tag{3}$$

where M_1, M_2, and M_4 are the non-dimensional masses of P_1, P_2, and P_4 respectively. Then, the position vectors of the S/C relative to the each of the primaries are expressed in the following component forms

$$\begin{aligned}
\bar{r}_{13} &= (X+A_1)\hat{X} + Y\hat{Y} + Z\hat{Z},\\
\bar{r}_{23} &= (X-A_2+a_1\cos\theta^*)\hat{X} + (Y+a_1\sin\theta^*)\hat{Y} + Z\hat{Z},\\
\bar{r}_{43} &= (X-A_2-a_2\cos\theta^*)\hat{X} + (Y-a_2\sin\theta^*)\hat{Y} + Z\hat{Z}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathcal{U}_X &= \Omega_1^2 X - \frac{M_1}{r_{13}^3}(X+A_1) - \frac{M_2}{r_{23}^3}(X-A_2+a_1\cos\theta^*)\\
&\quad -\frac{M_4}{r_{43}^3}(X-A_2-a_2\cos\theta^*), &\text{(4a)}\\
\mathcal{U}_Y &= \Omega_1^2 Y - \frac{M_1}{r_{13}^3}Y - \frac{M_2}{r_{23}^3}(Y+a_1\sin\theta^*)\\
&\quad -\frac{M_4}{r_{43}^3}(Y-a_2\sin\theta^*), &\text{(4b)}\\
\mathcal{U}_Z &= -\left(\frac{M_1}{r_{13}^3} + \frac{M_2}{r_{23}^3} + \frac{M_4}{r_{43}^3}\right)Z. &\text{(4c)}
\end{aligned}$$

Before analysis can begin, the motion of the primaries must be determined. Nevertheless, when using trigonometric series for obtaining approximate solutions, Equation (2) are an inconvenient set. *In particular, the time varying nature of the coefficients on the left side in this scalar set of equations makes it an inconvenient set for approximating a solution via trigonometric series.* Therefore, the relative formulation of the equations of motion is utilized to approximate the motion of the primaries and the S/C with trigonometric series.

RELATIVE FORMULATION

To approximate the motion of the primaries and the S/C, the relative formulation of the n-body problem is utilized. In this form, the vector (dimensional) equations of motion for P_i relative to P_q, and perturbed by the P_j's, can be written as,

$$\ddot{\bar{r}}_{qi} = \bar{\nabla}_i\left(U_i + \sum_{\substack{j=1\\ j\neq i,q}}^{n} R_j\right), \tag{5}$$

where $\bar{r}_{qi}$ is the position vector of i-th particle relative to the q-th particle, $\bar{\nabla}_i(\,) \triangleq \frac{\partial(\,)}{\partial \bar{r}_{qi}}$,

$$\begin{aligned}
U_i &= \frac{(M_i+M_q)}{r_{qi}},\\
R_j &= M_j\left(\frac{1}{r_{ij}} - \frac{\bar{r}_{qi}\cdot\bar{r}_{qj}}{r_{qj}^3}\right),
\end{aligned}$$

and where, M_i and M_q are the non-dimensional masses of P_i and P_q respectively. Of course, if $n=2$, then $R_j = 0$ for all j, and the problem reduces to the classical 2BP. The motion is readily available via conic sections. Consider cases where $n > 2$. If particle i is not assumed massless (denote this type of particle as a

primary), an n-body problem is obtained. That is, since particle i affects the motion of the other particles, a set of differential equations must be obtained for each particle i relative to particle q. Now, if particle i is massless, a restricted n-body problem is obtained. In this case, the motion of the primaries decouples as an $(n-1)$-body problem.

Form of the Approximate Solution

To obtain an approximate solution for Equation (5), let $q = 2$, i.e., P_2 is the reference particle. Then for each particle, P_i ($i \neq 2$), a perturbation approach is utilized. This, of course, implies that a reference or unperturbed motion must be assumed. For the primaries, assume that the unperturbed motion is circular two-body motion. Then, for the infinitesimal mass, assume that the unperturbed motion is CR3BP motion near the libration points. This allows the examination of perturbations to the libration points themselves and to the motion in their vicinity. Thus, for both the primaries and the spacecraft, the following form is utilized (ref. 4),

$$\bar{r}_{2i} = r^o_{2l}\left[f_1(\alpha)\hat{r}^o_{2l} + f_2(\beta)\hat{\theta}^o_{2l} + f_3(\gamma)\hat{h}^o_{2l}\right], \quad l \neq 2, j, \tag{6}$$

where i is either a primary or the spacecraft depending on the motion of interest, l is a primary other than 2 or j, and r^o_{2l} is the (constant) unperturbed distance between P_2 and P_l. The unit vectors ($\hat{r}^o_{2l}$, $\hat{\theta}^o_{2l}$, $\hat{h}^o_{2l}$) are, respectively, the (unperturbed P_2-P_l motion) radial, tangential, and normal directions. For example, consider the motion of P_1 relative to P_2, then

$$\bar{r}_{21} = r^o_{21}\left[f_1(\alpha)\hat{r}^o_{21} + f_2(\beta)\hat{\theta}^o_{21} + f_3(\gamma)\hat{h}^o_{21}\right].$$

Now, in this investigation, the form of the functions in Equation (6) is as follows,

$$f_1(\alpha) \triangleq \tilde{t} + \alpha, \tag{7a}$$
$$f_2(\beta) \triangleq \tilde{s} + \beta, \tag{7b}$$
$$f_3(\gamma) \triangleq \tilde{z} + \gamma, \tag{7c}$$

where $\tilde{t}$, $\tilde{s}$, and $\tilde{z}$ are constants and α, β, and γ are unknown functions of time. With this approximation, the inertial acceleration is (assuming the reference P_2-P_l motion is two-body circular),

$$\begin{aligned}\ddot{\bar{r}}_{2i} = r^o_{2l}\Big[&\left(\ddot{f}_1 - 2n_{2l}\dot{f}_2 - n^2_{2l}f_1\right)\hat{r}^o_{2l} \\ &+ \left(\ddot{f}_2 + 2n_{2l}\dot{f}_1 - n^2_{2l}f_2\right)\hat{\theta}^o_{2l} \\ &+ \left(\ddot{f}_3 \qquad\qquad\quad\right)\hat{h}^o_{2l}\Big],\end{aligned} \tag{8}$$

where n_{2l} is the unperturbed P_2-P_l mean motion. The $\bar{\nabla}_i$ operator in Equation (5) is evaluated as,

$$\bar{\nabla}_i(\,) = \frac{1}{r^o_{2l}}\left[\frac{\partial(\,)}{\partial f_1}\hat{r}^o_{2l} + \frac{\partial(\,)}{\partial f_2}\hat{\theta}^o_{2l} + \frac{\partial(\,)}{\partial f_3}\hat{h}^o_{2l}\right]. \tag{9}$$

Thus, the relative EOMs become

$$\ddot{f}_1 - 2n_{2l}\dot{f}_2 - f_1 n^2_{2l} = \left(\frac{1}{r^o_{2l}}\right)^2 \frac{\partial \mathcal{A}}{\partial f_1}, \tag{10a}$$

$$\ddot{f}_2 + 2n_{2l}\dot{f}_1 - f_2 n^2_{2l} = \left(\frac{1}{r^o_{2l}}\right)^2 \frac{\partial \mathcal{A}}{\partial f_2}, \tag{10b}$$

$$\ddot{f}_3 = \left(\frac{1}{r^o_{2l}}\right)^2 \frac{\partial \mathcal{A}}{\partial f_3}, \tag{10c}$$

where

$$\mathcal{A} \triangleq U_i + \sum_{\substack{j=1 \\ j\neq i,q}}^{n} R_j.$$

Then, it can easily be shown that, these equations can be written as

$$\ddot{f}_1 - 2n_{2l}\dot{f}_2 \underbrace{-3f_1}_{} = \tilde{U}_{f_1}, \tag{11a}$$

$$\ddot{f}_2 + 2n_{2l}\dot{f}_1 \underbrace{-3f_2}_{} = \tilde{U}_{f_2}, \tag{11b}$$

$$\ddot{f}_3 \underbrace{-3f_3}_{} = \tilde{U}_{f_3}, \tag{11c}$$

where

$$\tilde{U} \triangleq \frac{n_{2l}^2}{2}[f_1^2 + f_2^2] \underbrace{-\frac{3}{2}[f_1^2 + f_2^2 + f_3^2]}_{} + \left(\frac{1}{r_{2l}^o}\right)^2 \mathcal{A}, \tag{12}$$

and the terms with the underbraces have been added to alleviate problems with small divisors (ref. 8); their utility is exploited in the development used for harmonic balance (see ref. 13). The expression in Equation (11) represents a coupled set of autonomous, non-linear, non-homogeneous, ordinary differential equations.

PRIMARY MOTION

To determine periodic solutions in the R4BP, the mirror configurations must first be identified. A "mirror configuration" is defined as a set of position and velocities states such that all the particles have velocities perpendicular to every radius vector as measured from the system's stationary center of mass (ref. 14, 15). In the current case, this implies that all the particles align along the $\hat{X}$ axis. The term **"syzygy"** is utilized for this alignment condition. Relative to P_2, if P_2 is between P_1 and P_4, the syzygy is labeled opposition, otherwise, it is termed conjunction. Utilizing Roy and Ovenden's periodicity theorem — see ref. 14 — two mirror configurations (at two different epochs) are sufficient for periodicity. The restricted condition permits the search for mirror configurations in two steps: (1) for the primaries; (2) for the S/C. The primaries will satisfy the periodicity theorem at the syzygy alignments. Now, if the primary syzygies occur periodically, with period T_s, an approximation to the primary motion in terms of a Fourier series with period T_s can be utilized. In this case, the synodic period is given by $T_s = \frac{2\pi}{N}$, where $N = n_{24} - n_{21}$, and where n_{24} and n_{21} are, respectively, the mean motions of the P_2-P_4 and P_2-P_1 systems. The computation of these approximations is easily accomplished using the method of harmonic balance. This is detailed in references 8, 9, 12, 16.

SPACECRAFT MOTION

For the spacecraft, any periodic solutions must have periods commensurate with T_s. Therefore, the spacecraft periodic solutions occur as isolated solutions with well determined periods and *not* in families with continuously varying periods, as is the case in the CR3BP (ref. 17). (See also Wintner, ref. 18.) Again, two methodologies are used to obtain periodic solutions for the motion of the spacecraft: (1) differential corrections; and, (2) harmonic balance. In terms of computation time, the method of differential corrections is more efficient. Nevertheless, harmonic balance offers a semi-analytical approximation for the periodic orbit. With regards to the semi-analytical approximations, concern is not so much in obtaining a very accurate approximation but in obtaining an approximation that might be useful for further studies in spacecraft trajectory design. Furthermore, the approximation can be refined by differential corrections (see ref. 19). The availability of semi-analytical expressions provides a (computationally) quick and efficient starting point to generate solutions in the ephemeris model. In any case, in the coherent-R4BP, the period of any solution is predetermined; and, it is commensurate with the synodic period, that is, $T = \frac{p}{q}T_s = \frac{p}{q}\frac{2\pi}{N}$, where p and q are *integers* representing the number of P_1-P_2-P_4 synodic periods and the number of complete revolutions of P_3, respectively. Note that the P_3 orbit can be described as having period pT_s while undergoing q revolutions. Expansions developed for the CR3BP can be utilized as generating solutions. In particular, Richardson's approximation for periodic orbits is exploited (ref. 20). Nevertheless, since the periodic orbits are isolated and not in families, a quick search utilizing differential corrections is performed first. Still, differential corrections might not expose some periodic solutions that the method of harmonic balance could easily compute. Thus, the two methods truly complement each other.

All the examples considered in this paper are based in the Sun-Earth-Moon primary system. However, the methods are applicable to any Sun-Planet-Moon primary system. It should also be remarked that the search/computation of periodic orbits in this investigation is by no means an exhaustive one. The purpose is to illustrate the methodologies and to determine some of the possibilities in terms of spacecraft orbits. The orbit stability is also briefly examined, not to compare with the CR3BP, but for completeness. It is noted that a bifurcation analysis would also be of importance (see for example ref. 15 and ref. 12). This will uncover the likely existence of doubly-periodic solutions, triple-periodic, and so on.

Differential Corrections Approach

To quickly assess the available periodic orbits in the vicinity of the libration points, a search is performed by utilizing differential corrections. Specifically, in accordance with the mirror theorem, two perpendicular crossings of the X-Z plane, coincident with the primaries' syzygies, are sufficient for a periodic solution. Thus, an initial state vector of the form (X_o, 0, Z_o, 0, $\dot{Y}_o$, 0) is utilized. This produces a perpendicular crossing of the X-Z plane at $t = t_o$; then, a subsequent perpendicular crossing at $t = qT/2$ satisfies the periodicity theorem. The required initial state vector is computed via differential corrections.

L_1 Periodic Orbits

Utilizing a northern[2] L_1 family of halo orbits (from the CR3BP) as initial guesses, and with the values $p = 6$, $q = 1$, the periodic solutions illustrated in Figure 2 are computed in the coherent-R4BP. Note that the lunar orbit as well as the Earth also appear in the plot. In this system, the integer values of p and q indicate one revolution of the S/C relative to the B-B' rotating frame (see Figure 1) and six synodic periods of the Sun-Earth-Moon system, or a period of roughly six months. Three periodic orbits are actually obtained by using members of the L_1 halo family in the vicinity of the libration point as generating orbits. Two are planar orbits (orbit **6-1-a** and **6-1-c**) and one is a northern three-dimensional halo orbit (orbit **6-1-b**). Essentially, the computed orbits are the members of the Sun-Earth CR3BP halo family that, after the introduction of the lunar perturbation, possess a period commensurate with the synodic period of the Sun-Earth-Moon system. The same behavior is expected when computing periodic orbits using members of the southern halo family from the CR3BP as generating orbits, since the spacecraft equations of motion possess a reflection symmetry about the X-Y plane.

L_2 Periodic Orbits

To seek periodic orbits in the vicinity of L_2, a procedure is employed that is similar to the one just described for periodic orbits near L_1. The northern L_2 family of halo orbits (from the CR3BP) produces initial guesses in this case. Interestingly, with the values $p = 6$, $q = 1$, no periodic solutions emerge. A solution similar to the northern halo orbit computed near L_1 (Figure 2) was very close to convergence but did not meet the periodicity conditions, that is, the final state was not equal to the starting state. This result, of course, does not mean that there are no periodic solutions in the vicinity of L_2. Rather, periodicity may require a different period and/or multiple revolutions , $q > 1$. A more exhaustive search can be performed but is not within the scope of this investigation. Certainly, previous work suggests periodic orbits may exist near both L_1 and L_2 of a type corresponding to $p = 12$, $q = 2$ (ref. 9). However, another example of a periodic orbit with period equal to $12T_s$ in the vicinity of L_2 is computed in the next section.

Nearly Vertical Orbits

Initial guesses obtained from northern halo families (either L_1 or L_2) and fixing the desired period to be $12T_s$ ($p = 12$, $q = 1$), two periodic orbits emerge that appear in Figure 3. These orbits possess large out-of-plane excursions; the A_z amplitudes are approximately 7.9e+6 km. The characteristics, as viewed in the X-Y plane, can be appreciated in Figure 4 (the orbits are labeled as orbits **12-1-a** and **12-1-b**). These orbits are generally known as nearly-vertical orbits or orbits of the first type and were first sketched by Moulton (ref. 23).

[2]A class I (II), or northern (southern), halo orbit can be defined as one that has a maximum X component excursion, as measured from B', in the positive (negative) Z axis (ref. 21,22).

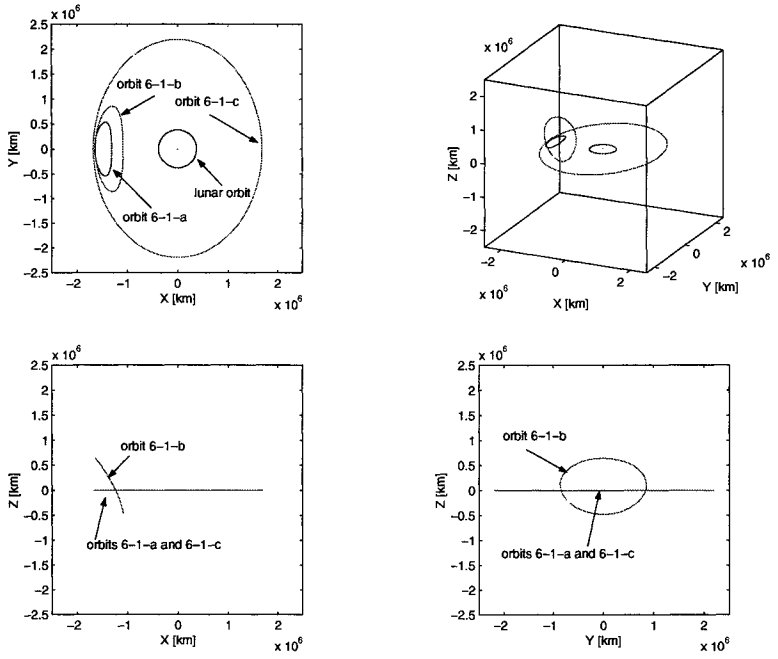


Figure 2: L_1 Periodic Orbits, Period = $6T_s$

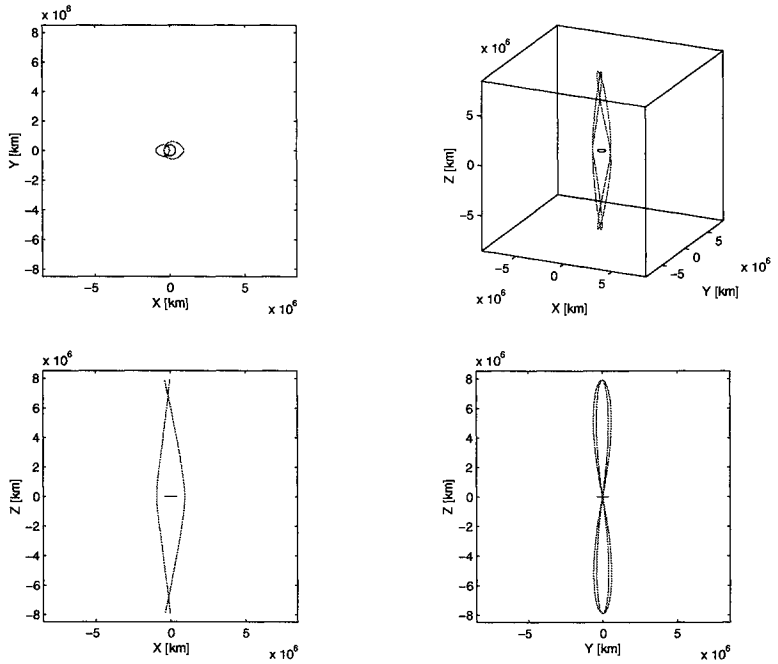


Figure 3: Nearly Vertical Orbits, Period = $12T_s$

Harmonic Balance Approach

The approach to approximating periodic solutions through a perturbation method and trigonometric series was first utilized by Kolenkiewicz and Carpenter (1967) to find planar periodic solutions in the vicinity of the Sun perturbed Earth-Moon L_4 and L_5 points (ref. 5). Again, this approach yields a semi-analytical approximation that will be useful when considering motion in the ephemeris model. Specifically, the functions α, β, and, γ in Equation (7) are approximated as follows,

$$\alpha = \sum_{k=0}^{K-1} \left[\alpha_k^{(c)} \cos(\omega_k t) + \alpha_k^{(s)} \sin(\omega_k t) \right], \tag{13a}$$

$$\beta = \sum_{k=0}^{K-1} \left[\beta_k^{(c)} \cos(\omega_k t) + \beta_k^{(s)} \sin(\omega_k t) \right], \tag{13b}$$

$$\gamma = \sum_{k=0}^{K-1} \left[\alpha_k^{(c)} \cos(\omega_k t) + \alpha_k^{(s)} \sin(\omega_k t) \right]. \tag{13c}$$

Now, since the solution of interest is periodic with period $T_{qp} = 2\pi/N_{qp}$, where $N_{qp} = Nq/p$, the frequencies ω_k are selected as follows: $\omega_k = kN_{qp}$. For periodic motion, this type of series is actually a *Fourier series* (with $K = \infty$). The error in satisfying the relative equations of motion is computed in a grid $\theta_{qp} \in [0,\ 2\pi]$ ($\theta_{qp} = N_{qp}t$) and is selected as the Euclidean norm of the difference between the left side and right side of Equation (5). Depending on the convergence properties of the orbit under consideration, different θ_{qp}-grids can be utilized. Arbitrarily, the θ_{qp}-grid is divided in 36 steps, that is, every 10 degrees. Then, the maximum error in the θ_{qp}-grid is selected as the reference value that defines the stopping condition, $E < \text{tol}$, where E is the error and tol is the user defined tolerance. Other sources of error exist and are dealt with appropriately (see ref. 16).

Mirror Theorem and Symmetry Considerations

When considering harmonic balance as a procedure to compute periodic orbits, several factors can speed the computations. First, note that at $\theta_{qp} = N_{qp}t = j\pi$, $j = 0, \pm 1, \pm 2, \ldots$, the primaries are in syzygies.

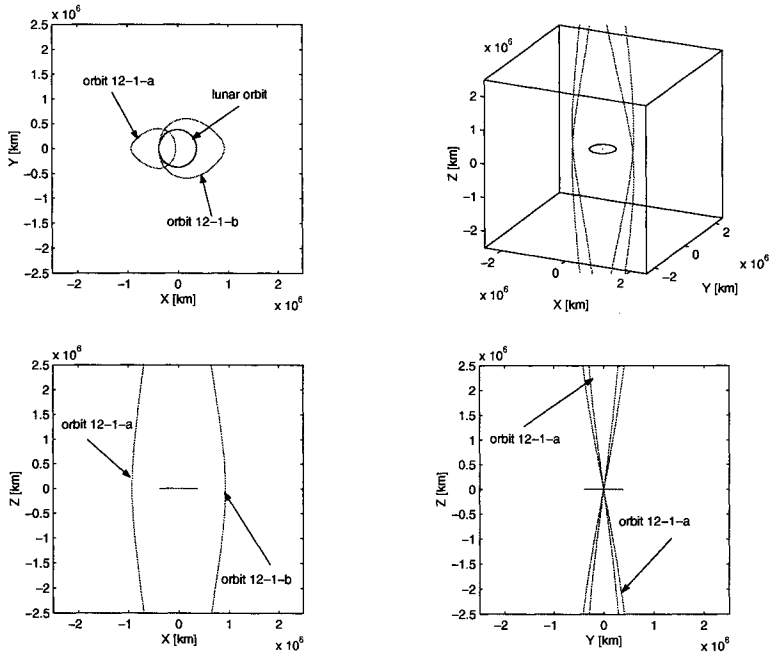


Figure 4: Zoom In of Nearly Vertical Orbits, Period = $12T_s$

At these instances, the unit vectors that correspond to the reference, or undisturbed motion, are parallel to the unit vectors that define the B-B' rotating frame (Figure 1). Thus, the perpendicular crossings can be examined in the context of the reference coordinate frames. Utilizing the approximation for the spacecraft (Equation (6)) and computing its inertial derivative,

$$\bar{r}_{23} = r_{21}^o \left[(\tilde{t}+\alpha)\hat{r}_{21}^o + (\tilde{t}+\beta)\hat{\theta}_{21}^o + (\tilde{t}+\gamma)\hat{h}_{21}^o \right], \tag{14a}$$

$$^I\dot{\bar{r}}_{23} = r_{21}^o \left[(\dot{\alpha} - n_{21}\beta)\hat{r}_{21}^o + (\dot{\beta} + n_{21}\alpha)\hat{\theta}_{21}^o + (\dot{\gamma})\hat{h}_{21}^o \right]. \tag{14b}$$

For a perpendicular crossing of the $\hat{X}$ axis at $\theta = j\pi$, $j = 0, \pm 1, \pm 2, \ldots$,

$$\dot{\alpha} - n_{21}\beta = 0, \tag{15a}$$

$$\dot{\gamma}\alpha = 0. \tag{15b}$$

The first condition above is automatically satisfied if all $\alpha_k^{(s)} = \beta_k^{(c)} = 0$. This result implies that α_{2i} is an even function and β_{2i} is odd. A similar examination of Equation (15) yields $\gamma_k^{(s)} = 0$, such that the equation is automatically satisfied. This implies that γ is an even function. Thus, orbits with these properties will automatically satisfy the mirror theorem. Moreover, in terms of computation time, the previous considerations allow the elimination of $3K$ terms and, thus, speed up the computations (ref. 16). Additionally, if the desired periodic orbit is planar, K more terms (the $\gamma_k^{(s)}$) are eliminated.

Libration Points in the Coherent-R4BP

Since interest is in spacecraft operating near the equilibrium points as originally defined in the CR3BP, analysis of the evolution of these points in the coherent-R4BP is useful. Specifically, the perturbation of the Sun-Earth libration points due to the presence of the Moon can be examined utilizing Equation (6). A result from the theory of ordinary differential equations ensures that if the perturbation is small enough and if the linear frequencies around the point are different[3] from $2\pi kN$, where k is an integer, then the equilibrium point is replaced by a periodic orbit of (synodic) frequency N (ref. 25). This orbit reduces to the equilibrium point when the perturbation is not present. These orbits are, of course, planar in the B-B' frame. In the Sun-Earth-Moon system $T_s \approx 29.49$ days.

For motion near the perturbed libration points, let $\tilde{t}$, $\tilde{s}$ and $\tilde{z}$ (see Equation (7)) be the coordinates of the (unperturbed) libration point. For L_i, $i = 1$, or 2, recall that the positions of the collinear libration points in the CR3BP are obtained by solving a quintic equation (ref. 26), specifically for L_1 and L_2,

$$\gamma_i^5 \mp (3-\mu)\gamma_i^4 + (3-2\mu)\gamma_i^3 - \mu\gamma_i^2 \pm 2\mu\gamma_i - \mu = 0, \tag{16}$$

where $\mu = M_2/(M_1+M_2)$ and γ_i is the non-dimensional distance from P_2 to either L_1 ($i = 1$) or L_2 ($i = 2$). Then, $\tilde{t} = \pm\gamma_i$ and $\tilde{s} = \tilde{z} = 0$. In both appearances of $\mp$ and $\pm$, the upper sign corresponds to L_1 and the lower to L_2.

Perturbed L_1

First, the Moon-perturbed L_1 point is examined, given utilizing Equation (6) with $\tilde{t} = \gamma_1$ and $\tilde{s} = \tilde{z} = 0$. With an initial guess corresponding to $\alpha = \beta = \gamma = 0$ and applying Newton's method (ref. 16), a periodic orbit for the "motion" of L_1 (Period$= T_s$) is computed (orbit **L_1 1-1-a**). Arbitrarily, the initial time is set to zero (this corresponds to a conjunction of the Sun-Earth-Moon system). With a truncated value of $K = 40$, a non-dimensional tolerance of 1.0e-12 is easily achieved (that is, the error in satisfying the relative equations of motion).[4] This planar orbit appears in both plots in Figure 5, relative to the B-B' rotating frame. Not surprisingly, this orbit is symmetric about the X-axis. Note that a point has also been indicated in both plots. This point, plotted for comparison, represents a different approximation (ref. 7) for the effect of the Moon in the location of the libration points. Specifically, it is the Sun-Earth/Moon barycenter L_1 point. To

[3] That is, the linear system is non-critical with respect to T_s-periodic functions. See Hale, ref. 24.

[4] In the Sun-Earth-Moon system, this number corresponds to an acceleration of 2.73e-15 m/s^2

Table 1: Coefficients of L_1 Periodic Planar Orbit, Period $= T_s$

k	$\alpha_k^{(c)}$	$\beta_k^{(s)}$
0	4.257928471021044e-005	5.843219081738828e-029
1	3.098471933080993e-005	3.144260512919833e-005
2	9.032412089776317e-008	-7.296757434881545e-008
3	5.879609021429410e-008	3.816957410917719e-008
4	1.863895939348030e-009	-1.509502782942379e-009
5	5.138325277174839e-010	-9.413079324364794e-011
6	6.347206583803272e-011	-5.420088541564582e-011
7	1.368158352217061e-011	-1.018609426111021e-011
8	2.658586135809054e-012	-2.351148076561475e-012
9	5.746530983243758e-013	-5.043552305855827e-013
10	1.245257055499692e-013	-1.126140175348783e-013
11	2.773845184707288e-014	-2.523116504988684e-014
12	6.250239243716285e-015	-5.740059648774831e-015
13	1.426715824055346e-015	-1.317780663850541e-015
14	3.287800697965160e-016	-3.053697113520435e-016

compute this point, referring to Figure 1, combine the mass of the Earth and the Moon and place it at the mean location of B', then compute the libration point location associated with the Sun and the combined Earth/Moon mass. *It is perhaps important to clarify that the artificial libration point is not the reference solution used to compute the periodic orbit.* The reference solution, the Sun-Earth L_1 point, is fixed in the Sun-Earth rotating frame. If plotted in the B-B' rotating frame, this point is a "circular" orbit.

Although both plots in Figure 5 reflect the same data, in the top plot the origin is selected as the location of B', while in the bottom plot, the origin is the position of the artificial L_1 point. Regarding this artificial libration point, in the coherent-R4BP, the actual distance between the Sun and the Earth-Moon barycenter varies periodically. Nevertheless, this variation has an amplitude of about 45.9 meters. As a result, the location of the periodic orbit in Figure 5, relative to the artificial libration point, does not vary significantly throughout a year. The first 15 coefficients of the trigonometric series are displayed in Table 1.

Perturbed L_2

For the same dynamical conditions, the perturbation of the Sun-Earth L_2 point due to the presence of the Moon is examined from Equation (6) with $\tilde{t} = -\gamma_2$ and $\tilde{s} = \tilde{z} = 0$. Corresponding to an initial guess of $\alpha = \beta = \gamma = 0$, and relying on Newton's method, a periodic orbit for the "motion" of L_2 (Period$= T_s$) is computed (orbit **L_2 1-1-a**). The truncated value of K is again $K = 40$, and a non-dimensional tolerance of

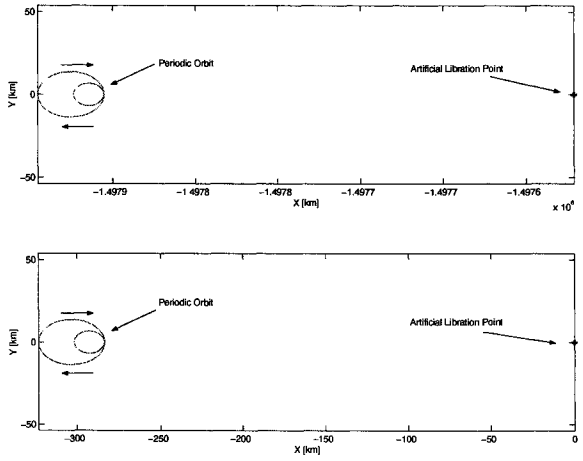


Figure 5: L_1 Periodic Planar Orbit, Period $= T_s$

Table 2: Coefficients of L_2 Periodic Planar Orbit, Period = T_s

k	$\alpha_k^{(c)}$	$\beta_k^{(s)}$
0	-4.308177853140367e-005	4.776383311993790e-029
1	3.098198446201324e-005	3.144398088429650e-005
2	-9.757720965177109e-008	6.109354400941634e-008
3	5.842262291212265e-008	3.846028012268063e-008
4	-1.803152086383687e-009	1.443930994091168e-009
5	4.997273355135388e-010	-8.230933298012393e-011
6	-6.036553242065950e-011	5.167640628376973e-011
7	1.300634868789042e-011	-9.595935821757570e-012
8	-2.503064572298579e-012	2.215347698754724e-012
9	5.386773226025647e-013	-4.720818324015804e-013
10	-1.160129836214154e-013	1.049248352060872e-013
11	2.570716699059446e-014	-2.337645533522387e-014
12	-5.760532585443235e-015	5.290039152257805e-015
13	1.307859285321864e-015	-1.207863761023478e-015
14	-2.997517335337430e-016	2.783916183765845e-016

1.0e-12 is easily achieved. This planar orbit appears in both plots in Figure 6 relative to the B-B' rotating frame. Again, note the position of an artificial L_2 point has been also been identified in both plots. The first 15 non-zero coefficients are displayed in Table 1. Note that the orbit is planar and symmetric.

Farquhar correctly noticed that utilizing the artificial L_2 point is indeed more accurate than using the Sun-Earth L_2 point (ref. 7). In fact, note that the "center" of the periodic orbit is shifted about 300 km from the artificial L_2 point. Relative to this "center" the maximum distance to the periodic path is about 20 km.

Three-Dimensional Periodic Orbits

The harmonic balance procedure is exploited to compute a three-dimensional periodic orbit in the vicinity of L_1. The initial guesses might come from approximations such as Richardson's (ref. 20) and/or numerical integrated trajectories. In the case of approximation such as Richardson's, the cosine and the sine coefficients just need to be placed in the appropriately in the approximation in Equation (6). In the case of numerically integrated trajectories, the following process is utilized: (1) Transform the position coordinates of $\bar{r}_{23}$ to the unperturbed frame, in this case the unperturbed frame has unit vectors ($\hat{r}_{21}^o$, $\hat{\theta}_{21}^o$, $\hat{h}_{21}^o$), (2) The values for the perturbations α_{23}, β_{23} and γ_{23} are easily obtained by utilizing dot products with Equation (6), and, (3) Once the values of α_{23}, β_{23} and γ_{23} have been computed for one period of the motion, the FFT algorithm

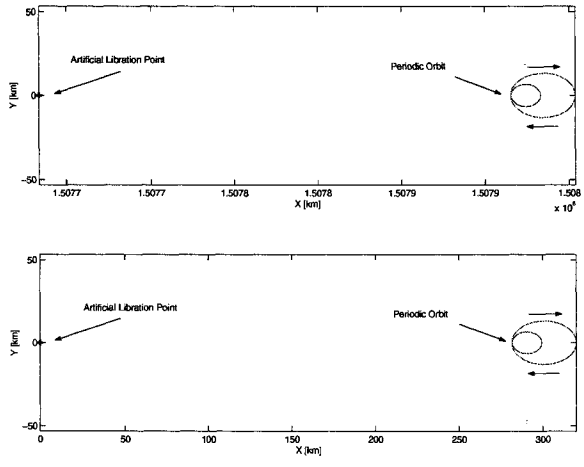


Figure 6: L_2 Periodic Planar Orbit, Period = T_s

can be used to obtain trigonometric series for each of them. The coefficients of these three series become the initial guess. Now, in the R4BP, the application of harmonic balance is complicated due to the fact that the period of the orbit must be specified and that a good initial guess is needed (since periodic orbits are isolated). Utilizing this procedure with the northern L_1 periodic orbit computed via differential corrections, with the values $p = 6$, $q = 1$, and with $K = 80$ an approximation for orbit 6-1-b is computed (see Figure 2) . The amplitudes are, in kilometers, A_x = 552,209, A_y = 860,510, and A_z = 600,022. The minimum, non-dimensional, error achieved was 6.633e-9. The first 15 coefficients are output in Table 3.

Table 3: First 15 Coefficients of L_1 Periodic Halo Orbit, Period = $6T_s$

k	$\alpha_k^{(c)}$	$\beta_k^{(s)}$	$\gamma_k^{(c)}$
0	0.000000000000000e+000	2.123621658099728e-024	0.000000000000000e+000
1	-1.872010605047259e-003	5.778838557474630e-003	-3.691446812490400e-003
2	1.328247333013201e-004	2.621096997603808e-004	-2.622855159042354e-004
3	2.461430018473915e-005	5.024912143715533e-005	-4.492213445304252e-005
4	5.832231639898343e-006	1.024569220748677e-005	-9.040806355724052e-006
5	1.418151309257614e-006	2.342612971454473e-006	-2.088073578087093e-006
6	3.129751102489206e-005	3.202595943912278e-005	-5.205178407880042e-007
7	1.000563745432675e-007	1.308183175944395e-007	-1.667987455374001e-007
8	4.531263161075272e-008	2.022595651049012e-009	-6.438258890445861e-008
9	5.287461851549687e-008	-4.804353291024348e-008	-4.537829566217946e-008
10	6.853106704694869e-008	-6.797191874935834e-008	-4.490080263590235e-008
11	7.186407183367793e-008	-6.115975756827563e-008	-3.586526465918923e-008
12	2.861803980737176e-008	-2.550397227724550e-008	-1.538742446656812e-008
13	-2.573618030012998e-008	2.267970270741595e-008	-2.423493558736808e-008
14	4.009245701445942e-009	-9.796902268764903e-009	-2.968895553400935e-009

Accuracy of the Approximations

Although not one of the goals of this investigation, it is important to check the accuracy of the approximations. The orbits are numerically integrated (and re-initialized) in time steps and the error between the approximation and the numerical integration is compared for one period of the motion. These errors are displayed in Table 4.

Table 4: Approximation versus Numerical Integration

Orbit	step interval [days]	max. position diff. [m]	max. velocity diff. [m/s]
6-1-b	0.1	3.4669e+3	8.0253e-1
L_1 1-1-a	2.0	1.2945e-2	1.5149e-7
L_2 1-1-a	2.0	1.3063e-2	1.4731e-7

It is clear that the approximation for the large amplitude, three-dimensional orbit 6-1-b is not very accurate. More terms are needed in the approximation. Nevertheless, this orbit can be efficiently stored (with only 80 coefficients) and quickly corrected to obtain an integrated trajectory.

Stability Information

To check the stability of the periodic orbits, the eigenvalues of the monodromy matrix are computed. To make sure that all the computations are free of error, all the orbits were numerically integrated along with the state transition matrix. These eigenvalues appear in Table 5 (the symbol $*$ denotes the complex conjugate).

Table 5: Stability: Periodic Orbits

Orbit	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
6-1-a	1.861e+3	5.374e-4	1.004e+0	9.963e-1	9.929e-1 + i(1.188e-1)	λ_5^*
6-1-b	1.085e+3	9.214e-4	1.011e+0	9.893e-1	8.553e-1 + i(5.182e-1)	λ_5^*
6-1-c	1.002e+0	9.976e-1	-6.501e-1 + i(7.599e-1)	λ_3^*	8.104e-2 + i(9.967e-1)	λ_5^*
12-1-a	6.165e+2	1.622e-3	1.768e+1	5.656e-2	1.117e+0	8.9541e-1
12-1-b	6.255e+2	1.599e-3	1.774e+1	5.639e-2	1.066e+0	9.3784e-1
L_1 1-1-a	3.615e+0	2.766e-1	4.903e-1 + i(8.716e-1)	λ_3^*	5.209e-1 + i(8.536e-1)	λ_5^*
L_2 1-1-a	3.528e+0	2.835e-1	5.032e-1 + i(8.642e-1)	λ_3^*	5.339e-1 + i(8.455e-1)	λ_5^*

All computed orbits are unstable. The eigenspaces for orbits in the vicinity of the collinear points retain the same dimension as the collinear points in the CR3BP. Note the strong hyperbolic character of orbits 6-1-a and 6-1-b. Finally, observe that the vertical orbits 12-1-a and 12-1-b are hyperbolic.[5]

SUMMARY

In this paper two procedures were utilized to compute periodic orbits in the context of a coherent-R4BP: (1) differential corrections, and (2) harmonic balance. Differential corrections are computationally faster but do not provide a semi-analytical description of the orbit. Harmonic balance, on the other hand provides semi-analytical expressions. The two procedures can be used together in an effort to search the phase space for periodic motion. It is noted that harmonic balance can provide an initial guesses for differential corrections and vice versa. It should also be remarked that the search/computation of periodic orbits in this investigation is by no means an exhaustive one. The purpose is mainly to illustrate the methodologies and to determine some of the possibilities in terms of nominal spacecraft orbits. The orbit stability was also briefly examined. With the computed periodic solutions, dynamical systems theory can then be applied to examine the phase space. New trajectory options can then be explored.

REFERENCES

[1] Su-Shu Huang. Very Restricted Four-Body Problem. Technical report, Goddard Space Flight Center, September 1960. NASA TN D-501.

[2] Su-Shu Huang. Some Dynamical Properties of the Natural and Artificial Satellites. Technical report, Goddard Space Flight Center, September 1960. NASA TN D-502.

[3] J.M.A. Danby. Inclusion of Extra Forces in the Problem of Three Bodies. *The Astronomical Journal*, 70(3):181–189, April 1965.

[4] P. Musen and L. Carpenter. On the General Planetary Perturbations in Rectangular Coordinates. *Journal of Geophysical Research*, 68(9):2727–2734, May 1963.

[5] R. Kolenkiewicz and L. Carpenter. Periodic Motion around the Triangular Libration Point in the Restricted Problem of Four Bodies. *The Astronomical Journal*, 72(2):180–183, March 1967.

[6] R. Kolenkiewicz and L. Carpenter. Stable Periodic Orbits about the Sun Perturbed Earth-Moon Triangular Points. *AIAA Journal*, 6(7):1301–1304, July 1968.

[7] R.W. Farquhar. The Moon's Influence on the Location of the Sun-Earth Exterior Libration Point. *Celestial Mechanics*, 2:131–133, July 1970.

[8] L. Carpenter. Periodic Orbits in Trigonometric Series. In G.E.O. Giacaglia, editor, *Periodic Orbits, Stability and Resonances*, pages 192–209. D. Reidel, Dordrecht-Holland, 1970.

[5] A periodic orbit in an autonomous system has always one eigenvalue equal to 1 (ref. 27). The coherent-R4BP, however, is a non-autonomous system.

[9] K.C. Howell and D.B. Spencer. Periodic Orbits in the Restricted Four-Body Problem. In *36th Congress*, Stockholm, Sweden, October 1985. International Astronautical Federation. IAF Paper 85-252.

[10] G. Gómez, A. Jorba, J. Masdemont, and C. Simó. Moon's Influence on the Transfer from the Earth to a Halo Orbit Around l_1. In Archie E. Roy, editor, *Predictability, Stability, and Chaos in N-Body Dynamical Systems*, pages 630–640. Plenum Press, New York, 1991.

[11] D.J. Scheeres. The Restricted Hill Four-Body Problem with Applications to the Earth-Moon-Sun System. *Celestial Mechanics*, 70:75–98, 1998.

[12] M.A. Andreu. *The Quasi-Bicircular Problem.* PhD thesis, Universitat de Barcelona, Barcelona, Spain, October 1998.

[13] K.C. Howell and J.J. Guzmán. Spacecraft Trajectory Design in the Context of a Coherent Restricted Four-Body Problem with Application to the MAP Mission. In *51st International Astronautical Congress*, Rio de Janeiro, Brazil, October 2000. IAF Paper 00-A.5.06.

[14] A.E. Roy and M.W. Ovenden. On the Occurrence of Commensurable Mean Motion in the Solar System II. The Mirror Theorem. *Monthly Notes of the Royal Astronomical Society*, page 300, 1955.

[15] E.T. Campbell. *Bifurcations from Families of Periodic Solutions in the Circular Restricted Problem with Application to Trajectory Design.* PhD thesis, Purdue University, West Lafayette, Indiana, August 1999.

[16] J.J. Guzmán. *Spacecraft Trajectory Design in the Context of a Coherent Restricted Four-Body Problem.* PhD thesis, Purdue University, West Lafayette, Indiana, May 2001.

[17] R. Broucke. Stability of Periodic Orbits in the Elliptic Restricted Three-Body Problem. *AIAA Journal*, 7(6):1003–1009, June 1969.

[18] A. Wintner. *The Analytical Foundations of Celestial Mechanics.* Princenton University Press, Princeton, New Jersey, 1947.

[19] K.C. Howell and H.J. Pernicka. Numerical Determination of Lissajous Trajectories in the Restricted Three Body Problem. *Celestial Mechanics*, 41:107–124, 1988.

[20] D.L. Richardson. Analytic Construction of Periodic Orbits About the Collinear Points. *Celestial Mechanics*, 22:241–253, 1980.

[21] R. W. Farquhar and A. A. Kamel. Quasi-periodic orbits about the translunar libration point. *Celestial Mechanics*, 7:458–473, 1973.

[22] J.V. Breakwell and J.V. Brown. The 'halo' family of 3-dimensional periodic orbits in the earth-moon restricted 3-body problem. *Celestial Mechanics*, 20:389–404, 1979.

[23] K.C. Howell. Families of Orbits in the Vicinity of the Collinear Libration Points. In *AAS/AIAA Astrodynamics Specialists Conference*, Boston, Massachusetts, August 1998. AAS Paper 98-4465; Journal of the Astronautical Sciences, Vol. 49, January-March 2001 (to appear).

[24] J.K. Hale. *Ordinary Differential Equations*, volume XXI of *Pure and Applied Mathematics.* John Wiley and Sons, Inc., New York, 1969.

[25] E. Castella and A. Jorba. On the Vertical Families of Two-Dimensional Tori Near the Triangular Points of the Bicircular Problem. *Preprint*, June 1999.

[26] V. Szebehely. *Theory of Orbits, The Restricted Problem of Three Bodies.* Academic Press, New York, 1967.

[27] T.S. Parker and L.O. Chua. *Practical Numerical Algorithms for Chaotic Systems.* Springer-Verlag, New York, 1989.

UNIQUE NON-KEPLERIAN ORBIT VANTAGE LOCATIONS FOR SUN-EARTH CONNECTION AND EARTH SCIENCE VISION ROADMAPS

David Folta
NASA Goddard Space Flight Center
Greenbelt, MD

Corissa Young
Aerospace Engineering
University of Colorado
Via GSFC's NASA Academy

Adam Ross
Physics & Astrophysics
Harvard University
Via GSFC's NASA Academy

ABSTRACT

The purpose of this investigation is to determine the feasibility of attaining and maintaining unique non-Keplerian orbit vantage locations in the Earth/Moon environment in order to obtain continuous scientific measurements. The principal difficulty associated with obtaining continuous measurements is the temporal nature of astrodynamics, i.e., classical orbits. This investigation demonstrates advanced trajectory designs to meet demanding science requirements which cannot be met following traditional orbital mechanic logic. Examples of continuous observer missions addressed include Earth pole-sitters and unique vertical libration orbits that address Sun-Earth Connection and Earth Science Vision roadmaps.

INTRODUCTION

The principal difficulty associated with obtaining continuous scientific measurements is the temporal nature of astrodynamics. As Earth and Space Science requirements become more variable, a new and innovative approach must be taken to enable unique vantagepoint missions. The Sun-Earth Connection's roadmap using a Magnetospheric Pole-Sitter and Earth Science Vision's roadmap using unique libration points orbits are examples of multiple continuous observers at unique vantage points. These requirements cannot be met using fundamental orbit design based on classical Keplerian mechanics.

To determine the feasibility of attaining and maintaining non-Keplerian orbits, an alternate mission design method based on advances in dynamical systems theory and modeling capabilities is investigated.[1] The results from this analysis provide an innovative set of previously undiscovered 'orbits' which meet demanding scientific objectives. An investigation of the feasibility of attaining and maintaining unique non-Keplerian orbit vantage locations in the Earth/Moon environment in order to obtain

continuous scientific measurements was proposed and awarded as a Director's Discretionary Fund activity. An association of the Goddard Space Flight Center's (GSFC) Flight Dynamics Analysis Branch with the NASA Academy and Purdue University accomplished the work performed. This analysis related to near-Earth and Libration orbit locations includes no-thrust and constant thrust using either a standard propulsion systems or a solar sail. A dynamical systems approach is utilized to initialize vertical libration orbits.[1] Since the objective is to provide a continuous measurement of the Polar Region; we considered the visibility duration of the Earth's Polar Regions as the main condition of success.

Application Context

There are many NASA customers seeking unique non-Keplerian orbits. In the Space Science arena: Origins, Solar Exploration of the Universe (SEU), and the Sun-Earth Connection (SEC) have unique orbit requirements that call for stationary polar locations. Earth Science Enterprise (ESE) customers interested in enabling continuous Earth measurements also require these types of observations. Note that space science teams involving the Earth Vision's Sensor Web which specifically calls for unique vantage locations may be unaware that the trajectory design capabilities in these unique regimes are relatively immature.

Objectives and Research Development Plan

The objective of this work was to provide NASA projects with an alternative to standard mission design with new and unique trajectories and orbits. We seek to establish the feasibility of non-Keplerian orbit applications to meet unique science requirements. Maintenance in these non-Keplerian trajectories is also addressed. Two distinct families of unique non-Keplerian locations and mission types are investigated: unique vertical lissajous orbits and non-Keplerian 'orbits' within the Earth's gravitational sphere of influence.

The vertical lissajous orbits which spend time over the Polar Regions include:

- Small vertical libration orbits with z-axis amplitudes of approximately 1 million km.
- Large vertical libration orbits (see Figure 1) with z-axis amplitudes of greater than 2 million km.

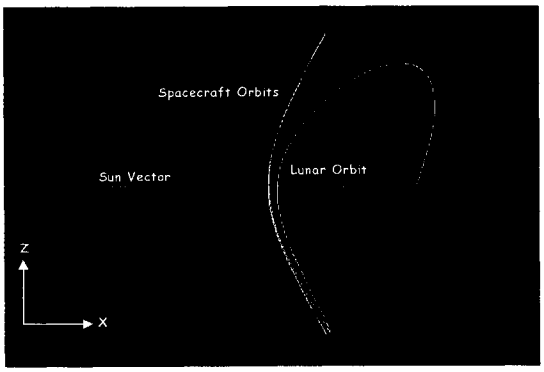


Figure 1. Large Vertical Libration Orbit

Non-Keplerian orbits within the Earth gravitational sphere of influence that maintained a constant location over the Earth's polar regions include:

- Stationary locations at a distance of 25,000 km.
- Stationary locations at a distance of 250,000 km.

A stationary 25,000km polar distance is shown in Figure 2.

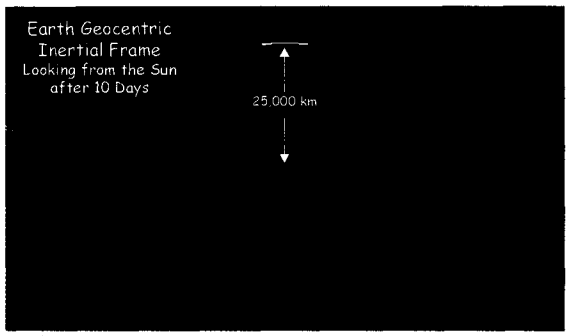


Figure 2. Pole Sitter Configuration

The analysis is performed in two phases. Phase 1 investigated different types of non-Keplerian orbits at the Sun-Earth L_1 point. We then derived comparable Keplerian orbits for contrast, such as a lunar distance circular orbit which used a lunar gravity assist and a more traditional highly elliptical orbit that permitted polar visibility.[2] This contrast determined the advantage if any of the vertical libration orbit in terms of polar coverage. Phase 2 addressed innovative propulsion scenarios using solar sails as a comparison to finite maneuvers to deal with the gravitational acceleration field at the Earth's Polar Region.

After initial orbital locations were identified, comparisons of continuous polar region accesses are made and the ΔV maintenance cost is addressed. Accelerations in the Earth potential well and the related solar sail mass to area ratios to counter these accelerations were studied. Maintenance requirements for spacecraft stationed at these locations may be reduced via judicious use of location selection, perturbation modeling, and enabling technologies using solar sails, electric propulsion, and autonomous control.

DDF and NASA Academy

During the analysis, as part of the Director's Discretionary Fund, NASA academy students were assigned and a grant was given to Purdue University. To examine the possibilities, they analyzed required accelerations for control of non-Keplerian trajectories at various vantage locations about the Earth. These vantage locations include geomagnetic regions (bowshock and tail), polar libration orbit distances, and polar lunar distances. To develop useful and accurate models, it is necessary to take into account all perturbations that affect the system.[3,4] These include all planetary gravitational interactions, solar radiation pressure, inhomogeneities in the Earth's gravity field, and atmospheric drag. Analysis using the GSFC program Swingby with high fidelity simulations including all perturbations is then performed.[2] The team analyzed solar sail concepts and incorporated a control method in Swingby via finite maneuvers. Geopotential accelerations and the related solar sail mass to area ratios required to counter these accelerations were studied. Purdue University concentrated on developing the initial conditions for the vertical libration orbits using a dynamical systems approach.

UNIQUE LIBRATION ORBITS

The Restricted 3-Body Problem is a good first approximation of the solar system and for our analysis. Five stable regions, known as Lagrange points, exist in the rotating frame of reference shown in Figure 3. The L_1 and L_2 co-linear libration points are used herein.

Initial Conditions

A set of initial conditions was obtained for the vertical orbits from Purdue University. These conditions are modeled in Swingby and the orbital parameters optimized. This created orbital data sets for each orbit to be used in comparative analysis.[6] Table 1 presents the initial conditions for a small and large vertical orbit about L_1. Using these conditions, simulations are run and maneuvers performed which keep the spacecraft within the proper location. The duration of the Polar Region coverage measured from the subsatellite location for these orbits is then computed over a North and South latitude range of 60 to 90 degrees. With the period of the libration orbit approximately 6 months, the insertion into the orbit should be phased so that the required polar region can be observed more easily for a give season. For example, a transfer and insertion can place the spacecraft at the most northern ecliptic location when the North pole of the Earth is tilted towards the Sun, thereby maximizing the exposure duration. Figure 4, 5, and 6 show the x, y, z views in a solar

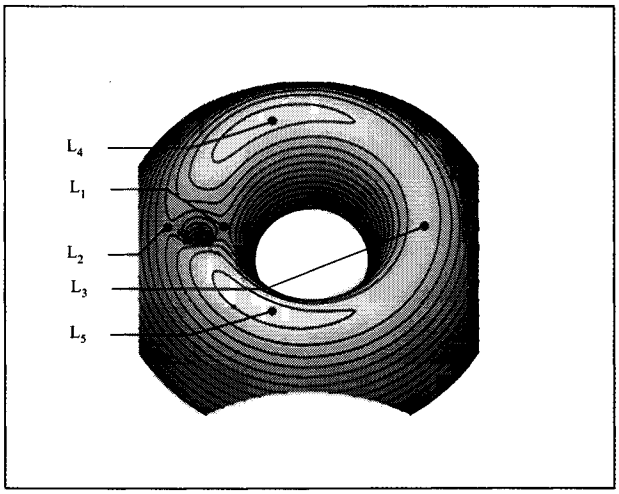


Figure 3. Libration Point Locations

Table 1. Initial Conditions for Vertical Libration Orbits

Small Vertical Orbit Initial Cond.	Large Vertical Orbit Initial Cond.
x = 7.004902317657201e+004 km	x = 1.317772822366962e+006 km
y = 0 km	y = 0 km
z =5 .376155536465669e+005 km	z = 2.688077768232865e+006 km
vx = 0 km/s	vx = 0 km/s
vy = -0.010798682045 km/s	vy = -0.206125010972 km/s
vz = 0 km/s	vz = 0 km/s

rotating coordinate frame of the small vertical orbit while Figure 7, 8, and 9 present the larger vertical orbit views.

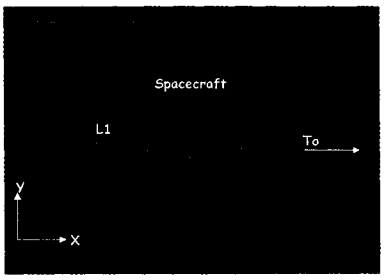


Figure 4. Small Vertical: Looking Down from the North Pole

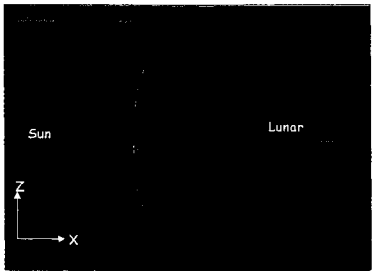


Figure 5. Small Vertical: Looking along the Ecliptic Plane

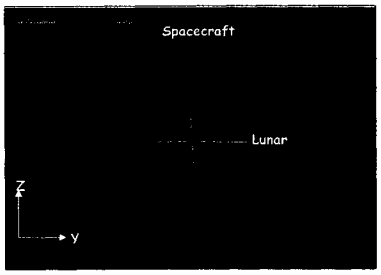


Figure 6. Small Vertical: Looking Back from the Sun

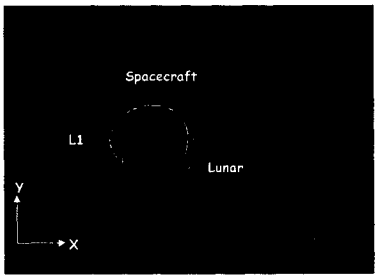


Figure 7. Large Vertical: Looking Down from the North Pole

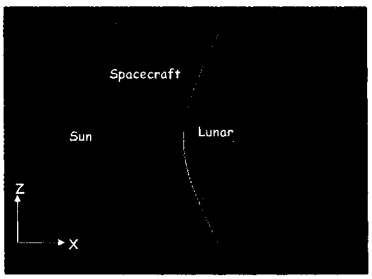


Figure 8 Large Vertical: Looking at the Ecliptic Plane

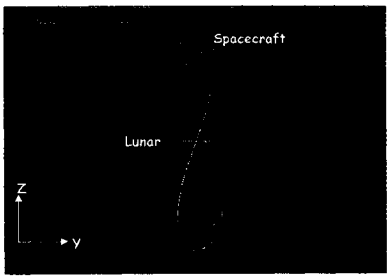


Figure 9. Large Vertical: Looking Back from the Sun

Highly Inclined Lunar and Elliptical Orbits

The next set of analysis considers orbits at a lunar orbit radius distance and an elliptical orbit for contrast to the lissajous orbit. To achieve the inclined lunar orbit, a lunar gravity assist is utilized[5]. This gravity assist along with an insertion maneuver provided an ecliptic inclination of almost 90 degrees. This inclination was chosen to meet the

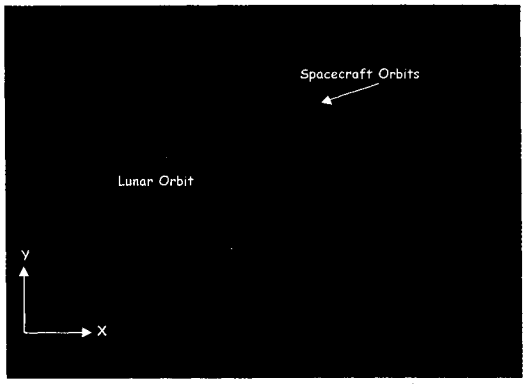


Figure 10. Lunar Orbit: Looking Down From North Pole In A Rotating Frame

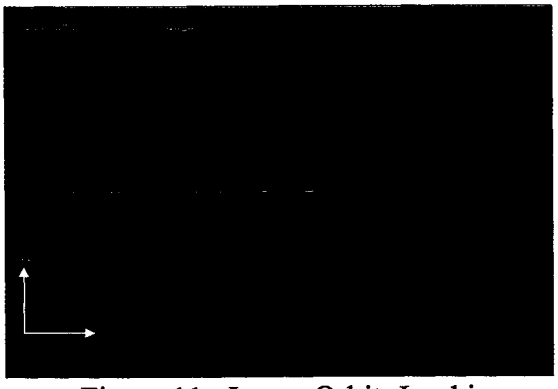


Figure 11. Lunar Orbit: Looking Along Ecliptic Plane In A Rotating Frame

requirement of viewing the Earth Polar Regions. This orbit provides good coverage, but is temporal, albeit slower, as with traditional orbits. The viewing coverage includes both hemispheres and allows a higher science return data rate than the libration orbit. Figures 10 and 11 present the lunar orbit in three views while Figures 12 and 13 present the elliptical orbit. All views are in a solar rotating frame. As with all elliptical orbits, the time at periapsis is limited therefore the orbit apses orientation is selected to place the apoapsis in the northern polar region. For best communication coverage a Molnyia orbit can be used.

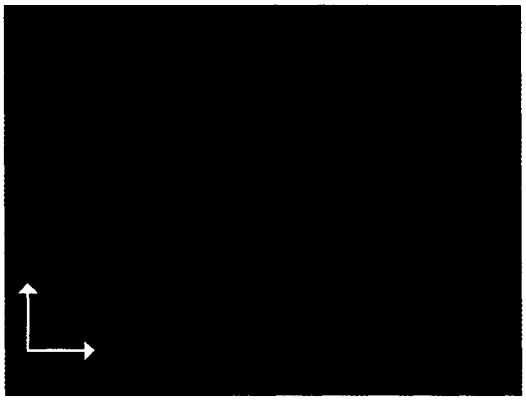


Figure 12. Elliptical Orbit: Looking Down From North Pole In A Rotating Frame

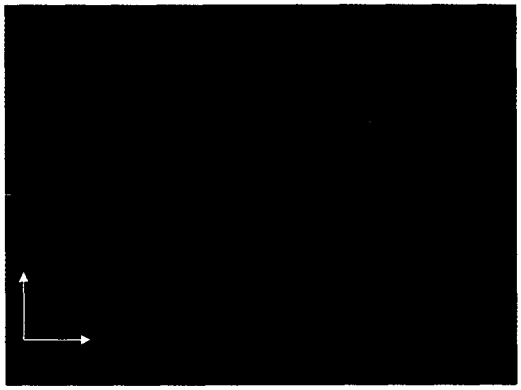


Figure 13. Elliptical Orbit: Looking Along Ecliptic Plane In A Rotating Frame

Phase 1 Results

For the cases studied, parameters of: distance from Earth over a 240 day period, the ΔV required for orbit maintenance, and the duration and percent of the time with views between 60 and 90 degrees North or South equatorial latitudes are computed and are shown in Figures 14 through 17. Using this information it can be seen that the large libration orbit provides the best viewing geometry in terms of duration, but requires the largest distance which impacts the communication and science resolution. The orbit at

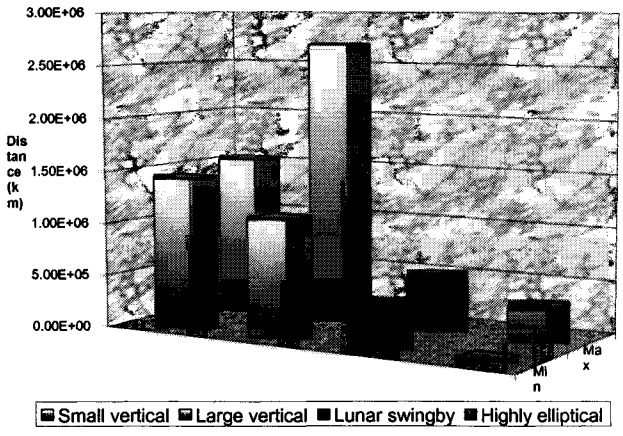


Figure 14. Min and Max distance from Earth over 240 days

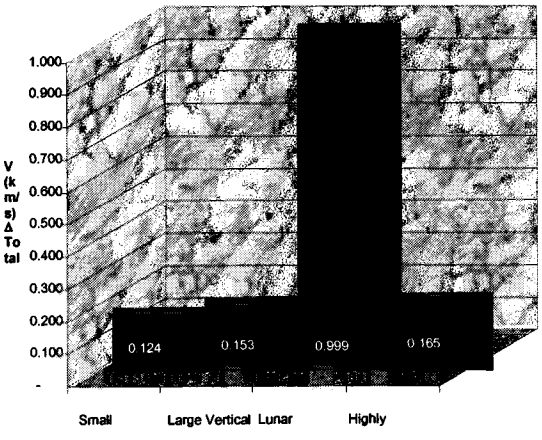


Figure 15. ΔV Requirements for Orbit Maintenance

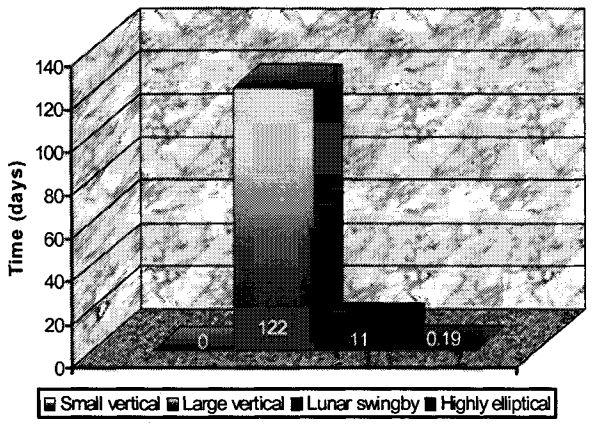


Figure 16. Duration with Views Between 60 and 90 degrees North

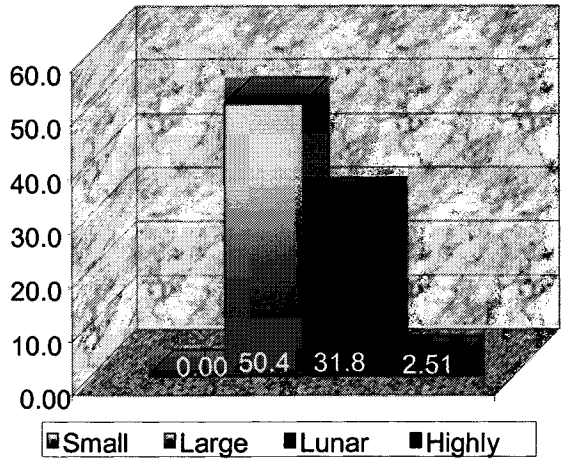


Figure 17. Percent of Time with Views Between 60 and 90 degrees

a lunar distance is the most expensive in fuel cost, requiring timing for a gravity assist as well as a large, approximately 1 km/s, ΔV for the 'insertion' into the mission orbit. The highly elliptical orbit is the worst in terms of temporal polar coverage. The ΔV requirements are all approximately 120 to 170 m/s per year with the exception of the additional lunar orbit insertion maneuver.

POLE SITTERS

The next section deals with the use of polar sitters to obtain continuous viewing.[6] This analysis looked at placements at two different polar distances and the associated ΔV or the accelerations required to maintain this unique non-Keplerian location. The distances selected are 25,000 km and 250,000 km directly over the Earth's North Pole. The initial conditions for these cases are presented in Table 2. The ideal condition would be one in which the orbit is stable and used minimum fuel. The stability duration was chosen to be 28 days (a lunar cycle). A view of these two cases is shown in Figures 18 and 19. The maintenance requirements were based on offsetting the accelerations in the

Table 2. Initial GCI Conditions for Pole Sitters

Close Sitter Orbit Initial Cond.	Far Sitter Orbit Initial Cond.
x = 0 km	x = 0 km
y = 0 km	y = 0 km
z =2.5e+004 km	z = 2.5e+005 km
vx = 0.5 km/s	vx = 0.005 km/s
vy = 0.5 km/s	vy = 0.005 km/s
vz = 0 km/s	vz = 0 km/s

vertical or Z-axis only. This method allowed us to use a maneuver that controlled the perturbations in the z-axis direction which would result in a Earth centered traditional orbit, while only monitoring the motion in the x and y axis directions.

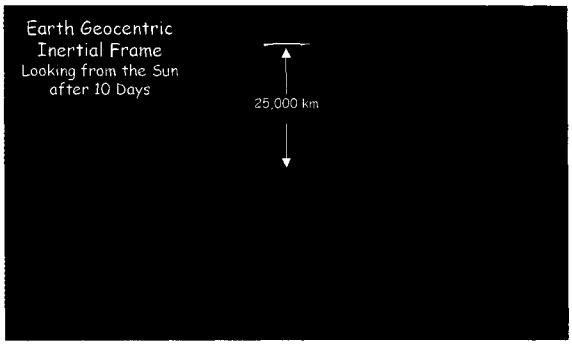


Figure 18. Lower Pole Sitter at 25,000km

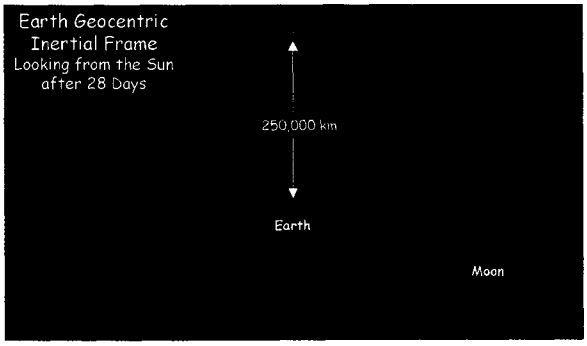


Figure 19. Higher Pole Sitter at 25,000km

Accelerations

The acceleration due to gravity in the z-axis direction can be found using the following two equations.

$$a_{grav,z} = \frac{GM_{\oplus} z}{\left(x^2 + y^2 + z^2\right)^{3/2}} \qquad \bar{a}_z = \frac{GM_{\oplus}}{2\alpha^2 R_{\oplus}^2} \frac{\left[2\lambda^2 - \frac{3}{2}\right]}{\lambda^4}$$

The acceleration in the vertical direction as a function of distance is shown in Figure 20. The acceleration under 50,000 km provides the most significant perturbation. In order to maintain this location, a maintenance strategy is employed using either a propulsion system or solar sail technology. To use solar sails, the acceleration due to the solar radiation pressure needs to cancel the gravitational accelerations. A simple model was constructed for an estimate of the required mass to area ration needed for the appropriate acceleration level.

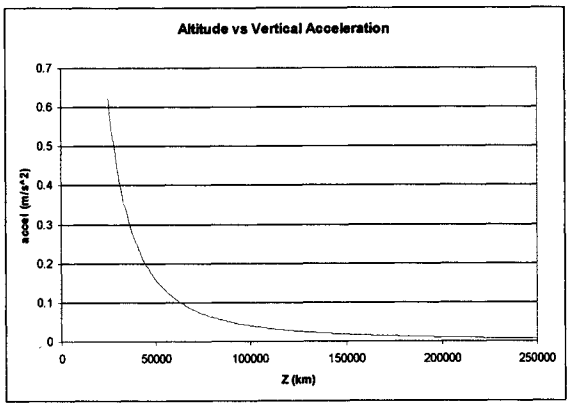


Figure 20 - Vertical Acceleration

Tilting the sail by an angle β causes a component of the solar pressure to cancel the downward pull of gravity. The optimal β is 35.26°. Using this estimate, a formula for the mass to area ratio, $\sigma_{s/c}$, for a solar sail is constructed from

$$\vec{a}_{sol} = 2\eta \frac{S_0}{c}\left(\frac{1}{R_{AU}}\right)^2 \frac{A}{m_{sc}} \cos^2\beta \cdot \vec{e}_n$$

$$\sigma_{sc} = \left[\frac{\lambda^4}{2\lambda^2 - \frac{3}{2}}\right] \underbrace{\left(\frac{\eta\alpha^2}{R^2}\right)\left\{\frac{4R_\oplus^2 S_0}{GM_\oplus c}\cos^2\beta_0 \sin\beta_0\right\}}_{\tilde{f}(\alpha,\eta,R)}$$

Where $\sigma_{sc} = \frac{m_{sc}}{A}$ $\lambda = \frac{z}{\alpha R_\oplus}$ $m_{sc} = m_{sail} + m_{sat}$

and R_o is the radius of the earth, S_0 the solar flux, GM the gravitational constant, β_0 the angle between the solar radiation and sail, c the speed of light, η the sail efficiency, R the distance, σ the mass to area ratio, A the area, m are the relative masses, λ the normalized distance. Calculating the solar sail acceleration required for the cancellation of the gravitational acceleration leads to the data in Figure 21 which shows the ratio for all distances between 25,000km and 250,000km.

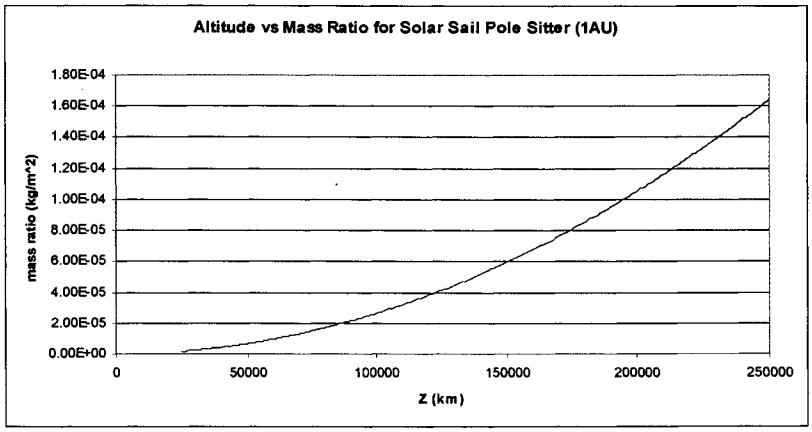


Figure 21. Distance vs. Mass/Area Ratio

Propulsion Maintenance

A ΔV estimate is computed using the Swingby program with a continuous ΔV in the z-axis direction to cancel the vertical accelerations similar to the solar sail case. The propulsion system requirements are as follows; for the 250,000km distance a ΔV of 0.0064 m/s per second duration leading to approximately 0.5 km/s per day. At the 25,000km distance, a ΔV of 0.596 m/s per second duration is required leading to over 50km/s per day. This ΔV needs to be computed in terms of a fuel mass fraction as an impact statement into the selection of the payload mass and launch capability. The subsatellite pattern shown in Figures 22 and 23 for the 25,000km and 250,000km case, respectively, is presented in a rotating Earth frame.

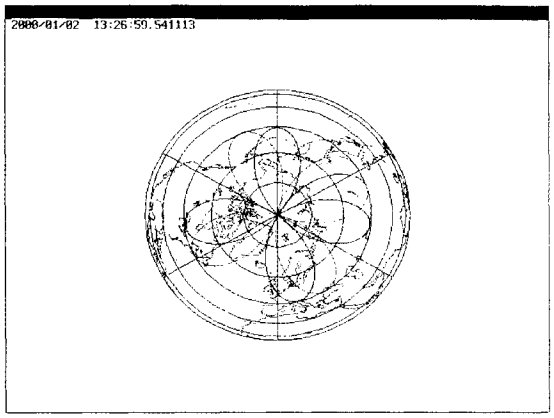


Figure 22 - Subsatellite Pattern for 25,000km Distance

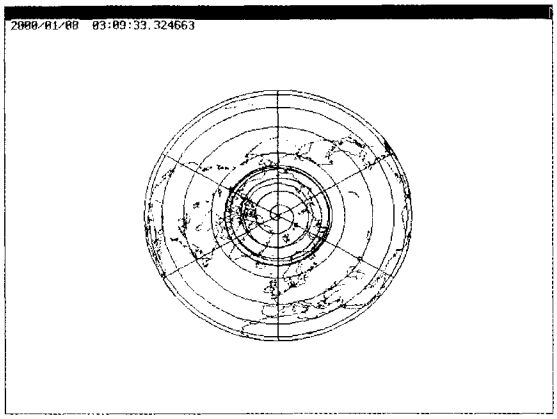


Figure 23 - Subsatellite Pattern for 250,000km Distance

SUMMARY

The feasibility of attaining and maintaining unique non-Keplerian orbit vantage locations in the Earth/Moon environment in order to obtain continuous scientific measurements has been demonstrated. This investigation demonstrated vertical libration orbits, polar orbits at a lunar distance, and stationary polar sitters which require a maintenance strategy. A dynamical systems approach was utilized to initialize the vertical libration orbits. The maintenance of the polar sitters is demonstrated using both a long duration ΔV and solar sail concept. It was shown that maintenance fuel for spacecraft stationed at these locations can be reduced thus enabling technology for innovative missions such as solar sails, electric propulsion, and autonomous control. The results from this analysis provide an innovative set of 'orbits' which meet demanding scientific objectives.

CONCLUSION

An example of continuous observer missions which include Earth pole-sitters and unique vertical libration orbits that meet Sun-Earth Connection and Earth Science Vision roadmaps has been demonstrated. Continuous polar coverage requirements can be met using a non-Keplerian approach to maximize the viewing requirement. These results provide NASA projects with an alternative to standard mission design and the related impacts. Results provide NASA projects with new and unique non-Keplerian 'orbits' and vantage locations as alternative to standard mission design.

REFERENCES

1. Guzman, J., Cooley, S., Howell, K, Folta D, "Trajectory Design Strategies that Incorporate Invariant Manifolds and Swingby", AAS/GSFC International Symposium on Space Flight Dynamics, Greenbelt, Maryland, May 1998. AAS Paper 98-349.

2. Carrico, J. et al, "An Interactive Tool for Design and Support of Lunar, Gravity Assist, and Libration Point Trajectories", AIAA/AHS/ASEE, AeroSpace Design Conference, Irvine, Ca, February 1993 Paper 96-3642.

3. Howell, K.C., and Campbell, E. T., "Families of Periodic Orbits that Bifurcate from Halo Families in the Circular Restricted Three-Body Problem," AAS/AIAA Spaceflight Mechanics Conference, Breckenridge, Colorado, February 1999.

4. Howell, K.C., and Barden, B.T., "Trajectory Design and Stationkeeping for Multiple Spacecraft in a Fixed Configuration Near the Sun-Earth L_1 Point," 50^{th} International Astronautical Federation Congress, Amsterdam, The Netherlands, October 1999

5. Folta, D.C.,"Multiple Lunar Flyby Targeting for the WIND Mission", AAS/AIAA Space Flight Mechanics Meeting, Austin, TX., 1996. AIAA

6. Young, C., and Ross, A., "Unique Non-Keplerian Orbit Vantage Locations", Final Presentation to the NASA Academy, August 2000

COLLISION PROBABILITY ANALYSES FOR EARTH-ORBITING SATELLITES

Ken Chan

Aerospace Corporation
15049 Conference Center Drive
Chantilly, VA 20151

ABSTRACT

This paper is first concerned with a detailed calculation of the collision probability between the primary spacecraft and a single secondary orbiting object. The analysis is then extended to the case in which the secondary objects comprise a class having a specific altitude and inclination, the latter parameter not necessarily the same as that of the primary spacecraft. All the secondary objects in this class have ascending nodes uniformly distributed around the equator, and have phase angles uniformly distributed within an orbital plane. For a pairwise encounter, the starting point need not be a 3-dimensional Gaussian distribution but is rather a 2-dimensional Gaussian distribution defined in the plane perpendicular to the relative velocity vector at the point of closest approach. The ellipses of constant probability density function (pdf) are replaced by circles of constant pdf having the same area. This makes the Gaussian distribution isotropic in the plane of interest. The rationale for this is that the relative position vector at closest approach is random in direction in that plane, when one considers any member of the secondary objects encountering the primary spacecraft. The consequence is that the 2-dimensional Gaussian distribution is replaced by the 1-dimensional Rician distribution which involves the modified Bessel function of the first kind of zero order.

Expansion of this function yields a series which is termwise dominated by the exponential series. Moreover, truncation of the Bessel series after any finite number of terms yields an error which is bounded by the product of a power and an exponential series. Thus, the probability of collision for a pairwise encounter may be expressed in an analytical expression with error bounds which can be easily computed. Typical examples show that the error bound for truncation after the first term appears on the third or fourth significant digit. Next, by considering encounters between the primary spacecraft and all the secondary orbiting objects in the same altitude and inclination class, one introduces an integral whose integrand depends on the orbital characteristics of both the primary and secondary orbiting objects. This integrand becomes singular at the inclination latitudes of the primary and the secondary objects. Despite this singularity, it is proved that the integral exists. Detailed expressions for this integral are given for the three cases in which the inclination latitude of the primary spacecraft is less than, equal to, or greater than the inclination latitude of the secondary orbiting objects.

An example is worked out for the case of a primary spacecraft with semimajor axis 7,000 km and inclination 86.4 deg encountering presently orbiting objects within ± 160 km and belonging to the inclination classes of 33, 53, 63, 74, 81, 90, 98, and 106 deg that are tracked by NORAD. It is found that the number of collisions per year is approximately 2E-4.

INTRODUCTION

There is great interest in the problem of a primary spacecraft colliding with secondary orbiting space objects that are either active spacecraft or inactive spacecraft or their remnants from previous missions. Most previous analyses on collision probability have been based on the Poisson distribution. More recently,[1,2,3,4,5,6] the model has utilized the Gaussian distribution because the primary spacecraft and secondary space objects can be tracked and their orbits are determined to within the errors associated with the corresponding error ellipsoids, as illustrated in Figure 1. Thus, the probability for a pair to collide would be formulated in terms of a combined covariance matrix leading to a general Gaussian distribution in three dimensions.

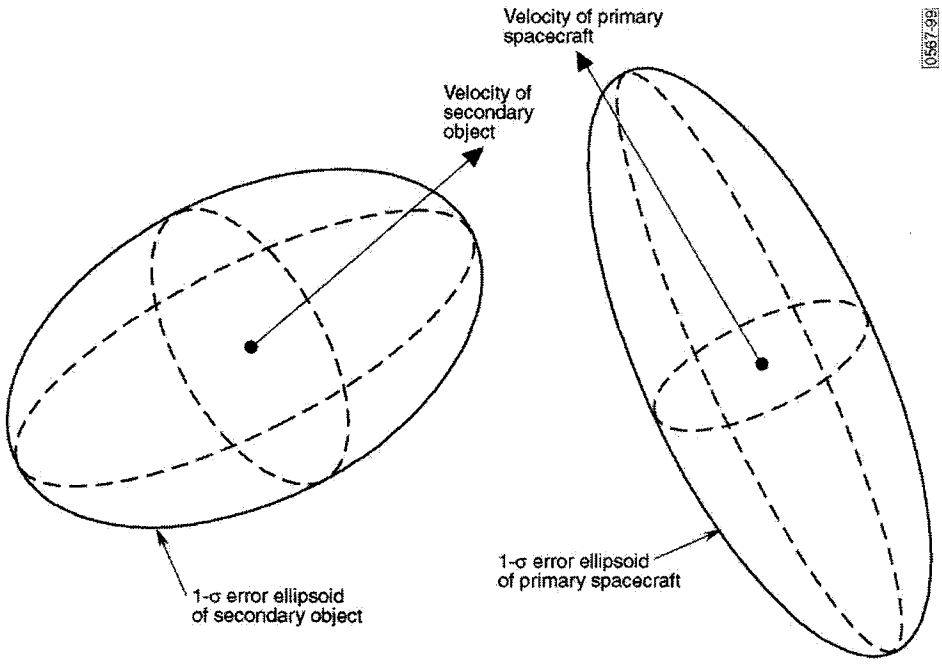


Figure 1. Error ellipsoids of primary spacecraft and secondary space object.

At a conjunction involving a fly-by (which holds for most encounters except possibly for some infrequent ones involving geosynchronous satellites), an encounter coordinate system is defined as follows: Let $\bar{v}_{ps}$ denote the relative velocity of the primary spacecraft with respect to the secondary object. In the encounter coordinate system, let the y-axis be defined along this vector. Then, the (x, z)-plane is normal to this vector. At the instant of closest approach, let the x-axis be defined such that the primary spacecraft is at $(x_e, 0, 0)$. Thus, we have defined the encounter coordinate system whose unit basis vectors are $\hat{x}$, $\hat{y}$, and $\hat{z}$. This is illustrated in Figure 2. If the two covariance matrices are obtained from independent or (less stringent) uncorrelated measurements, then the combined covariance matrix is given by the sum of the individual covariance matrices, all expressed in the same coordinate system.

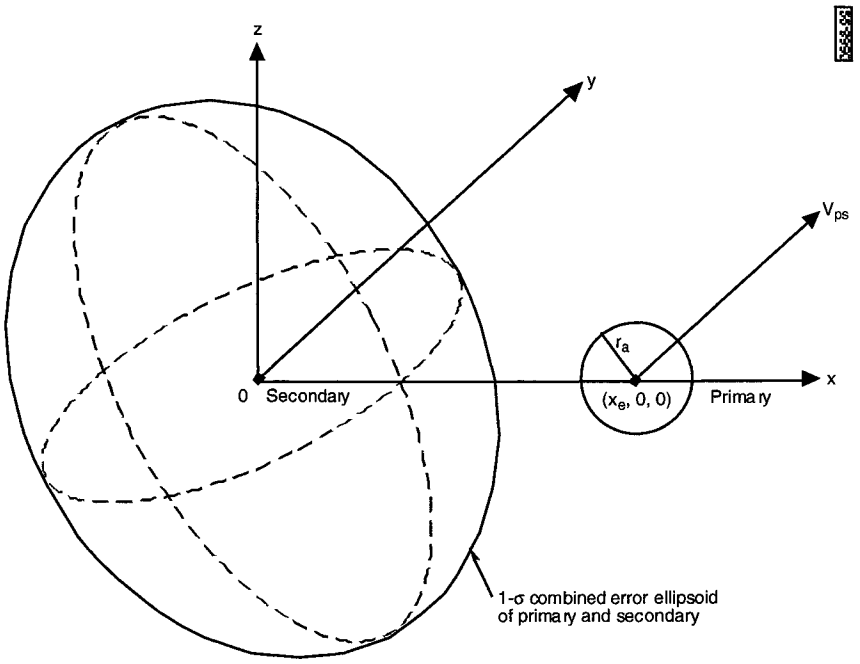


Figure 2. Combined error ellipsoid and collision cross-section in encounter coordinate system.

Let C be the combined covariance matrix in the encounter system. Then, the probability density function (pdf) of the uncertainty of the separation between the primary and secondary is

$$f_3(x, y, z) = \frac{1}{\sqrt{(2\pi)^3 |C|}} e^{-\frac{1}{2}\bar{r}^* C^{-1} \bar{r}} \quad . \tag{1}$$

Let r_p denote the radius of the sphere, which just circumscribes the primary spacecraft, the center of the sphere being located at the center of mass. Similarly, let r_s denote the radius of the circumscribing sphere of the secondary object. Let us now assume that both objects are spherical, the reasons being that we simplify the problem and also we do not account for their orientation. Then, if the secondary comes within a sphere with radius $r_a = r_p + r_s$ centered at the primary, there will be a collision between the two. Therefore, the probability of collision is

$$P = \iiint_V f_3(x, y, z)\, dx\, dy\, dz \tag{2}$$

where V is the volume swept by the sphere of radius r_a centered at the primary.

However, a little consideration reveals that we do not have to deal with this 3-dimensional integral. The reason is that, at the instant of closest approach, the primary crosses the (x, z)-plane and its relative position uncertainty is described by the bivariate Gaussian pdf

$$f_2(x,z) = \frac{1}{2\pi\sigma_x\sigma_z\sqrt{1-\rho_{xz}^2}} e^{-\left[\left(\frac{x}{\sigma_x}\right)^2 - 2\rho_{xz}\left(\frac{x}{\sigma_x}\right)\left(\frac{z}{\sigma_z}\right) + \left(\frac{z}{\sigma_z}\right)^2\right] \Big/ 2\left(1-\rho_{xz}^2\right)} . \quad (3)$$

Hence, the probability of collision is given the 2-dimensional integral

$$P = \iint_A f_2(x,z)\, dx\, dz \quad (4)$$

where A is the area of a circle with radius r_a centered at the primary.

For the simple case in which the collision cross-section radius r_a is much less than the distance of closest approach x_e, the integrand (3) may be replaced by the value at the center (x_e, 0) of the collision cross-sectional area A. Then, the collision probability is approximately given by

$$P \cong \frac{r_a^2}{2\sigma_x\sigma_z\sqrt{1-\rho_{xz}^2}} e^{-x_e^2/2\sigma_x^2(1-\rho_{xz}^2)} . \quad (5)$$

This expression may be used in some cases and should be used with extreme caution since it can yield a result an order (or several orders) of magnitude different from the exact value. Moreover, if $\rho_{xz} = 0$, σ_x = 100 m, σ_z = 10 m, x_e = 100 m, r_a = 60 m, we obtain the absurd answer of $P \cong 1.09$. For more general cases, we need to develop expressions that are not dependent on such simplifying assumptions. Thus, it is imperative to find an alternative approach which makes evaluating the integral (4) more tractable. Even if it involves an approximation, it must be justifiable and sufficiently accurate with small error bounds which, preferably, can be expressed analytically.

ENCOUNTER WITH A SINGLE ONCOMING SATELLITE

Approximation by Isotropic Gaussian Distribution

In seeking to approximate the bivariate Gaussian pdf (3) by an isotropic Gaussian pdf, we do not simply and merely set $\sigma_x = \sigma_z$ and $\rho_{xz} = 0$ in equation (3). This would be too naïve and unsatisfying. In order to accomplish the objective meaningfully, we rotate the coordinate system (x, z) to a new coordinate system (x′, z′) such that the off-diagonal terms no longer appear in the pdf (3). That is, in this new coordinate system, the major and minor axes of the ellipse associated with the covariance matrix C will be along the coordinate axes, which are the principal directions. For what we need, it turns out to be unnecessary to perform the actual coordinate rotation to obtain the principal directions. Rather, we proceed as follows: The covariance matrix C is given by

$$C = \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix} . \quad (6)$$

Rotation of coordinates to principal directions yields the covariance matrix C′ given by

$$C' = \begin{bmatrix} \sigma_{x'}^2 & 0 \\ 0 & \sigma_{z'}^2 \end{bmatrix} . \quad (7)$$

Since the determinant remains invariant under a rotation, therefore

$$|C'| = |C| \tag{8}$$

which yields

$$\sigma_{x'}\sigma_{z'} = \sqrt{1-\rho_{xz}^2}\,\sigma_x\sigma_z \ . \tag{9}$$

Moreover, since the trace also remains invariant, therefore

$$\sigma_{x'}^2 + \sigma_{z'}^2 = \sigma_x^2 + \sigma_z^2 \ . \tag{10}$$

In the (x', z') system, the pdf (3) becomes

$$f_e(x', z') = \frac{1}{2\pi\,\sigma_{x'}\sigma_{z'}} e^{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_{x'}}\right)^2 + \left(\frac{z'}{\sigma_{z'}}\right)^2\right]} \tag{11}$$

and the collision probability (4) is given by

$$P_e = \frac{1}{2\pi\,\sigma_{x'}\sigma_{z'}} \iint_A e^{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_{x'}}\right)^2 + \left(\frac{z'}{\sigma_{z'}}\right)^2\right]} dx'dz' \ . \tag{12}$$

We next wish to replace the ellipses of constant pdf

$$\frac{x'^2}{\sigma_{x'}^2} + \frac{z'^2}{\sigma_{z'}^2} = \alpha^2 \tag{13}$$

with circles of constant pdf having the same area $\pi\sigma_{x'}\sigma_{z'}\alpha^2$

$$\frac{x'^2}{\sigma^2} + \frac{z'^2}{\sigma^2} = \alpha^2 \tag{14}$$

where α is a constant which labels the circle and its corresponding ellipse. Since the area of the circle is $\pi\sigma^2\alpha^2$, therefore

$$\sigma^2 = \sigma_{x'}\sigma_{z'} \ . \tag{15}$$

The pdf (11) is approximated by the following 2-dimensional isotropic Gaussian pdf

$$f_c(x', z') = \frac{1}{2\pi\,\sigma^2} e^{-(x'^2 + z'^2)/2\sigma^2} \tag{16}$$

and the collision probability corresponding to (12) becomes

$$P_c = \frac{1}{2\pi\sigma^2}\iint_A e^{-\left(x'^2+z'^2\right)/2\sigma^2}\,dx'\,dz' \quad . \tag{17}$$

It is easy to show that the integral of the pdf (11) within the ellipse α is exactly the same as the integral of the pdf (16) within the circle α. We shall refer to this as replacing the ellipses by equivalent circles.

However, the integral of the pdf (11) within the circular cross-section A is not the same as the integral of the pdf (16) within the circular cross-section A. It is proved[10] that the error introduced in replacing the integral (12) by the integral (17) is

$$e^{-\lambda\sigma_{x'}} < \frac{P_e}{P_c} < e^{\lambda\sigma_{z'}} \tag{18}$$

$$\lambda = \frac{1}{2\sigma^4}(\sigma_{x'} - \sigma_{z'})(x_e + r_a)^2 \quad \text{where} \quad \sigma_{x'} \geq \sigma_{z'} \,. \tag{19}$$

From equations (18) and (19), it is seen that if the pdf contours in the encounter plane are almost circular, then the approximation of P_e by P_c is extremely accurate. The question that remains is: How good is this approximation? We note that σ is of the order of several hundred meters or one km, and r_a is of the order of a few meters. Obviously, for specified values of $\sigma_{x'}$ and $\sigma_{z'}$, it is heavily dependent on the ratio of x_e to σ since r_a is generally small compared to x_e. It is easy to verify that the approximation is also extremely accurate provided that x_e is small compared to $\sigma_{x'}$ and $\sigma_{z'}$.

The following table illustrates the upper and lower bounds for this collision integral ratio (P_e / P_c) for typical values of $\sigma_{x'}$, $\sigma_{z'}$, x_e, and r_a, all expressed in kilometers.

Table 1. Bounds of the Collision Integral Ratio (P_e / P_c) for Combined Covariance of Two Identical Ellipsoids with Aspect Ratios 1 : 4 : 1

	Case 1	Case 2	Case 3	Case 4
Sigma $\sigma_{x'}$	3.0	3.0	3.0	3.0
Sigma $\sigma_{z'}$	1.5	1.0	1.5	1.0
Miss Distance x_e	0.1	0.1	1.0	1.0
Radius r_a	0.01	0.01	0.01	0.01
Lower Bound	0.9987	0.996	0.893	0.71
Upper Bound	1.0007	1.001	1.058	1.12

To illustrate the approximation of P_e by P_c, consider two satellites of equal velocities approaching each other at an angle θ. For simplicity, assume that each covariance matrix has only diagonal terms, with crosstrack variance a^2, intrack variance b^2, and radial variance c^2, all units of distance being in km for specificity. Then, it may be shown that the combined covariance matrix C in the encounter system is given by

$$C = T^* \begin{bmatrix} (a^2(1+\cos^2\theta)+b^2\sin^2\theta) & \frac{1}{2}(a^2-b^2)\sin 2\theta & 0 \\ \frac{1}{2}(a^2-b^2)\sin 2\theta & (a^2\sin^2\theta+b^2(1+\cos^2\theta)) & 0 \\ 0 & 0 & 2c^2 \end{bmatrix} T \tag{20}$$

where

$$T = \begin{bmatrix} \sin\theta/2 & -\cos\theta/2 & 0 \\ \cos\theta/2 & \sin\theta/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{21}$$

(Note that because the two velocities are equal in magnitude, the case of $\theta = 0$ is omitted from consideration since there can be no direction for the zero relative velocity to uniquely define the y-axis of the encounter system. However, this pathological case cannot happen in a realistic situation for the two velocity magnitudes cannot be mathematically equal. Any difference between these two, no matter how small, will yield a relative velocity either parallel or anti-parallel to their direction.) From equations (20) and (21), it follows that we have

$$C|_{\theta=\pi} = \begin{bmatrix} 2a^2 & 0 & 0 \\ 0 & 2b^2 & 0 \\ 0 & 0 & 2c^2 \end{bmatrix} \tag{22}$$

$$C|_{\theta=\pi/2} = \begin{bmatrix} (a^2+b^2) & 0 & 0 \\ 0 & (a^2+b^2) & 0 \\ 0 & 0 & 2c^2 \end{bmatrix} . \tag{23}$$

If we now take the simple average of these two, we obtain

$$C_{ave} = \begin{bmatrix} (3a^2+b^2)/2 & 0 & 0 \\ 0 & (a^2+3b^2)/2 & 0 \\ 0 & 0 & 2c^2 \end{bmatrix} . \tag{24}$$

First, consider the case of the two velocities at right angles ($\theta = \pi/2$). Then, the relative velocity is at 45^0 to these directions. For simplicity, assume that a = c = 1 and b =$\sqrt{17}$. These are comparatively elongated ellipsoids with aspect ratios of approximately 1 : 4 : 1. Then, by equation (23), the combined error ellipsoid has $\sigma_{x'} = 3\sqrt{2}$ km and $\sigma_{z'} = \sqrt{2}$ km, which yield pdf elliptical contours with 3 : 1 aspect ratio in the encounter plane. At the conjunction, suppose the miss distance is 1 km. Thus, from Case 4 in the above table, the approximation of P_e by P_c is accurate to within 29% (or less). However, if the miss distance is 0.1 km, then Case 2 reveals that this approximation is within 0.4%. On the other hand, if the two satellites approach each other directly ($\theta = \pi$) but offset by any (non-zero) miss distance at conjunction, then by equation (22) the combined error ellipsoid has $\sigma_{x'} = \sqrt{2}$ km and $\sigma_{z'} = \sqrt{2}$ km, which yield circular pdf contours in the encounter plane. Thus, there is no need to approximate P_e by P_c. That is, the approximation error is zero. For an intermediate encounter angle, the accuracy of approximation is between the two extreme values corresponding to broadside encounter and direct approach, and may be obtained from equations (20) and (21). Of special interest is the simple average case since it gives an approximation of the average situation. For this, we observe from equation (24) that $\sigma_{x'} = \sqrt{10}$ km and $\sigma_{z'} = \sqrt{2}$ km, which yield pdf elliptical contours with 2.2 : 1 aspect ratio in the encounter plane. Thus, from Case 3 in the above table, the approximation of P_e by P_c is accurate to within 11% for miss distance of 1 km. However, if the miss distance is 0.1 km, then Case1 reveals that this approximation is within 0.2%. Note that all these results are for comparatively elongated ellipsoids with aspect ratios of approximately 1 : 4 : 1.

Approximation by Analytical Expression

Another benefit from the previous approximation is that the 2-dimensional integral (21) can be transformed into the following 1-dimensional integral involving the Rician pdf[7]

$$P_c = \int_0^{r_a} \frac{r}{\sigma^2} e^{-\left(r^2+x_e^2\right)/2\sigma^2} I_0\left(\frac{r\,x_e}{\sigma^2}\right) dr \tag{25}$$

where $I_0(\,)$ denotes the modified Bessel function of the first kind of order zero. This pdf arises frequently in signal processing, especially in the detection of signals in the presence of noise.[8] When x_e is equal to zero, it is precisely the Rayleigh pdf. If x_e is large, it tends to be symmetrical in the interval of several units of σ containing the maximum point at $r \cong x_e$ and resembles the Gaussian pdf in this interval (keeping in mind that it is defined from 0 to ∞ whereas the latter is defined from $-\infty$ to $+\infty$). Approximations of the Rician pdf have been made usually under some simplifying assumptions in order to obtain analytical expressions.[9] Alternatively, the integral (25) has also been available numerically through the Marcum Q-function which is the complementary integral evaluated from r_a to ∞. However, the general results obtained by the author[10] are new. It is proved that this collision integral may be expressed analytically as

$$P_c = e^{-v/2}(1 - e^{-u/2}) + S_1 \tag{26}$$

where S_1 satisfies the inequality

$$S_1 < \frac{1}{8} u^2 v\, e^{-v/2} e^{uv/4} \quad . \tag{27}$$

The variables u and v are defined by

$$u = \left(\frac{r_a}{\sigma} \right)^2 > 0 \tag{28}$$

$$v = \left(\frac{x_e}{\sigma} \right)^2 > 0 \ . \tag{29}$$

We may refer to S_1 as the truncation error. From the above equations, it is seen that this term depends on the ratios of x_e and r_a to σ which are essentially u and v. Again, since r_a is of the order of meters, u is extremely small. It is easy to show that whether x_e is small or large compared to σ, the upper bound of the truncation error is also small. Thus, in general, the approximation of P_c by retaining only the explicit term in equation (26) is extremely accurate.

The following table illustrates the bound for this truncation error ratio (S_1 / P_c) for typical values of σ, x_e, and r_a, all expressed in kilometers

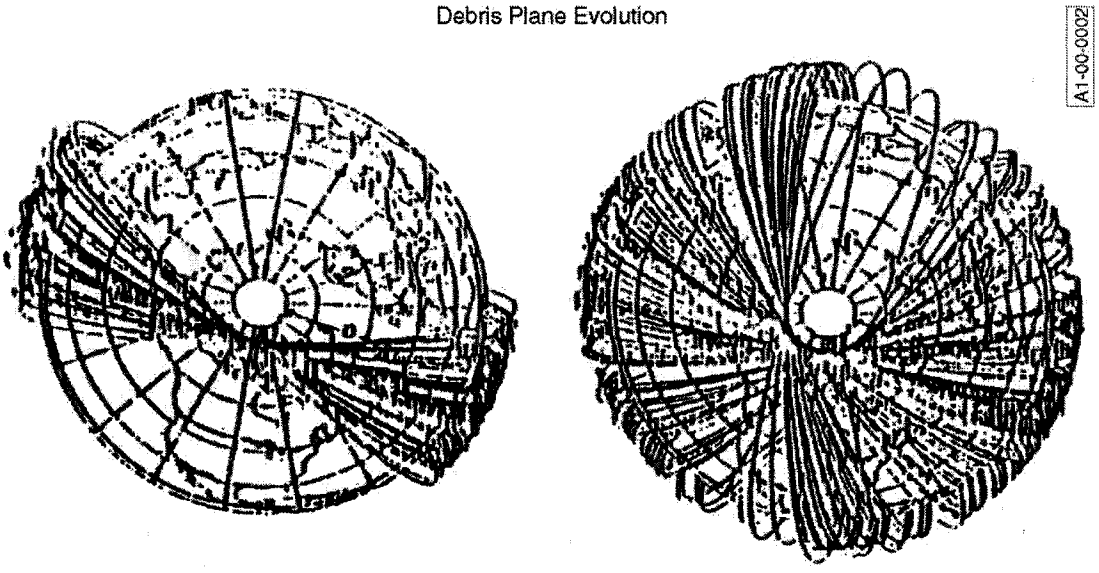
Table 2. Upper Bounds of the Truncation Error Ratio (S_1 / P_c)

	Case 1	Case 2	Case 3	Case 4
Conjunction	Very Close	Close	Close	Far
Sigma σ	0.05	2.0	1.732	2.0
Miss Distance x_e	0.01	1.5	1.0	10.0
Radius r_a	0.005	0.01	0.01	0.01
Upper Bound	E-4	E-6	E-6	E-4

The first case is, by almost any standard, an extremely close conjunction of two satellites with highly accurate orbit determination. The second and third cases are considered close encounters by most acceptable standards to warrant considering a collision avoidance maneuver. The fourth case illustrates an encounter that would normally not cause any concern. Thus, for a wide range of encounter scenarios, it is seen that the error ratio (S_1 / P_c) incurred is of the order of E-4 or less. This means that we may justifiably set S_1 to zero and retain only the explicit term in the expression for the collision integral P_c in equation (26) and still have sufficient precision for our computational purposes. Thus, we have a fairly accurate analytical expression of P_c in terms of σ, x_e, and r_a. Finally, the third case in Table 2 is derived from Case 4 in Table 1. From these two cases, we note that the predominant error lies in the approximation of P_e by P_c and not in the analytical representation of P_c. However, the upper and lower bounds given in equation (18) are for the ratio (P_e / P_c) without regard to the angular position of the point of closest approach in the (x′, z′) plane. It is possible to obtain an expression for this ratio (as a function of the angular position) accurate to one significant figure (1% to 10%) or better for an extremely wide range of the collision parameter ratios $1 \le (\sigma_{x'}/\sigma_{z'}) \le 10$, $10^{-3} \le (x_e/\sigma_{z'}) \le 5$, and $10^{-3} \le (r_a/\sigma_{z'}) \le 1$.

ENCOUNTER WITH MANY ONCOMING SATELLITES

In the previous section, the collision probability was derived for the case of the primary satellite encountering a single secondary satellite. In this section, the analysis is extended to the case of collision when there are many oncoming satellites. It is assumed that the oncoming satellites have circular orbits (with radius R and velocity magnitude V) and can be grouped into classes, each class being characterized by the orbital inclination and orbital altitude. It is further assumed that, in a given orbital plane, the phase is statistically uniform; and the ascending node is also statistically uniform in the equatorial plane. (This is already evident in the dispersion of the debris fragments from the Ariane breakup as shown in Figure 3.)



Dispersion of debris fragments from Nov. 86 Ariane Flight 18 upper stage breakup. Each line indicates only the orbital track of a single small fragment—not a solid band of debris.

Figure 3. Dispersion of debris fragments from Ariane breakup.

Consider satellites of a given class. Suppose the satellites of this class at latitude θ are uniformly distributed over a layer of thickness h. Let $\bar{V}_{rel}$ denote the average velocity of the oncoming secondary satellites relative to the primary satellite, defined within the region of interest. In order to derive the relevant equations for the encounters, we have to consider direct and retrograde orbits of the two objects, whether they are ascending or descending (with respect to the equator), and whether they are engaged in an approaching or a trailing collision. Thus, there are eight cases to consider and the configurations can be very perplexing because we have to account for acute or obtuse angles. For simplicity of illustration, let ψ_1 denote the acute angle between the orbit plane of the primary satellite and the meridian plane at latitude θ. Similarly, let ψ_2 denote the corresponding acute angle for the class of oncoming secondary satellites. This is illustrated in Figure 4.

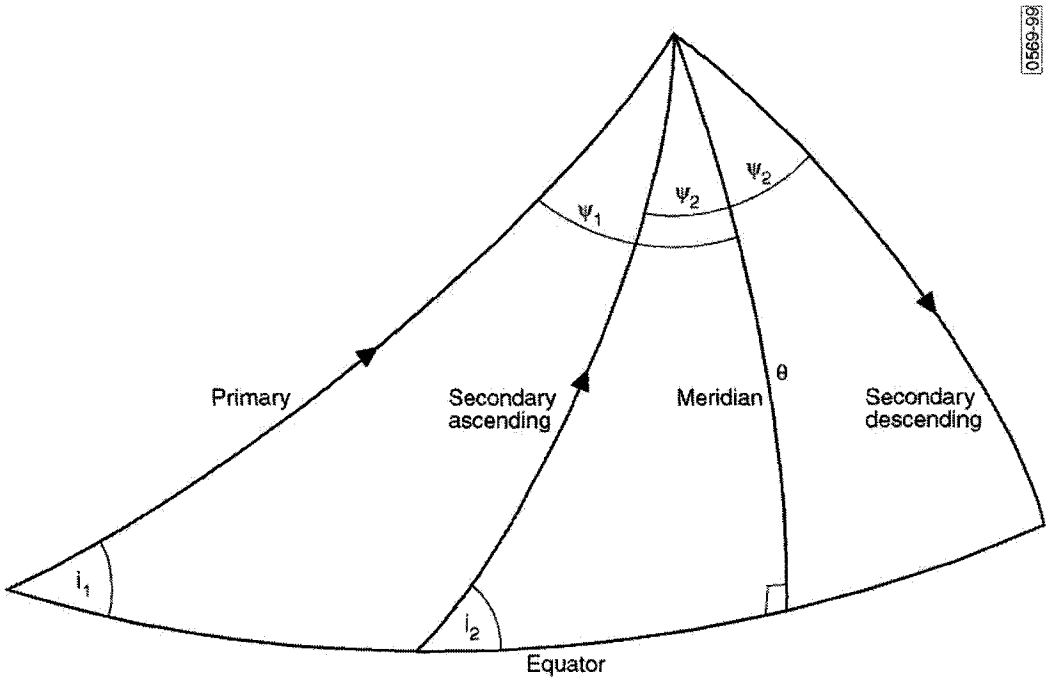


Figure 4. Orbital planes of primary spacecraft and secondary object.

Then, from spherical trigonometry, we have

$$\cos i_1 = \cos\theta \sin\psi_1 \tag{30}$$

$$\cos i_2 = \cos\theta \sin\psi_2 \tag{31}$$

so that

$$\psi_1 = \sin^{-1}\left(\frac{\cos i_1}{\cos\theta}\right) \quad \text{where} \quad -\pi/2 \le \psi_1 \le \pi/2 \tag{32}$$

$$\psi_2 = \sin^{-1}\left(\frac{\cos i_2}{\cos\theta}\right) \quad \text{where} \quad -\pi/2 \le \psi_2 \le \pi/2 \ . \tag{33}$$

There are two cases to consider:

Approaching collision: There are two subcases: (a) both orbits are direct or both orbits are retrograde, and one satellite is in ascending motion (relative to the equator) and the other satellite is descending; (b) one orbit is direct and the other is retrograde, one is ascending and the other is descending. Let us assume that the magnitudes of the two velocity vectors are the same since the satellites are essentially at the same altitude. Let ψ_A and $V_{rel,A}$ denote the angle and magnitude of the relative velocity between these two velocity vectors. Then, we may show that, in general, ψ_A and $V_{rel,A}$ are respectively given by

$$\psi_A = \pi - (\psi_1 + \psi_2) \tag{34}$$

$$V_{rel,A} = 2V\cos\left(\frac{\psi_1 + \psi_2}{2}\right) . \tag{35}$$

Trailing collision: There are two subcases: (a) both orbits are direct or both orbits are retrograde, and both are ascending or both are descending; (b) one orbit is direct and the other is retrograde, both are ascending or both are descending. Let ψ_T and $V_{rel,T}$ denote the angle and the magnitude of the relative velocity between these two velocity vectors. Then, we may show that, in general $\psi_{rel,T}$ and $V_{rel,T}$ are respectively given by

$$\psi_T = |\psi_1 - \psi_2| \tag{36}$$

$$V_{rel,T} = 2V\sin\left|\frac{\psi_1 - \psi_2}{2}\right| . \tag{37}$$

Then, by using equation (26) for P_c without the truncation error S_1, we may obtain the probability of collision of the two objects. The error of this approximation is given by equation (27) and is accurate to 4 or 5 significant figures for typical examples. Next, since the relative position vector at closest approach for any one encounter is random in direction in the plane perpendicular to the relative velocity vector, the approximation of P_e by P_c averaged over all directions is even more accurate that indicated by equation (18). We may show after much work that the total number of collisions per orbit of the primary satellite is given by[1]

$$N_{total,orbit} = K\,J\,I \tag{38}$$

where, for convenience, we define

$$K = \frac{4\sigma^2 n_{total}}{\pi R h}\left[1 - e^{-\frac{1}{2}\left(\frac{r_a}{\sigma}\right)^2}\right] \tag{39}$$

$$J = \left[1 - e^{-\frac{1}{2}\left(\frac{r_e''}{\sigma}\right)^2}\right] \tag{40}$$

$$I = \int_0^{\theta'} \frac{\left[\cos\left(\frac{\psi_1 + \psi_2}{2}\right) + \sin\left|\frac{\psi_1 - \psi_2}{2}\right|\right]\cos\theta}{\sqrt{\left[(\sin^2 i_1 - \sin^2\theta)(\sin^2 i_2 - \sin^2\theta)\right]}} d\theta . \tag{41}$$

It may be shown that equation (41) is applicable for <u>all</u> combinations of direct and/or retrograde orbits for the primary and secondary satellites. The cosine term corresponds to <u>approaching</u> collisions; and the sine term corresponds to <u>trailing</u> collisions.

In the above equations, n_{total} denotes the number of satellites in a given class and r_e'' is the maximum flux radius taken to be h/2. Thus, the K term contains the collision cross-section, the number of satellites of a given class, orbit determination uncertainties, orbital radius and the thickness of the altitude band for that class. The J term contains the orbit determination uncertainties and the maximum flux radius.

Finally, the I term contains only the effects of the orbital characteristics of both the primary satellite and the secondary orbiting object. It is noted that the integrand becomes singular at the inclination latitudes of the primary and the secondary objects. Despite this singularity, it is proved[1] that the integral exists. Detailed expressions for this integral are given for the three cases in which the inclination latitude of the primary spacecraft is less than, equal to, or greater than the inclination latitude of the secondary orbiting objects.

DISCUSSION AND CONCLUSION

The main results are given by equation (38) for the total number of collisions $N_{total,orbit}$ per orbit. The expression for **K, given by equation (39)** involves the number n_{total} of secondary objects of a particular class, the orbital radius R, the layer thickness h, the collision cross-section radius r_a, and the combined standard deviation σ of the isotropic Gaussian distribution. The expression for **J, given by equation (40)**, involves σ and the maximum flux radius r_e'' which is bounded by h. The expression for **I, given by equation (41),** involves only the inclinations i_1 of the primary spacecraft and i_2 of the secondary objects. Furthermore, as indicated by equation (41), the integral I comprises two terms: the first term I_A for the approaching collisions and the second I_T for the trailing collisions. Thus, we may compute I for various combinations of i_1 and i_2, independent of altitude.

Table 3 gives the values for I_A, I_T, and I for i_1 = 86.4 deg and i_2 = 33, 53, 63, 74, 81, 90, 98, and 106 deg. Because of the singularity of the integrand, the integrals I_A and I_T are each decomposed into two terms. The first term corresponds to integration from 0 to $(\theta'' - \varepsilon)$, being free of singularities, and is evaluated numerically. The second term corresponds to integration from $(\theta'' - \varepsilon)$ to θ'' and is evaluated analytically using an upper bound[1]. The value of ε is chosen to be 2 deg, which is 0.0349 radians. The general expression for J is given by equation (40) and is shown to be very close to unity when $r_e'' > 4\sigma$. This inequality is certainly satisfied for reasonable values of σ and the values of h since we can take r_e up to be h/2. It is safe to take σ up to be 5 km. For this, we need to choose h to be at least 40 km.

Table 4 shows the inclination and total number of secondary objects having perigee **and/or** apogee between 457 and 780 km altitude. (**Remark:** This data was extracted by Dr. Tom Mullikin in September 1995 from the NORAD orbital tracking reports. It is seen that the "average" density in the layer (618.5 ± 161.5) km is 490/323 = 1.517 objects/km of altitude. This density includes orbiting objects that spend partial time in the layer. In May 1996, Capt. Pam Neuman of the U.S. Air Force furnished the value of 45 objects having perigee **and** apogee in the layer (630 ± 20) km, which yields an average density of 45/40 = 1.125 objects/km of altitude.) Thus, we have R = (6,378.1 + 618.5) = 6,996.6 km and h = 323 km. The collision cross-section radius r_a is taken to be (5 + 5) = 10 m. These values are used to compute K, via equation (39), for each class of objects. Together with the values of I in Table 3 and taking J = 1, equation (38) is used to compute the number of collisions per orbit and per year for each of the eight classes. We conclude that the total number of collisions per year experienced by the primary spacecraft is approximately 2E-4.

This result is in agreement with those obtained using time-consuming Monte Carlo simulations. If the input parameters are changed, these simulations have to be repeated. On the other hand, by using equations (38) through (41) we need only recompute some or all of the three terms I, J, and K as necessary. Moreover, we can also observe the dependence of the input parameters analytically.

Table 3. Evaluation of Improper Integral I Defined by Equation (41)

Satellite Class	Inclination i_2 (deg)	Approaching Collision Integral $I_A(0,\theta') + I_A(\theta',\theta'') = I_A$	Trailing Collision Integral $I_T(0,\theta') + I_T(\theta',\theta'') = I_T$	Total Integral $I_A + I_T = I$
1	33	1.094 + 0.287 = 1.381	0.665 + 0.244 = 0.909	2.290
2	53	1.437 + 0.283 = 1.720	0.612 + 0.223 = 0.834	2.554
3	63	1.680 + 0.309 = 1.989	0.562 + 0.230 = 0.791	2.780
4	74	2.080 + 0.169 = 2.249	0.461 + 0.022 = 0.483	2.731
5	81	2.506 + 0.503 = 3.009	0.317 + 0.251 = 0.568	3.577
6	90	3.113 + 0.841 = 3.954	0.356 + 0.406 = 0.762	4.716
7	98	2.696 + 0.691 = 3.387	0.697 + 0.502 = 1.199	4.586
8	106	2.156 + 0.446 = 2.603	0.696 + 0.338 = 1.034	3.637

Remarks: (1) θ'' is defined as the upper limit of integration. Since $i_1 = 86.4$ deg, therefore θ'' is given by the minimum of 86.4 and i_2 (if i_2 is acute) or the minimum of 86.4 and the supplement of i_2 (if i_2 is obtuse).

(2) The value of θ' is chosen to be 2 deg (or 0.0349 radians) less than θ''.

Table 4. Expected Number of Collisions Experienced by Primary Spacecraft Using Equation (38)

Satellite Class	Inclination i_2 (deg)	Number of Space Objects	Value of K for Each Class	Number of Collisions per Orbit	Number of Collisions per Year
1	33	20	5.634 E-10	1.290 E-9	6.991 E-6
2	53	30	8.451 E-10	2.158 E-9	11.695 E-6
3	63	30	8.451 E-10	2.349 E-9	12.729 E-6
4	74	150	42.255 E-10	11.540 E-9	62.526 E-6
5	81	160	45.072 E-10	16.122 E-9	87.345 E-6
6	90	10	2.817 E-10	1.329 E-9	7.198 E-6
7	98	80	22.536 E-10	10.335 E-9	55.998 E-6
8	106	10	2.817 E-10	1.025 E-9	5.551 E-6
		490		46.148 E-9	250.042 E-6

Remarks: (1) The orbital radius R = 6,996.6 km, the layer thickness h = 323 km, and the collision cross-section radius r_a = 10 m are used in equation (39) for computing K.

(2) We observe that the total number of collisions per year is approximately 2.5 E-4. However, if we consider the more accurate density of 1.125 objects/km altitude (per Capt. Neuman) instead of 1.517 objects/km altitude (per Dr. Mullikin), we obtain the value of 1.8 E-4 collisions per year.

REFERENCES

1. F. K. Chan, "Collision Probability Analyses for Earth-Orbiting Satellites," Proceedings of the 7th International Space Conference of Pacific Basin Societies, Nagasaki, Japan, July 1997 (Abridged). Aerospace Corporation Report No. ATR-2000(8456)-1, September 2000 (Unabridged).

2. K. T. Alfriend, et al., "Probability of Collision Error Analysis," J. Space Debris, **1**, 21-35, 1999.

3. R. A. LeClair, "Probability of Collision in the Geosynchronous Orbit," Space Control Conference, MIT Lincoln Laboratory, Lexington, Massachusetts, April 2000.

4. R. P. Patera, "A General Method for Calculating Satellite Collision Probability," AAS/AIAA Space Flight Mechanics Meeting, Clearwater, Florida, January 2000.

5. M. R. Akella, and K. T. Alfriend, "Probability of Collision between Space Objects," J. Guidance, Control, and Dynamics, **23**, 5, 769-772, 2000.

6. R. P. Patera, "A Method for Calculating Collision Probability between a Satellite and a Space Tether," AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, California, February 2001.

7. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 195–196 (1965).

8. M. I. Skolnik, *Introduction to Radar Systems,*McGraw-Hill, New York, 24-29 , 2nd edition (1980).

9. S. O. Rice, "Mathematical Analysis of Random Noise," Bell System Tech. J., **23**, 282-332 (1944) and **24**, 46-156 (1945).

10. F. K. Chan, "Analytical Expressions for Computing Spacecraft Collision Probabilities," AAS/AIAA Spaceflight Mechanics Meeting, Santa Barbara, California, February 2001.

HIGH EARTH ORBIT DESIGN FOR LUNAR-ASSISTED MEDIUM CLASS EXPLORER MISSIONS

Daniel A. McGiffin and Michael Mathews
Computer Sciences Corporation
Lanham-Seabrook, Maryland USA 20706

Steven Cooley
NASA Goddard Space Flight Center
Greenbelt, Maryland USA 20771

ABSTRACT

This study investigates the application of high-Earth orbit (HEO) trajectories to missions requiring long on-target integration times, avoidance of the Earth's radiation belt, and minimal effects of Earth and Lunar shadow periods which could cause thermal/mechanical stresses on the science instruments. As used here, a HEO trajectory is a particular solution to the restricted 3-body problem in the Earth-Moon system with the orbit period being either ½ of, or ¼ of, the lunar sidereal period. A primary mission design goal is to find HEO trajectories where, for a five-year mission duration, the minimum perigee radius is greater than 7 Earth radii (R_E). This minimum perigee radius is chosen so that, for the duration of the mission, the perigee is always above the relatively heavily populated geosynchronous radius of 6.6 R_E. A secondary goal is to maintain as high an ecliptic inclination as possible for the duration of the mission to keep the apsis points well out of the Ecliptic plane. Mission design analysis was completed for launch dates in the month of June 2003, using both direct transfer and phasing loop transfer techniques, to a lunar swingby for final insertion into a HEO. Also provided are analysis results of eclipse patterns for the trajectories studied, as well as the effects of launch vehicle errors and launch delays.

1 – INTRODUCTION

A paper was presented at the 1994 Goddard Space Flight Center's Flight Mechanics Estimation Theory (FMET) Symposium titled *High Earth Orbit Design for Lunar Assisted Small Explorer Class Missions* (Reference 1). It described a lunar gravity-assisted method for achieving a class of stable HEO trajectories for which apogee is near the distance (|R|) of the Moon and perigee is in the range of 8 to 25 R_E. Lunar and solar perturbations usually cause dramatic changes in the elements of such highly elliptical orbits, but it was found possible to adjust the orbit period and phase with respect to the Moon so that the result is a quasi-periodic solution to the restricted three-body problem in the Earth-Moon-spacecraft system. This solution possesses great stability (in the sense that secular changes in key design parameters such as inclination and perigee height are eliminated, although these parameters may oscillate about some average value) and orbit lifetime without requiring spacecraft stationkeeping.

In 1994, it was noted that the advantages of the HEO orbit were freedom from passage through the Earth's radiation belts, small gravity gradient effects, and no atmospheric torques. Perigee altitudes remaining above geosynchronous altitude (~6.6 R_E), excellent coverage by a single ground station, and modest launch vehicle and spacecraft propulsion system requirements, as a result of the lunar gravity assist feature, seemed likely to make this an attractive orbit type for future missions. Additionally, a HEO has more launch opportunities per month than, for example, a lunar swingby Lagrange L2 point mission because there is greater flexibility in the choice of Moon location at lunar swingby. A drawback of the HEO, however, is that it is very difficult to analyze because of the complexity in modeling lunar and solar perturbations. One has to be careful that a chosen orbit meets all mission requirements. For example, one must verify that Transfer Trajectory Injection (TTI) dispersions and launch delays do not cause a nominal orbit to become unsuitable. Another possible disadvantage of the HEO orbit is the spacecraft disposal problem at the mission end of life (EOL). For example, this could be of concern if the long-term evolution of the orbit carried the spacecraft repeatedly through the relatively highly populated geosynchronous altitude.

Since 1994, however, the HEO orbit received little attention—perhaps because of the number of new missions utilizing libration point orbits about the Earth-Sun Lagrange L1 and L2 points. Yet HEO trajectories possess several advantages over the libration point orbits including the following: they are stable and much closer to the Earth; their station keeping requirements are negligible; and, it is much easier to achieve high data transmission rates than is the case for the libration point orbits. Consequently, several potential missions (Kronos, SNAP, Constellation-X, etc.) are now considering a HEO as their baseline trajectory design. It is beyond the scope of this paper to discuss the details and science objectives of these missions. Those interested in such matters are referred to the World Wide Web pages for those various missions (Reference 2). Most of the analysis presented here was performed for the Kronos mission but it is sufficiently general to be applicable to any proposed HEO mission. Various mission requirements might differ (e.g., lifetime, minimum perigee altitude, ecliptic inclination, etc.) but the behavior of the trajectories and methods of achieving them is similar in all cases.

Analysis Goals

The primary goal of this study is to find HEO trajectories where, for the entire five-year duration of the mission, the minimum perigee radius is always greater than 7 R_E, so that perigee is always well above the relatively heavily populated geosynchronous radius of 6.6 R_E. A secondary goal is to maintain the ecliptic inclination (ECL INC) as high as possible for the duration of the mission for shadow minimization.

This analysis will address the following properties of the proposed Kronos orbit:

- Dependence of delta velocity (delta-V or ΔV) at the post lunar encounter perigee as a function of orbit parameters and injection dispersions
- Characterization of launch window (days per month)
- Ecliptic Inclination Evolution
- Perigee Evolution
- Eclipse Overview
- Mission lifetime

A full parametric study of the proposed Kronos trajectory would require considerable effort because of the lack of analytic solutions to the 4-body (Sun, Earth, Moon, Kronos spacecraft (SC)) problem. To determine seasonal as well as monthly trends in ΔV, perigee radius, ecliptic inclination, etc., would require the development of three fully integrated and fully targeted 5-year trajectories for each launch day (nominal and ± 3 sigma (σ) launch vehicle injection energy error cases) as well as launch window analysis for each of the cases. This was beyond the scope of this preliminary analysis.

Modeling Assumptions

All trajectories were computed in Earth-centered Mean-of-J2000 coordinates, with the Mission Analysis and Design Tool, commonly known by the name, 'Swingby' (Reference 3). The Earth centered trajectories were numerically integrated with an eight-order Runge-Kutta-Nystrom algorithm with adaptive step size control through sixth order, and had perturbation models consisting of an 8 x 8 Joint Gravity Model-2 (JGM-2) geopotential field, solar radiation pressure and point mass gravities assumed for the Sun and Moon (Reference 3). Planetary gravitation was not invoked for this preliminary work. All parameter targeting was performed using the differential correction trajectory-targeting scheme in Swingby. Swingby was used to calculate the trajectory from launch site to parking orbit insertion (POI). Manufacturer's data for the Delta II 7325 launch vehicle was used to give the insertion state into a 185-kilometer (km) parking orbit with a 28.7 degree (°, deg) inclination from a launch assumed to be from the Eastern Range. Swingby was also used to model the propagation (coast time (C)) from the POI to the TTI point. The mass of the Kronos spacecraft used in all propagations was 600 kilogram (kg), and all delta-V's used to simulate maneuvers by the launch vehicle (TTI) and spacecraft were modeled as impulsive.

2 – TRAJECTORY DESIGN APPROACH

The HEO can be established via either a direct transfer to the lunar gravity assist or via a series of phasing loops culminating in a lunar gravity assist. Detailed results for both methods will be discussed later. In either case, the spacecraft is launched into a highly elliptical orbit with apogee near the lunar distance. A close flyby of the Moon is used to provide an effective perigee raising delta-V (typically equivalent to 700 to 900 meters per second (m/s)) and also to target the desired perigee altitude, ecliptic inclination, and time of flight to the post-lunar encounter perigee (PLEP) passage. At the Moon, time of arrival and B-plane (B•T, B•R) parameters are used as targeting controls to establish a preliminary HEO trajectory (Reference 4). At the PLEP, a period adjust maneuver (PAM) is performed to adjust the mission orbit period to one-half the sidereal month. (A sidereal month is approximately 27.3 days.)

For both the direct and phasing loop techniques, an attempt was made to determine if the phase of the Moon on the launch day had any affect on the trajectory. Table 1 below lists the phases of the Moon for the month of interest, i.e., June 2003 (Reference 5).

Table 1: Start Times of Lunar Phases for June 2003

Lunar Phase	New Moon (NM)	First Quarter (FQ)	Full Moon (FM)	Last Quarter (LQ)
Start for June 2003	2003/05/31 04:20	2003/06/07 20:28	2003/06/14 11:16	2003/06/21 14:46

As an example of the influence of lunar phase at launch, we note that if the Earth-spacecraft-Moon (E-SC-M) angle is near 180° when the PAM is performed at PLEP, then the geometry shown in Figure 1 results. Figure 1 depicts a fixed-size phasing loop transfer to a HEO, displayed in an Earth-Moon Rotating coordinate frame. Subsequent apogee passages (where the Moon perturbs perigee height most strongly) occur when the Moon-Earth-Spacecraft angle is approximately 90° and the apogee passages are alternately on the leading, then trailing, side of the Moon's orbit. In this orbit, the lunar secular perturbations average, roughly, to zero resulting in significant long-term "stability." Solar perturbations do impact the trajectory as well, resulting in cyclic variations in the ecliptic inclination and perigee altitude. As will be shown, careful choice of initial conditions can yield a trajectory that maintains perigee above geosynchronous altitude and an inclination high enough to keep apogee out of the ecliptic plane to minimize the impact of eclipses.

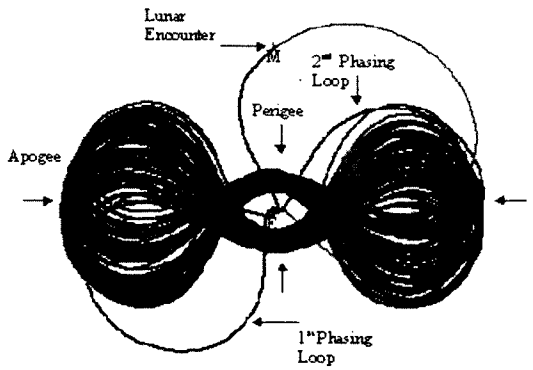


Figure 1: Establishing the HEO for Phasing Loop Transfer in Earth-Moon Rotating Coordinates

2.1 – Direct Transfer Technique

Direct transfer solutions with launch dates in June 2003 were sought. For these launch dates, the objective was to determine the number of nominal launch opportunities available while letting mission related goal values float. In this approach, we wished to obtain a PLEP inclination between 40° and 50° and a PLEP radius value between 7 R_E and 25 R_E. The aforementioned mission design goals are attainable by using the targeting variables, launch epoch,

parking orbit coast time, and TTI delta-V magnitude (direction assumed along the instantaneous velocity vector). The TTI is assumed to correspond to the second cutoff of the 2nd stage, SECO2) of a Delta II launch vehicle for this class of mission. Before a trajectory for a given launch day can be calculated, an initial estimate of the launch vehicle launch epoch and low Earth parking orbit coast time, or duration, must be computed. A launch epoch and either a long coast (≈ 4500 sec) or a short coast (≈ 1500 sec) solution was calculated using the following:

- Launch location, launch vehicle outgoing asymptote angles, launch energy (C_3)
- Right ascension (RA) and declination (DEC) of the Moon approximately 3.5 days after launch (or at approximately 30 days for a phasing loop transfer to the lunar swingby point).

Table 2 gives an outline of the targeting approach used in designing a direct transfer to the Moon to achieve a nominal HEO mission trajectory. It lists the targeting variables that were used to achieve mission goals at specific points in the Kronos mission trajectory profile.

Table 2: Kronos Mission Targeting Event Sequence for a Direct Transfer Technique

Targeting Event	Initial Conditions	Variables	Goals
Establish transfer trajectory with right energy to reach the Moon.	Launch (L) Epoch Coast (C) Duration Coast to TTI	Impulsive TTI ΔV magnitude (along velocity)	$C_3 \approx -2.0$ km^2/s^2
Achieve desired RA and DEC of Moon at L+3.5 days.	Propagate to Lunar Distance (370,000 km from Earth)	L, C and TTI ΔV magnitude	Lunar RA and DEC, Lunar Arrival Time
Perform Lunar gravity assist	Propagate to Periselene (closest approach to Moon)	Variables: L, C	B•T, B•R
Evaluate the E-SC-M angle, |R| and Ecliptic Inclination at the PLEP.	Propagate to the post-lunar encounter-descending node in the lunar orbital plane.	L, C and TTI ΔV magnitude	PLEP |R| > 7 R_E, ECL INC > 40°, Delta RA = 0°, E-V-M ≈ 180°
Establish orbit period ≈ ½ of Lunar sidereal period	Propagate from PLEP to Post Lunar Encounter Apogee (PLEA)	PAM ΔV (along velocity)	Period ≈ 13.66 days
Examine orbit stability and ability to meet mission requirements.	Propagate for five years	-----	-----

Figure 2 is sample plot of an initial 1-month segment of a direct transfer trajectory viewed in solar rotating coordinates that was targeted for June 1, 2003 launch and then propagated for 5 years. The direct transfer to lunar swingby followed by the return to the PLEP (the PAM location) and a half-revolution of the resulting HEO is shown.

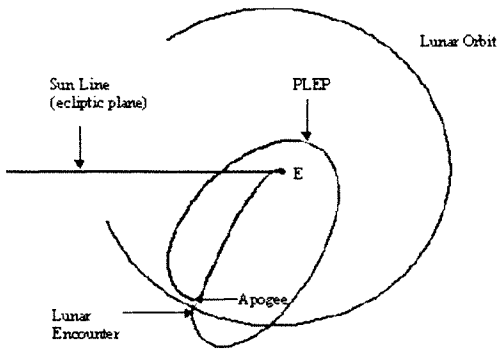


Figure 2: Direct Transfer to HEO

2.2 – Phasing Loop Transfer Technique

The phasing loop case involves completing 2.5 phasing orbits, or "loops" (so-called because consecutive revolutions do not lie on top of each other when plotted in a solar rotating reference frame as is typical in work with these orbits), before the lunar encounter occurs approximately 25 to 30 days after the time of launch. To properly phase the satellite trajectory for a Lunar Swingby without the use of maneuvers at perigee in the phasing loops, both phasing loop apogee radii must extend out to lunar orbit distance. Because the Moon's synodic period is approximately 29.5 days, the Sun-Earth-Moon geometry will be similar at encounter for both the direct transfer and phasing loop cases for launch on a given day of the month. That is, if a direct transfer launch took place at New Moon, the encounter would take place about 3 days after New Moon. A phasing loop launch at new Moon would encounter the Moon anywhere from a few days before the next New Moon (i.e., a synodic month later) to a day after the next New Moon. In other words, if encounter takes place with the Moon near the Sunline for a direct transfer, then a phasing loop transfer launched on the same day will also encounter the Moon when it is near the Sunline. This encounter geometry similarity is because the time spent in the phasing loops is nearly a complete month. The implication here is that a widened launch window could not be obtained by changing strategies by perhaps initially planning a phasing loop trajectory and then hoping to switch to a direct transfer scenario to buy some more time after the phasing loop window had passed. Other possible phasing loop designs such as the "stunted" 3.5 phasing loop scenario being used for the Microwave Anisotropy Probe (MAP) mission might be useful but have not been considered in this preliminary analysis.

The same analysis strategies (specific mission goal values) employed in the direct transfer case were also applied for the phasing loop cases in the month of June 2003. The targeting scheme for the phasing loop case involves one additional item; a maneuver at the first apogee (A1) may be needed to prevent both perigee radii in the phasing orbits from dropping below the Earth's surface. This apogee maneuver is unnecessary for a direct transfer scenario. This A1 maneuver maintains the perigee radius above an assumed minimum acceptable altitude of approximately 450 km.

Table 3 gives an outline of the approach that was taken in designing a phasing loop transfer to the Moon to achieve a nominal HEO mission trajectory. It lists the targeting controls used to achieve mission goals at specific points in the Kronos mission trajectory profile. In this outline, P1 is the first perigee after A1; P2 is the final perigee before the lunar encounter; and, the PLEP occurs at P3.

Table 3: Kronos Mission Targeting Event Sequence for a Phasing Loop Transfer Technique

Targeting Event	Initial Conditions	Variables	Goals
Establish transfer trajectory with right energy to reach the Moon.	Launch (L) Epoch Coast (C) Duration Coast to TTI	Impulsive TTI ΔV magnitude (along velocity)	$C_3 \approx -2.0$ km^2/s^2
Establish 1st phasing loop geometry	Propagate to A1	L, C, TTI ΔV magnitude	$\|R\| \approx 400{,}000$ km (from Earth), RA, DEC of the Moon
Maintain Perigee Radius ($\|R\|$) > 7000 km	Propagate to P1 and P2	A1 ΔV	$\|R\| > 7000$ km
Achieve desired RA and DEC of Moon at L+30 days.	Propagate to near Lunar Distance (370,000 km from Earth)	L, C and TTI ΔV magnitude	Lunar RA and DEC, Lunar Arrival Time
Perform Lunar gravity assist	Propagate to Periselene (closest approach to Moon)	Variables: L, C	B•T, B•R
Evaluate the E-SC-M angle, Radius and Ecliptic inclination at the PLEP.	Propagate to the Descending Node in the Lunar Orbital Plane.	L, C and TTI ΔV magnitude	PLEP $\|R\| > 7$ R_E, ECL INC > 40°, Delta RA = 0°, E-V-M ≈ 180°
Establish orbit period ≈ ½ of Lunar	Propagate from PLEP to PLEA	PAM ΔV	Period

sidereal period		(along velocity)	≈ 13.66 days
Examine orbit stability and ability to meet mission requirements.	Propagate for five years	-----	-----

Figure 3 is a sample plot of a phasing loop trajectory that was targeted for a June 1, 2003 launch and propagated for 5 years as viewed in solar rotating coordinates. The figure depicts the initial portion of the mission from launch through the 2.5 phasing loop transfer to the Moon, on through the lunar swingby to the PLEP, and finally a half revolution in the HEO ending at the PLEA following the PLEP. The time of flight in this trajectory segment is approximately two months.

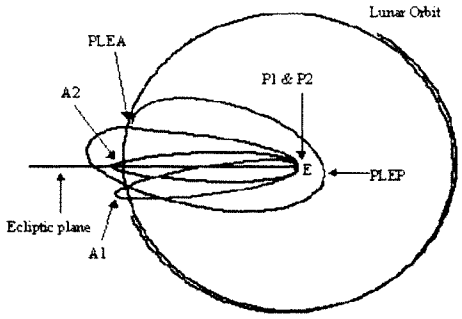


Figure 3: Phasing Loop Transfer to HEO

2.3 – Eclipse Analysis Approach

One important goal of the Kronos mission is to survive eclipses. Considering the geometry of the HEO, the trajectory crosses the ecliptic plane twice per orbit and it is around these nodal crossings that an eclipse is most likely to occur. Though eclipses cannot be designed out of the mission entirely, attempts can be made to minimize the eclipse frequency and duration. The easiest way to reduce eclipses is to orient the line-of-apsides out of the ecliptic plane. The current method uses the Moon's gravity to torque the orbit plane out of the ecliptic. This is an essential part of the trajectory design. Another option is to perform a pair of line-of-apsides maneuvers, which can rotate apogee and perigee out of the ecliptic plane. This method was not investigated for this analysis, as it is a plainly costly option in terms of delta-V.

2.4 – Launch Vehicle Errors and Launch Window Approach

Three sigma launch vehicle energy errors were examined for three direct transfer cases and three phasing loop transfer cases with launch dates in June 2003. In this analysis, a ±10.5 m/s (corresponding to $\pm 3\sigma$ errors) magnitude impulsive ΔV error is applied after TTI for each of the nominal cases examined. These numbers are based on analysis for the MAP project for a similar launch vehicle. The following variables used in the baseline trajectories were frozen for the launch error cases:

- Launch Epoch
- Coast Duration
- Impulsive nominal TTI ΔV magnitude and direction

For the direct transfer cases, a mid-course correction ΔV introduced only 7 hours or more after TTI will be available as a control to target to the B-Plane incoming asymptote. One needs to wait this long after TTI to perform the maneuver in order for an accurate Orbit Determination (OD) solution to become available. The direct transfer cases that were analyzed used an energy correction impulsive ΔV 7 hours after TTI. The phasing loop cases use impulsive maneuvers at each perigee (P1 and P2) in the phasing loops to control the timing and encounter geometry of the lunar swingby. The perigee delta-Vs are applied tangentially to the orbit and affect the apogee distance and period of the subsequent phasing loop.

A launch window analysis was performed for each of three selected nominal direct transfer cases and each of three selected nominal phasing loop cases for June 2003. For a given trajectory, a number of cases are analyzed as part of the launch window analysis. One has to analyze the nominal (no TTI dispersions) trajectory as well as the $\pm 3\sigma$ TTI dispersion cases. Additionally, for each of these three (nominal TTI, $\pm 3\sigma$ TTI) cases for a given trajectory, one has to introduce launch delays of varying magnitude in order to determine the affect of both TTI dispersions and launch delay on the final mission orbit.

2.5 – Solar and Lunar Perturbation Analysis

High apogee, highly eccentric orbits are extremely sensitive to lunisolar perturbations. In many cases, they cause perigee to decay to the point of impact with the Earth's surface. Hence, there is frequently the need to do a perigee-raising maneuver at A1 for the pre-swingby phasing loops. To test the conjecture that the chosen HEO geometry eliminates such mission terminating secular perturbations, however, each HEO trajectory was propagated to the end of the 5-year mission and plots of perigee radius, ecliptic inclination, declination of apogee, etc. vs. time were generated in each case. In addition, to determine the relative impact of the Sun and Moon individually on the trajectory, we numerically integrated representative HEO trajectories with full force modeling of Sun, Earth, and Moon and then integrated the trajectories again with the Moon perturbing-body effects eliminated from the force modeling. The time variations of the orbital elements in both cases were compared graphically. Finally, anticipating that end of life disposal questions will inevitably arise, we looked at the long term variation of the HEO orbital elements by propagating several of the trajectories for 100 years to see just how stable the HEO orbits are.

3 – TRAJECTORY DESIGN RESULTS

3.1 – Direct Transfer and Phasing Loop Technique

As previously discussed, the goal for the direct transfer and phasing loop transfer launch dates is to find an initial PLEP ecliptic inclination and PLEP radius that yield perigee radii above 7 R_E and maximum ecliptic inclination over the mission lifetime. After calculating numerous direct and phasing loop transfer trajectories, it was determined that it is not possible to maintain an ecliptic inclination between 40° to 70° for an entire 5-year mission. Perturbations due to the Sun and Moon force the ecliptic inclination down to a value as low as 20°. Only with the use of plane change maneuvers of prohibitive size can the trajectory inclination be maintained between 40° and 70°. The strategy used to achieve the PLEP ecliptic inclination and radii goals was to target to orbits with initial PLEP ecliptic inclinations greater than or equal to 40° and PLEP radii between 10 and 15 R_E. These desired goals were achieved for 21 launch dates in June 2003 (Reference 6). The combination of the direct and phasing loop transfer techniques provided launch opportunities from the beginning of New Moon phase, to the beginning of Full Moon phase and covering the duration of Last Quarter. No direct transfer designed orbits meeting the aforementioned criteria were found for Full Moon phase and no phasing loop designed orbits meeting the aforementioned criteria were found for Last Quarter phase, as shown in Figure 4.

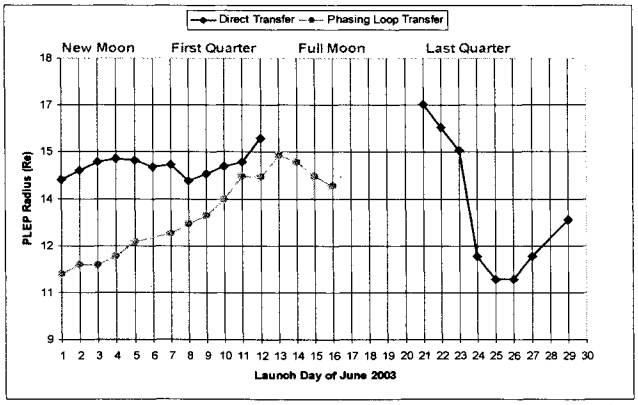


Figure 4: Nominal June 2003 Direct and Phasing Loop Transfer PLEP Radius Summary

For the direct transfer designed orbits launched during New Moon and First Quarter, a varying long parking orbit coast solution was used to achieve mission goals while a short coast solution was determined to be optimal for the launch dates occurring in Last Quarter. For the phasing loop transfer designed orbits launched during New Moon and First Quarter, a varying short coast solution was used to achieve mission goals. Long coast solutions were found for only three launch dates occurring around Full Moon. However, these trajectories barely met the ecliptic inclination constraint and violated the E-SC-M angle constraint of $150.0° < \text{E-SC-M} \leq 180°$. It was found that maintaining the E-SC-M goal keeps the lunar perturbation effects down to a minimum.

The launch dates attempted near or during New Moon tend to have the highest total mission delta-V. It was seen in the results that the higher PAM ΔVs were correlated with a higher PLEP radius. Additionally, the PLEP radius value is a determining factor in maintaining an altitude above 7 R_E for the duration of the mission. Figure 5 shows the total delta-v results for nominal direct and phasing loop transfer launch dates during New Moon through Last Quarter, respectively. Notice that the total delta-V for the phasing loop transfer trajectories is greater than the direct transfer cases. This difference is attributed to the A1 maneuver that is needed to maintain the perigee radii in the phasing loops above a safe altitude (greater than 450 km). Figure 6 shows that the Sun's perturbations effectively act as a negative delta-V when the spacecraft is near apogee and the HEO is oriented as at point A. In this case, the perturbation is opposite the direction of the spacecraft velocity and the P1 altitude will decrease. At the beginning of New Moon or just after point C from Figure 6, the A1 ΔV is at 12 m/s. It then reaches a peak of 27 m/s for the June 7 launch date and then decreases to a minimum of 5 m/s on the June 14 launch date during the beginning of Full Moon. The situation is just reversed at point B where the Sun's perturbation is in the direction of the velocity and acts as a positive apogee delta-V, thus increasing the perigee altitude. At points C and D, the Sun has a minimal effect on the P1 height.

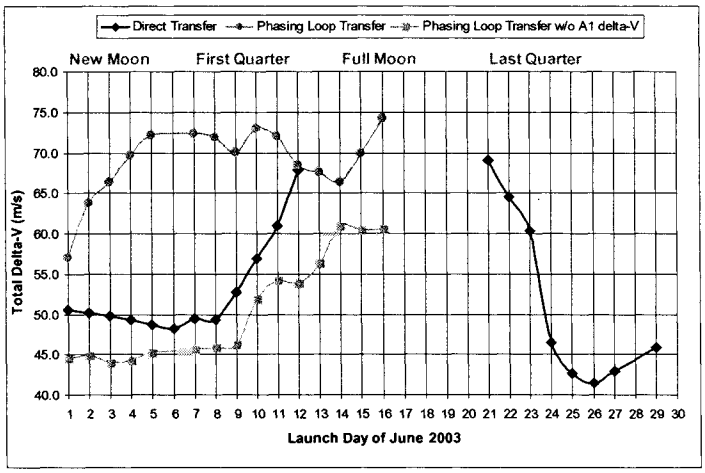


Figure 5: Nominal June 2003 Direct and Phasing Loop Transfer PLEP Radius Summary

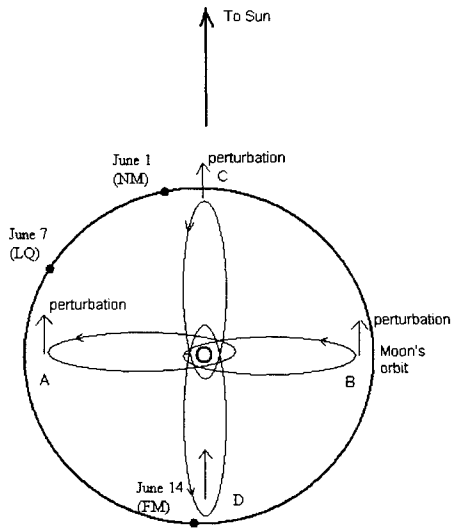


Figure 6: Effect of Solar Perturbations on 1st Phasing Loop Perigee Pass Height

The ecliptic inclination and the perigee and apogee radius oscillate over time but the semimajor axis remains constant. Figures 7 and 8 show the evolution of ecliptic inclination, perigee and apogee radius after lunar encounter for a mission lifetime of six years for the complete set of nominal trajectories calculated for the month of June. Figure 7 demonstrates that the ecliptic inclination values achieve a minimum in the 20 to 25° range, but start to increase after 2 years and eventually reach values between 50 and 65°. Once the ecliptic inclination and PLEP radius reach a minimum (apogee radius is at a maximum) as shown in Figure 8, both begin to increase and then eventually surpass the initial PLEP values. During this latter time, apogee radius decreases to a minimum. The apogee radii values during the first 2 years increase to a peak value of 70 R_E and then decrease to a range of 50 to 60

R_E toward the end. The perigee radii values initially decrease to 7 R_E after 2 years but ultimately peak at approximately 28 R_E after six years.

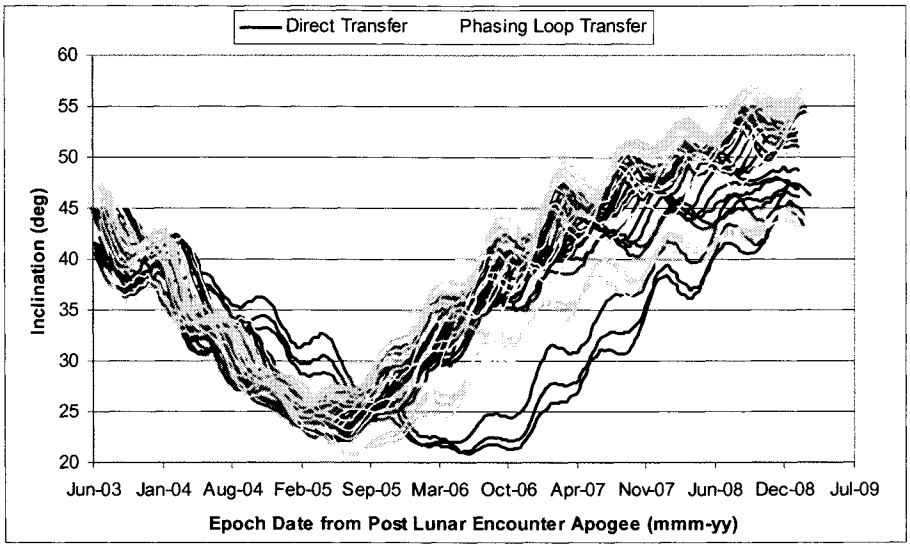


Figure 7: Evolution of Ecliptic Inclination for Nominal Trajectories

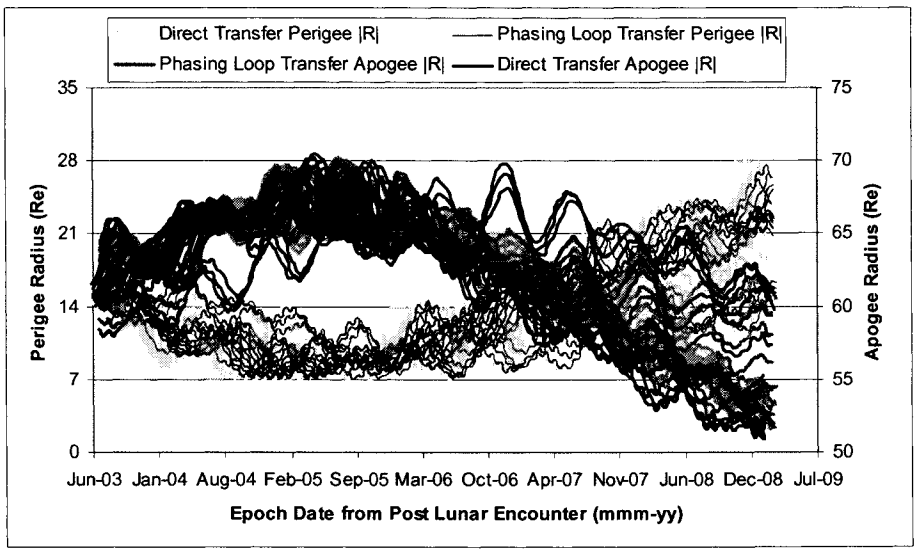


Figure 8: Evolution of Apogee and Perigee Radius for Nominal Trajectories

3.2 – Eclipse Event Results

Figure 9 shows the eclipse duration during a 6-year mission for all the computed launch dates in June. The eclipse events are mostly Earth shadow events with only a couple of lunar eclipse (penumbra) events appearing during each six-year mission. The eclipse cycle for each of the computed trajectories ranged from two to three per year during their six year mission propagations. Note that usually any eclipse event exceeding 60 minutes in duration will

involve an umbra cone (except cases involving the Moon, which will only have penumbras). The majority of eclipse events have durations of 200 minutes or less and mostly occur at true anomalies approximately 90° off of apogee, or where the trajectory plane intersects the ecliptic plane. It was observed that when the line-of-apsides of the HEO is close to the Earth-Sun line and the ecliptic declination is below 25 °, the eclipse durations have exceeded 200 minutes and tended to occur less than five days away from apogee. Figure 10 shows the evolution of the ecliptic declination at apogee for all the computed trajectories in June. Also, if the ecliptic declination at apogee approaches zero, it increases the likelihood of long duration eclipses that could exceed a half-day. Figure 11 is a sample plot of a direct transfer trajectory that was targeted for a June 26, 2003 launch and propagated for one orbit starting after PLEA. The figure depicts the geometry of the line-of apsides with respect to the ecliptic plane and even though the ecliptic inclination of this orbit is ≈ 45°, the line-of-apsides is only ≈ 20° off the ecliptic plane.

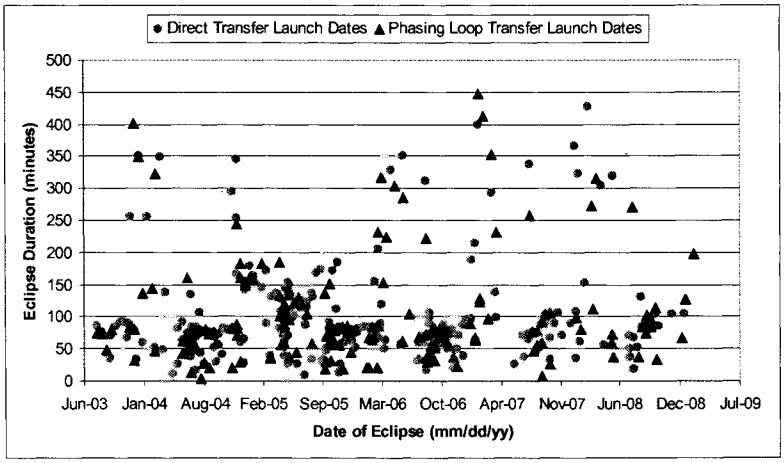


Figure 9: Eclipse Duration for Nominal Trajectories in June 2003

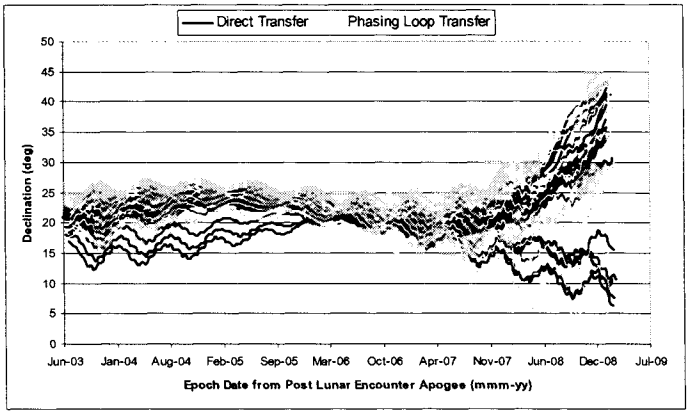


Figure 10: Evolution of Ecliptic Declination for Nominal Trajectories in June 2003

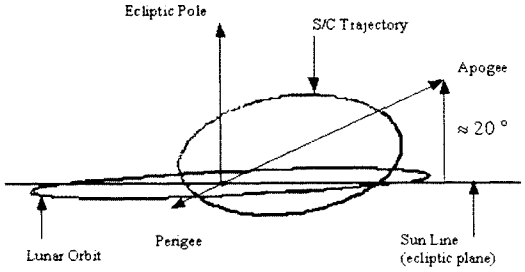


Figure 11: Geometry of the Line-of-Apsides of a Nominal June 26, 2003 Direct Transfer

3.3 – Launch Vehicle Errors and Launch Window Results

For this analysis, one starts with a nominal trajectory that meets the baseline mission constraints. One then first introduces TTI energy dispersion errors and then combinations of energy dispersion errors and launch delays. Launch delays of 5, 10, and 20 minutes were analyzed. If a given launch delay results in a mission orbit that does not meet the previous derived constraints, then larger launch delays are not analyzed. Note that the PLEP targets were allowed to float, but not to the extent of constraint violation, for the dispersion and launch delay cases, if needed, to maintain a suitable mission orbit.

A set of six nominal trajectories, three from the direct transfer design and three from the phasing loop design, were investigated for this analysis. It was assumed that if all 6-baseline cases can meet all mission conditions that were established earlier, then it should be possible to establish several additional launch opportunities near the epoch of each nominal case examined.

Table 4 shows a summary of mission total delta-V results for the direct and phasing loop-designed nominal, $\pm 3\sigma$ error, and launch window delay (+5, +10, +20 minute) cases together with the corresponding lunar phase at launch for the selected launch dates in June 2003. Both the direct transfer design and the phasing loop design methods established the ability to maintain a 20-minute launch window, as delta-V results were acceptable in all cases. Notice the phasing loop approach generally requires less ΔV to correct launch vehicle errors and allows more time to correct potential spacecraft problems. It does, however, take longer to get to the Moon and requires two or more passes through the Earth's radiation belts. Note that the launch dates, which occur when the Moon is near conjunction with the Sun (New Moon or Last Quarter), yield the most favorable launch windows.

Table 4: TTI Error Dispersions and Launch Delay Delta-V Summary

Error Case Event	Direct Transfer (m/s)	Lunar Phase	Phasing Loop Transfer (m/s)	Lunar Phase
Nominal Trajectory Maximum Total ΔV	69.1	LQ	74.2	FM
Nominal Trajectory Minimum Total ΔV	41.5	LQ	57.1	NM
+3-Sigma Trajectory Maximum Total ΔV	110.7	FQ	121.9	NM
+3-Sigma Trajectory Minimum Total ΔV	97.6	NM	112.7	FQ
-3-Sigma Trajectory Maximum Total ΔV	114.5	FQ	95.1	FQ
-3-Sigma Trajectory Minimum Total ΔV	87.8	LQ	91.8	NM
5-Minute Delay Trajectory Maximum Total ΔV	89.0	NM	76.0	FQ
5-Minute Delay Trajectory Minimum Total ΔV	57.2	LQ	67.4	NM
+3-Sigma/5-Minute Delay Trajectory Maximum Total ΔV	126.6	NM	120.3	NM
+3-Sigma/5-Minute Delay Trajectory Minimum Total ΔV	99.0	LQ	114.9	FQ
-3-Sigma/5-Minute Delay Trajectory Maximum Total ΔV	110.6	FQ	100.4	FQ
-3-Sigma/5-Minute Delay Trajectory Minimum Total ΔV	81.6	NM	92.5	NM
10-Minute Delay Trajectory Maximum Total ΔV	128.1	NM	82.3	FQ
+3-Sigma/10-Minute Delay Trajectory Minimum Total ΔV	163.1	NM	117.9	FQ
-3-Sigma/10-Minute Delay Trajectory Maximum Total ΔV	100.9	NM	105.8	FQ
20-Minute Delay Trajectory Maximum Total ΔV	85.1	LQ	71.6	NM
+3-Sigma/20-Minute Delay Trajectory Minimum Total ΔV	135.5	LQ	117.3	NM
-3-Sigma/20-Minute Delay Trajectory Maximum Total ΔV	94.6	LQ	95.9	NM

3.4 – Solar and Lunar Perturbation Analysis Results

It is well known that perturbations due to the Sun on the Earth-Moon system cause both secular and cyclic changes in the Moon's orbital elements. For example, the Moon's line-of-nodes regresses continuously completing a full cycle in about 18.6 years and the line-of-apsides rotates through a full 360° in approximately half this time. There

are cyclic variations in the Moon's inclination and eccentricity. For the HEO, we expect the Sun to affect the trajectories in a similar manner but we expect no secular changes in perigee radius and inclination because of our choice of initial Sun-Earth-Moon-SC geometry and orbital period. Figures 12, 13, and 14 show the time variation of perigee radius, ecliptic inclination, and argument of perigee for a typical HEO propagated both with and without lunar gravity.

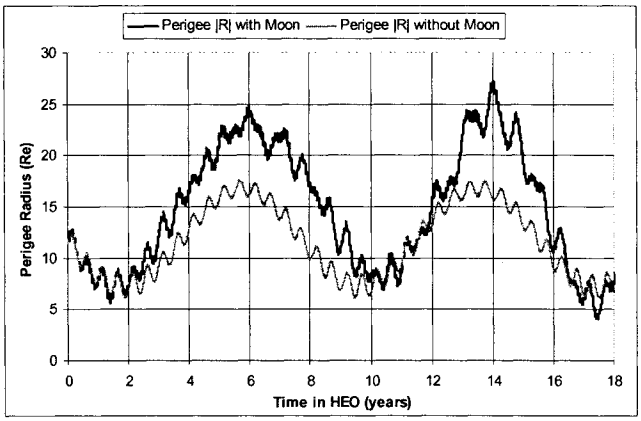


Figure 12: Evolution of Perigee Radius with and without Lunar Perturbations

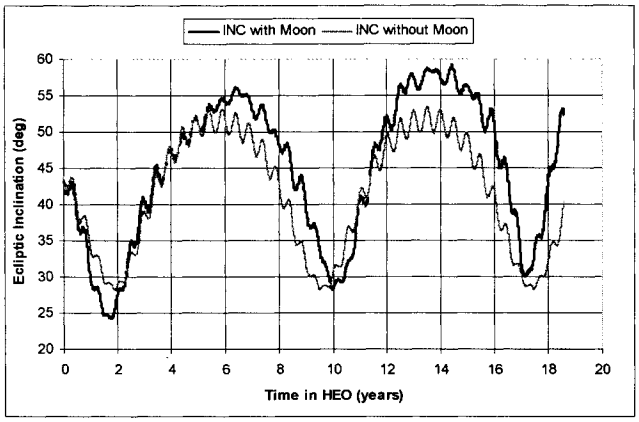


Figure 13: Evolution of Ecliptic Inclination with and without Lunar Perturbations

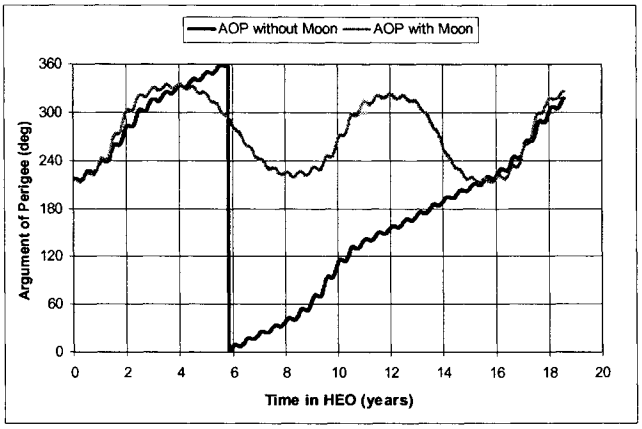


Figure 14: Evolution of Argument of Perigee with and without Lunar Perturbations

As was expected, the Sun dominates the behavior of the orbit but the Moon adds some interesting complications. The Moon affects the inclination very little but tends to increase the amplitude of the perigee radius oscillation. The minimum perigee radius is almost unaffected but the maximum increases.

Interestingly, the right ascension of the ascending node regresses in a manner unsurprisingly similar to that of the Moon, completing one revolution in about 19 years. The argument of perigee, however, oscillates about a mean value of approximately 275° with an amplitude of 50° to 75° (there is apparently a long-term variation in the amplitude) and a period of approximately 8 years. This would seem to indicate that, at least in this case, the line-of-apsides does not advance secularly. The implication of this for our purpose is that with judicious choice of launch day or with a line-of-apsides rotating maneuver, the apogee position can be controlled within limits. This could be

useful, for example, if it were desirable to keep apogee in the Northern Hemisphere for communications or other purposes.

Unfortunately, nature is seldom simple and when we examined a case with a very low ecliptic inclination, this oscillatory behavior of the line-of-apsides was not observed. Instead, the line-of-apsides rotated secularly at approximately the same rate as the nodal regression. At this point, we are not certain whether there is some critical initial inclination at which the line-of-apsides begins to oscillate instead of rotating or whether some other factor is responsible for this behavior. In this preliminary investigation, we did not complete a parametric study for a range of initial inclinations.

Another fascinating phenomenon was observed when we attempted to examine the long-term stability of the HEO by propagating the trajectories for 100 years. Figures 15 and 16 show the behavior of the argument of perigee vs. time for both the initial ecliptic inclination cases of 6° and 44°. For the 6° case, the variation is quite regular for the entire 100 years except that there is a long-term increase in the average value of inclination and its amplitude. For the 44° case, however, the behavior is unremarkable until approximately 75 to 80 years into the propagation. At this point, the nodal regression rate slows by roughly one-half, the line-of-apsides stops oscillating and begins advancing secularly at about one rotation every 8 years, and the ecliptic inclination becomes nearly constant near its maximum value as does the perigee radius also. At this time, we can offer no hypotheses about this anomaly nor can we, with any degree of assurance, suggest it might be a software problem. It is a curious problem awaiting future investigation and resolution.

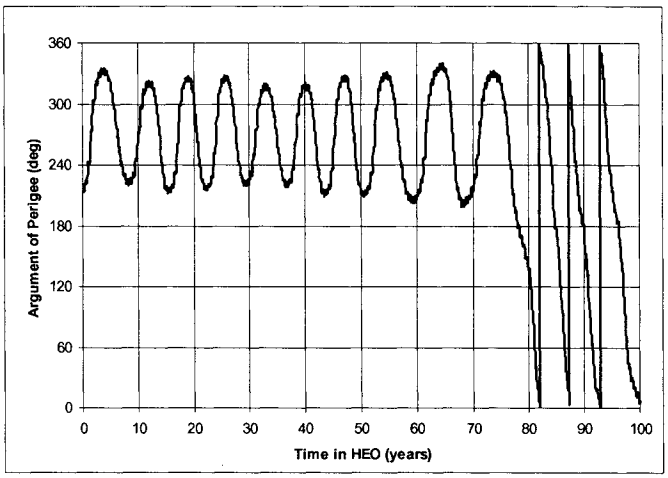


Figure 15: Evolution of Argument of Perigee with Lunar Perturbations for 44° Inclined Orbit

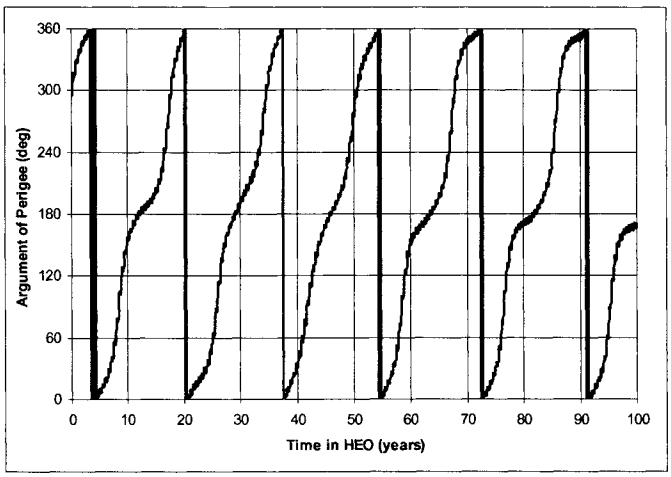


Figure 16: Evolution of Argument of Perigee with Lunar Perturbations for 6° Inclined Orbit

4 – CONCLUSIONS AND FUTURE DIRECTIONS

Approximately 150 individual trajectories were designed and examined to look for trends and general conclusions that could be used to establish guidelines for designing the Kronos baseline mission trajectory. Unfortunately, the lunisolar perturbations have greatly complicated the analysis so that it has been difficult to be as incisive as one might like. It is apparent that both the direct and phasing loop transfer techniques are feasible. We have found that suitable trajectories can found for 11 to 13 days of each month, assuming the month of June to be reasonably representative.

There are advantages and disadvantages for both the direct and phasing loop transfers. The ΔV budget for the direct transfer design is slightly higher (~ 140 to 170 m/s) because of the need to correct the 3σ launch vehicle energy errors within approximately seven hours of launch. However, it should be pointed out that these delta-V numbers represent impulsive delta-Vs. Possible finite burn losses have not been investigated. The phasing loop transfer does

not require such a severe time-critical error correction maneuver but it does introduce the necessity of performing perigee raising maneuvers at A1 for many possible launch dates because solar and lunar perturbations can drive perigee down sufficiently to result in spacecraft reentry. The ΔV budget for the phasing loop scenario is approximately 120 m/s.

No trajectories could maintain a high ecliptic inclination for the entire five-year mission lifetime. Typically, the inclination varies between 23° to 25° and 55° to 65°. For the June trajectories, judicious choice of launch date and PLEP orbit parameters, however, can give trajectories that maintain perigee radius above the geosynchronous radius for the entire mission lifetime. It should be noted that even though perigee radius and ecliptic inclination oscillates between certain limits, these trajectories are actually very stable. One case was propagated for 250 years to determine if it would eventually leave the HEO orbit and perhaps reenter the Earth's atmosphere. It did not; in point of fact, this trajectory maintained a perigee radius greater than 7 R_E for the first 28 years.

No matter what choice of PLEP targets is used, it appears impossible to avoid eclipses. The Sun must pass through the orbit plane twice per year regardless. Eclipses with the spacecraft near perigee will be less frequent and of shorter duration, but when the Sun is eclipsed with the spacecraft near apogee, shadow periods of several hours' duration were observed.

Several HEO orbits with ¼ lunar period orbits were examined and showed essentially the same behavior as the ½ lunar period orbits. The only significant difference was the approximately four to five times greater PAM ΔV needed to establish the proper phasing with the Moon (References 1 and 6).

It would be desirable to take this analysis further in several respects. Launch window and error analysis should be performed for more than the few cases examined so far in order to pin down the ΔV budgets more precisely. Any possible finite burn losses should also be investigated. The maximum eclipse duration should be determined to see if it might impact spacecraft design or conflict with mission requirements. Further analysis of the effects of the lunisolar perturbations should be performed to better predict optimal launch opportunities and initial orbital elements.

REFERENCES

1. Mathews, M., Hametz, M., Cooley, J. and Skillman, D., "High Earth Orbit Design for Lunar Assisted Small Explorer Class Missions," 1994 GSFC Flight Mechanics and Estimation Theory Symposium, May 1994.

2. http://www.astronomy.ohio-state.edu/~kronos/, http://snap.lbl.gov/, http://constellation.gsfc.nasa.gov/

3. McGiffin, D., et al., Mission Analysis and Design Tool (Swingby) Mathematical Principles, Revision 1, draft version, CSC/TR-92/6091R1UD0, Computer Sciences Corporation, September 1995.

4. Kizner, W., "A Method for Describing Miss Distances for Lunar and Interplanetary Trajectories," Ballistic Missile and Space technology, III, 1961.

5. Naval Observatory website where the lunar phase was obtained: http://aa.usno.navy.mil/AA/data/docs/WebMICA_2.html

6. McGiffin, D., Mathews, M., "Mission Feasibility Study for Kronos High Earth Orbit," Flight Dynamics Navigation, Attitude, and Information Technology, CSC-96-968-19, Computer Sciences Corporation, August 2000.

Session 2: Constellation Design and Formation Flying

ALTERNATE FORMS OF RELATIVE ATTITUDE KINEMATICS AND DYNAMICS EQUATIONS*

Guang Q. Xing† and Shabbir A. Parvez‡
Space Products and Applications, Inc.

ABSTRACT

In this paper the alternate forms of the relative attitude kinematics and relative dynamics equations are presented. These developments are different from the earlier developments that have been presented in other publications. The current forms of equations have the advantage of being simpler than earlier ones. These equations are applied in developing the necessary kinematics and dynamics for relative navigation in formation flying and virtual platforms. These equations also have application in the implementation of nonlinear full state feedback and nonlinear output feedback control for large attitude angle acquisition and tracking. This paper presents simulations from such a full state feedback control application.

INTRODUCTION

Navigation and control for spacecraft flying in formation requires the relative attitude information. Since the attitude rotation matrices and their kinematics relationships are generally defined in the inertial reference frame, Ref.1 was the first attempt at completely defining the kinematics and relative attitude dynamics when the attitude matrix is defined in a non-inertial frame. While the concept of relative attitude had earlier been addressed in Refs.2-4, Ref.1 was the first publication to provide the complete development that is necessary for formation flying control. This paper continues these developments, providing the relative attitude kinematics and dynamics that are alternative to the forms developed in Ref.1.

The attitude of a rigid-body with respect to the inertial frame is determined by a rotation transformation matrix from the inertial frame to body frame. This rotation matrix is referred to as the attitude matrix. In practical design, as noted in Ref.1, the attitude matrix is parameterized to be 4-dimension parameters such as axis/angle variables and quaternion, and 3-dimension parameters such as Euler angles, Rodrigues (Gibbs vector) and modified Rodrigues. Similarly, rotation matrix of the body-frame with respect to a non-inertial frame can be defined as the relative attitude matrix.

In developing the kinematics and dynamics of the relative attitude, Ref.1 addressed the following:

i. Kinematics equations for relative attitude matrix
ii. Kinematics equations for relative attitude parameters
iii. Relative attitude dynamics equations

This paper proceeds along the same line, providing alternate solutions to the above items. The primary difference in this approach results from a different definition of the relative angular velocity. This alternate form provides simpler relative kinematics and dynamics equations than those in Ref.1.

SYMBOLS

[dp1, dp2, dp3]	the three components of relative modified Rodrigues parameters
[dw1, dw2, dw3]	the three components of the relative angular velocity (rad/sec)
g	Rodrigues parameter
p	modified Rodrigues parameter

* This work was supported by NASA Goddard Space Flight Center, Greenbelt, Maryland, under Contract NAS5-99163
† Principal Scientist, Space Products and Applications, Inc., 3900 Jermantown Road, Suite 300, Fairfax, Virginia 22030. Email: xingspa@erols.com
‡ President, Space Products and Applications, Inc., Email: parvzspa@erols.com

q	quaternion parameter
R_{BD}	relative attitude matrix (transformation matrix) of reference frame $\mathcal{F}_B$ with respect to $\mathcal{F}_D$.
ω	angular rate
$d(.)/dt_B$	vector time derivative in $\mathcal{F}_B$.
$d(.)/dt_D$	vector derivative in $\mathcal{F}_D$
$\mathcal{F}_I$	vectrice of the inertial reference system
$\mathcal{F}_T$	vectrice of the local horizontal reference system for target satellite
$\mathcal{F}_S$	vectrice of the local horizontal reference system of the chasing spacecraft
$\mathcal{F}_B$	vectrice of the body-fixed reference system of the chasing spacecraft
$\mathcal{F}_D$	vectrice of the body-fixed reference system of the target attitude
$\otimes$	nonlinear operator

REPRESENTATION OF RELATIVE ATTITUDE

Relative Attitude Matrix

It is assumed that any reference frame can be denoted by a vectrix[7]. The following series of reference frames can then be defined.

$$\mathcal{F}_I = \begin{bmatrix} \underline{i}_I \\ \underline{j}_I \\ \underline{k}_I \end{bmatrix} \quad \mathcal{F}_T = \begin{bmatrix} \underline{i}_T \\ \underline{j}_T \\ \underline{k}_T \end{bmatrix} \quad \mathcal{F}_S = \begin{bmatrix} \underline{i}_S \\ \underline{j}_S \\ \underline{k}_S \end{bmatrix} \quad \mathcal{F}_B = \begin{bmatrix} \underline{i}_B \\ \underline{j}_B \\ \underline{k}_B \end{bmatrix} \quad \mathcal{F}_D = \begin{bmatrix} \underline{i}_D \\ \underline{j}_D \\ \underline{k}_D \end{bmatrix} \tag{1}$$

The attitude of a rigid spacecraft is the orientation of the reference frame with respect to another frame. The most convenient reference frame is a dextral, orthogonal triad which is fixed with the rigid body of spacecraft. The other reference frame can be an inertial reference frame, or it can be a moveable reference frame which is fixed to another body. The attitude with respect to the inertial reference frame is the absolute attitude. The attitude with respect to a movable rotting reference frame is the relative attitude. which is respect to a movable rotating reference frame is named the relative attitude. For example in order to study the orientation relationship between two rotating reference frames, the relative attitude can be defined as a transformation matrix between two rotating reference frames.

$$\mathcal{F}_B = R_{BD}\mathcal{F}_D \tag{2}$$

Parametric Representation of the Relative Attitude Matrix

As with parametric representation of absolute attitude matrix (Refs.6-8), the relative attitude matrix can also be represented by the attitude parameters such as Axis/angle, Quaternion, Rodrigues parameters (Gibbs vector), modified Rodrigues parameters and Euler angles. The relationship between relative attitude matrix and absolute attitude matrix is

$$R_{BD}(p_{bd}) = R_{BI}(p_b)\ R_{DI}^{-1}(p_d) \tag{3}$$

It is obvious that the relative attitude parameters g_{bd} and absolute attitude parameters g_b , g_d have a nonlinear relationship. This can be represented by the following unified notation:

$$p_{bd} = p_b \otimes p_d^{-1} \tag{4}$$

where the rule of the nonlinear operator ⊗ is determined by composition rotation rule of the attitude parameters [6]. This operation will, therefore, be different for different attitude parameters.

For the modified Rodrigues parameters:

$$p_{bd} = p_b \otimes p_d^{-1} = \frac{p_d(p_b^T p_b - 1) + p_b(1 - p_d^T p_d) - 2[p_d^{\times}]p_b}{1 + p_d^T p_d p_b^T p_b + 2 p_d^T p_b} \tag{5}$$

For Rodigues parameters (Gibbs vector):

$$g_{bd} = g_b \otimes g_d^{-1} = \frac{g_b - g_d + [g_b^{\times}]g_d}{1 + g_b^T g_d} \tag{6}$$

For Quaternion parameters:

$$\bar{q}_{bd} = \bar{q}_b \otimes \bar{q}_d^{-1} = \begin{bmatrix} q_{d4}q_b - q_{b4}q_d + [q_b^{\times}]q_d \\ q_{b4}q_{d4} + q_b^T q_d \end{bmatrix} \tag{7}$$

where

$$\bar{q}_{bd} = \begin{bmatrix} q_{bd} \\ q_{bd4} \end{bmatrix} \quad \bar{q}_b = \begin{bmatrix} q_b \\ q_{b4} \end{bmatrix} \quad \bar{q}_d = \begin{bmatrix} q_d \\ q_{d4} \end{bmatrix} \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad [a^{\times}] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{8}$$

RELATIVE ATTITUDE KINEMATICS FOR ATTITUDE MATRIX

Equation (2) can be written as

$$\mathcal{F}_D^T = \mathcal{F}_B^T R_{BD} \tag{9}$$

Taking derivative with respect to the inertial reference frame for both sides,

$$\omega_d^{\times} \mathcal{F}_D^T = \frac{d(\mathcal{F}_B^T R_{BD})}{dt_B} + \omega_b^{\times} \mathcal{F}_B^T R_{BD} \tag{10}$$

Based on the differential rules for the vectrices (Ref.7), Eq.(10) can be developed as

$$\omega_d^T R_{DB} \mathcal{F}_B \times \mathcal{F}_B^T R_{BD} = \mathcal{F}_B^T \frac{dR_{BD}}{dt_B} + \omega_b^T \mathcal{F}_B \times \mathcal{F}_B^T R_{BD} \tag{11}$$

In the body reference frame $\mathcal{F}_B$, the kinematics equation for relative attitude matrix R_{BD} will be

$$\frac{dR_{BD}}{dt_B} = [(R_{BD}\omega_d)^\times] R_{BD} - [\omega_b^\times] R_{BD} = [(R_{BD}\omega_d - \omega_b)^\times] R_{BD} \tag{12}$$

If relative angular velocity is defined as

$$\underline{\omega}_{bd} = \underline{\omega}_b - \underline{\omega}_d = \mathcal{F}_B^T(\omega_b - R_{BD}\omega_d) \qquad \omega_{bd}^b = (\omega_b - R_{BD}\omega_d) \tag{13}$$

the attitude kinematics equation for the relative attitude matrix (12) can be written the compact form in the body reference frame $\mathcal{F}_B$

$$\frac{dR_{BD}}{dt_B} = -[(\omega_{bd}^b)^\times] R_{BD} \tag{14}$$

Considering

$$R_{BD} R_{DB} = I \tag{15}$$

the relative kinematics equation for R_{DB} in Body reference frame $\mathcal{F}_B$ becomes

$$\frac{dR_{DB}}{dt_B} = R_{DB}[(\omega_{bd}^b)^\times] \tag{16}$$

In the similar way, the kinematics equation for R_{DB} in reference frame $\mathcal{F}_D$ can be obtained:

$$\frac{dR_{DB}}{dt_D} = -[(\omega_{db}^d)^\times] R_{DB} \qquad \frac{dR_{BD}}{dt_D} = R_{BD}[(\omega_{db}^d)^\times] \tag{17}$$

where

$$\omega_{db}^d = \omega_d - R_{DB}\omega_b \tag{18}$$

Since

$$\frac{dR_{BD}}{dt_B}=\frac{dR_{BD}}{dt_D}=\frac{dR_{BD}}{dt}=\dot{R}_{BD} \qquad \frac{dR_{DB}}{dt_B}=\frac{dR_{DB}}{dt_D}=\frac{dR_{DB}}{dt}=\dot{R}_{DB} \tag{19}$$

these kinematics equations for relative attitude above can be represented in form of a theorem of relative attitude kinematics:

Theorem of Relative Attitude Kinematics

The angular velocities of $\mathcal{F}_B$, $\mathcal{F}_D$ with respect to the inertial reference frame are are ω_b ,ω_d , and

$$\underline{\omega}_b = \mathcal{F}_B^T\omega_b \qquad \underline{\omega}_d = \mathcal{F}_D^T\omega_d \tag{20}$$

The relative attitude between the reference frames $\mathcal{F}_I$, $\mathcal{F}_B$, $\mathcal{F}_D$ are:

$$\mathcal{F}_B = R_{BI}\mathcal{F}_I \qquad \mathcal{F}_D = R_{DI}\mathcal{F}_I \qquad \mathcal{F}_B = R_{BD}\mathcal{F}_D \tag{21}$$

The kinematics equations for R_{BD} and R_{DB} are:

$$\dot{R}_{BD} = -[(\omega_{bd}^b)^\times]R_{BD}=[(\omega_{db}^b)^\times]R_{BD}=R_{BD}[(\omega_{db}^d)^\times]=-R_{BD}[(\omega_{bd}^d)^\times] \tag{22}$$

$$\dot{R}_{DB} = -R_{DB}[(\omega_{db}^b)^\times]=R_{DB}[(\omega_{bd}^b)^\times]=R_{DB}[(\omega_{bd}^d)^\times]=-R_{DB}[(\omega_{db}^d)^\times] \tag{23}$$

where

$$\omega_{bd}^b =(\omega_b-R_{BD}\omega_d)=-\omega_{db}^b=-R_{BD}\omega_{db}^d=R_{BD}\omega_{bd}^d \tag{24}$$

RELATIVE ATTITUDE KINEMATICS FOR ATTITUDE PARAMETERS

The First Kind of Relative Attitude Kinematics Equations

Based on the definitions of the relative attitude and relative angular velocity given above the relative attitude kinematics equations for various attitude representation have been developed in Ref.1. These equations are referred to as the first kind of relative attitude kinematics equations, and are summarized as follows:

Quaternion

Relative kinematics equation

$$\dot{q}_{bd} = \frac{1}{2}([q_{bd}^\times]+q_{bd4}I)\omega_{bd}^b \qquad \dot{q}_{bd4} = -\frac{1}{2}(\omega_{bd}^b)^Tq_{bd} \tag{25}$$

Relative velocity

$$\omega_{bd}^{b} = 2(q_{bd4}\dot{q}_{bd} - \dot{q}_{bd4}q_{bd} - q_{bd}\times\dot{q}_{bd}) \tag{26}$$

Attitude transformation matrix

$$R_{BD} = (q_{bd4}^{2} - q_{bd}^{T}q_{bd})I + 2q_{bd}q_{bd}^{T} - 2q_{bd4}[q_{bd}\times] \tag{27}$$

Rodrigues Parameters

Relative attitude kinematics equations

$$\dot{g}_{bd} = \frac{1}{2}([g_{bd}^{\times}] + g_{bd}g_{bd}^{T} + I)\omega_{bd}^{b} \tag{28}$$

Relative velocity

$$\omega_{bd}^{b} = \frac{2}{(1+g_{bd}^{T}g_{bd})}(I - [g_{bd}^{\times}])\dot{g}_{bd} \tag{29}$$

Attitude transformation matrix

$$R_{BD} = \frac{1}{(1+g_{bd}^{T}g_{bd})}((1-g_{bd}^{T}g_{bd})I + 2g_{bd}g_{bd}^{T} - 2[g_{bd}\times]) \tag{30}$$

Modified Rodrigues Parameter

Relative attitude kinematics equation

$$\dot{p}_{bd} = \frac{1}{4}[(1-p_{bd}^{T}p_{bd})I + 2[p_{bd}^{\times}] + 2p_{bd}p_{bd}^{T}]\omega_{bd}^{b} \tag{31}$$

Relative velocity

$$\omega_{bd}^{b} = \frac{4}{(1+p_{bd}^{T}p_{bd})^{2}}[(1-p_{bd}^{T}p_{bd}) + 2p_{bd}p_{bd}^{T} - 2[p_{bd}^{\times}]]\ \dot{p}_{bd} \tag{32}$$

Attitude transformation matrix

$$R_{BD} = \frac{1}{(1+p_{bd}^{T}p_{bd})^{2}}[(1-6p_{bd}^{T}p_{bd} + (p_{bd}^{T}p_{bd})^{2})I + 8p_{bd}p_{bd}^{T} - 4(1-p_{bd}^{T}p_{bd})[p_{bd}^{\times}]] \tag{33}$$

The Second Kind of Relative Attitude Kinematics Equations

The relative kinematics equations are dictated by the definition of the relative angular velocity. Using the following (alternate) definition of angular velocity

$$\Delta\omega=\omega_b-\omega_d$$

an alternate set of relative attitude kinematics are developed. These may be defined as the alternate (second) form of relative attitude kinematics equations. It should be noted that in this case, the subscript of relative attitude are replaced by the use of Δ.

Quaternion

Using the new notation, the first kind of relative kinematics equation in Quaternion can be written as

$$\frac{d\Delta\bar{q}}{dt}=M\omega_{bd}^{b} \tag{35}$$

where

$$M=\frac{1}{2}\begin{bmatrix}[\Delta q^{\times}]+\Delta q_4 I\\ -\Delta q^{T}\end{bmatrix} \tag{36}$$

Since

$$\omega_{bd}=(\omega_b-R_{BD}\omega_d)=\Delta\omega+(I-R_{BD})\omega_d \tag{37}$$

the relative attitude kinematics Eq.(35) can be written as

$$\frac{d\Delta\bar{q}}{dt}=M\Delta\omega+M(I-R_{BD})\omega_d \tag{38}$$

where

$$R_{BD}=(\Delta q_4^2-\Delta q^{T}\Delta q)I+2\Delta q\Delta q^{T}-2\Delta q_4[\Delta q\times] \tag{39}$$

Because

$$M(I-R_{BD})=\frac{1}{2}\begin{bmatrix}[\Delta q^{\times}]+\Delta q_4 I\\ -\Delta q^{T}\end{bmatrix}[(1-\Delta q_4^2+\Delta q^{T}\Delta q)I-2\Delta q\Delta q^{T}+2\Delta q_4[\Delta q\times]] \tag{40}$$

Considering the following equality relationships for any 3 dimensional vector p:

$$[p^{\times}][p^{\times}]+p^{T}pI-pp^{T}=0_{3\times3}\ , \qquad p^{T}ppp^{T}=pp^{T}pp^{T} \tag{41}$$

$$pp^{T}[p^{\times}]=0_{3\times3}, \qquad [p^{\times}]pp^{T}=0_{3\times3} \tag{42}$$

Eq.(40) can be simplified as follows:

$$M(I-R_{BD}) = \begin{bmatrix}[\Delta q^{\times}] \\ 0_{1\times3}\end{bmatrix} \tag{43}$$

Finally Eq.(38) can be simplified to obtain the second kind of relative attitude kinematics equations in Quaternion :

$$\frac{d\Delta\bar{q}}{dt} = \frac{1}{2}\begin{bmatrix}[\Delta q^{\times}]+\Delta q_4 I \\ -\Delta q^{T}\end{bmatrix}\Delta\omega+\begin{bmatrix}[\Delta q^{\times}] \\ 0_{1\times3}\end{bmatrix}\omega_d \tag{44}$$

Relative angular velocity:

$$\Delta\omega = 2(\Delta q_4\Delta\dot{q}-\Delta\dot{q}_4\Delta q-\Delta q\times\Delta\dot{q})+(R_{BD}(\Delta\bar{q})-I)\omega_d \tag{45}$$

Attitude transformation matrix:

$$R_{BD}(\Delta\bar{q}) = (\Delta q_4^2-\Delta q^{T}\Delta q)I+2\Delta q\Delta q^{T}-2\Delta q_4[\Delta q\times] \tag{46}$$

Rodrigues Parameters

Using new notation, the relative attitude kinematics equations can be written as

$$\frac{d\Delta g}{dt} = M\Delta\omega+M(I-R_{BD})\omega_d \tag{47}$$

where

$$R_{BD} = \frac{1}{(1+\Delta g^T \Delta g)}((1-\Delta g^T \Delta g)I+2\Delta g\Delta g^T-2[\Delta g^\times]) \quad \textbf{(48)}$$

Applying equality relationship Eq.(41-42) simplifies the equation as

$$M(I-R_{BD}) = \frac{1}{2}([\Delta g^\times] + \Delta g\Delta g^T + I)(I - \frac{1}{(1+\Delta g^T \Delta g)}((1-\Delta g^T \Delta g)I+2\Delta g\Delta g^T-2[\Delta g^\times])) \quad \textbf{(49)}$$
$$= [\Delta g^\times]$$

Substituting Eq.(49) into Eq.(47), the second kind of relative attitude kinematics equation in Rodrigues parameters are obtained:

$$\frac{d\Delta g}{dt} = \frac{1}{2}([\Delta g^\times] + \Delta g\Delta g^T + I)\Delta\omega+[\Delta g^\times]\omega_d \quad \textbf{(50)}$$

Relative velocity:

$$\Delta\omega = \frac{2}{(1+\Delta g^T \Delta g)}(I-[\Delta g^\times])\Delta\dot{g}+(R_{BD}(\Delta g)-I)\omega_d \quad \textbf{(51)}$$

Attitude transformation matrix:

$$R_{BD}(\Delta g) = \frac{1}{(1+\Delta g^T \Delta g)}((1-\Delta g^T \Delta g)I+2\Delta g\Delta g^T-2[\Delta g\times]) \quad \textbf{(52)}$$

Modified Rodrigues Parameter

In a similar way, the second kind of relative attitude kinematics equations in modified Rodrigues parameters can be written as

$$\frac{d\Delta p}{dt} = \frac{1}{4}[(1-\Delta p^T \Delta p)I+2[\Delta p^\times]+2\Delta p\Delta p^T]\Delta\omega+[\Delta p^\times]\omega_d \quad \textbf{(53)}$$

Relative velocity:

$$\Delta\omega = \frac{4}{(1+\Delta p^T \Delta p)^2}[(1-\Delta p^T \Delta p)+2\Delta p\Delta p^T-2[\Delta p^\times]]\ \Delta\dot{p}+[R_{BD}(\Delta p)-I]\omega_d \quad \textbf{(54)}$$

Attitude transformation matrix:

$$R_{BD}(\Delta p) = \frac{1}{(1+\Delta p^T \Delta p)^2}[(1-6\Delta p^T \Delta p+(\Delta p^T \Delta p)^2)I + 8\Delta p \Delta p^T - 4(1-\Delta p^T \Delta p)[\Delta p^\times]] \quad (55)$$

RELATIVE ATTITUDE DYNAMICS EQUATIONS

The First Kind of Relative Attitude Dynamics Equations

It is assumed that $\underline{J}_b(t)$, $\underline{\omega}_b(t)$, $\underline{h}_b(t)$ and $\underline{L}_b$ are the inertial dyadics, angular velocity, angular momentum and the applied angular moment of the chase satellite. Similarly, $\underline{J}_d(t)$, $\underline{\omega}_d(t)$, $\underline{h}_d(t)$ and $\underline{L}_d$ are the corresponding items for the target satellite. The relative angular velocity of the chase satellite body reference frame $\mathcal{F}_B$ with respect to the target satellite reference frame $\mathcal{F}_D$ is defined as follows:

$$\underline{\omega}_{bd}=\underline{\omega}_b-\underline{\omega}_d = \mathcal{F}_B^T \omega_{bd}^b \qquad \omega_{db}^b=(\omega_d-R_{DB}\omega_b) \quad (56)$$

The relative attitude dynamics equation in pursuer satellite body reference frame $\mathcal{F}_B$ developed in Ref.1 is

$$\begin{aligned} &J_b\dot{\omega}_{bd}^b+[(\omega_{bd}^b)^\times]J_b\omega_{bd}^b+[(R_{BD}\omega_d)^\times]J_b\omega_{bd}^b+[(\omega_{bd}^b)^\times]J_bR_{BD}\omega_d-J_b[(\omega_{bd}^b)^\times]R_{BD}\omega_d \\ &= L_b-R_{BD}L_d-R_{BD}(\Delta J_d\dot{\omega}_d+[\omega_d^\times]\Delta J_d\omega_d) \end{aligned} \quad (57)$$

where

$$J_d\dot{\omega}_d + [\omega_d^\times]J_d\omega_d = L_d \qquad \Delta J_d = R_{DB}J_bR_{BD}-J_d \quad (58)$$

The Second Kind of Relative Attitude Dynamics Equations

Using the alternate definition of relative angular velocity:

$$\Delta\omega=\omega_b-\omega_d \quad (59)$$

The relationship between ω_{bd} and $\Delta\omega$ (the two definitions of relative angular velocity) is

$$\omega_{bd}^b=\Delta\omega+(I-R_{Bd})\omega_d \quad (60)$$

Considering Eq.(59) and (60), the left hand terms of Eq.(57) can be developed as:

$$J_b\dot{\omega}_{bd}^b=J_b\Delta\dot{\omega}+J_b(I-R_{BD})\dot{\omega}_d+J_b[\Delta\omega^\times]R_{BD}\omega_d+J_b[\omega_d^\times]R_{BD}\omega_d \quad (61)$$

$$[(\omega_{bd}^{b})^{\times}]J_b\omega_{bd}=[\Delta\omega^{\times}]J_b\Delta\omega+[\omega_d^{\times}]J_b\Delta\omega-[(R_{BD}\omega_d)^{\times}]J_b\Delta\omega+[\Delta\omega^{\times}]J_b\omega_d-[\Delta\omega^{\times}]J_bR_{BD}\omega_d$$
$$+[\omega_d^{\times}]J_b\omega_d-[\omega_d^{\times}]J_bR_{BD}\omega_d-[(R_{BD}\omega_d)^{\times}]J_b\omega_d+[(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d \quad \textbf{(62)}$$

$$[(\omega_{bd}^{b})^{\times}]J_bR_{BD}\omega_d=[\Delta\omega^{\times}]J_bR_{BD}\omega_d+[\omega_d^{\times}]J_bR_{BD}\omega_d-[(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d \quad \textbf{(63)}$$

$$[(R_{BD}\omega_d)^{\times}]J_b\omega_{bd}^{b}=[(R_{BD}\omega_d)^{\times}]J_b\Delta\omega+[(R_{BD}\omega_d)^{\times}]J_b\omega_d-[(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d \quad \textbf{(64)}$$

$$J_b[(\omega_{bd}^{b})^{\times}]R_{BD}\omega_d=J_b[\Delta\omega^{\times}]R_{BD}\omega_d+J_b[\omega_d^{\times}]R_{BD}\omega_d-J_b[(R_{BD}\omega_d)^{\times}]R_{BD}\omega_d \quad \textbf{(65)}$$

Substituting Eqs. (61-65) into Eq.(57) and re-arranging

$$J_b\Delta\dot{\omega}+[\Delta\omega^{\times}]J_b\Delta\omega+[\Delta\omega^{\times}]J_b\omega_d=J_b(R_{BD}-I)\dot{\omega}_d-[\omega_d^{\times}]J_b\Delta\omega-[\omega_d^{\times}]J_b\omega_d$$
$$+\ [(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d+L_b-R_{BD}L_d-R_{BD}(\Delta J_d\dot{\omega}_d+[\omega_d^{\times}]\Delta J_d\omega_d) \quad \textbf{(66)}$$

where

$$J_d\dot{\omega}_d\ +\ [\omega_d^{\times}]J_d\omega_d\ =\ L_d \quad \textbf{(67)}$$

These are the second kind of relative attitude dynamics equation. The two equations may be summarized in terms of a relative attitude dynamics theorem:

Theorem of Relative Attitude Dynamics

If the relative angular velocity is defined as

$$\omega_{bd}^{b}=\omega_b-R_{BD}\omega_d \quad \textbf{(68)}$$

then the first kind of relative attitude dynamics equation in the chase satellite body reference frame $\mathcal{F}_B$ is

$$J_b\dot{\omega}_{bd}^{b}+[(\omega_{bd}^{b})^{\times}]J_b\omega_{bd}^{b}+[(\omega_{bd}^{b})^{\times}]J_bR_{BD}\omega_d$$
$$=J_b[(\omega_{bd}^{b})^{\times}]R_{BD}\omega_d-[(R_{BD}\omega_d)^{\times}]J_b\omega_{bd}^{b}+L_b-R_{BD}L_d-R_{BD}(\Delta J_d\dot{\omega}_d+[\omega_d^{\times}]\Delta J_d\omega_d) \quad \textbf{(69)}$$

If the relative angular velocity is defined as

$$\Delta\omega=\omega_b-\omega_d \tag{70}$$

then the second kind of relative attitude dynamics equation in the $\mathcal{F}_B$ frame is

$$\begin{aligned} J_b\Delta\dot{\omega}+[\Delta\omega^\times]J_b\Delta\omega+[\Delta\omega^\times]J_b\omega_d=J_b(R_{BD}-I)\dot{\omega}_d-[\omega_d^\times]J_b\Delta\omega-[\omega_d^\times]J_b\omega_d \\ +\ [(R_{BD}\omega_d)^\times]J_bR_{BD}\omega_d+L_b-R_{BD}L_d-R_{BD}(\Delta J_d\dot{\omega}_d+[\omega_d^\times]\Delta J_d\omega_d) \end{aligned} \tag{71}$$

where

$$\Delta J_d\ =\ R_{DB}J_bR_{BD}-J_d \tag{72}$$

$$J_d\dot{\omega}_d\ +\ [\omega_d^\times]J_d\omega_d\ =\ L_d \tag{73}$$

APPLICATION TO LARGE ATTITUDE ANGLE ACQUISITION AND TRACKING CONTROL

An application of relative attitude kinematics and dynamics equations developed in this paper is the large angle acquisition and tracking control. This section provides the simulations of such an application. The advantage in using relative attitude is that the tracking control problem is converted into a regulator problem, simplifying the control system design.

It is assumed that the second kind of relative kinematics and dynamics equations in modified Rodrigues parameters are used for the acquisition and tracking control. The equations are as follows:

Relative kinematics equations:

$$\frac{d\Delta p}{dt}\ =\ M\Delta\omega+[\Delta p^\times]\omega_d \tag{74}$$

where

$$M\ =\ \frac{1}{4}[(1-\Delta p^{\,T}\Delta p)I+2[\Delta p^\times]+2\Delta p\Delta p^{\,T}] \tag{75}$$

The relative dynamics equations:

$$J\Delta\dot{\omega}+\Delta\omega\times J\omega_d+\Delta\omega\times J\Delta\omega=L \tag{76}$$

where

$$L = J_b(R_{BD} - I)\dot{\omega}_d - [\omega_d^{\times}]J_b\Delta\omega - [\omega_d^{\times}]J_b\omega_d$$
$$+ [(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d + L_b - R_{BD}L_d - R_{BD}(\Delta J_d\dot{\omega}_d + [\omega_d^{\times}]\Delta J_d\omega_d) \quad (77)$$

The Lyapunov function V can be selected to be

$$V = K_p\Delta p^{\,T}\Delta p + \frac{1}{2}\Delta\omega^T J\Delta\omega \quad (78)$$

where K_p is a positive constant. The time-derivative of Lyapunov function V is

$$\dot{V} = 2K_p\Delta p^{\,T}\Delta\dot{p} + \Delta\omega^T J\Delta\dot{\omega} = \Delta\omega^T(2K_pM^{\,T}\Delta p + L) \quad (79)$$

If

$$L = -2K_pM^{\,T}\Delta p - K_dJ\Delta\omega \quad (80)$$

then

$$\dot{V} = -K_d\Delta\omega^T J\Delta\omega \leq 0 \quad (81)$$

where K_d is a positive constant. The closed loop system dynamics equation is

$$J\Delta\dot{\omega} + \Delta\omega\times J\omega_d + \Delta\omega\times J\Delta\omega = -K_dJ\Delta\omega - 2K_pM^{\,T}\Delta p \quad (82)$$

Equation (81) implies that $V(t) \leq V(0)$, and therefore, that Δp and $\Delta\omega$ are bounded. In addition, from Eq.(81)

$$\ddot{V} = -2K_d\Delta\omega J\Delta\dot{\omega} \quad (83)$$

so it can be seen that d^2V/dt^2 is bounded. Hence dV/dt is uniformly continuous [9]. Application of Barbalat's lemma [9] then indicates that $\Delta\omega \to 0$ as $t \to \infty$. Considering the closed loop equation (82), it can be obtained that $\Delta p \to 0$ as $t \to \infty$. Therefore, we have that $(\Delta p, \Delta\omega) \to (0,0)$ as $t \to \infty$. It means that the nonlinear control law given by Eq.(84)

$$L_b = R_{BD}L_d - J_b(R_{BD} - I)\dot{\omega}_d + [\omega_d^{\times}]J_b\Delta\omega + [\omega_d^{\times}]J_b\omega_d$$
$$-[(R_{BD}\omega_d)^{\times}]J_bR_{BD}\omega_d + R_{BD}(\Delta J_d\dot{\omega}_d + [\omega_d^{\times}]\Delta J_d\omega_d) - 2K_pM^{\,T}\Delta p - K_dJ\Delta\omega \quad (84)$$

is a global asymptotically stable control law for the system given by Eqs.(74-77).

SIMULATIONS

The Lyapunov nonlinear attitude control law has been used for large attitude angle acquisition and tracking control for the EO-1/LandSat 7 formation. For this simulation the control is implemented using full state feedback only. The simulations using measurement output feedback will be presented in future papers.

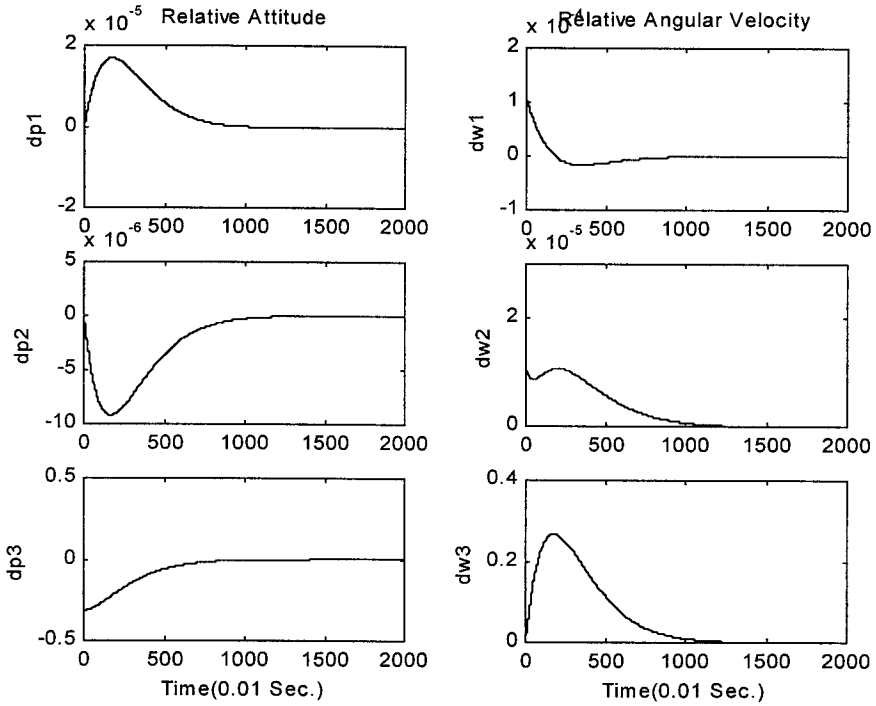


Figure 1. Large Attitude Angle Acquisition and Tracking Control for EO-1/LandSat7 Formation using Full State Feedback

The initial conditions for this full state feedback simulations are as follows:

Maneuver angle = 70 deg.
Target satellite (Landsat 7) angular velocity = $(0\ 0\ \omega_T)^T$
Pursuer satellite (EO-1) angular velocity = $(0.1\omega_T\ \ 0.01\omega_T\ \ 2\omega_T)^T$
ω_T =angular velocity of the 705 km circular orbit

[dp1, dp2, dp3] in Fig 1 are the three components of the relative attitude represented in the relative modified Rodrigues parameters, and [dw1, dw2, dw3] are the three components of the relative angular velocity (rad/sec). In this simulation, it is assumed that the attitude state vector is given and there no attitude measurement information available for feedback control.

CONCLUSIONS

In this paper the alternate relative attitude kinematics and dynamics equations are developed for the various attitude parametric representations. Compared with the first kind of relative attitude kinematics and dynamics equations (Ref.1), this has the advantage of being simpler. These developments will find ready application in the problems of relative attitude determination and control, and will be very useful for spacecraft formation control and relative navigation. As an example of such application, the Lyapunov nonlinear control law for large attitude angle acquisition and tracking has been developed and simulated for the EO-1/LandSat 7 formation. This simulation

implemented the full state feedback control.

REFERENCES

1. Xing, G.Q., and Parvez, S.A., "Relative Attitude Kinematics and Dynamics Equations and Its Applications to Large Attitude Angle Tracking Maneuvers," Proceeding of the 1999 Space Control Conference, Ed. L.B. Spence, MIT Lincoln Laboratory, Lexington, MA, April, 1999.

2. Wen, J.T., and Kreutz-Delgado, K. K., "The Attitude Control Problem," *IEEE Transactions On Automatic Control*, Vol. 36, No.10, Oct. 1991, pp.1148-1162.

3. Fjellstad, O., and Fossen, T.I., "Comments on the Attitude Control Problem," *IEEE Transactions on Automatic Control*, Vol. 36, No. 3, March 1994, pp. 699-700.

4. Wang, P. K.C. and Hadaegh. F, "Coordination and Control of Multiple Micro Spacecraft Moving in Formation," *The Journal of the Astronautical Sciences*, Vol. 44, No. 3, July-Sept. 1996, pp. 315-355.

5. Isidori, A., *Nonlinear Control System,* Third Edition, Springer-Verlag London Ltd., 1995.

6. Shuster, M..D., "A Survey of Attitude Representations," *The Journal of the Astronautical Sciences*, Vol. 41, No. 4, Oct-Dec., 1993, pp. 439-517.

7. Peter C. Hughes, *Spacecraft Attitude Dynamics*, John Wiley & Sons, Inc., 1986.

8. Wertz, J. R., *Spacecraft Attitude Determination and Control*, Kluwer Academic Publishers, Dordrecht/Boston/London, 1980.

9. Slotine, J.E. and Li, W., *Applied Nonlinear Control*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.

EVALUATION OF RELATIVE NAVIGATION ALGORITHMS FOR FORMATION-FLYING SATELLITES

David Kelbel, Taesul Lee, and Anne Long
Computer Sciences Corporation
Lanham-Seabrook, Maryland USA 20706

J. Russell Carpenter and Cheryl Gramling
NASA Goddard Space Flight Center
Greenbelt, Maryland USA 20771

ABSTRACT

Goddard Space Flight Center is currently developing advanced spacecraft systems to provide autonomous navigation and control of formation flyers. This paper discusses autonomous relative navigation performance for formations in eccentric, medium and high-altitude Earth orbits using Global Positioning System (GPS) Standard Positioning Service (SPS) and intersatellite range measurements. The performance of several candidate relative navigation approaches is evaluated. These analyses indicate that the relative navigation accuracy is primarily a function of the frequency of acquisition and tracking of the GPS signals. A relative navigation position accuracy of 0.5 meters root-mean-square (RMS) can be achieved for formations in medium-attitude eccentric orbits that can continuously track at least one GPS signal. A relative navigation position accuracy of better than 75 meters RMS can be achieved for formations in high-altitude eccentric orbits that have sparse tracking of the GPS signals. The addition of round-trip intersatellite range measurements can significantly improve relative navigation accuracy for formations with sparse tracking of the GPS signals.

1 – INTRODUCTION

Formation-flying techniques and satellite autonomy will revolutionize space and Earth science missions and enable many small, inexpensive satellites to fly in formation and gather concurrent science data. The Guidance, Navigation, and Control Center (GNCC) at Goddard Space Flight Center (GSFC) has successfully developed high-accuracy autonomous satellite navigation systems using the National Aeronautics and Space Administration's (NASA's) space and ground communications systems and the Global Positioning System (GPS) (References 1 and 2). Recently, the GNCC has leveraged this experience to develop advanced spacecraft systems that provide autonomous navigation and control of formation flyers.

To support this effort, the GNCC is assessing the relative navigation accuracy achievable for proposed formations using GPS and intersatellite range measurements. Several universities and corporations are developing GPS transceivers that support this tracking concept for NASA and the Air Force Research Laboratory; these include Johns Hopkins Applied Physics Laboratory, International Telephone and Telegraph, Honeywell, Motorola, Jet Propulsion Laboratory, Cincinnati Electronics, and Stanford University (Reference 3). This paper evaluates the performance of several candidate relative navigation algorithms for missions with more than two vehicles maintaining a relatively tight formation, in a relatively eccentric orbit.

High-fidelity simulations were performed to study two proposed formation-flying missions. One is a mission designed to study the Earth's aurora. This medium-altitude formation consists of four satellites maintained in Earth orbits with approximately 500x7000 kilometer altitudes. To support autonomous planning of the formation-flying maneuvers to maintain the initial 10-kilometer separation at apogee, the total relative position and velocity accuracy must be about 100 meters and 20 millimeters per second, respectively. Later in the mission, when the separation is reduced to about 500 meters, the total relative position accuracy requirement reduces to 5 meters.

The other formation-flying mission is the initial phase of the Magnetospheric Multiscale (MMS) mission. The MMS formation consists of four satellites in orbits of approximately 1.2x12 Earth radii. The intersatellite separation at apogee varies from 10 kilometers to 0.1 Earth-radii over the life of the mission. The absolute position knowledge requirement is 100 kilometers and the intersatellite position knowledge requirement is 1 percent of the actual separation.

Previously, the authors investigated the relative navigation accuracy that could be achieved for the 500x7000 kilometer formation by differencing independently-estimated state vectors (Reference 4). That analysis indicated that an autonomous relative navigation position accuracy of 1 meter root-mean-square (RMS) can be achieved by differencing high-accuracy filtered solutions if only measurements from common GPS space vehicles (SVs) are used in the independently-estimated solutions. This paper quantifies the relative navigation accuracy improvements achievable using a high-accuracy multi-satellite filter to simultaneously estimate the satellite state vectors. Improvements to be achieved through the estimation of GPS biases, differencing of GPS measurements, and addition of intersatellite range measurements are evaluated. This research was supported by the NASA Space Operations and Management Office and NASA Research Announcement 98-OSS-10: Technology Development for Explorer Missions.

2 – RELATIVE NAVIGATION ALGORITHMS

The most straightforward relative navigation approach computes the satellite relative positions by differencing the absolute position vectors of each satellite in the formation. The state differencing method can be used to support decentralized, centralized, or hierarchical formation control strategies. Figure 1 illustrates one possible configuration for using this approach to support decentralized control of a distributed satellite formation. In this case, each satellite independently computes its absolute state vector using GPS and possibly intersatellite measurements and transfers this state vector via an intersatellite communications link to every other satellite in the formation. Each satellite computes its relative position to the other satellites by state vector differencing and uses this relative state to plan and execute formation maintenance maneuvers to maintain its desired position within the formation. Reference 5 discusses a recent investigation of decentralized formation control strategies.

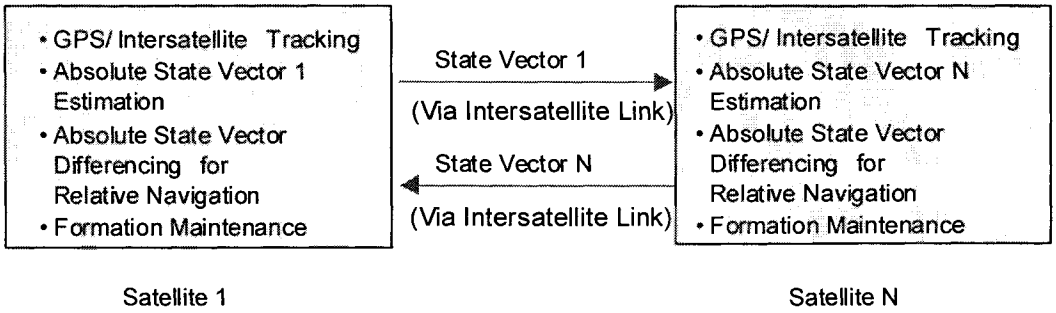


Figure 1. Independent State Vector Differencing Configuration with Decentralized Formation Control

The absolute state vector computation can be performed using either an instantaneous point solution method or a real-time filtered algorithm. The analysis presented in Reference 4 indicates that relative navigation by differencing absolute state vectors obtained using the point solution method is not suitable for continuous real-time navigation of formations of satellites in orbits for which fewer than six GPS SVs are visible during significant portions of the orbit. In addition, because point solutions do not provide accurate velocity estimates, they are not suitable for applications in which state vector information must be predicted ahead in time, e.g., to support autonomous maneuver planning.

A real-time filtered algorithm, such as that implemented in the GPS Enhanced Orbit Determination (GEODE) software (Reference 6), reduces the impact of the measurement errors by using an extended Kalman filter in conjunction with a high-fidelity orbital dynamics model. In addition to the differencing of independently-estimated state vectors, the real-time filtered approach can support more complex relative navigation approaches that simultaneously estimate the state vectors of all satellites in the formation. Figure 2 illustrates one possible configuration for using the simultaneous estimation approach to support decentralized control of a distributed satellite formation. In this configuration, each satellite computes the absolute state vectors of all satellites in the

formation using GPS measurements to all satellites and possibly intersatellite measurements. Each satellite transfers its GPS measurements via an intersatellite communications link to every other satellite in the formation. Each satellite computes its relative position to the other satellites by state vector differencing and uses this relative state to plan and execute formation maintenance maneuvers to maintain its desired position within the formation.

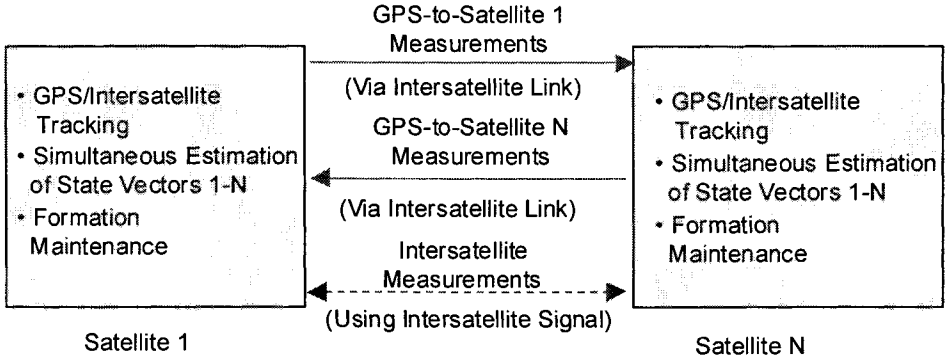


Figure 2. Simultaneous State Vector Estimation Configuration with Decentralized Formation Control

Table 1 lists the characteristics of the navigation algorithms evaluated in this paper. In this table, "local" refers to the satellite on which a specific copy of the navigation software resides. "Remote" refers to the other satellites in the formation. In the decentralized satellite configuration, each satellite hosts the same navigation software but does not necessarily process the same set of measurements.

Table 1. Navigation Algorithms

Navigation Algorithm	Absolute State Vectors Estimated	Measurement Types	Data Transferred From Remotes to Local
Independent	Local satellite	GPS pseudorange to local	Remote satellite state vectors
Simultaneous with standard GPS	Local and remote satellites	GPS pseudorange to local and remotes	GPS measurements for remotes
Simultaneous with standard GPS and bias estimation	Local and remote satellites Random walk measurement biases for each GPS SV	GPS pseudorange to local and remotes	GPS measurements for remotes
Simultaneous with differenced GPS	Local and remote satellites	Singly differenced GPS pseudorange between local and remotes	GPS measurements for remotes
Simultaneous with standard and differenced GPS	Local and remote satellites	GPS pseudorange to the local Singly differenced GPS pseudorange between local and remotes	GPS measurements for remotes
Simultaneous with standard GPS and 1-way intersatellite	Local and remote satellites Random walk measurement biases for each remote to local intersatellite link	GPS pseudorange to local and remotes Intersatellite pseudorange from remote to local	GPS measurements for remotes
Simultaneous with standard GPS and round-trip intersatellite	Local and remote satellites	GPS pseudorange to local and remotes Intersatellite range from local-to-remote-to-local	GPS measurements for remotes

3 – PERFORMANCE SIMULATION PROCEDURE

To quantify the level of relative navigation performance that is achievable for formations of satellites in medium-altitude Earth orbits (MEOs) or high-altitude Earth orbits (HEOs), realistic simulations were performed for two representative missions. The MEO formation studied consists of four satellites maintained in Earth orbits at an

inclination of 80 degrees with altitudes of approximately 500 kilometers at perigee by 7000 kilometers at apogee. The orbital period is approximately 3 hours. All satellites have nearly identical surface areas of 0.6613 meters^2 and masses of 200 kilograms. The intersatellite separations in this tetrahedral formation range from 10 kilometers at apogee to 30 kilometers at perigee, with three satellites in the same orbit plane and one out of plane.

The HEO formation consists of four satellites maintained in 1.2-Earth-radii by 12-Earth-radii orbits at an inclination of 10 degrees, which is similar to the formation proposed for the first phase of the Magnetospheric Mapping mission. The orbital period in this mission phase is 1 day. All satellites have nearly identical surface areas of 1.12 meters^2 and masses of 220 kilograms. The intersatellite separations in this tetrahedral formation range from 10 kilometers at apogee to 150 kilometers at perigee, with planar separations of less than 0.1 degree.

3.1 – Measurement Simulation

Realistic GPS pseudorange and remote-to-local and round-trip intersatellite range measurements were simulated for each MEO/HEO satellite using high-fidelity truth ephemerides and the measurement simulation options listed in Table 2. The truth ephemerides were generated using the Goddard Trajectory Determination System (GTDS) with the high-accuracy force model, which included a Joint Goddard Model (JGM) for nonspherical gravity forces, Jet Propulsion Laboratory Definitive Ephemeris 200 for solar and lunar gravitational forces, atmospheric drag, and solar radiation pressure forces. GTDS is the primary orbit determination program used for operational satellite support at GSFC.

Table 2. GPS Measurement Simulation Parameters

Parameter	Value
Measurement data rate	GPS: every 60 seconds from all visible GPS SVs Intersatellite: every 60 seconds from all remote SVs for MEO every 60 seconds from all remote SVs for 10 minutes per hour for HEO
GPS SV ephemerides	Broadcast ephemerides for June 21-26, 1998
GPS SV characteristics:	
Ephemeris and clock errors	2 meter (1-sigma)
Transmitting antenna pattern	GPS L-band pattern, modeled from 0 to 90 degrees down from boresight
Transmitted power	29.8 dB-watts in maximum gain direction
User antenna models:	Hemispherical antenna : Maximum gain : 3.5 dBic for MEO, 4.9 dBic for HEO Horizon mask: 90 degrees from boresight
Visibility constraints	• Earth blockage with 500 km altitude tropospheric mask • GPS SV transmitting antenna beamwidth and receiving antenna horizon masks • Received signal-to-noise ratio above tracking threshold
GPS receiver characteristics	• Receiver noise figure: 2.9 dB for MEO, 4.9 dB for HEO • System noise temperature: Earth-point antenna: 300K Otherwise: 190K • 24-channels • 35 dB-Hertz receiver acquisition threshold for MEO • 30 dB-Hertz receiver acquisition threshold for HEO
Ionospheric delays	MEO: 26 meters at 500 km height 5 meters at 1000 km HEO: 32 meters at 400 km height 3 meters at 1000 km
Receiver clock bias white noise spectral density	9.616×10^{-20} seconds^2 per second
Receiver clock drift rate white noise spectral density	1.043×10^{-27} seconds^2 per seconds^3
Random measurement errors	GPS pseudorange: 2 meters (1-sigma) Intersatellite pseudorange: 2 meters (1-sigma)

The GPS constellation configuration was based on the GPS broadcast messages for the epoch date. The GPS signal strength at the GPS receiver's location was modeled assuming the nominal GPS Block II signal antenna pattern (including both the main and side lobes). Each MEO satellite had one hemispherical GPS antenna, with a zenith-pointing boresight. The MEO's GPS receiver's signal-to-noise ratio acquisition threshold was 35 dB-hertz, consistent with the performance of most space receivers. Each HEO satellite had identical hemispherical antennas, pointing in the zenith and nadir directions. The HEO formation pseudorange measurements were created based on a GPS receiver with a reduced acquisition threshold of 30 dB-Hertz. The GPS SV signal attenuation model that was used provides realistic signal acquisition predictions (Reference 7). The number of simultaneous measurements was not restricted.

GPS SV ephemeris and clock errors were applied at a 2-meter (1-sigma) level, using the Lear4 autoregressive integrated moving average time series model (Reference 8). When the satellites were below the Earth's ionosphere, ionospheric delays were modeled as a function of the height of the signal path above the Earth, which was based on ionospheric delays computed for the test orbit using the Bent ionospheric model available in GTDS. In the case of signals with long paths "over-the-Earth-limb", ionospheric delays were modeled using an exponential function of the height of ray path (HORP) above the Earth, which was based on ionospheric delays computed for each test orbit using the Bent ionospheric model. Receiver clock noise was simulated assuming a highly-stable crystal oscillator with a 1-second root Allan variance of $0.16(10^{-9})$. A twice-integrated random walk model, which is based on Reference 9, was used to simulate the clock bias and clock drift noise contributions to the GPS and intersatellite measurement errors.

Intersatellite pseudorange measurements have potentially large biases due to the transmitter's and receiver's clock biases. Several strategies could be used to reduce or eliminate these biases. To reduce these biases, each transmitting satellite could estimate its clock offset from GPS time and frequency offset from nominal based on GPS measurements and steer its clock to be synchronized to within 100 nanoseconds (30 m) with GPS time. Summation of the remote-to-local and local-to-remote pseudoranges would cancel the clock bias contributions. Measurement of the round-trip intersatellite range eliminates the clock bias contributions. In the simulations reported in this paper, the remote-to-local (1-way) interstatellite measurements are biased by the difference between the simulated transmitter and receiver clock biases. The simulated round-trip intersatellite measurements are unbiased.

3.2 – Navigation Performance Analysis Procedure

Monte Carlo simulations were performed for each formation to quantify the expected distribution in the absolute and relative solution errors as a function of variations in the random measurement errors. The following ensemble error statistics were accumulated for the ensemble of navigation solutions obtained by processing 25 sets of simulated GPS pseudorange measurements that were created by varying the random number seeds used for the GPS ephemeris and clock, receiver clock, and random measurement errors:

- The ensemble RMS/maximum error, which is the RMS/maximum of the true error (difference between the estimated and the true state) at each time computed across all Monte Carlo solutions.
- The steady-state time-wise ensemble RMS/maximum error, which is the RMS/maximum of the ensemble true errors computed along the time axis, omitting the initial convergence period.

The extended Kalman filter algorithm available in the GEODE flight software was used to process these measurement sets. The filter was "tuned" by adjusting the process noise parameters and measurement standard deviation to produce an estimated state error root variance that was consistent with the ensemble RMS state error obtained in the Monte Carlo analysis. Table 3 lists the GEODE processing parameters common to all cases. Atmospheric drag and solar radiation pressure forces were included in the state propagation using atmospheric drag and solar radiation pressure coefficients that were offset by 10 percent and 5 percent respectively from the values used in the truth ephemeris generation.

The absolute navigation errors were computed by differencing the truth and estimated absolute state vectors. The estimated relative state vectors were computed by differencing the estimated absolute state vectors for the two satellites. The relative navigation errors were computed by differencing the true relative state vectors and the estimated relative state vectors.

Table 3. GEODE Processing Parameters

Parameter	Value
Nonspherical Earth Gravity model	MEO: 30x30 JGM-2 HEO: 8x8 JGM-2
Solar and lunar ephemeris	High-precision analytical ephemeris
Initial position error in each component	100 meters for MEO, 5000 meters for HEO
Initial velocity error in each component	0.1 meter per second
Initial solar radiation pressure coefficient error	0.07 (5 percent)
Atmospheric drag coefficient error	0.22 (10 percent)
Initial receiver time bias error	100 meters
Initial receiver time bias rate error	0.1 meter per second
Estimated state (local and remote satellites)	• position and velocity • GPS receiver time bias and time bias drift • Atmospheric drag coefficient correction (MEO only) • Intersatellite bias with 1-way measurements
GPS SV ephemerides	Broadcast ephemerides for June 21-26, 1998
Ionospheric editing	500 kilometer minimum ray path height

4 – RELATIVE NAVIGATION PERFORMANCE FOR MEDIUM ALTITUDE FORMATION

This section presents the absolute and relative navigation results for the MEO formation. Figure 3 illustrates the geometry of the MEO spacecraft with respect to the primary beam of a single GPS satellite (ignoring the effects of the differences in inclinations). The MEO lies well below the GPS constellation altitude.

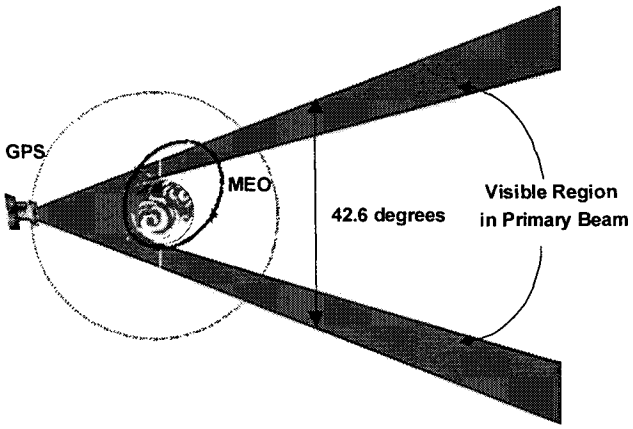


Figure 3. MEO Tracking Geometry

Figure 4 shows the number of GPS SVs visible as a function of time and altitude. The satellite's single zenith-pointing hemispherical antenna considerably limits GPS visibility at high altitudes. The periods of lowest visibility (4 or fewer GPS SVs) occur when the satellites are at altitudes above 5500 kilometers, where the visibility is highly dependent on the exact position of the GPS SVs within each orbit plane. The periods of best visibility with 6 or more visible GPS SVs occur when the satellites are within the main lobe of the GPS signal (i.e. below approximately 3000 kilometers).

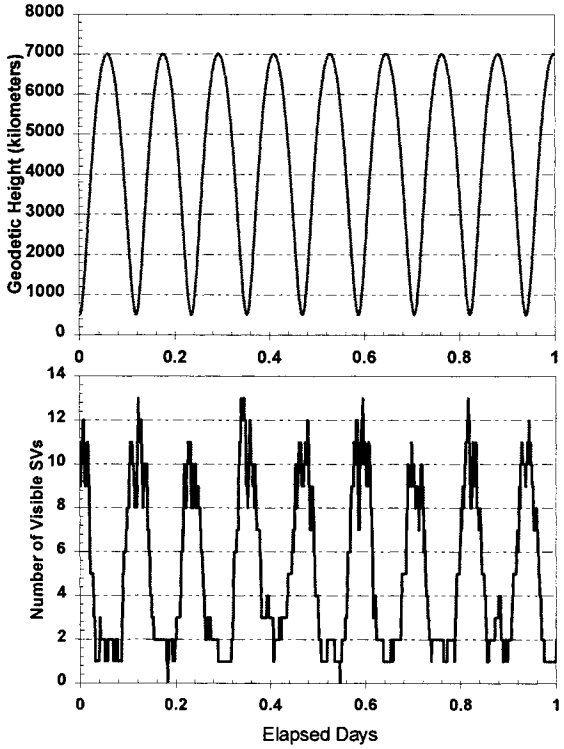


Figure 4. GPS SV Visibility as a Function Time and Altitude for MEO Formation

Figure 5 shows the ensemble RMS and maximum absolute position and clock errors for the local satellite over the 2-day estimation time span followed by a 1-day prediction time span, from a Monte Carlo simulation in which all satellites were independently estimated using standard GPS measurements. The estimation process reached steady state after 3 hours of processing, immediately following the second perigee passage. The variation in the absolute and relative error statistics for all satellites in the formation is less than 2 and 3 percent; respectively. The steady-state (i.e., omitting the initial convergence period) time-wise ensemble RMS of the absolute errors were 3.35 meters position, 1.7 millimeters per second velocity, and 8 meters (0.03 microseconds) clock bias, which were very consistent with the state errors predicted by the estimator. The maximum errors occur following apogee. For the Monte Carlo simulation, the maximum absolute errors encountered were less than 9.3 meters, 6.8 millimeters per second velocity, and 26 meters (0.09 microseconds). During a 1-day prediction using a converged filter solution, the absolute errors remain under 17 meters in position, 17 millimeters per second in velocity, and 1412 meters (4.71 microseconds) in clock bias. The primary sources of the absolute navigation error are dynamic modeling errors, unmodeled ionospheric delay, GPS SV ephemeris and clock errors, and receiver clock errors.

Figure 6 shows the ensemble RMS and maximum relative position error for the local satellite and one remote satellite, from the same Monte Carlo simulation. When the absolute solutions are differenced, the error contributions from correlated measurement and dynamic errors cancel and the relative navigation accuracy is significantly better than the absolute errors. Since the satellites are in tight formation in nearly the same orbits, the dynamic errors are highly correlated, and a large percentage of the dynamic error contribution cancels in all of these cases. In addition, since the satellites are in close formation and track common GPS SVs 99.8 percent of the time, the ionospheric delay and GPS SV ephemeris and clock errors are highly correlated, and a large percentage of these error contributions cancels in all of these cases. Therefore, the uncorrelated measurement noise and receiver clock errors are the primary contributors to the relative navigation error for this formation. For two MEO satellites separated by approximately 10 kilometers, the Monte-Carlo simulations yielded a steady-state time-wise ensemble RMS relative accuracy of approximately 0.43 meters in position and 0.25 millimeters per second in velocity, with maximums below 2.2 meters in position and 1.4 millimeters per second in velocity. The relative accuracy is well within the 100 meter and 20-millimeter-per-second mission requirements for a 10-kilometer separation. During a 1-day prediction using a converged filter solution, the relative errors remain under 9 meters in position and 8.5 millimeters per second in velocity.

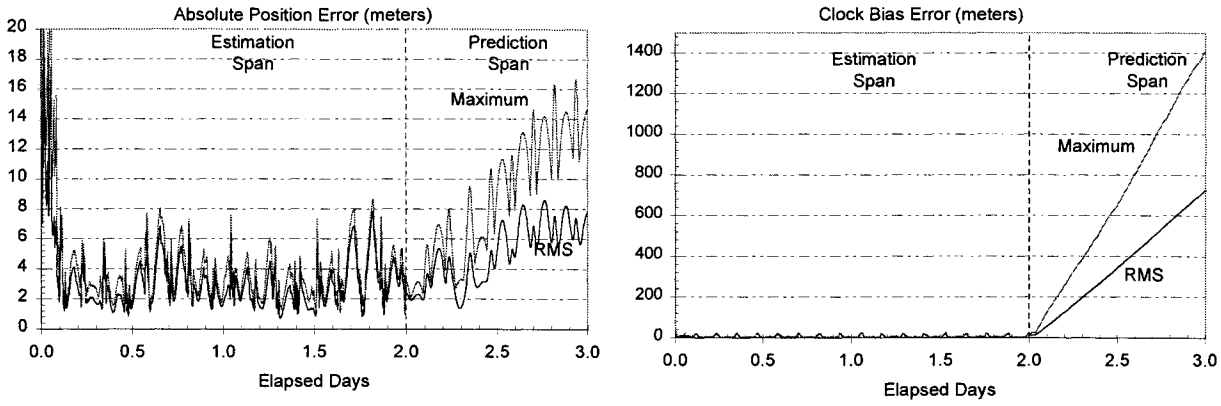


Figure 5. Ensemble Absolute Position and Clock Bias Errors for MEO Formation Using Independently-Estimated Solutions with Standard GPS Measurements

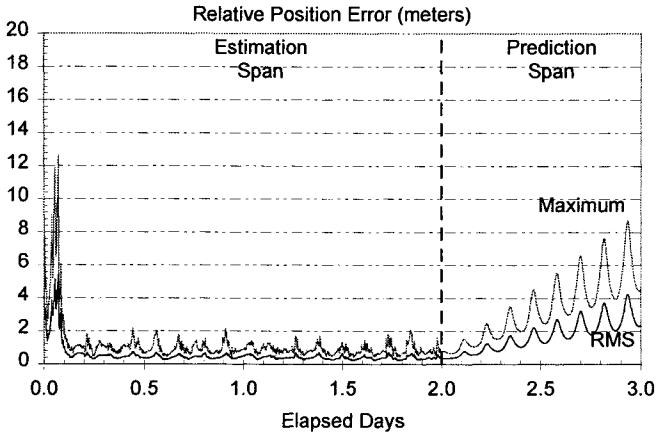


Figure 6. Ensemble Relative Position Errors for MEO Formation Using Independently-Estimated Solutions with Standard GPS Measurements

Simulations were performed to evaluate the relative performance of the navigation algorithms listed in Table 1, using a representative measurement set from the Monte Carlo simulation. Table 4 compares the steady-state absolute error statistics for the local satellite for each of the navigation algorithms evaluated. Table 5 compares the steady-state relative error statistics for the local satellite and one remote satellite for each of the navigation algorithms evaluated. The variation in the absolute error statistics for all satellites in the formation is about 5 percent. The variation in the relative error statistics for all satellites in the formation is less than 30 percent. These comparisons produced the following results:

- Simultaneously estimating the local and remote satellites produced identical absolute and relative results to those obtained by independently estimating each of the satellites when there was no explicit correlation between the satellite state vectors.
- Estimation of a pseudorange bias for each GPS SV can be used to absorb much of the unmodeled ionospheric delay errors producing smoother clock bias estimates. However, due to the high correlation between the pseudorange and clock biases, the clock state improvement is very sensitive to the initial pseudorange bias value that is used and the process noise parameters used for the clock bias and pseudorange bias. When tuned to provide improved clock bias estimates, the absolute navigation position and velocity errors increased significantly and the relative navigation errors increased slightly as compared with the case where the biases were not estimated. In this analysis, the pseudorange bias was modeled as a random walk variable; this model had been selected to reflect the behavior of large Selective Availability

biases. Improved absolute results might be achieved if the pseudorange bias model better reflected the physical characteristics of the expected ionospheric delay.

- When only singly-differenced GPS measurements were processed, the absolute state vector solutions were unstable. The inclusion of standard GPS measurements for only the local satellite, in addition to singly-differenced GPS pseudorange measurements between the remote and local satellites, stabilized the absolute solutions. The absolute navigation accuracy was about 16 centimeters larger for the local satellite and 35 centimeters larger for the remote satellites as compared with results from processing only standard GPS measurements. The relative navigation errors increased by about 30 centimeters, probably due to a reduction in the cancellation of the absolute errors associated with dynamic errors and correlated measurement errors. The magnitude of this increase was found to be sensitive to the relative weighting of the two measurement types, with equal weighting producing the smallest relative error.
- Inclusion of 1-way (remote-to-local) intersatellite pseudorange in addition to standard GPS pseudorange measurements for the remote and local satellites did not improve the absolute navigation results obtained without the 1-way measurements. In this case, the relative navigation errors increased by about 5 percent.
- Inclusion of roundtrip intersatellite range between the local and remote satellites in addition to standard GPS pseudorange measurements for the remote and local satellites produced comparable absolute results to those obtained without the round-trip measurements. However, the relative navigation errors decreased by about 25 percent due to a reduction in the impact of the uncorrelated receiver clock biases associated with the standard GPS measurements.

Table 4. Steady-State Absolute Error Statistics for the MEO Formation

Navigation Algorithm	Position Error (meters)		Velocity Error (millimeters per second)		Clock Error (meters)	
	RMS	Maximum	RMS	Maximum	RMS	Maximum
Independent	3.36	8	1.7	5.8	8.0	25
Simultaneous with standard GPS	3.36	8	1.7	5.8	8.0	25
Simultaneous with standard GPS and bias estimation	13.8	23	5.9	12.6	6.7	20
Simultaneous with standard and differenced GPS	3.52	9	1.7	5.7	8	24
Simultaneous with standard GPS and 1-way intersatellite pseudorange	3.37	8	1.7	5.8	8	25
Simultaneous with standard GPS and round-trip intersatellite range	3.36	8	1.7	5.9	8	25

Table 5. Steady-State Relative Error Statistics for the MEO Formation

Navigation Algorithm	Position Error (meters)		Velocity Error (millimeters per second)	
	RMS	Maximum	RMS	Maximum
Independent	0.42	1.3	0.25	0.93
Simultaneous with standard GPS	0.42	1.3	0.25	0.93
Simultaneous with standard GPS and bias estimation	0.66	2.1	0.34	1.01
Simultaneous with standard and differenced GPS	0.71	2.4	0.40	1.17
Simultaneous with standard GPS and 1-way intersatellite pseudorange	0.45	1.4	0.26	1.06
Simultaneous with standard GPS and round-trip intersatellite range	0.31	0.8	0.18	0.60

5 – RELATIVE NAVIGATION PERFORMANCE FOR HIGH ALTITUDE FORMATION

This section presents the absolute and relative navigation results for the HEO formation. Figure 7 illustrates the geometry of the HEO spacecraft with respect to the primary and first side lobe of the signal of a single GPS SV (ignoring the effects of the differences in inclinations). The formation is above the GPS constellation and outside the primary beam for a large portion of its orbit.

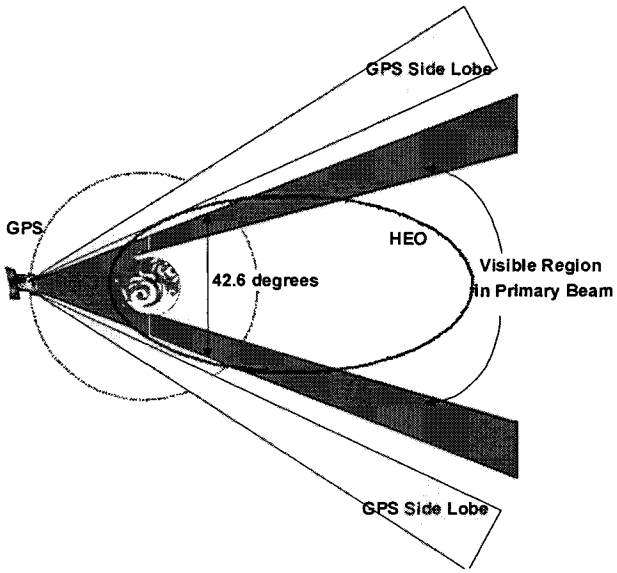


Figure 7. HEO Tracking Geometry

Figure 8 shows the number of GPS SVs that can be acquired and tracked by a GPS receiver on a 1.2x12 Earth-radii HEO, assuming a 30-dB-Hertz receiver acquisition and tracking threshold. This reduced threshold was selected because preliminary analysis performed using a receiver with a standard 35-dB-Hertz threshold did not meet the relative navigation requirements for the initial phase of the MMS mission. The 30-dB-Hertz tracking threshold can be achieved if the receiver employs weak signal tracking strategies to track the weaker signals in the side lobes of the GPS antenna pattern (Reference 7). In this case, the receiver can acquire and track at least one GPS SV about 27 percent of the time and can acquire a maximum number of 22 GPS SVs at perigee. The HEO receiving antenna model consists of two hemispherical antennas, one located on the top face of the satellite and one located on the bottom face. Use of a nadir-pointing high gain GPS antenna would further improve GPS signal acquisition and tracking at high altitudes.

Figure 9 shows the ensemble RMS and maximum absolute position and clock bias errors for the local satellite over the 3.5-day estimation time span followed by a 1-day prediction time span, from Monte Carlo simulations in which all satellites were independently estimated using standard GPS measurements. The estimation process reached steady state immediately following the first perigee passage. The steady-state time-wise ensemble RMS of the absolute errors were 75 meters position, 3.8 millimeters per second velocity, and 49 meters (0.16 microseconds) clock bias, which were consistent with the state errors predicted by the estimator. The maximum errors occur following apogee. The absolute position accuracy is well within the 100-kilometer mission requirement. The maximum absolute errors encountered were less than 365 meters in position, 27 millimeters per second in velocity, and 265 meters (0.88 microseconds) in clock bias. The primary sources of the absolute navigation error are dynamic modeling errors, unmodeled ionospheric delay, and receiver clock errors. During a 1-day prediction using a converged filter solution, the absolute errors remain under 3.3 kilometers in position, 2.3 meters per second in velocity, and 400 meters (1.3 microseconds) in clock bias for all Monte Carlo cases that were run.

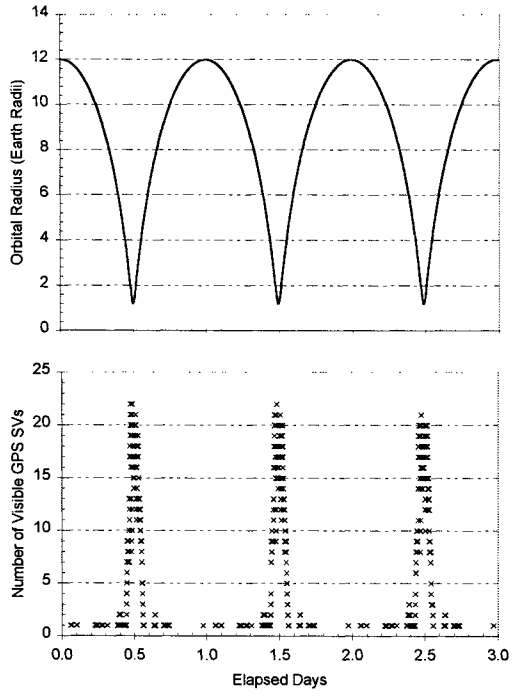


Figure 8. GPS SV Visibility as a Function Time and Altitude for HEO Formation

Figure 10 shows the ensemble RMS and maximum relative position errors for the local satellite and one remote satellite, from Monte Carlo simulations in which all satellites were independently estimated using standard GPS measurements. The dynamic modeling errors and ionospheric errors will nearly cancel when the absolute state vectors are differenced. In addition, since the satellites are in close formation and track common GPS SVs 98.7 percent of the time, the ionospheric delay and GPS SV ephemeris and clock errors are highly correlated, and a large percentage of the these error contributions cancels in all of these cases. Therefore, the uncorrelated measurement noise and receiver clock errors are the primary contributors to the relative navigation error for this formation. For this HEO satellite formation, the Monte-Carlo simulations yielded a steady-state time-wise ensemble RMS relative position accuracy of approximately 76 meters, with a maximum below 400 meters. During a 1-day prediction using a converged filter solution, the relative errors remain under 3.4 kilometers in position and 2.3 meters per second in velocity.

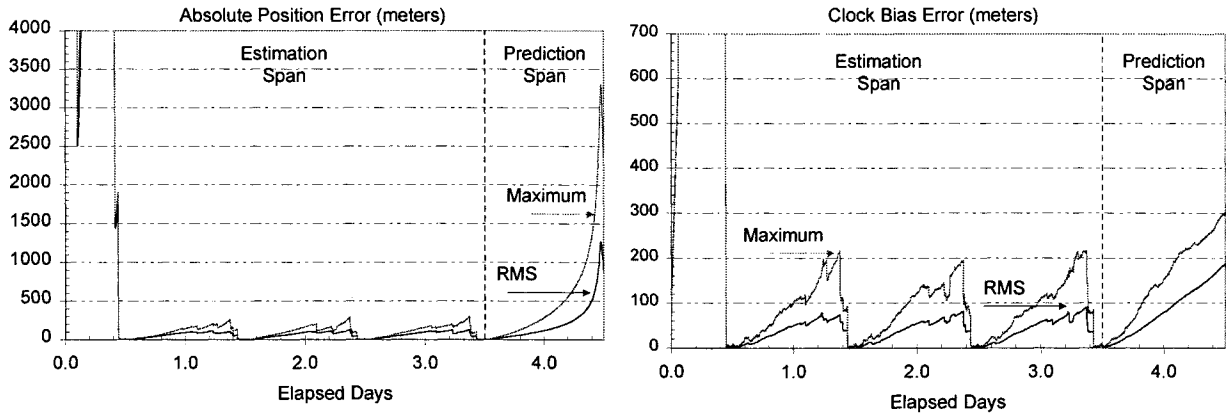


Figure 9. Ensemble Absolute Position and Clock Bias Errors for HEO Formation Using Independently-Estimated Solutions with Standard GPS Measurements

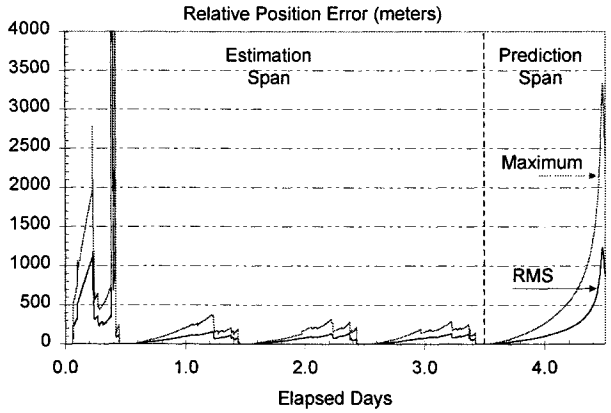


Figure 10. Ensemble RMS and Maximum Relative Position Errors for HEO Formation Using Independently-Estimated Solutions with Standard GPS Measurements

Additional Monte-Carlo simulations were performed to evaluate the relative performance of several of the navigation algorithms listed in Table 1. Tables 6 and 7 compare the steady-state time-wise ensemble absolute and relative error statistics for each of the navigation algorithms evaluated, respectively. The statistics listed are the largest errors obtained for all satellites in the formation. The variation in the error statistics for all satellites in the formation is less than 10 percent for the absolute errors and 50 percent for the relative errors. These comparisons produced the following results:

- Simultaneously estimating the local and remote satellites produced identical absolute and relative results to those obtained by independently estimating each of the satellites when there was no explicit correlation between the satellite state vectors.
- Inclusion of singly-differenced GPS pseudorange measurements between the remote and local satellites, in addition to standard GPS measurements for the local satellite, produced a small increase in the absolute errors. The relative navigation errors increased by about 10 percent, probably due to a reduction in the cancellation of the absolute errors associated with dynamic errors and correlated measurement errors.
- Inclusion of 1-way (remote-to-local) intersatellite pseudorange in addition to standard GPS pseudorange measurements for the remote and local satellites produced comparable absolute and relative results to those obtained without the 1-way measurements.
- Inclusion of round-trip intersatellite range between the local and remote satellites, in addition to standard GPS pseudorange measurements for the remote and local satellites, produced a significant reduction in both the absolute and relative navigation errors (Figures 11 and 12) due to a reduction in the impact of the uncorrelated receiver clock biases associated with the standard GPS measurements.

Table 6. Steady-State Time-Wise Ensemble Absolute Error Statistics for the HEO Formation

Navigation Algorithm	Position Error (meters)		Velocity Error (milli-meters per second)		Clock Error (meters)	
	RMS	Maximum	RMS	Maximum	RMS	Maximum
Independent	75	365	3.8	27	49	265
Simultaneous with standard GPS	75	365	3.8	27	49	265
Simultaneous with standard and differenced GPS	76	442	3.9	27	51	297
Simultaneous with standard GPS and 1-way intersatellite range	76	351	4.0	26	48	266
Simultaneous with standard GPS and round-trip intersatellite range	56	182	2.8	20	31	136
Simultaneous with standard GPS and round-trip intersatellite range (with 35-db_Hertz threshold)	134	1045	7.5	76	79	413

Table 7. Steady-State Time-Wise Ensemble Relative Error Statistics for the HEO Formation

Navigation Algorithm	Position Error (meters)		Velocity Error (millimeters per second)	
	RMS	Maximum	RMS	Maximum
Independent	76	391	4.1	30
Simultaneous with standard GPS	76	391	4.1	30
Simultaneous with standard and differenced GPS	81	517	4.5	39
Simultaneous with standard GPS and 1-way intersatellite pseudorange	76	485	4.2	36
Simultaneous with standard GPS and round-trip intersatellite range	24	117	1.3	6.2
Simultaneous with standard GPS and round-trip intersatellite range (with 35-db_Hertz threshold)	19	101	1.2	6.3

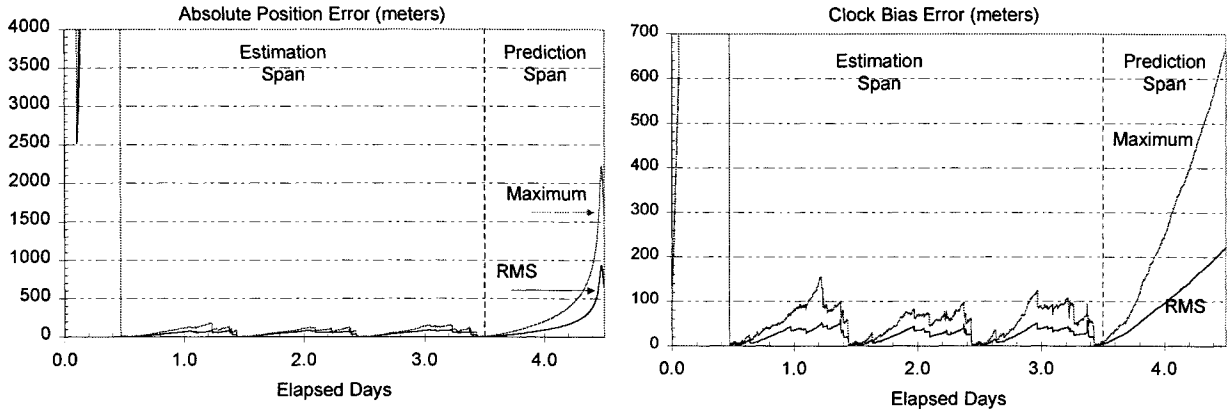


Figure 11. Ensemble Absolute Position and Clock Bias Errors for HEO Formation Using Simultaneously-Estimated Solutions with Standard GPS and Round-Trip Intersatellite Measurements

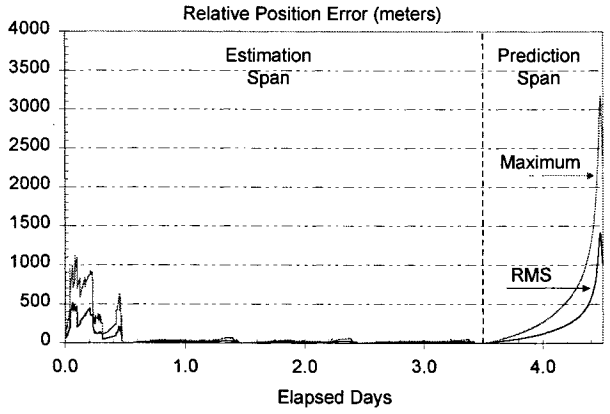


Figure 12. Ensemble RMS and Maximum Relative Position Errors for HEO Formation Using Simultaneously-Estimated Solutions with Standard GPS and Round-Trip Intersatellite Measurements

Because of the improvement obtained by including round-trip intersatellite range, the navigation performance was investigated that could be achieved by processing round-trip intersatellite measurements in addition to GPS measurements from a more standard space receiver with a 35-dB-Hertz tracking and acquisition threshold. Reduction in the number of standard GPS measurements increased the absolute navigation errors but the relative navigation errors decreased due to a reduction in the impact of the uncorrelated receiver clock biases associated with the standard GPS measurements.

6 – CONCLUSIONS AND FUTURE DIRECTIONS

This study assessed the relative navigation accuracy achievable using a high-accuracy multi-satellite filter to simultaneously estimate the satellite state vectors for a 500x7000 kilometer altitude MEO and a 1.2x12 Earth radii HEO formation. The satellites in both formations are in nearly co-planar orbits with intersatellite separations sufficiently small that the satellites acquire and track the same GPS SVs more than 98.5 percent of the time. Improvements to be achieved through the estimation of GPS biases, differencing of GPS measurements, and addition of intersatellite range measurements were evaluated.

For both formations using only GPS measurements:

- The frequency of acquisition and tracking of signals from common GPS SVs was found to be the primary factor driving the relative navigation accuracy.
- The uncorrelated measurement noise and receiver clock errors were found to be the next largest contributors to the relative navigation error.
- Comparable absolute and relative navigation accuracies were obtained using absolute state vectors estimated either independently or simultaneously based on standard GPS pseudorange measurements when there was no explicit correlation between the satellite state vectors.
- GPS bias estimation was found to improve clock estimation in some cases but did not improve the relative navigation performance.
- The processing of only singly-differenced GPS measurements did not provide a stable absolute solution. The processing of singly-differenced GPS measurements in combination with standard GPS pseudorange to the local satellite did not provide any accuracy improvement.

For the MEO formation, which has nearly continuous tracking of the GPS signals, the differencing of absolute state vectors can provide a relative navigation accuracy of better than 2.5 meters in position and 1.5 millimeters per second in velocity. For the HEO formation, which has continuous tracking of the GPS signals only near perigee, the differencing of absolute state vectors can provide a relative navigation accuracy of better than 400 meters in position and 30 millimeters per second in velocity, using a GPS receiver with weak signal tracking improvements and a highly stable clock.

When only GPS measurements were processed for these close formations, the operationally-more-complex simultaneous estimation algorithm did not provide any accuracy benefit over the operationally-simpler and more efficient approach of differencing independently-estimated state vectors. This conclusion appears to be in conflict with the following results achieved in real and simulated flight demonstrations using real measurements provided by GPS receivers:

- Flight data results published by one of the authors (Reference 10) show superior performance of a simultaneous estimator that estimated common measurement biases. However, that conclusion was based on the assumption of zero correlation of the absolute errors in the independently-estimated solutions, a very conservative assumption corresponding to the case in which the receivers are tracking no common GPS SVs. For the formations studied in this paper, which can track the same GPS SVs more than 98.5 percent of the time, much higher cancellation of the absolute errors would be expected.
- Reference 11 evaluated relative navigation performance using two 8-channel GPS receivers running in a simulated flight environment. In that experiment, simultaneous estimation algorithms that processed pseudorange measurements from only common GPS SVs provided superior performance as compared with differencing independently-estimated state vectors computed using all observed measurements from each receiver. The author concluded that the poorer performance of the latter method was primarily because a significant number of measurements from uncommon GPS SVs were processed in the independent estimators. However, he observed that the performance of the simultaneous estimators could be improved by introducing correlated process noise or by processing standard GPS measurements to estimate the local satellite and singly-differenced measurements to estimate the relative state of the remote satellite with respect to the local satellite.

Comparison of the results presented in the current paper with the real and simulated flight data results shows the sensitivity of the relative navigation accuracy to the percentage of measurements from common GPS SVs that are

processed, a result the authors also demonstrated in Reference 4. The real data results point out that even though all satellites in the formation may be able to acquire and track a very high percentage of common GPS SVs, the actual percentage of common SVs tracked may be significantly smaller. Under more realistic flight conditions, it will be difficult to synchronize independent estimators running on each satellite in the formation to process measurements from only common GPS SVs. The processing of measurements from only common GPS SVs can be better controlled using a simultaneous estimation algorithm.

For both formations, the inclusion of accurate round-trip intersatellite range measurements in the simultaneous estimation algorithm was found to improve relative navigation accuracy by reducing the impact of the GPS receiver clock biases. For the HEO formation, the inclusion of accurate round-trip intersatellite range was found to significantly improve both absolute and relative navigation accuracy. However, it should be noted that the results presented for round-trip intersatellite range are optimistic with respect to the elimination of all biases on the round-trip range.

Future directions will focus on refinements to the simultaneous estimation algorithms, the impact on the navigation performance of using less accurate clocks, a more detailed investigation of the navigation accuracy that can be achieved by including intersatellite measurements, and investigations of formations flying in different orbital configurations. The relative navigation version of GEODE will be integrated into a low cost GPS satellite receiver being developed by the GSFC GNCC. This formation-flying receiver will be used to demonstrate end-to-end performance in GNCC's formation-flying test bed.

REFERENCES

1. C. Gramling et al., "Preliminary Operational Results of the TDRSS Onboard Navigation System (TONS) for the Terra Mission," Paper MS00/32, *Proceedings of the International Symposium on Space Dynamics*, CNES, Biarritz , France, June 26-30, 2000

2. R. Hart, A. Long, and T. Lee, "Autonomous Navigation of the SSTI/Lewis Spacecraft Using the Global Positioning System (GPS), *Proceedings of the Flight Mechanics Symposium 1997*, NASA Conference Publication 3345, May 19-21, 1997, pp.123-133

3. F. Bauer et al., "Enabling Spacecraft Formation Flying Through Spacebourne GPS and Enhanced Autonomy Technology," *Proceedings of the ION GPS-99*, Nashville, TN, September 1999.

4. A. Long et al., "Autonomous Relative Navigation for Formation-Flying Satellites Using GPS," Paper MS00/18, *Proceedings of the International Symposium on Space Dynamics*, CNES, Biarritz , France, June 26-30, 2000.

5. J. R. Carpenter, "A Preliminary Investigation of Decentralized Control for Satellite Formations," Paper 264 presented at the IEEE Aerospace Conference, Big Sky Montana, March 20-22, 2000.

6. Goddard Space Flight Center, CSC-5506-06R0UD0, *Global Positioning System (GPS) Enhanced Orbit Determination (GEODE) Mathematical Specifications, Version 5, Update 3*, T. Lee and A. Long (CSC), prepared by Computer Sciences Corporation, March 2001.

7. M. Moreau et al., "GPS Receiver Architecture and Expected Performance for Autonomous GPS Navigation in High Earth Orbits," *NAVIGATION: Journal of the Institute of Navigation,* Vol. 47, No. 3, Fall 2000, pp. 191-204.

8. Johnson Space Center, *Range Bias Models for GPS Navigation Filters*, William M. Lear, Charles Stark Draper Laboratory, June 1993.

9. R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, Third Edition, John Wiley and Sons, 1997.

10. J. R. Carpenter and E. R. Schiesser, "Semi-Major Axis Knowledge and GPS Orbit Determination," accepted for publication in *NAVIGATION: Journal of the Institute of Navigation.*

11. P. W. Binning, *Absolute and Relative Satellite to Satellite Navigation Using GPS*, Ph.D. Dissertation, Colorado Center for Astrodynamics Research, University of Colorado, May 1997

RELATIVE NAVIGATION FOR FORMATION FLYING OF SPACECRAFT*

Roberto Alonso, Ju-Young Du, Declan Hughes, John L. Junkins
Department of Aerospace Engineering
Texas A&M University
College Station, TX 77843-3141

John L. Crassidis
Department of Mechanical & Aerospace Engineering
University at Buffalo, The State University of New York
Amherst, NY 14260

ABSTRACT

This paper presents a robust and efficient approach for relative navigation and attitude estimation of spacecraft flying in formation. This approach uses measurements from a new optical sensor that provides a line of sight vector from the master spacecraft to the secondary satellite. The overall system provides a novel, reliable, and autonomous relative navigation and attitude determination system, employing relatively simple electronic circuits with modest digital signal processing requirements and is fully independent of any external systems. Experimental calibration results are presented, which are used to achieve accurate line of sight measurements. State estimation for formation flying is achieved through an optimal observer design. Also, because the rotational and translational motions are coupled through the observation vectors, three approaches are suggested to separate both signals just for stability analysis. Simulation and experimental results indicate that the combined sensor/estimator approach provides accurate relative position and attitude estimates.

INTRODUCTION

The vision-based navigation (VISNAV) system described in this paper for formation flying applications comprises an optical sensor of a new kind combined with specific light sources (beacons) in order to achieve a selective or "intelligent" vision. The VISNAV sensor[1,2] is made up of a Position Sensing Diode (PSD) placed in the focal plane of a wide angle lens. When the rectangular silicon area of the PSD is illuminated by energy from a beacon focused by the lens, it generates electrical currents in four directions that can be processed with appropriate electronic equipment to estimate the energy centroid of the image. While the individual currents depend on the intensity of the light, their imbalances are weakly dependent on the intensity and are almost linearly proportional to the location of the centroid of the energy incident on the PSD. The idea behind the concept of intelligent vision is that the PSD can be used to see only specific light sources, accomplished by frequency domain modulation of the target lights and some relatively simple analog signal processing (demodulation). The light is produced by LEDs (beacons) modulated at an arbitrary known frequency while the currents generated are driven through an active filter set on the same frequency. Calculating the current imbalances then yields two analog signals directly related to the coordinates locating the centroid of that beacon's energy distribution on the PSD, in a quasi-linear fashion, and therefore to the incident direction of this light on the wide-angle lens (which gives a line of sight vector). A more detailed description of the VISNAV system can be found in Refs. [1-2].

Because the beacons are offset from the mass center of the secondary satellite, the observed line of sight couples the rotational and the translational motion. The traditional Kalman filter uses this raw information to update the attitude and the position equations without any discrimination about the nature of the signal. This approach is effective in most of the cases, but it is very difficult for stability analysis because of the complexity of the system in hand and of the way that the Kalman gain is calculated. The approach presented in this paper is based on two

*This work was supported under a NASA grant (NCC 5-448), and from a Texas Higher Education Coordinating Board Grant (000512-0004-1999) and Air Force Office of Sponsored Research grant (32525-57200).

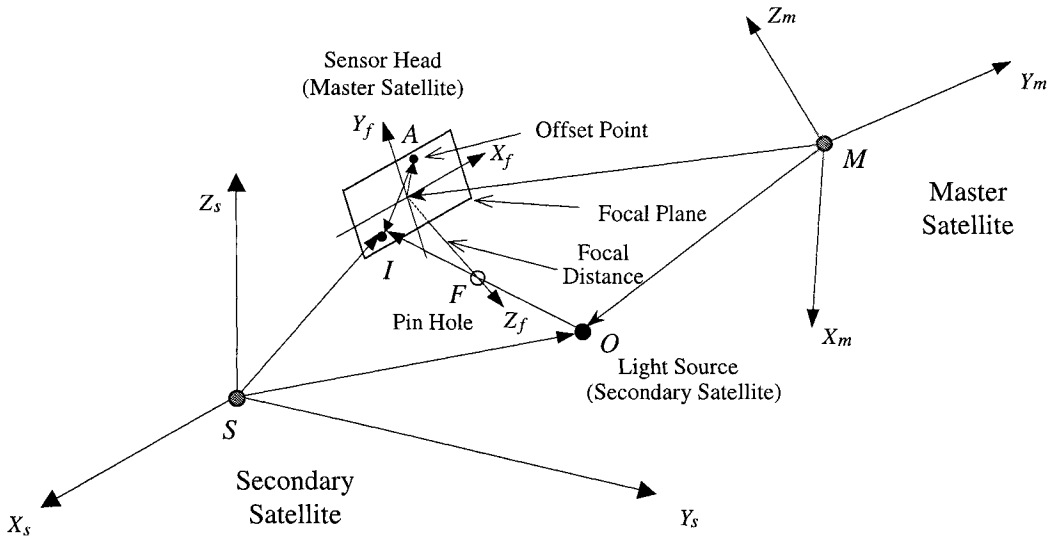


Figure 1: Focal Plane Measurement from One Light Source

special characteristics: the observer uses a constant gain for each parameter to be estimated (suboptimal filter) and the incoming signal is split according to the translational and the rotational dynamics. The use of constant gains avoids dealing with nonlinear time-varying systems, and the signal separation allows two independent plants where the stability analysis is feasible on each one using different approaches.

The presence of four or more beacons in the sensor head field of view (FOV) assures a deterministic solution for the navigation and the attitude problem as well. In this case, the sensor redundancy produces a *natural* division of both dynamics, without any additional mathematical procedures. The filtering stage uses this deterministic solution as the VISNAV measurements so that the plant models only play the role of increasing the sensor accuracy. In this particular case, the stability can be proven by adopting a constant gain in the updating equation suggested in the present paper. Three methods for signal separation are given, one in time domain and two in the frequency domain. The first method is based on the signal magnitude difference produced by the navigation and the attitude at the sensor head location. The other two approaches use the frequency differences of both motions.

The organization of this paper proceeds as follows. First, the basic equations for the VISNAV system are given. Next, the experimental calibration procedure is shown. Then, the relative attitude equations are derived, followed by a presentation of the orbital equations of motion. The suggested methods for the signal separation process are then presented. Next, the observer design for relative attitude and position estimation is shown. Finally, simulation results for formation flying applications are presented.

BASIC EQUATIONS

In this section the mathematical models are presented in the context of the particular problem related to relative position and attitude estimation from line of sight observations. The notation[3] used in the derivations is briefly revisited for the sake of clarification. The angular velocity of the α frame with respect to the β frame is represented by the physical vector $\boldsymbol{\omega}_{\beta\alpha}$ (*physical* denotes that the vector is independent of the frame, whereas *mathematical* denotes the physical vector components expressed in some frame). The vector $\boldsymbol{\omega}^{\gamma}_{\beta\alpha}$ is the mathematical vector made up of the components of $\boldsymbol{\omega}_{\beta\alpha}$ taken in the γ frame. The derivative with respect to time is indicated by the operator p, where $p_{\alpha}\mathbf{R}$ is the rate of change of the vector $\mathbf{R}$ relative to the frame α, and $p\mathbf{R}^{\alpha}$ is the time derivative of the vector expressed in the α frame.

Measurement Equation

Figure 1 shows the focal plane measurement of the VISNAV system for a master and secondary satellite system using one light source from a beacon (see Ref. [1] for more details). Three frames are used to describe the orientation and

position of the master and secondary satellites. The first one, denoted by (X_s, Y_s, Z_s), is fixed on the secondary satellite, with the LED beacons firmly attached to the body of the satellite, and having known positions in the (X_s, Y_s, Z_s) frame. This frame is also the reference frame for the attitude problem. We assume that this frame is centered at the mass center of this spacecraft, and is denoted using the superscript s on the mathematical vectors. The second reference system, denoted by (X_f, Y_f, Z_f), is fixed on the master satellite, where the focal plane of the VISNAV system is located. We assume that the Z_f axis is along the boresight, passing through the input pin hole which is at a distance $Z_f = +f$ from the focal plane. The axes X_f and Y_f are arbitrary, but fixed in the VISNAV sensor. This frame is denoted as the f frame. The third frame, denoted by (X_m, Y_m, Z_m), is fixed to the mass center of the master satellite. The position and orientation of this frame with respect to the focal frame is assumed to be known. The vectors for the master frame are identified with the superscript m.

The point S is the origin of the frame s. The point O is the location of each light beacon in the secondary satellite; normally there are several beacons to assure continuous tracking of the satellite and for redundancy. The point I is sometimes referred as the *image center* since it is the intersection of each light beam from the beacon with the focal plane, where position of I with respect to the focal reference system is used to form a line of sight observation. The point denoted as F in Figure 1 is the pinhole which is coincident with the sensor principal point. Three vectors are now defined: $\vec{SO}$ (the vector from the center S of the s frame to the beacon location O), $\vec{SI}$ (the vector from the the center S of the s frame to the image center I), and $\vec{OI}$ (the vector from the beacon location O to the image center I, with the constraint equation given by $\vec{OI} = \vec{SI} - \vec{SO}$.

The orientation between the secondary and master frames is denoted by the (unknown) rotation matrix C_s^m which transforms a vector expressed in the secondary frame s to the primary frame m. The rotation matrix C_f^m between the focal and the master frames is known by ground calibration. Expressing the vectors $\vec{SI}$, $\vec{OI}$ and $\vec{SO}$ in frame components gives the following relation (colinearity equations)[4]

$$C_s^m \left(\vec{SI} - \vec{SO}\right)^s \equiv C_s^m \mathbf{v}^s = \mathbf{v}^m \equiv \left(\vec{OI}\right)^m \tag{1}$$

where $\mathbf{v}^s = \xi^{-1/2} \left[X_I - x_O, \quad Y_I - y_O, \quad Z_I - z_O\right]^T$ and $\xi \equiv (X_I - x_O)^2 + (Y_I - y_O)^2 + (Z_I - z_O)^2$. The quantity (X_O, Y_O, Z_O) represents the known beacon location, and (X_I, Y_I, Z_I) is the unknown position with respect to the secondary satellite. The measurements x_I and y_I in the focal frame can be expressed in unit vector form by

$$\mathbf{v}^f = \frac{1}{\sqrt{x_I^2 + y_I^2 + f^2}} \begin{bmatrix} x_I \\ y_I \\ -f \end{bmatrix} \tag{2}$$

where f is the known focal distance. This unit vector in the master frame is expressed using the fixed rotation matrix between the sensor plane frame and the master satellite reference frame, with $\mathbf{v}^m = C_f^m \mathbf{v}^f$. A bias offset in the measurement is also accounted for in the model (denoted by A in Figure 1). The bias vector is a constant error vector induced by an unbalance of the horizontal and vertical gains in the focal plane detector relative to the particular coordinate system associated with the detector at calibration. Essentially this is the same offset between the "electrical center" (zero voltage imbalance) and the geometrical center associated with the optical boresight and sensor coordinate system. This vector is denoted by $\mathbf{v}_a$ and is normally referenced in the focal plane frame:

$$\mathbf{v}_a^m = C_f^m \mathbf{v}_a^f = C_f^m \begin{bmatrix} x_a \\ y_a \\ 0 \end{bmatrix}^f \tag{3}$$

Finally, the measurement equation for each light source from a beacon, placed on the secondary satellite, is as follows:

$$\mathbf{v}_j^m = C_s^m \mathbf{v}_j^s + \mathbf{v}_a^m \quad \text{for } j = 1, \ldots, N \tag{4}$$

where N is the number of LED beacons.

Small separations between light beams from multiple LEDs reduces the discrimination of each beacon, which ultimately produces a dilution of precision in the position and attitude solution. A larger distance between the satellites also leads to a dilution of precision since the beacons ultimately approach angular co-location. If the relative position between satellites is known then only two non-colinear line of sight vectors are required to determine an attitude solution. In a similar fashion for the position navigation only problem, where the satellite is considered to be a "mass point" (in other words without attitude), two line of sight vectors are only required. A covariance analysis

shows that when the relative position and attitude both are unknown then two line of sight vectors provide only one axis of attitude and one axis of position information.[5] Furthermore, an observability analysis using two line of sight observations indicates that the beacon that is closest to the target provides the most attitude information but has the least position information, and the beacon that is farthest to the target provides the most position information but has the least attitude information. In order to find a deterministic solution for the position and velocity at least four vector observations are required.

VISNAV SENSOR CALIBRATION

In order to use the sensor effectively, an accurate mapping between the normalized voltages (V_x, V_y) returned by the sensor and the location of the light centroid on the image plane (x, y) must be known. The projection equations in Eq. (4) represent the ideal case for a pin-hole camera model. However, in practice the lens and PSD detector nonlinearities cause any camera to depart from this ideal model. Therefore a correction of all non-ideal effects must be applied into a calibration process that is implicitly constrained to be consistent with the colinearity equations. The PSD sensor experiment configuration for the calibration process places the camera at many known positions and attitudes $(X_{cj}, Y_{cj}, Z_{cj}, \phi_j, \theta_j, \psi_j)$ relative to an array of targets located at (X_i, Y_i, Z_i). Unfortunately, we must also consider the realistic uncertainty in the camera position, orientation and target location since only certain levels of precision in the laboratory setting can be achieved. The current experimental configurations are as follows: the focal length of PSD sensor $f = 0.01$m; beacon position = $[1.5716m, 0, 0]$, sensor position = [0, 0, 0]; roll = $[-50° : 1° : 50°]$, pitch = $[-50° : 1° : 50°]$, yaw = 0°; total test points = 10201, total test points in the FOV = 6635. Figure 2(i) represents the ideal location of the light centroid corresponding to each test point configuration. The acquisition of the calibration data is computer-automated: a precision two axis air bearing permits the sweeping of the sensor over known angles and image-fixed targets. The actual experimental results in Figure 2(ii) show the presence of high systematic distortions associated both with a wide angle lens and electronic nonlinearities, which must be calibrated.

One method to accomplish the calibration involves a direct interpolation using the experimental data set. Another way involves determining the global calibration function mapping of the normalized voltages (V_x, V_y) into the corresponding known ideal locations of the light centroid on the image plane (x, y) for the whole FOV data. Yet other way is to determine the local calibration functions available for the corresponding local part of the FOV region. The local calibration functions can capture the fine structure of the distortion. In order to compare the optimality of the proposed methods, it is necessary to compare the calibration accuracy, memory requirement and computational load. Here, we discuss how to determine the global calibration functions. The problem involves determining the optimal coefficient sets a_{ij}, b_{ij} for the bivariate calibration functions

$$x = f_x(V_x, V_y) \approx \sum_{i=0}^{n} \sum_{j=0}^{i} a_{ij} \phi_{ij}(V_x, V_y) = \sum_{i=0}^{n} \sum_{j=0}^{i} a_{ij} T_{i-j}(V_x) T_j(V_y) \tag{5a}$$

$$y = f_y(V_x, V_y) \approx \sum_{i=0}^{n} \sum_{j=0}^{i} b_{ij} \phi_{ij}(V_x, V_y) = \sum_{i=0}^{n} \sum_{j=0}^{i} b_{ij} T_{i-j}(V_x) T_j(V_y) \tag{5b}$$

where n is the order of the univariate element polynomials, $\phi_{ij}(V_x, V_y)$ is a bivariate polynomial basis, and $T_i(V_x)$, $T_i(V_y)$ are the univariate first type Chebyshev polynomials. Let the vector form of the complete set of the basis function be given by

$$\phi(V_x, V_y) = [\phi_{00}(V_x, V_y), \phi_{10}(V_x, V_y), \phi_{11}(V_x, V_y), \phi_{20}(V_x, V_y) ..., \phi_{nn}(V_x, V_y)]^T \tag{6}$$

Then, our goal is to determine the two coefficient sets

$$\mathbf{a} = [a_{00}, a_{10}, a_{11}, a_{20}, \ldots, a_{nn}]^T \tag{7a}$$

$$\mathbf{b} = [b_{00}, b_{10}, b_{11}, b_{20}, \ldots, b_{nn}]^T \tag{7b}$$

to minimize the weighted least squares magnitude of the residual vectors $\mathbf{R}_x$ and $\mathbf{R}_y$:

$$\mathbf{R}_x = (\mathbf{g}_x - H\mathbf{a}) \tag{8a}$$

$$\mathbf{R}_y = (\mathbf{g}_y - H\mathbf{b}) \tag{8b}$$

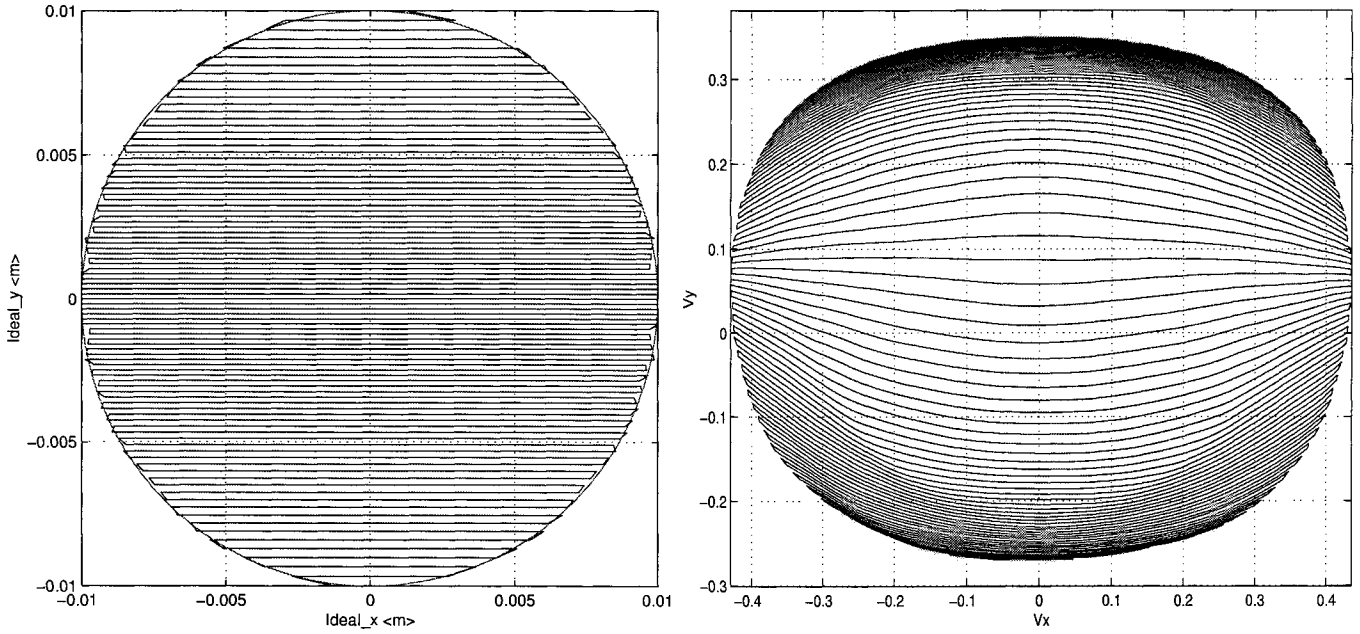
where $H = [\phi(V_{x1}, V_{y1}) \vdots \phi(V_{x2}, V_{y2}) \vdots \cdots \vdots \phi(V_x, V_y)]^T$ and $\mathbf{g}_x$, $\mathbf{g}_y$ are the vector forms of the ideal location of the light centroid calculated by the colinearity equations. Consequently, the best coefficient sets are determined by

$$\hat{\mathbf{a}} = (H^T W H)^{-1} H^T W \mathbf{g}_x \tag{9a}$$

$$\hat{\mathbf{b}} = (H^T W H)^{-1} H^T W \mathbf{g}_y \tag{9b}$$

where W is typically the inverse of the measurement covariance matrix.

Thereafter, only the coefficient sets are needed to evaluate (x, y) from the PSD sensor output (V_x, V_y). Figure 3 shows that the results of the 35^{th}-order calibration functions. The standard deviation of the x errors is $6.77 \times 10^{-6} m$ and the standard deviation of the y errors is $5.95 \times 10^{-6} m$, which indicates an accurate calibration. Figure 4 shows that the calibration error decreases as the order of polynomial functions for the calibration function increases. These experimental results indicate that the sensor can be accurately calibrated using polynomial functions.



(i) Ideal Location of (x, y)

(ii) Actual PSD Sensor Output (V_x, V_y)

Figure 2: Ideal Location and Actual PSD Sensor Output

Relative Attitude Equations

In this section the governing equations for the relative attitude kinematics between two bodies are reviewed. The kinematic equations presented here are derived using non-inertial reference frames, however only minor changes are required from the standard formulation.[3] Starting from Eq. (1) and taking derivative of each vector with respect to the same frame in which they are expressed gives the following expressions

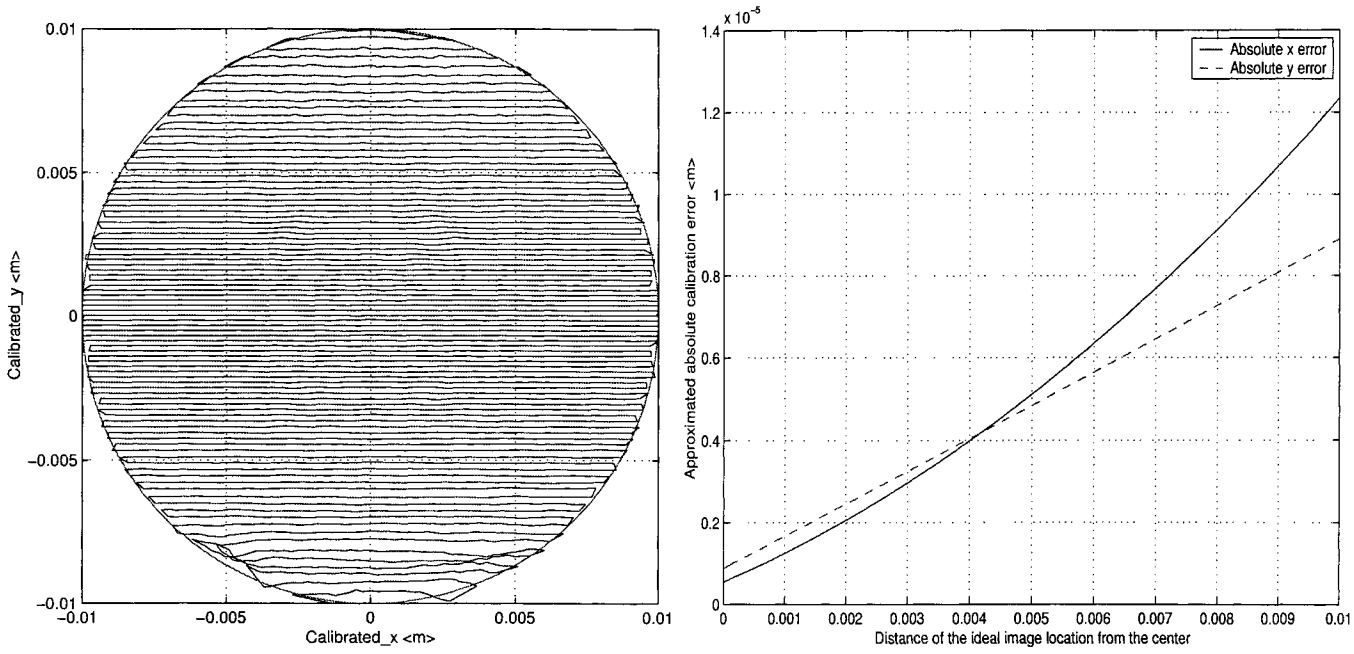
$$p\mathbf{v}^m = C_s^m p\mathbf{v}^s + pC_s^m \mathbf{v}^s = C_s^p \left(p\mathbf{v}^s + C_m^s pC_s^m \mathbf{v}^s\right) \tag{10}$$

The bias in Eq. (1) is considered to be a constant, so it's derivative is zero. The same expression in Eq. (10) can be derived by the application of the transport theorem, which yields the following expressions

$$p_m\mathbf{v} = p_s\mathbf{v} + \boldsymbol{\omega}_{ms} \times \mathbf{v} \tag{11a}$$

$$p\mathbf{v}^m = C_s^m \left(p\mathbf{v}^s + [\boldsymbol{\omega}_{ms}^s \times] \mathbf{v}^s\right) \tag{11b}$$

where the matrix $[\cdot\times]$ denotes the cross product matrix.[6]



(i) Calibrated (x, y) **(ii)** Approximated Absolute Calibration Error

Figure 3: Experimental Calibration Results

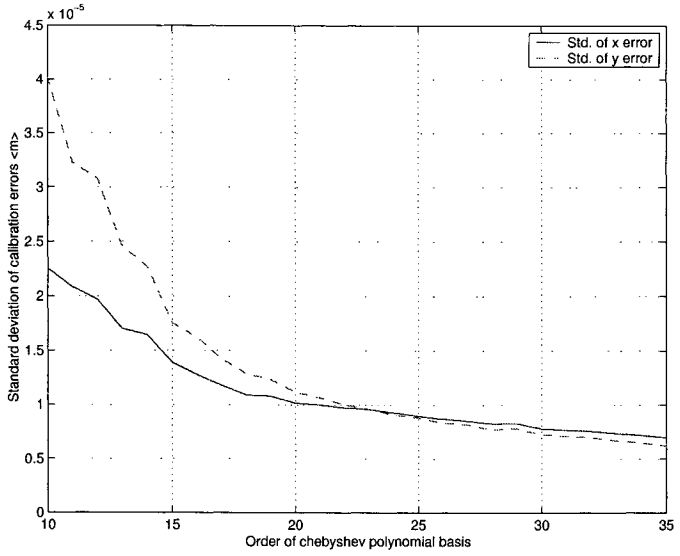


Figure 4: Order of Polynomials Versus Calibration Accuracy

Both expressions, (10) and (11), must be equivalent. Setting these equations equal to each other yields the time rate of change of the attitude matrix, given by

$$C_m^s pC_s^m = [\omega_{ms}^s \times] \tag{12a}$$

$$pC_s^m = C_s^m[\omega_{ms}^s \times] = -[\omega_{sp}^m \times]C_s^m \tag{12b}$$

The relative attitude kinematics are described by the expression in Eq. (12) in terms of attitude matrix and the angular velocity between both frames.

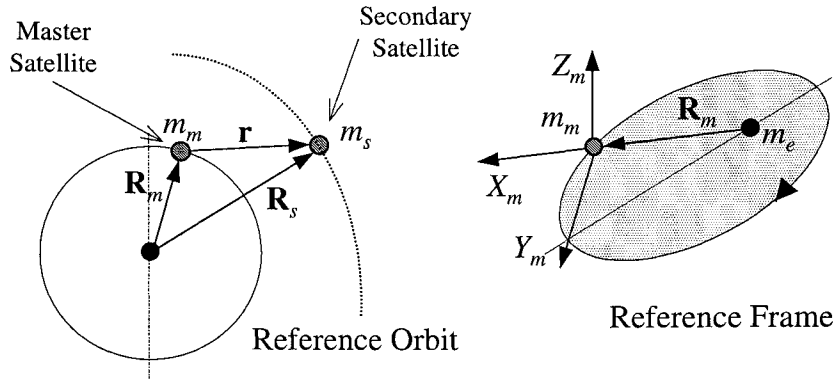


Figure 5: Relative Navigation: Master and Secondary Satellite Orbits

We now write the expression in Eq. (12) in terms of the corresponding quaternions.[7] Toward this end, the quaternion is expressed as, $\mathbf{q}_m^s = \left[\mathbf{e}^T \sin\frac{\alpha}{2} \quad \cos\frac{\alpha}{2}\right]^T$, where $\mathbf{e}$ is the eigenaxis between both frames and α is the rotation angle measured from frame m to frame s. The quaternion is a vector with the same components in both the m and s frames, and can be expressed in any external frame as an arbitrary (i.e. general) vector. This has an advantage over the rotation matrix formulation, which is fixed to the reference system s and m in this case.

An infinitesimal rotation is expressed in terms of the quaternion as $d\mathbf{q}_m^s = 1 + \frac{1}{2}\boldsymbol{\omega}_{sm}dt$, where dt is the time differential. Multiplying by the quaternion $\mathbf{q}_m^s$ and taking the first-order infinitesimal part, the following differential equation is given

$$p\mathbf{q}_m^s = \frac{1}{2}\mathbf{q}_m^s \otimes \boldsymbol{\omega}_{sm}^m = \frac{1}{2}\begin{bmatrix} \varrho & \vdots & q_o I_{3\times 3} + [\varrho\times] \\ \cdots & \vdots & \cdots\cdots\cdots \\ q_o & \vdots & -\varrho^T \end{bmatrix}\begin{bmatrix} \boldsymbol{\omega}_{sm} \\ \cdots \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} q_o\boldsymbol{\omega}_{sm} + [\varrho\times]\boldsymbol{\omega}_{sm} \\ \cdots\cdots\cdots \\ -\varrho^T \cdot \boldsymbol{\omega}_{sm} \end{bmatrix} \tag{13}$$

where the quaternion $\mathbf{q}_m^s$ is decomposed into a scalar and a vector part as $\mathbf{q}_m^s = \left[(\varrho_m^s)^T \quad q_0\right]^T$, and $[\varrho\times]$ is the skew symmetric cross product matrix. Both the attitude matrix and quaternion formulations will be used in the definition of the observer feedback error, but the quaternion formulation is used in the actual implementation of the observer.

Relative Navigation and Dynamics Equations

From basic orbit theory,[8] the equations of motion are written assuming that each satellite is referenced with respect to the same inertial frame. The vectors are described in the Figure 5. The relative orbit is described by the difference between both vectors, $\mathbf{r} = \mathbf{R}_m - \mathbf{R}_s$. If the master satellite position vector is written as $\mathbf{R}_m = R_m[1,0,0]^T$, the expression can be simplified. The frame with this property is the *Local Vertical Local Horizontal* (LVLH) reference frame,[9] which is widely used to reference Earth Pointing satellites. The vector $\mathbf{r}$ is decomposed in m frame components and takes the final expression given by

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = -\mu\begin{bmatrix} \dfrac{R_m + x}{\|\mathbf{R}_m + \mathbf{r}\|^3} - \dfrac{R_m}{\|\mathbf{R}_m\|^3} \\ \dfrac{y}{\|\mathbf{R}_m + \mathbf{r}\|^3} \\ \dfrac{z}{\|\mathbf{R}_m + \mathbf{r}\|^3} \end{bmatrix} - \dot{\omega}\begin{bmatrix} -y \\ x \\ 0 \end{bmatrix} - \omega\begin{bmatrix} -2\dot{y} - \omega x \\ 2\dot{x} - \omega y \\ 0 \end{bmatrix} + \begin{bmatrix} \Delta a_x \\ \Delta a_y \\ \Delta a_z \end{bmatrix} \tag{14}$$

where $\Delta\mathbf{a}$ is the relative acceleration. The forcing part along the X axis component has the following structure: $[f(\mathbf{R}_m + \mathbf{r}_{12}) - f(\mathbf{R}_m)]^m$ which is not robust from a numerical point of view. This expression is maintained for compactness and will be used in the observer analysis, but for practical implementations it is convenient to re-write it avoiding the subtraction of two large numbers.[8] Equation (14) expresses the dynamical model for relative navigation between the secondary satellite with respect to the master satellite. We note that the number of master satellite orbit

parameters computed on the ground and to be used in Eq. (14) is at most 3. For the general case, the magnitude R_m, the angular velocity ω, and angular acceleration $\dot{\omega}$ are just needed. For the special case involving circular orbits, only the position magnitude is necessary.

In the attitude problem, Euler's equation or the measured gyro outputs are the starting point for the derivation of the rotational dynamics equation to obtain the angular velocity between the inertial frame and a body frame. For the relative attitude problem, Euler's equation must be applied in a differential mode, similar in fashion as the orbit case. However, we seek an expression without an additional "third" frame (inertial one included), in addition to the m and the s frames, so that the system is independent of the extra reference frame's choice. In other words, the relative navigation and the relative attitude must be a function only of the definition of the master and secondary frames and completely independent of the particular choice of the inertial frame or any other frame other than m and s. This simple fact is common in control theory, where the error or its derivative is only defined by the current and the desired state independent of any other frame choice.

In the two body problem previously derived, the equation for $\mathbf{r}$ is very accurate because it is supported by well known models for almost all involved forces in hand, with any remaining small perturbation bounded. In the relative attitude dynamics the presence of internal torques, which are normally unmodeled with an unbounded time integral, plays an important role in the model equations. We assume that each satellite in the constellation has an attitude control subsystem able to maintain the desired satellite orientation inside of some allowable bound. The last hypothesis is a qualitative one. We assume that the measurements are available frequently enough to use simpler propagation models (to be derived) as a function of the sampling interval. A detailed derivation of the relative attitude dynamics equations can be found in Ref. [6].

SIGNAL SEPARATION

Due to the slow moving dynamics in orbital motion, the translational and rotational dynamics are almost independent (in fact there is a second order coupling thought the angular velocity which appears in both systems, but for this analysis we can assume a negligible coupling between both motions). This section presents three approaches for signal separation of the translational and rotational dynamics:

1. *Time Domain Approach*: This approach assumes that the contribution of the angular rotation is smaller than the translational motion contribution on the sensor output. In other words the distance between the beacon location and the secondary satellite mass center is much smaller than the distance between the mass center of both satellites.

2. *Spline Wavelets Filter*: This approach is a frequency band filter which can run in quasi-real time, and uses splines as the basis functions for which the incoming signal is represented.

3. *Standard Band Pass Filter*: This is the standard approach from basic signal processing theory and in general does not run in real time. It uses the different frequency scales between the rotational and translational motion to separate them via one or two band filters.

Time Domain Approach

Consider $\mathbf{r}$ as the *effective position vector* between the sensor head and the beacon, which couples the orbit and the attitude dynamics:

$$\mathbf{r} = \mathbf{r}^o + \mathbf{r}^a \tag{15}$$

where the superscripts $(\cdot)^o$ and $(\cdot)^a$ denote orbit and attitude respectively. The first and second derivatives of above equation are given by $\dot{\mathbf{r}} = \dot{\mathbf{r}}^o + \dot{\mathbf{r}}^a$ and $\ddot{\mathbf{r}} = \ddot{\mathbf{r}}^o + \ddot{\mathbf{r}}^a$. For notation simplicity $\dot{(*)}$ denotes the time derivative relative to the frame where the forcing function is expressed. It is assumed that all vectors are expressed in the master satellite body frame without loss in generality. The solution of this differential equation is given by

$$\mathbf{r}(t_k) = \mathbf{r}(t_j) + \dot{\mathbf{r}}(t_j)(t_k - t_j) + \int_{t_j}^{tk} \mathbf{F}(\dot{\mathbf{r}}^o, \mathbf{r}^o, t)(t_k - t_j)\,dt + \int_{t_j}^{tk} \mathbf{P}(\dot{\mathbf{r}}^a, \mathbf{r}^a, t)(t_k - t_j)\,dt \tag{16}$$

where $\mathbf{F}(\cdot^o)$ represents the orbit dynamics and $\mathbf{P}(\cdot^a)$ denotes the attitude dynamics which is considered as a perturbation of the main motion. In other words, the contribution of the navigation signal is much important in magnitude than the rotational motion, which allows us to split one from the other.

The *reference problem* is computed as

$$\mathbf{r}^r(t_k) = \mathbf{r}^r(t_j) + \dot{\mathbf{r}}^r(t_j)(t_k - t_j) + \int_{t_j}^{tk} \mathbf{F}(\dot{\mathbf{r}}^{ro}, \mathbf{r}^{ro}, t)(t_k - t_j)\, dt \tag{17}$$

where the superscript $(\cdot)^r$ denotes the *reference* value. In order to solve the signal separation problem an extra (weak) condition is required: the reference and the effective vectors are coincident at each sampling time, i.e., $\mathbf{r} = \mathbf{r}^r$ and $\dot{\mathbf{r}} = \dot{\mathbf{r}}^r$. The distance of the i^{th} component of $(\mathbf{r} - \mathbf{r}^r)$ is expressed as the following integral

$$(r - r^r)_i = \int_{t_j}^{t_k} \Phi(t)(t_k - t_j)\, dt. \tag{18}$$

The disturbance function is defined as

$$\Phi_{t_j}(t) = \mathcal{F}(\mathbf{r}^o, \dot{\mathbf{r}}^o, t) - \mathcal{F}(\mathbf{r}^{r\,o}, \dot{\mathbf{r}}^{r\,o}, t) + \mathcal{P}(\mathbf{r}^a, \dot{\mathbf{r}}^a, t) \equiv \Delta\mathcal{F}_{t_j \to t}(\cdot) + \mathcal{P}_t(\cdot). \tag{19}$$

where $\mathcal{F}$ and $\mathcal{P}$ represent the i^{th} component of $\mathbf{F}$ and $\mathbf{P}$, respectively. The residual function $R_{t_j \to t_k}$ is now defined as

$$R_{t_j \to t_k} = \frac{\int_{t_j}^{t_k} \Phi_{t_j}(t)\, dt}{t_k - t_j}. \tag{20}$$

There are several methods to compute the residual function, but for a small sampling time the overall error is almost independent of the approach selected. A quadratic linear function is used because for simplicity, given by

$$\Phi_{t_j}(t) = a + b(t_j - t) + c(t_j - t)^2 \tag{21}$$

where a, b and c are the linearly embedded coefficients to be computed. The polynomial coefficients can be calculated using three points, denoted at times t_{k-1}, t_k and t_{k+1}. The approximate functions are written as

$$\begin{bmatrix} R_{t_k \to t_{k-1}} \\ R_{t_{k-1} \to t_k} \\ R_{t_{k+1} \to t_k} \\ R_{t_k \to t_{k+1}} \end{bmatrix} = \frac{1}{\Delta t^2} \begin{bmatrix} 1/8\Phi_{t_k}(t_{k-1}) & 5/12\Phi_{t_k}(t_k) & -1/24\Phi_{t_k}(t_{k+1}) \\ 7/24\Phi_{t_{k+1}}(t_{k-1}) & 1/4\Phi_{t_{k-1}}(t_k) & -1/24\Phi_{t_{k-1}}(t_{k+1}) \\ -1/24\Phi_{t_{k+1}}(t_{k-1}) & 1/4\Phi_{t_{k+1}}(t_k) & 7/24\Phi_{t_{k+1}}(t_{k+1}) \\ -1/24\Phi_{t_k}(t_{k-1}) & 5/12\Phi_{t_k}(t_k) & 1/8\Phi_{t_k}(t_{k+1}) \end{bmatrix} \tag{22}$$

We now have

$$\begin{bmatrix} 1/8 & 5/12 & -1/24 \\ 7/24 & 1/4 & -1/24 \\ -1/24 & 1/4 & 7/24 \\ -1/24 & 5/12 & 1/8 \end{bmatrix} \begin{bmatrix} \mathcal{P}(\Delta r, t_{k-1}) \\ \mathcal{P}(\Delta r, t_k) \\ \mathcal{P}(\Delta r, t_{k+1}) \end{bmatrix} = \mathcal{M} \tag{23}$$

where the matrix $\mathcal{M}$ is given by

$$\mathcal{M} = \begin{bmatrix} \Delta t^{-2} R_{t_k \to t_{k-1}} - 1/8\Delta\mathcal{F}_{t_k \to t_{k-1}} + 1/24\Delta\mathcal{F}_{t_k \to t_{k+1}} \\ \Delta t^{-2} R_{t_{k-1} \to t_k} - 1/4\Delta\mathcal{F}_{t_{k-1} \to t_k} + 1/24\Delta\mathcal{F}_{t_{k-1} \to t_{k+1}} \\ \Delta t^{-2} R_{t_{k+1} \to t_k} + 1/24\Delta\mathcal{F}_{t_{k+1} \to t_{k-1}} - 1/4\Delta\mathcal{F}_{t_{k+1} \to t_k} \\ \Delta t^{-2} R_{t_k \to t_{k+1}} + 1/24\Delta\mathcal{F}_{t_k \to t_{k-1}} - 1/8\Delta\mathcal{F}_{t_k \to t_{k+1}} \end{bmatrix} \tag{24}$$

The minimum norm solution applied to each one of the above vector equations gives the following results

$$\mathcal{A}_{pv}^{+}\mathcal{M} = \mathcal{P} \quad \text{where } \mathcal{A}_{pv}^{+} = \begin{bmatrix} -0.9 & 3.7 & 1.3 & -2.1 \\ 1.5 & -0.5 & -0.5 & 1.5 \\ -2.1 & 1.3 & 3.7 & -0.9 \end{bmatrix} \tag{25}$$

It is well known from the interpolation theory that the minimum error can be expected at the middle point, where "the mass center of the data" is allocated, therefore, only $\mathcal{P}(t_k)$ is calculated from the above expression to reduce the integration error. After the calculation of the vector $\mathcal{P}(t)$, the signal proportional only to the attitude motion is obtained and the angular displacement can be computed.

Spline Wavelets Filter

In this section, we use B-splines N_m to be the scaling functions. Let N_1 be the characteristic function, which is zero everywhere except in the finite support $[0,1)$ where it takes unit value. For each positive integer m, the m^{th}-order cardinal B-spline is defined, inductively, by

$$N_m(t) = (N_{m-1} * N_1)(t) = \int_0^1 N_{m-1}(t-x)\,dx \tag{26}$$

A fast computation of $N_m(t)$ can be achieved by using the following formula recursively until we arrive at the first order B-spline N_1:[10]

$$N_m(t) = \frac{t}{m-1} N_{m-1}(t) + \frac{m-t}{m-1} N_{m-1}(t-1) \tag{27}$$

One of the most important properties of the B-spline is "total positivity" (see Ref. [11]) by virtue of which the function in terms of a B-spline series follows the shape of the data. For example, if $g(t) = \sum_j [\alpha_j N_m(t-j)]$, then $\alpha_j \geq 0 \,\forall j \Rightarrow g(t) \geq 0$, α_j (increasing) $\Rightarrow g(t)$ (increasing), and α_j (convex) $\Rightarrow g(t)$ (convex). Furthermore, the number of sign changes of $g(t)$ does not exceed that of the coefficient sequence $\{\alpha_j\}$.

Any function $f(t) \in L^2$ can be mapped into a spline space of order m as

$$f(t) \;\Rightarrow\; f_M(t) = \sum_k c_k^M N_m(2^M t - k) \;\in V_M \tag{28}$$

For $m = 1$, we have an orthonormal basis (see Ref. [10] for a complete theory). The scaling function coefficients c_k^M are computed from the signal samples using

$$c_k^M = 2^M \int_\infty^\infty f_M(t)\, \bar{\tilde{N}}_m\left(2^M t - k\right) dt \tag{29}$$

where $\tilde{N}_m(t)$ is the dual of $N_m(t)$ and $\bar{(\cdot)}$ is the complex conjugate. In practice the signal $f(t)$ is known at some discrete points. The given time step determines the scale M to which the function can be mapped. Because of the interpolatory representation $f(t) = f_M(t)$ at the sampling points, and because of the polynomial reproducibility, the representation is exact at every point for a polynomial of order m if the basis is N_m. In addition, since they are a local basis, the representation (28) is also local, which means we only need a few values of the function to obtain the coefficients c_k^M for some k.

For a given B-spline N_m we can construct the corresponding wavelet function ψ_m with minimal support (for more details on wavelet theory see Refs. [11-12]). A function $f(t)$ can be represented in terms of ψ as

$$f(t) = \sum_{j,k} d_k^j\, \psi\left(2^j t - k\right) \tag{30}$$

where

$$d_k^j = 2^j \int_\infty^\infty f(t)\, \bar{\tilde{\phi}}\left(2^j t - k\right) dt \tag{31}$$

Unlike the total positivity of splines, wavelets have the so-called "complete oscillation" property by virtue of which the wavelet coefficients help to detect any change in the function behavior. The algorithm based on a multiresolution analysis (MRA) has the following relations

$$\begin{aligned}
x_{M+1}(t) \in A_{M+1}\,, \quad &\Rightarrow \quad x_{M+1} = \sum_k c_k^{M+1} N_m\left(2^{M+1} t - k\right) \\
x_M(t) \in A_M\,, \quad &\Rightarrow \quad x_M = \sum_k c_k^M N_m\left(2^M t - k\right) \\
y_M(t) \in W_M\,, \quad &\Rightarrow \quad y_M = \sum_k d_k^M \psi\left(2^M t - k\right)
\end{aligned}$$

Since the MRA requires[11]

$$A_{M+1} = A_M + W_M \tag{32}$$

we have

$$x_{M+1}(t) = x_M + y_M \tag{33a}$$

$$\sum_k c_k^{M+1} N_m\left(2^{M+1}t - k\right) = \sum_k c_k^M N_m\left(2^M t - k\right) + \sum_k d_k^M \psi\left(2^M t - k\right) \tag{33b}$$

Using the decomposition relation[12] gives

$$N_m\left(2^{M+1}t - l\right) = \sum_k \left\{h_0\left[2k - l\right] N_m\left(2^M t - k\right) + h_1\left[2k - l\right]\psi\left(2^M t - k\right)\right\} \tag{34}$$

Substituting (34) into (33b) gives an equation in the same resolution M. Finally, the coefficients are computed as

$$c_k^M = \sum_l h_0\left[2k - l\right] c_l^{M+1} \tag{35a}$$

$$d_k^M = \sum_l h_1\left[2k - l\right] c_l^{M+1} \tag{35b}$$

where h_0 represents a low pass filter and h_1 a high pass filter. The input signal is projected in two subspaces A_M and W_M via the filter h_0 and h_1. In our particular case, these subspaces are the navigation plane (low frequency) and the attitude plane (high frequency). By repeating this algorithm, a signal decomposition at various frequency octaves is obtained, which is the standard MRA decomposition of any arbitrary function. The objective of this analysis is obtain two functions which may or may not be related to octave frequencies. To deal with this problem, as *interoctave parameter* is defined, given by[12]

$$\alpha_n = \alpha_n, N = \frac{2^N}{n + 2^N} \quad N > 0 \text{ and } n = 1, \ldots, 2^N - 1 \tag{36}$$

which gives $2^N - 1$ additional levels between any two consecutive octave levels.

The original problem of mapping x_M to x_M^n is similar to the standard MRA decomposition. For the special case of a B-spline, with $m = 2$ we have

$$x_M^n(t) = \sum_k c_{k,n}^M N_2\left(2^M \alpha_n(t) t - k\right) \tag{37}$$

with

$$c_{k,n}^M = x_M\left(\frac{k+1}{2^M \alpha_n}\right) \tag{38}$$

The proper selection of n and N allows the separation of any frequency band that is desired. The important fact of this interoctave algorithm is the original signal can be projected on each subspace without intersection (i.e., the attitude subspace does not receive any navigation signal).

Standard Band Pass Filter

In general the navigation motion is dominated by the orbital period of the master satellite which is around 6000 seconds for Low Earth Orbit (LEO). The attitude motion is related to the internal closed-loop design which is in general around a couple of hundred seconds. This frequency distance can be used to filter one motion with respect to the other a using standard band filter design. This approach saves previous data to sequentially apply the band filters at each time. Therefore, in general, it is more suitable for batch processing than for real time application.

OBSERVER DESIGN

In this section an observer is designed to estimate the relative attitude and angular velocity as well as the relative position and the linear velocity. In Ref. [5] the information matrix of the attitude and position estimation errors is explicitly calculated for two line of sight observations. The information matrix is divided into four partitions, where the two main diagonal elements correspond to the attitude and position information matrices that have the identical structure if each problem (i.e., attitude or position) was considered independent of each other. The off-diagonal partitions couple the attitude and the position errors. A diagonalization (i.e., a decoupling of the attitude and

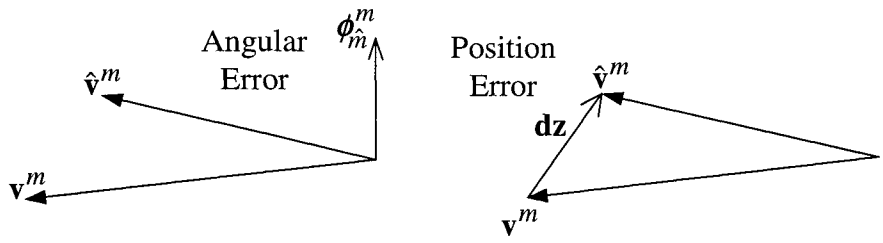


Figure 6: Angular and Position Error Visualization

position) of the information matrix occurs only in very special cases (the presence of a deterministic solution for example). Therefore, the entire problem which includes both attitude and position estimation should be considered in the observer design. Toward this end, the signal is separated using the previous methods, and both observers (attitude and position) are designed independently.

Attitude and Navigation Observer

The observer design treats the attitude portion by representing the residual (measurement minus estimate) error through a quaternion vector formulation, and treats the position portion of the residual in a straightforward position vector formulation. The angular error between the measured ($\mathbf{v}^m$) and the estimated ($\hat{\mathbf{v}}^m$) vectors in master frame can be "visualized" by a rotation axes normal to plane that contains both vectors. This axis ($\phi_{\hat{m}}^m$) can be interpreted as the vector part of the quaternion error, and the rotation angle between both vectors is the scalar part of the quaternion. The position error ($\mathbf{dz}$) is simple vector difference between the estimated and measured vectors. Figure 6 shows both approaches.

Before continuing with this concept, the following matrix relation is first written

$$C_s^m = C_{\hat{m}}^m \hat{C}_s^{\hat{m}} = \Delta C \hat{C}_s^{\hat{m}} \tag{39}$$

where the estimated vector, matrix or frames are noted with the superscript $\hat{(\cdot)}$, and $C_{\hat{m}}^m \equiv \Delta C$. The rotation error matrix between the estimated and measured quantities can be written in terms of the quaternion as $\Delta C = I + 2q_o\boldsymbol{\varrho}^k + 2[\boldsymbol{\varrho}\times]^2$. To simplify the notation this matrix is simply defined as $\Delta C \equiv (I + [\boldsymbol{\delta}\times])_{\hat{m}}^m$. Equation (1) can now be re-written as

$$\mathbf{v}^m = (I + [\boldsymbol{\delta}\times])_{\hat{m}}^m \hat{C}_s^{\hat{m}} \hat{\mathbf{v}}^s \tag{40}$$

where $\hat{\mathbf{v}}^s$ is an estimated vector,which depends on only of the angular motion (after signal separation). Equation (40) can be re-written in residual form as

$$\mathbf{v}^m - \hat{\mathbf{v}}^m = [\boldsymbol{\delta}\times]_{\hat{m}}^m \hat{\mathbf{v}}^{\hat{m}} \tag{41}$$

Using the multiplicative property of the cross product matrix the right hand side of Eq. (41) can be expressed in a more convenient form as

$$\hat{\mathbf{v}}^m - \mathbf{v}^m = [\hat{\mathbf{v}}\times]^{\hat{m}} (\boldsymbol{\phi}_{\hat{m}}^m)^{\hat{m}} \tag{42}$$

where the vector $\boldsymbol{\phi}_{\hat{m}}^m$ is expressed as the vector part of a quaternion in any frame. As stated previously this is one advantage of using the quaternion parameterization over the rotation matrix in the observer. The left hand side of Eq. (42) is denoted by $\mathbf{dz} \equiv \hat{\mathbf{v}}^m - \mathbf{v}^m$ for simplicity.

The number of measured line of sight vectors is generally greater than one, and the processing of this information can be done in the least square sense. Each estimated vector cross product is stacked into a matrix as

$$\hat{V}_{\hat{m}} = \begin{bmatrix} [\hat{\mathbf{v}}_1\times]^{\hat{m}} \\ \vdots \\ [\hat{\mathbf{v}}_N\times]^{\hat{m}} \end{bmatrix} \tag{43}$$

In this case the pseudoinverse is computed using all available information. Therefore, the quaternion error is computed by

$$\hat{V}_{\hat{m}}^+ \mathbf{dz} = \boldsymbol{\phi}_s^m \rightarrow [\boldsymbol{\delta\varrho}, \delta q_o] \tag{44}$$

Table 1: Orbital Elements of the Master Satellite

Semimajor axis	$a = 6,878$ km
Eccentricity	$e = 0$
Inclination	$i = 50°$
Node right ascension	$\Omega = 5°$
Argument of perigee	$\omega = 10°$
Mean anomaly	$M = 8°$

where $\hat{V}_{\hat{m}}^{+}$ is the pseudoinverse of $\hat{V}_{\hat{m}}$. The computation of the quaternion error is comparable to the algorithm presented in Ref. [13], but the scalar part of the quaternion (δq_o) is assumed to always be equal to $+1$. However, the scheme presented in this section maintains all four elements of the quaternion error because the sign of the scalar part is used in the design of the observer.

The nonlinear observer presented in Ref. [14] is used for attitude estimation; however, two slight modifications are introduced. The first one incorporates an angular velocity model, and the second includes a model of a potential bias, represented by $\mathbf{b}^m$, in the quaternion differential equation to include any offset of the sensor, which may even be the computation of the focal distance. The dynamics of the observer are given by

$$\begin{aligned} \dot{\hat{\boldsymbol{\varrho}}}_s^m &= \frac{1}{2}\left[\hat{q}_o I_{3\times 3} + [\hat{\boldsymbol{\varrho}}_s^m \times]\right] \\ &\quad \times \left[\hat{\boldsymbol{\omega}}_{ms}^m + \hat{\mathbf{b}} + K_v \delta\boldsymbol{\varrho}\,\mathrm{sign}\,(\delta q_o)\right] \end{aligned} \tag{45}$$

$$\dot{\hat{q}}_o = -\frac{1}{2}\left(\hat{\boldsymbol{\varrho}}_s^m\right)^T \left[\hat{\boldsymbol{\omega}}_{ms}^m + \hat{\mathbf{b}} + K_v \delta\boldsymbol{\varrho}\,\mathrm{sign}\,(\delta q_o)\right] \tag{46}$$

$$\dot{\hat{\boldsymbol{\omega}}}_{ms}^m = -\Lambda \hat{\boldsymbol{\omega}}_{ms}^m + \Delta\hat{\mathbf{T}}^m + K_p \delta\boldsymbol{\varrho}\,\mathrm{sign}\,(\delta q_o) \tag{47}$$

$$\Delta\dot{\hat{\mathbf{T}}}^m = -H\Delta\hat{\mathbf{T}}^m + K_T \delta\boldsymbol{\varrho}\,\mathrm{sign}\,(\delta q_o) \tag{48}$$

$$\dot{\hat{\mathbf{b}}}^m = -M\hat{\mathbf{b}}^m + K_b \delta\boldsymbol{\varrho}\,\mathrm{sign}\,(\delta q_o) \tag{49}$$

where $\hat{\mathbf{T}}$ is the torque difference estimate, $\hat{\boldsymbol{\omega}}_{ms}^m$ is the relative angular velocity, M is a diagonal positive definite matrix which represents the time constant of the process, and K_v, K_p, K_T and K_b are positive definite matrices. The sign function ensures that the smallest possible angle is chosen between the two equivalent rotations angles described by ϕ and $2\pi - \phi$. The stability proof of this observer can be found in Ref. [6].

The observers for the relative position and relative linear velocity are given by

$$\dot{\hat{\mathbf{r}}} = \hat{\mathbf{v}} - K_p\,\mathbf{dz} \tag{50}$$

$$\dot{\hat{\mathbf{v}}} = \mathbf{f}\,(\hat{\mathbf{r}}, \hat{\mathbf{v}}) - K_v\,\mathbf{dz} \tag{51}$$

where $\mathbf{f}\,(\cdot)$ is the right hand side of Eq. (14), $\mathbf{r}$ is the relative position vector, and $\mathbf{v}$ is the relative linear velocity vector. The minus signs in Eqs. (50) and (51) are due to the definition of $\mathbf{dz}$. The constant gains K_p and K_v are positive definite matrices (usually diagonal). The stability proof of this observer can be found in Ref. [6].

SIMULATION

The orbital elements used in the simulation of the master satellite are shown in Table 1. A small initial condition perturbation of these elements is used to simulate the motion of the secondary satellite. The true inertia matrices of both satellites is given by

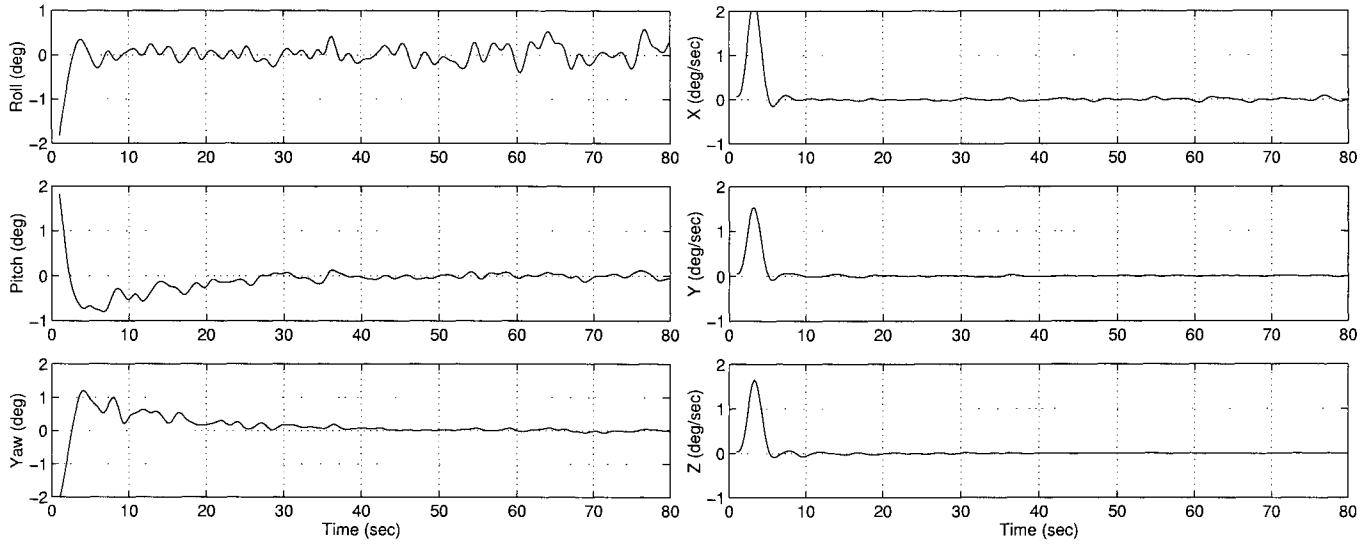
$$I_s = I_m = \mathrm{diag}[100,\ 120,\ 130]\ \text{N-m-s}^2 \tag{52}$$

In the observer the following inertia matrices are used:

$$I_s = I_m = \mathrm{diag}[110,\ 115,\ 140]\ \text{N-m-s}^2 \tag{53}$$

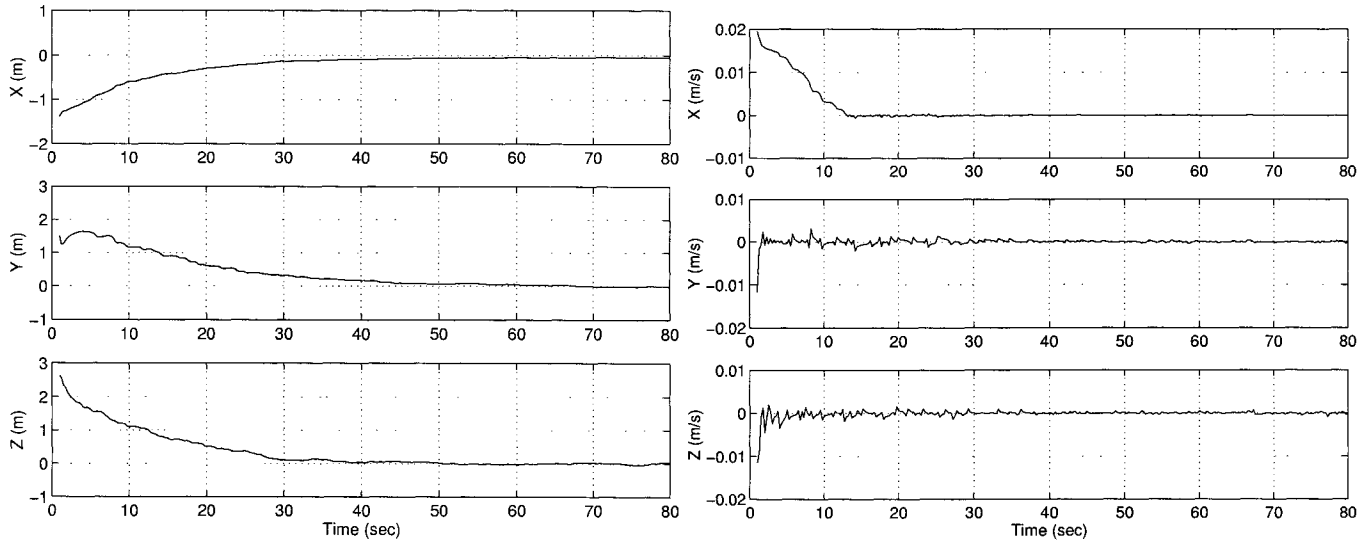
The true relative initial angular velocity is constant, and given by

$$\boldsymbol{\omega} = [0.065,\ 0.048,\ 0.03]^T\ \text{deg/sec} \tag{54}$$



(i) Attitude Errors

(ii) Angular Velocity Errors



(iii) Position Errors

(iv) Linear Velocity Errors

Figure 7: Attitude and Angular Velocity Errors, and Position and Linear Velocity Errors

The relative angular velocity trajectory is computed by integrating the following equation

$$\dot{\omega} = \lambda I_s^{-1} \omega \tag{55}$$

where I_s is the true inertia and $\lambda = 0.02$. A noise of $0.01/3000$ is assumed for each measurement on the focal plane. Four beacons are placed on the secondary satellite at a distance of 1 meter from the mass center along each coordinate axis. The fourth beacon is placed at $[1, 1, 1]^T$ in the secondary frame.

The observer described in the last section is implemented for state estimation from the line of sight measurements. The initial condition angular error is a rotation of about 15° along each of the coordinates axes. The initial angular velocity has 50 percent errors from Eq. (54). The initial position condition 10 percent from the true value and the initial linear velocity condition is 30 percent from the true value. The sampling rate is 4 Hz. The plots in Figure 7 show attitude and angular velocity errors, and position and linear velocity errors for the estimator. The relative

distance along the X axis is almost three times the distance along the other two axes (around 94 meters against 30 meters). This difference can be observed in the oscillation of the attitude error in roll, which is intuitively correct. The roll angle error is within 0.3 degrees, and the pitch and yaw angles are within 0.05 degrees. The position error in all three axis is within 1 cm. Also, the velocities are well estimated using the observer.

CONCLUSION

A novel vision-based sensor involving LED beacons and position sensing technology in the focal plane has been introduced for formation flying applications. In order to achieve an accurate line of sight measurement from this sensor a calibration procedure has been shown. Experimental results indicate that the calibration provides accurate results. Also, an observer based system has been presented as an alternative to the extended Kalman filter for formation flying navigation of spacecraft. Simulation results have shown that accurate relative attitude and position estimation is possible.

REFERENCES

[1] Junkins, J. L., Hughes, D. C., Wazni, K. P., and Pariyapong, V., "Vision-Based Navigation for Rendezvous, Docking and Proximity Operations," *22nd Annual AAS Guidance and Control Conference*, Breckenridge, CO, February 1999, AAS Paper 99-021.

[2] Junkins, J. L., Hughes, D., and Schaub, H., "A Laser Position Sensor," *U.S. Patent Pending*, October 2000, Disclosure 60-101,386.

[3] Wrigley, W., Hollister, W., and Denhard, W., *Gyroscopic Theory, Design, and Instrumentation*, MIT Press, Cambridge, MA, 1969.

[4] Light, D. L., "Satellite Photogrammetry," *Manual of Photogrammetry*, edited by C. C. Slama, chap. 17, American Society of Photogrammetry, Falls Church, VA, 4th ed., 1980.

[5] Crassidis, J. L., Alonso, R., and Junkins, J. L., "Optimal Attitude and Position Determination From Line of Sight Measurements," *The Richard H. Battin Astrodynamics Conference*, College Station, TX, March 2000, AAS Paper 00-268.

[6] Alonso, R., Crassidis, J. L., and Junkins, J. L., "Vision-Based Relative Navigation For Formation Flying of Spacecraft," *AIAA Guidance, Navigation and Control Conference*, Denver, CO, August 2000, AIAA Paper 2000-4439.

[7] Shuster, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, Oct.-Dec. 1993, pp. 439–517.

[8] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, American Institute of Aeronautics and Astronautics, Inc., New York, NY, 1987.

[9] Bate, R. R., Mueller, D. D., and White, J. E., *Fundamentals of Astrodynamics*, Dover Publications, New York, NY, 1971.

[10] Chui, C. K., "Multivariable Splines," *CBMS-NSF Series in Applied Mathematics*, No. 54, SIAM, Philadelphia, PA, 1988.

[11] Chui, C. K., *An Introduction to Wavelets*, Academic Press, Boston, MA, 1992.

[12] Goswami, J. C. and Chan, A. K., *Fundamentals of Wavelets: Theory, Algorithms, and Applications*, John Wiley & Sons, Inc., New York, NY, 1999.

[13] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, Sept.-Oct. 1982, pp. 417–429.

[14] Salcudean, S., "A Globally Convergent Angular Velocity Observer for Rigid Body Motion," *IEEE Transactions on Automatic Control*, Vol. 36, No. 12, Dec. 1991, pp. 1493–1497.

A PRELIMINARY FORMATION FLYING ORBIT DYNAMICS ANALYSIS FOR LEONARDO-BRDF

Steven P. Hughes
Flight Dynamics Analysis Branch
NASA Goddard Space Flight Center
Greenbelt, MD 20771

Laurie M. Mailhe
a.i. solutions, Inc.
Lanham, MD 20706

ABSTRACT

Leonardo-BRDF is a NASA mission concept proposed to allow the investigation of radiative transfer and its effect on the Earth's climate and atmospheric phenomenon. Enabled by the recent developments in small-satellite and formation flying technology, the mission is envisioned to be composed of an array of spacecraft in carefully designed orbits. The different perspectives provided by a distributed array of spacecraft offer a unique advantage to study the Earth's albedo. This paper presents the orbit dynamics analysis performed in the context of the Leonardo-BRDF science requirements. First, the albedo integral is investigated and the effect of viewing geometry on science return is studied. The method used in this paper, based on Gauss quadrature, provides the optimal formation geometry to ensure that the value of the integral is accurately approximated. An orbit design approach is presented to achieve specific relative orbit geometries while simultaneously satisfying orbit dynamics constraints to reduce formation-keeping fuel expenditure. The relative geometry afforded by the design is discussed in terms of mission requirements. An optimal two-burn initialization scheme is presented with the required ΔV to distribute all spacecraft from a common parking orbit into their appropriate orbits in the formation. Finally, formation-keeping strategies are developed and the associated ΔV's are calculated to maintain the formation in the presence of perturbations.

REDUCING FORMATION-KEEPING MANEUVER COSTS FOR FORMATION FLYING SATELLITES IN LOW-EARTH ORBIT

Nicholas Hamilton*

ABSTRACT

Several techniques are used to synthesize the formation-keeping control law for a three-satellite formation in low-earth orbit. The objective is to minimize maneuver cost and position tracking error. Initial reductions are found for a one-satellite case by tuning the state-weighting matrix within the linear-quadratic-Gaussian framework. Further savings come from adjusting the maneuver interval. Scenarios examined include cases with and without process noise. These results are then applied to a three-satellite formation. For both the one-satellite and three-satellite cases, increasing the maneuver interval yields a decrease in maneuver cost and an increase in position tracking error. A maneuver interval of 8-10 minutes provides a good trade-off between maneuver cost and position tracking error. An analysis of the closed-loop poles with respect to varying maneuver intervals explains the effectiveness of the chosen maneuver interval.

INTRODUCTION

Currently, formation flying spacecraft control is being extensively researched. This paper presents a strategy to reduce the amount of control needed for formation keeping within the framework of a linear-quadratic-Gaussian (LQG) controller. By varying the maneuver interval, the trade off between maneuver cost and position tracking error is discovered.

Speyer[1] first introduced a decentralized LQG control method. Carpenter[2,3] applied this work to formation flying satellites, and further expanded it to deal with both time-invariant and time-varying systems. Carpenter, Folta, and Quinn[4] investigated the decentralized framework for the applicability of autonomous formation flying control for the EO-1 mission to follow Landsat-7. In addition, Sparks[5] studied the long-term ΔV for a relative circular formation at an 800 km altitude orbit. Orbital rendezvous is a related problem to formation flying. Kluever and Tanck[6] looked at reducing ΔV for a geosynchronous orbit rendezvous problem using constant thrust magnitudes and varying thrust lengths, as well as varying maneuver intervals.

This paper will address tuning an LQG controller by adjusting the cost of the state tracking error in an effort to minimize the maneuver cost. These results will then be applied to a scenario that drives one satellite from random initial conditions to the origin. Further reductions to maneuver cost will be achieved by varying the maneuver interval. Finally, the one-satellite results will be tested on a three-satellite formation simulation.

---

* Second Lieutenant, United States Air Force, with NASA-Goddard Space Flight Center as part of the Program for Research and Education in Space Technology

SYSTEM MODEL AND CONTROLLER DESIGN

Both Kaplan[7] and Carpenter[3] describe the dynamics for formations of closely spaced satellites in low-Earth orbits. In summary, for each formation, an imaginary satellite, or hub, is in a circular low-Earth orbit. This hub defines a reference frame with radial, in-track, and cross-track components. The radial component is in the direction from the central body (Earth) to the hub, the in-track component is in the direction of the hub's motion, and the cross-track component is in the direction of the orbit normal. Hill's equations give the mathematics of the relative motion between the real satellites and the hub.

Controller Design

Carpenter built a standard LQG controller for the relative motion of the formation with respect to the hub. The cost function to be minimized for this problem is

$$J = \frac{1}{2} E\left[\begin{array}{l} \left\{X(t_N) - X^R(t_N)\right\}^T S_N \left\{X(t_N) - X^R(t_N)\right\} + \\ \int_{t1}^{t_N} \left\{ \begin{array}{l} \left[X(t) - X^R(t)\right]^T W(t) \left[X(t) - X^R(t)\right] + \\ \sum_{j=1}^{K} \left[U^j(t) - U^{Rj}(t)\right]^T V^j(t) \left[U^j(t) - U^{Rj}(t)\right] \end{array} \right\} dt \end{array} \right], \tag{1}$$

subject to the dynamic constraint

$$\dot{X} = AX + BU + w, \tag{2}$$

where X is the state vector consisting of the positions and velocities for each satellite in the formation, U is the control vector, and w is the process noise that has power spectral density

$$Q = \begin{bmatrix} Q^j & & \\ & \ddots & \\ & & Q^j \end{bmatrix}$$

$$Q^j = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9.81e-6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9.81e-6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 9.81e-6 \end{bmatrix}. \tag{3-a,b}$$

U-U^R is the maneuver cost, and X-X^R is the state tracking error. Specifically for this application,

$$\begin{aligned} U &= \Delta V \\ U^R &= 0 \\ S_N &= 0 \end{aligned} . \tag{4-a,b,c}$$

Of interest to this paper are the W and V matrices. In Equation 1 they can be time varying, but are constant matrices in this study. W is the continuous state-weighting

matrix, and V is the continuous control-weighting matrix. For this study, Equation 1 can be simplified to

$$J = \frac{1}{2} E\left[\int_{t_1}^{t_N}\left\{\left[X(t) - X^R(t)\right]^T W\left[X(t) - X^R(t)\right] + \sum_{j=1}^{K}\left[U^j(t)^T V^j U^j(t)\right]\right\}dt\right]. \tag{5}$$

Carpenter[3] also gives the discrete forms

$$W_d(t_{i+1},t_i) = \int_{t_i}^{t_{i+1}} \Phi(t,t_i)^T W(t)\Phi(t,t_i)dt$$

$$V_d^j(t_{i+1},t_i) = \int_{t_i}^{t_{i+1}}\left[B_d^j(t,t_i)^T W(t)B_d^j(t,t_i) + V^j(t)\right]dt \tag{6-a,b}$$

where Φ is the state transition matrix, and B^j is the continuous control mapping whose discrete form is given by

$$B_d^j(t,t_i) = \int_{t_i} \Phi(t,\tau)^T B^j(\tau)d\tau\,. \tag{7}$$

The j in the above equations is the notation describing each node (or satellite) in the formation. The integrals in Equations 6a, 6b, and 7 are approximated as

$$\begin{aligned} W_d(t_{i+1},t_i) &\approx \left[\Phi(t_{i+1},t_i)^T W\Phi(t_{i+1},t_i)\right]\cdot\left[t_{i+1} - t_i\right] \\ V_d^j(t_{i+1},t_i) &\approx \left[B_d^j(t_{i+1},t_i)^T W B_d^j(t_{i+1},t_i) + V^j\right]\cdot\left[t_{i+1} - t_i\right], \\ B_d^j(t_{i+1},t_i) &\approx \left[\Phi(t_{i+1},t_i)^T B^j(t_i)\right]\cdot\left[t_{i+1} - t_i\right] \end{aligned} \tag{8-a,b,c}$$

except as noted below.

In this study, I simulate the control and tracking of both one-satellite and three-satellite formations with varying maneuver intervals. For all scenarios, the simulation runs for two revolutions of the hub around the earth. The maneuvers are considered ideal and impulsive.

LQR Controller Tuning

One way to reduce control effort is to tune the state weighting matrix and/or the control-weighting matrix. The relationship between these two matrices is what matters, rather than their individual values. Therefore, the control-weighting matrix can be kept constant, at identity, while the state-weighting matrix is varied. The controller is tuned without noise.

The state weighting matrix is a block diagonal matrix with each block relating to an individual satellite. All satellites are assumed to be identical, so

$$W = \begin{bmatrix} W_j & & \\ & \ddots & \\ & & W_j \end{bmatrix}. \tag{9}$$

Because the satellites are assumed identical, tuning can be done on one satellite and then applied similarly to others if necessary. For the one satellite case,

$$W = W_j\,. \tag{10}$$

Denoting the position weights by a and the velocity weights by b, I assume a diagonal W_j:

$$W_j = \begin{bmatrix} a & & & & & \\ & a & & & & \\ & & a & & & \\ & & & b & & \\ & & & & b & \\ & & & & & b \end{bmatrix}, \tag{11}$$

where, as a starting point,

$$\begin{aligned} a &= 0.04 \\ b &= 40000 \end{aligned}. \tag{12-a,b}$$

To tune W_j, initial conditions are chosen, an a is chosen, the simulation is run, and the process is repeated until the minimum ΔV is found for that set of initial conditions. In the simulation, the satellite is driven from its initial conditions to the reference orbit. The reference orbit for one satellite is simply zero with respect to the hub. In other words, the satellite is driven from an initial offset to the origin. Total ΔV in this case is the sum of the absolute values of the control for every maneuver. I found that the magnitude of the initial displacements has no effect on determining which a is best. However, because the radial and in-track states are coupled, the relationship between the initial conditions on those two displacements does have an effect on which a is best. I performed two investigations. In the first, I studied only in-plane initial conditions, and in the second, only out-of-plane initial conditions.

For the in-plane study,

$$\begin{aligned} X_0 &= \begin{bmatrix} r_0 & i_0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \alpha &= \left| r_0 / i_0 \right| \end{aligned}, \tag{13-a,b}$$

where r_0 is the radial initial condition and i_0 is the in-track initial condition. Table 1 shows how the "best" a varies with α:

Table 1

α	∞	100	10	5	2	1	0.9	0.8	0.7	0.6	0.5
Best a	0.16	0.16	0.14	0.13	0.09	0.05	0.05	0.04	0.03	0.03	0.02

By choosing the best a, the total ΔV for 2 revolutions can be reduced by up to 1 m/s for some cases when the initial displacements are on the order of 500 meters. As α approaches zero (the in-track initial displacement is much greater than the radial initial displacement), the minimum ΔV occurs at a value of a less than 0.02. However, the satellite does not converge on the reference orbit within one revolution. Choosing a to be 0.02 for values of α less than 0.5 ensures convergence within one revolution, even though the ΔV is not a minimum.

Rather than implement a table lookup in the software to determine the a for a given initial state vector, two scenarios will be investigated. The first case sets α equal to one. This is a reasonable choice applicable for running random or semi-random initial state vectors. Choosing the weighting $a = 0.05$ ($\alpha = 1$), a similar process to finding the best a can be done to determine the best b. The total ΔV savings for differing b are very small, on the order of 0.01 m/s for two revolutions. Nevertheless, the best b is 39000 for

$a = 0.05$. The second case is α equal to one half. This is chosen based on the three-satellite formation reference orbit proposed by Alfriend, Schaub, and Gim[8] for the TechSat21 program. Choosing the weighting $a = 0.02$ ($\alpha = 0.5$), the best b is once again 39000.

The process of finding the best a is repeated for the uncoupled cross-track initial displacements as well. For this case,

$$X_0 = [0 \quad 0 \quad c \quad 0 \quad 0 \quad 0]. \tag{14}$$

I found the best a to be 0.03, and the corresponding best b to be 36000, regardless of the initial cross-track displacement. For the cross-track offset of 500 meters, altering a and b only resulted in total ΔV savings of 0.05 m/s for two revolutions.

PERFORMANCE EVALUATION

One-Satellite Simulation

Next, I used simulations with random initial conditions with and without process noise to determine the performance of the controller. For these cases, the state weighting matrix is chosen to be

$$W = W_j = \begin{bmatrix} 0.05 & & & & & \\ & 0.05 & & & & \\ & & 0.05 & & & \\ & & & 40000 & & \\ & & & & 40000 & \\ & & & & & 40000 \end{bmatrix}. \tag{15}$$

This choice is not the best for the cross-track displacements, but the effect is negligible once noise is introduced. First, 15 "semi-random" initial displacements are chosen in the radial and in-track directions. The offsets range from –500 meters to +500 meters in both directions. The set of runs over these 15 points yields statistical results that are used to determine trends. The 15 points are shown below in Figure 1 and remain the same for all subsequent cases.

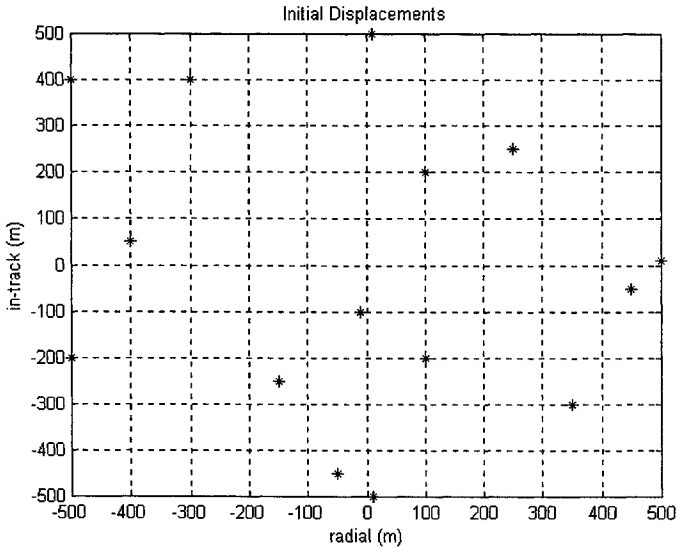


Figure 1: Assorted initial conditions for one satellite simulation

The total ΔV is defined as the sum of the absolute value of all control effort. The total ΔV is calculated for simulations of each of the 15 initial conditions, with all noise turned off. The RMS is then calculated over the set of 15 total ΔV's with no noise, and found to be 2.2634 m/s. The maneuver interval in this case is one minute.

Next, the process noise is turned on and the simulation is run again for each of the 15 points. In addition to determining the total ΔV, the steady-state tracking error is calculated as well. Steady-state tracking error is measured by defining Δx, which is a statistical determination of how far the satellite is from its reference orbit (*during the second half of the simulation.*)

Let

$$\Delta x^j = RSS(\Delta x_n^j), \tag{16}$$

where

$$\Delta x_n^j = \sqrt{(r_n - r_n^{ref})^2 + (i_n - i_n^{ref})^2 + (c_n - c_n^{ref})^2}\,. \tag{17}$$

r_n, i_n, and c_n are the radial, in-track, and cross-track positions of each satellite at some time n; and r_n^{ref}, i_n^{ref}, and c_n^{ref} are the radial, in-track, and cross-track reference positions for each satellite at some time n. Figure 2 shows Δx_n plotted against time for the one-satellite scenario at different maneuver intervals.

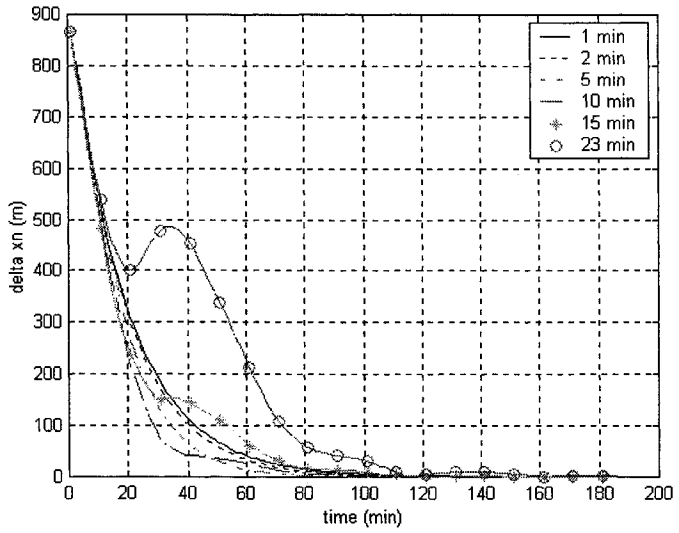


Figure 2: Position tracking error (Δx_n) versus time

The RMS is calculated over the set of 15 ΔV's and 15 Δx's for each maneuver interval. Table 2 and Figure 3 summarize the results of the simulations using various maneuver intervals.

Table 2: Data corresponding to Figure 3

Maneuver interval	1 min	2 min	5 min	10 min	15 min	23 min
RMS ΔV (m/s)	8.4514	5.8260	3.3272	2.1793	1.6827	1.3920
RMS Δx (m)	58.3374	74.2114	143.2634	317.4274	646.8495	1058.5662

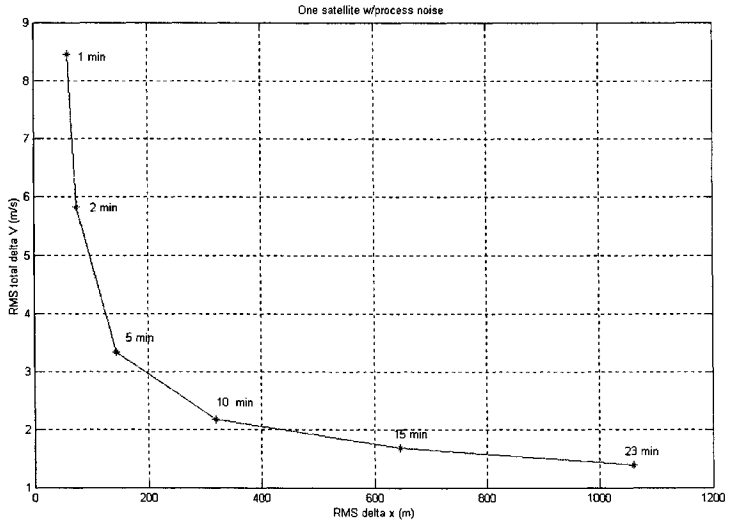


Figure 3: RMS of total ΔV versus the RMS of Δx for one satellite with process noise

As one can see, with process noise on, the amount of control or total ΔV needed is higher (8.4514 versus 2.2634 m/s). One way to reduce total ΔV is to reduce the number of maneuvers performed. The trade off for reducing ΔV using this method is that the Δx increases. In other words, reducing the number of maneuvers decreases the amount of control needed, but the displacement error due to the noise increases.

Three-Satellite Simulation

Next, I tested the controller design in a three-satellite simulation. For this study, I used the "best" a and b determined above for the state weighting matrix.

$$W_j = \begin{bmatrix} 0.02 & & & & & \\ & 0.02 & & & & \\ & & 0.03 & & & \\ & & & 39000 & & \\ & & & & 39000 & \\ & & & & & 36000 \end{bmatrix} \qquad (18\text{-a,b})$$

$$W = \begin{bmatrix} W_j & & \\ & W_j & \\ & & W_j \end{bmatrix}$$

The reference orbit is the "circular horizontal plane" formation proposed by Alfriend, Schaub, and Gim[8]. The horizontal plane is formed by the in-track and cross-track basis vectors. The radial projection is half the length of the in-track projection. This ellipse is then inclined out-of-plane so the projection on the horizontal plane is circular. Finally, the satellites are arranged such that their projections on the horizontal plane are always spaced 120 degrees apart. The formation appears to be a rotating equilateral triangle in the circular horizontal plane with the hub at the center of the triangle. For initial conditions, all satellites are started at the hub with zero velocity:

$$X_0^j = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]. \qquad (19)$$

One difference between this study and the previous studies is in the calculation of the discrete cost weighting matrices, W_d and V_d. Using the MATLAB symbolic toolbox, the state transition matrix can be expressed symbolically. With this, the exact definite integrals shown in Equations 6a, 6b, and 7 can be calculated exactly rather than approximated by Equations 8a, 8b, and 8c.

With noise turned off, the simulation is run at the different maneuver intervals. In addition to Δx, another determination of tracking error is calculated which I will call μ_x. μ_x is similar to Δx except that instead of taking the time-wise RSS of the Δx_n vector, the mean is taken. This is still done for the second half of the simulation, after the initial convergence. Table 3 and the Figure 4 illustrate the results.

Table 3: Data corresponding to Figure 4

Satellite # 1

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
ΔV (m/s)	9.007	8.179	5.660	4.654	4.452	4.310	3.881	3.454
μ_x (m)	10.224	1.676	1.463	0.979	0.883	0.907	3.851	17.108

Satellite # 2

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
ΔV (m/s)	14.121	12.63	8.609	6.732	6.363	6.052	5.044	4.650
μ_x (m)	6.713	2.071	1.662	1.606	1.646	1.902	7.688	33.768

Satellite # 3

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
ΔV (m/s)	10.807	9.370	7.027	6.066	5.863	5.696	5.142	4.694
μ_x (m)	5.857	2.505	0.471	0.795	1.087	1.693	9.442	35.049

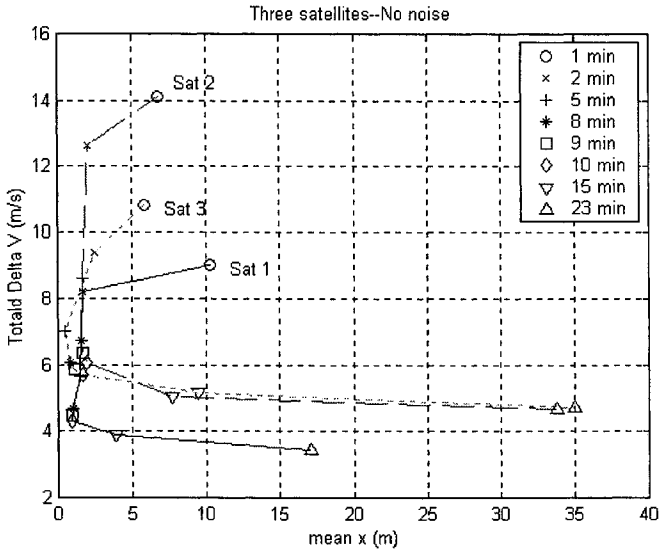


Figure 4: Total ΔV versus μ_x for three-satellite formation with no noise

As expected, with fewer maneuvers, less total ΔV is required. However, μ_x actually is at a minimum around the 8-10 minute maneuver intervals. This differs from the one satellite case where the position tracking error was smaller with more maneuvers.

Figure 5 shows the closed-loop poles in the polar plane (z-plane) as they vary with increasing maneuver intervals. From Phillips and Nagle[9], the poles are of the form

$$z = e^{sT}, \tag{20}$$

where z is the discrete closed-loop pole, s is the continuous closed-loop pole, and T is the maneuver interval.

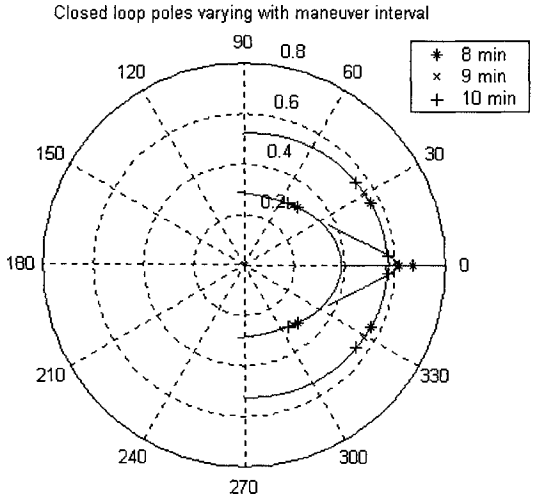


Figure 5: Closed-loop poles varying with maneuver interval (one satellite)

Poles within the unit circle are stable. Specifically looking at the set of poles on the far right in Figure 5, at the 8 minute maneuver interval they are still on the real axis; at the 9 minute maneuver interval they have just left the real axis; and at the 10 minute maneuver interval they have further diverged. Converting this set of poles back to the s-plane, the settling time for a second order system is approximated by Nise[10] as

$$T_s = -8/(s1 + s2), \tag{21}$$

where $s1$ and $s2$ are the corresponding s-plane poles at a given maneuver interval. A second order system assumption is valid because these poles are the most dominant. Figure 6 shows the relationship between settling time and maneuver interval.

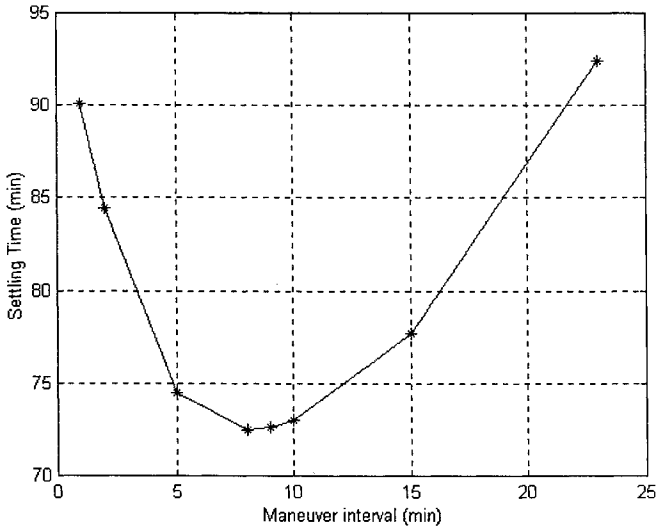


Figure 6: Settling time versus maneuver interval

Note that around the 8-10 minute maneuver interval, a minimum settling time occurs. This helps to explain the minimum μ_x around the same maneuver interval range.

Next, the process noise is turned on and the simulation is run five times for each of the maneuver intervals. The RMS is taken for the total ΔV and μ_x for each satellite at each maneuver interval. Table 4 and Figure 7 show the results.

Table 4: Data corresponding to Figure 7

Satellite # 1

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
RMS ΔV (m/s)	13.734	10.140	6.653	5.283	5.107	4.888	4.334	3.734
RMS μ_x (m)	11.226	13.565	14.303	27.780	27.877	27.119	56.006	95.106

Satellite # 2

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
RMS ΔV (m/s)	18.005	14.276	9.258	7.217	6.597	6.261	5.110	4.767
RMS μ_x (m)	9.021	11.779	17.701	27.796	27.505	36.165	44.963	111.545

Satellite # 3

Maneuver Interval	1 min	2 min	5 min	8 min	9 min	10 min	15 min	23 min
RMS ΔV (m/s)	14.857	11.784	8.217	6.950	6.637	6.341	5.446	4.875
RMS μ_x (m)	7.962	7.816	12.014	20.882	26.240	22.216	39.782	117.646

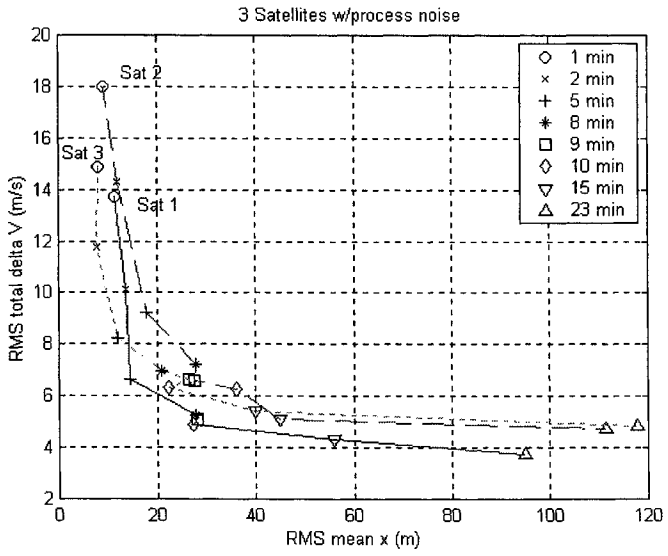


Figure 7: RMS of total ΔV versus RMS of μ_x for three-satellite formation with process noise

These results show the same trend as the one satellite case. Fewer maneuvers correspond to less total ΔV, but also lead to increases in position tracking error. A small discrepancy remains around the 8-10 minute maneuver intervals. I attribute this to only having five runs from which to take the statistics. I believe more runs would smooth these curves out.

CONCLUSION

A method for reducing formation-keeping maneuver cost has been developed. Tuning the state-weighting matrix of a single satellite yields a relationship between initial conditions and total ΔV. These results are applied to the three-satellite formation based on the desired geometry. By altering the maneuver interval, a relationship has been found between total ΔV and position tracking error. For all cases, fewer maneuvers require less ΔV. However, fewer maneuvers also tend to increase the position tracking error. Based on a closed-loop pole analysis neglecting noise, a minimum mean position tracking error is found to correspond to an 8 minute maneuver interval. This analysis is verified by the simulation. Once process noise is included, position tracking error continually increases as fewer maneuvers are performed and is inversely proportional to the ΔV needed. An 8-10 minute maneuver interval appears to be a good tradeoff between maneuver cost and position tracking error.

REFERENCES

1. Speyer, JL. "Computation and Transmission Requirements for a Decentralized Linear-Quadratic-Gaussian Control Problem." *IEEE Transactions on Automatic Control* 1979; AC-24(2): 266-269.

2. Carpenter, JR. "A Preliminary Investigation of Decentralized Control for Satellite Formations." Proceedings of the 2000 IEEE Aerospace Conference, March 18-25, 2000.

3. Carpenter, JR. "Decentralized Control of Satellite Formations." To appear in *International Journal of Robust and Nonlinear Control*, John Wiley & Sons, Inc: New York.

4. Carpenter, JR, Folta, DC, and Quinn, DA. "Integration of Decentralized Linear-Quadratic-Gaussian Control into GSFC's Universal 3-D Autonomous Formation Flying Algorithm." Paper, AIAA 99-4269, AIAA Guidance Navigation & Control, Modeling & Simulation Technologies and Atmospheric Flight Mechanics Conference and Exhibit, August 9-11, 1999, Portland, OR.

5. Sparks, A. "Linear Control of Spacecraft Formation Flying." Paper, AIAA 2000-4438, AIAA Guidance, Navigation, and Control Conference, August 14-17, 2000, Denver, CO.

6. Kluever, CA and Tanck, GS. "A Feedback Guidance Scheme for Orbital Rendezvous." *The Journal of the Astronautical Sciences*, Vol. 47, Nos. 3 and 4, July-December 1999: 229-237.

7. Kaplan, MH. *Modern Spacecraft Dynamics & Control.* John Wiley & Sons, Inc: New York, 1976: 108-112.

8. Alfriend, KT, Schaub, H, and Gim, DW. "Gravitational Perturbations, Nonlinearity and Circular Orbit Assumption Effects on Formation Flying Control Strategies." Paper, AAS 00-012, 23rd Annual American Astronautical Society Guidance and Control Conference, February 2-6, 2000, Breckenridge, CO.

9. Phillips, CL and Nagle, HT. *Digital Control System Analysis and Design.* Third Edition, Prentice-Hall, Inc: New Jersey, 1995.

10. Nise, NS. *Control Systems Engineering.* Second Edition, Addison-Wesley Publishing Company: Menlo Park, CA, 1995.

AN INVESTIGATION OF ATMOSPHERIC PENETRATION TO MODIFY THE RIGHT ASCENSION OF ORBITAL PLANES

Joseph R. Schultz and Darryll J. Pines
University of Maryland - College Park

ABSTRACT

Formation flying satellites positioned in similar orbits that vary only by different longitudes of the ascending node, Ω, create a regular, distributive constellation to cover the Earth. If the satellites in the constellation need to be launched by the same launch vehicle, there are two standard ways to transfer them to their proper position. The first approach involves transferring the satellites with a propulsive impulse maneuver that is quick but requires a large amount of fuel. The second approach uses the Earth's J2 gravitational effect to position the satellites into their respective orbits without using much fuel, however, it may take many months to do so leading to higher operational costs. One novel approach that may lead to overall lower costs than both standard methods is to use atmospheric penetration maneuvers. Such maneuvers would be used to "dip" into the atmosphere and take advantage of aerodynamic lift that can be designed into the satellite to achieve the new desired Ω. Depending on the orbit and satellite characteristics, this maneuver requires less fuel than the impulsive case and can be conducted on the order of days as compared to months over the J2 method. This paper details the atmospheric penetration maneuver and the conditions in which this transfer method will provide cost savings over the two approaches that are commonly used today.

INTRODUCTION

There are a variety of scientific and military benefits for changing the right ascension of a low-earth-orbiting satellite. (See Figure 1 for a review of the orbital elements used in this paper.) Some of these benefits include:

- Ability to change the satellite's ground track to monitor different regions of the earth.
- Ability to monitor different regions of space for science applications such as measuring the magnetosphere around the Earth.
- Construction of a synchronous constellation from multiple satellites for concurrent science measurements.

The first two benefits are mainly isolated to the current measurement needs of a single satellite. For example, a single Earth-sensing satellite will have a ground track over a certain region of the Earth. If the satellite wishes to view other Earth regions on a different ground track, a change in the orbit's Ω parameter is one way to do this. Or perhaps a satellite previously measuring the magnetosphere on the Earth's sun-facing side wishes to maneuver to now measure the magnetosphere in the region of the Earth's shadow. An orbital maneuver to change Ω will be required. Some of the ideas and strategies developed in this paper can be used to enhance the performance of these two benefits.

This paper will concentrate, however, on the third benefit listed: construction of a synchronous constellation architecture. This benefit concerns itself not only with the immediate location requirements of a single satellite, but on the overall location needs of multiple satellites working in unison to address complex mission requirements.

Satellite constellations provide advantages over single satellites in that their science instruments can collectively measure different regions of the Earth or space (celestial objects or fields) at the same time. A synchronous constellation architecture can provide the advantage of allowing regular concurrent measurements over evenly spaced regions. This allows scientists to take "snapshots" of a reading in a

predictable and useful manner. This paper will concentrate on a synchronous constellation where all of the satellites have the same orbital parameters (same apogee and perigee height, inclination, and argument of periapsis) except for their value of Ω.

Despite the advantages, construction of a constellation is neither cheap nor easy. It is expensive to launch a large number of satellites into different orbits. To combat these costs, NASA and industry are leaning towards smaller and smaller satellites in order to keep the launch costs down. With the development of smaller satellites, launch costs can be kept low if several satellites of a constellation are launched from the same launch vehicle. Once the satellites are launched together into the same orbit, however, the problem becomes: How does one separate the satellites into the desired constellation? (See Figure 2.)

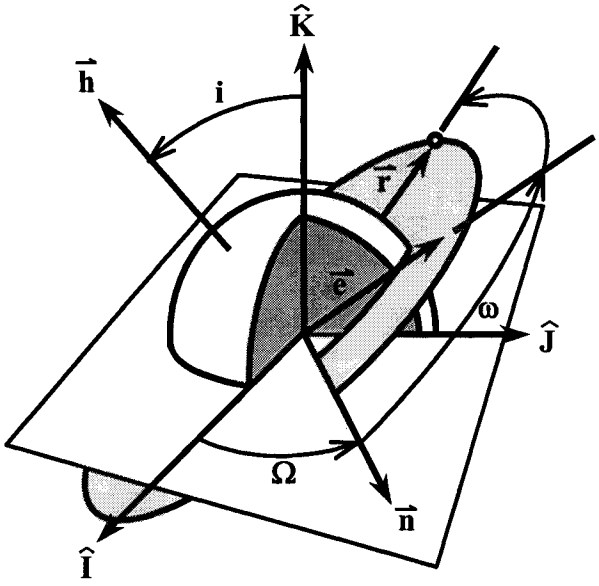


Figure 1. Summary of Orbital Parameters

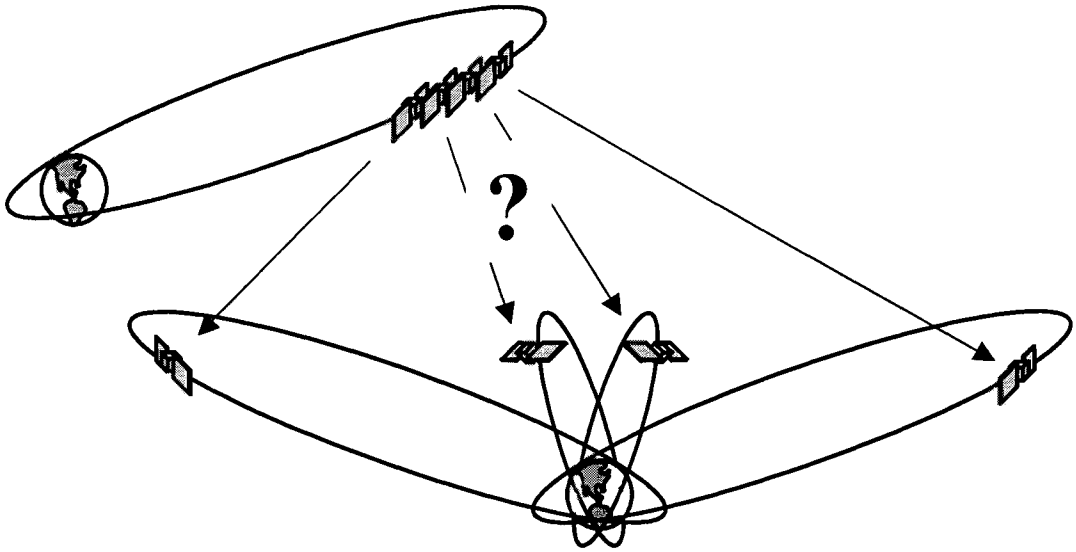


Figure 2. How does one create the desired constellation?

Two methods are currently used in practice to change the right ascension, Ω, of a satellite's orbit. They are: 1) the use of the Earth's oblateness via the J2 perturbation effect, and 2) the use of impulsive change-of-velocity (ΔV) maneuvers. Both methods will be briefly reviewed in this paper.

An alternate approach to change the Ω parameter of a satellite's orbit is the use of atmospheric penetration. This method entails sending the satellite into the atmosphere and taking advantage of aerodynamic forces to change orbital parameters. While this method will also require some propulsive maneuvering, under certain conditions the total fuel, or ΔV cost, is less than the pure impulsive ΔV maneuver.

This paper will explore the use of atmospheric penetration as an effective method for changing the right ascension of an orbiting satellite. Issues to be addressed include the number of atmospheric passes needed, the effect of the lift-to-drag (L/D) ratio and ballistic coefficients (BCs) on the total ΔV cost, and how the argument of periapsis, ω, can affect the aeroassist results.

SYMBOLS

a	- orbit's semi-major axis; acceleration
BC	- ballistic coefficient
C_L	- lift coefficient
C_D	- drag coefficient
e	- orbit's eccentricity; eccentricity vector
F	- force
velocity	
h	- angular momentum vector
i	- orbit's inclination
J_2	- zonal harmonic perturbation coefficient
L/D	- lift to drag ratio
m	- mass of the satellite
M_o	- mean anomaly of satellite
n	- mean motion; line of nodes vector
p	- orbit's semi-parameter
$p_{transfer}$	- transfer orbit's semi-parameter
r	- satellite's radius
R_E	- Earth's radius
S	- surface area of the satellite
t	- time
u	- argument of latitude of satellite
v, V	- velocity of the satellite
ΔV	- amount of velocity change require for a maneuver
$V_{transfer}$	- velocity at transfer orbit
ΔV_{total}	- total ΔV of maneuver
X	- state space of satellite's radius and vector
ϕ	- lift direction angle from the orbital plane
μ	- gravitational coefficient
ν	- true anomaly of satellite
ω	- orbit's argument of periapsis
Ω	- orbit's longitude of the ascending node
$\Delta\Omega$	- change in the longitude of the ascending node
ρ	- atmospheric density
Subscripts:	
I,J,K	- components in IJK frame
R,S,W	- components in RSW frame
1,2,3	- location points

THEORETICAL DEVELOPMENT

J2 Effect

Background

It is well-documented that the oblateness of the Earth (bulging at the equator) causes changes in a satellite's orbit similar to the way a spinning toy top precesses on its axis.[1] A disturbing-potential function can represent the perturbing effect of the non-spherical Earth on an orbiting satellite, and the "J2" zonal harmonic is clearly the largest contributor of this effect. For this paper, secondary effects of other zonal harmonics, sectorial harmonics, and tesseral harmonics will not be considered. The orbital parameters of both the longitude of the ascending node, Ω, and the argument of periapsis, ω, will change depending on the inclination of the orbit and its length of the semi-major axis.

Using the J2 effect to change Ω doesn't require any ΔV costs directly, but it can take on the order of months to separate a pair of satellites by a $\Delta\Omega$ amount of 180 degrees. As far as costs are concerned, the maneuver will require small ΔV costs to implement, but one would have to spend many months of operational costs before the satellite is in the right orbit to conduct the desired science.

Nodal Regression Rate

Equations to calculate the amount of Ω change due to the J2 effect can be found in most standard orbital mechanics texts. The amount of Ω change (also commonly referred to as the "nodal regression rate") is given by Equation 1.

$$\dot{\Omega} = -\frac{3nR_E^2 J_2}{2p^2}\cos(i) \qquad (1)$$

As an example, for one satellite in a 500 km circular orbit inclined by 30 degrees, the Ω rate change is approximately 6.6 deg/day. To change Ω for a single satellite by the maximum of 180 degrees, it would take a little over 27 days to make the change.

If there are two satellites in the same 500 km circular orbit, they would both change their Ω with the same rate of 6.6 deg/day, but they would not separate from each other. In order to separate one from the other so that they have the same orbital parameters but with a different Ω, one of the satellites would have to be sent into a different orbit where the Ω rate change is different.

For instance, if the second satellite were sent into a 500km x 1000km orbit, then its Ω rate change would be 6.3 deg/day. The two satellites would separate by a relative distance of 0.3 deg/day. At that rate, it would take 600 days to achieve a 180-degree separation from each other. If a greater Ω rate change is desired, it would be necessary to send the second satellite into an even higher orbit, but the transfer time would still likely be on the order of months.

In order to send the second satellite into another orbit, it would take, at minimum, two fuel burns. The complete maneuver would be conducted as follows: Both satellites start in the same orbit. One satellite does a fuel burn to send it into another orbit. After some time, the desired relative ΔΩ separation is achieved. Finally, the separated satellite does a second fuel burn to bring it back to the same orbit with a new right ascension, Ω.

ΔV Cost of Change

The amount of time required for the desired separation is related to both the ΔV cost needed to separate the satellites from the original orbit and the method used to achieve the separation. The time vs. cost for a sample orbit is shown in Figure 3.

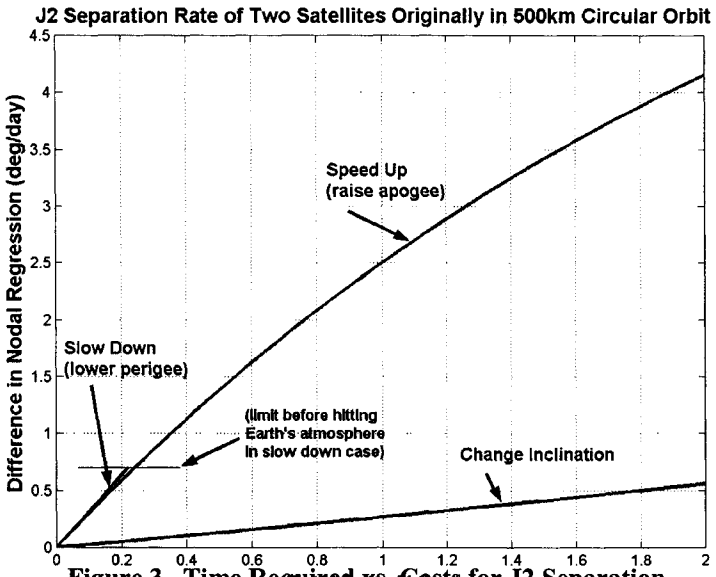


Figure 3. Time Required vs. Costs for J2 Separation

From a circular orbit, the three most fuel-efficient ways to change into another orbit are to slow down and lower the apogee, speed up and raise the apogee, or change the inclination. Lowering the perigee is the most efficient way to change Ω since it produces the greatest ΔΩ change per ΔV expenditure, but this way is limited by lowering too far and into the Earth's atmosphere. Invariably, one has to raise the apogee to achieve any reasonable Ω separation rate change.

Note that timing issues to re-synchronize the satellites at the end of the maneuver are not considered. Smart planning or other additional orbits would address any timing issues.

Impulsive ΔV Maneuver

Background

To change the orbit in a much faster time frame, the only method currently available is via a propulsive impulse maneuver. These maneuvers can be completed, if needed, within only a few orbital

periods instead of months as in the J2 method. The disadvantage, however, is that the ΔV cost required to complete the maneuver is much larger.

Optimal ΔV Maneuvers

There are many ways in which one can move from one orbit to another using propulsive impulse maneuvers.[2] Finding the optimal solution that gives the least amount of fuel usage cannot be obtained in closed form, but the equations are not difficult to solve numerically. The transfer methods possible can generally be grouped into the number of burns used for the transfer. For the low eccentricity orbits examined in this paper, it will be more often the case that the two-impulse transfer is the most efficient. Thus, for simplicity, this paper will limit itself to just the two impulse case for comparison to the J2 maneuver and aeroassist. A more detailed analysis could include multiple-impulse transfers.

The two-impulse transfer uses an intermediate transfer orbit to get from the initial orbit to the final desired orbit. This method is outlined in Figure 4.

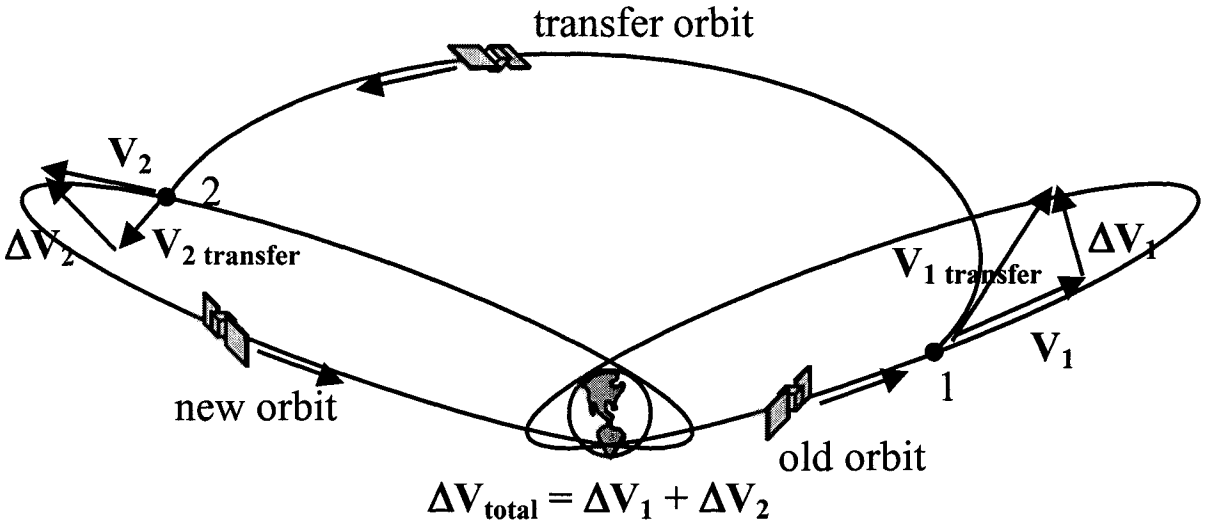


Figure 4. Two Impulse Orbital Transfer

Referring to Figure 4, at point 1 in the original orbit, the satellite is traveling with velocity V_1. An impulsive burn of velocity ΔV_1 is then made at this point to send the satellite along the transfer orbit with its new velocity of $V_{1transfer} = V_1 + \Delta V_1$. When the satellite reaches point 2, it will have a velocity of $V_{2transfer}$. A second impulsive burn of velocity ΔV_2 is made to result in a velocity at that point of $V_2 = V_{2transfer} + \Delta V_2$ that matches the velocity needed to place it in the final orbit. The total velocity required for this maneuver is thus $\Delta V_{total} = \Delta V_1 + \Delta V_2$.

ΔV Cost of Change

The equations describing the optimal transfer orbit are generally not solvable analytically. So to find the lowest ΔV required to transfer from one orbit to the next, numerical methods were used. Reference 3 provides a useful algorithm one could use.

The actual amount of ΔV required depends on how much change in Ω is desired. Figure 5 shows the minimum ΔV cost for a desired Ω change. As would be guessed, for a larger desired value of ΔΩ, more ΔV must be spent to achieve it.

Note that the ΔV in this propulsive impulse case is generally far larger than the ΔV required for the same ΔΩ change in the J2 effect case. However, the maneuver will be completed on the order of one orbit's time period instead of months.

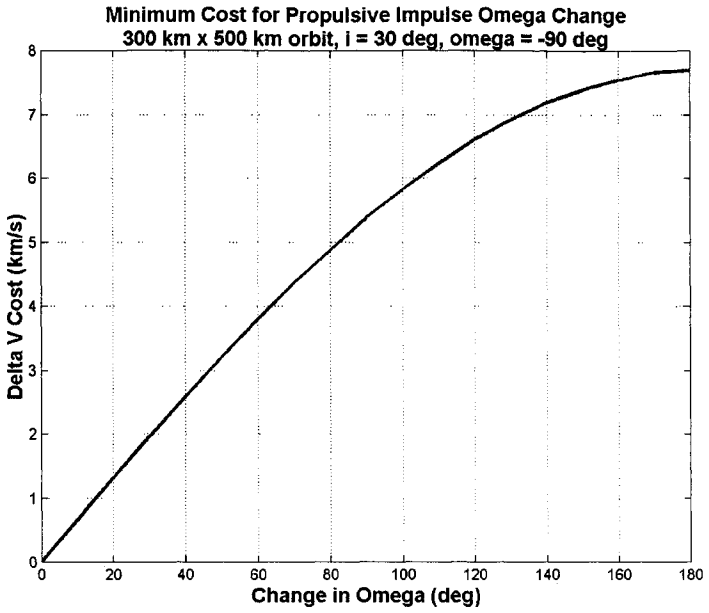


Figure 5. ΔV Cost Per Omega Change

ΔV Cost of Change

The ΔV cost to complete an aeroassist maneuver consists of three parts: 1) the burn required to enter into the atmosphere, 2) the impulse required to overcome atmospheric drag so that the satellite will be able to exit out of the atmosphere, and 3) the amount of fuel needed to either return the satellite into the new orbit or ready it for another atmospheric pass.

Aeroassist Analysis and Solution Method

Aeroassisted Maneuver

Background

The concept of using an aeroassisted maneuver has been around since the 1960's.[4] The general idea of the aeroassisted maneuver is to "dip" the satellite into the atmosphere from an initial parking orbit and use the aerodynamic forces of lift and drag to modify the satellite's orbit. See Figure 6. To take advantage of these aerodynamic forces, the satellite will have to be specially modified or configured, such as utilizing an aeroshell, using solar panels as "wings", or building the satellite into an aerodynamic shape. For this paper, the exact details of this modification will not be addressed.

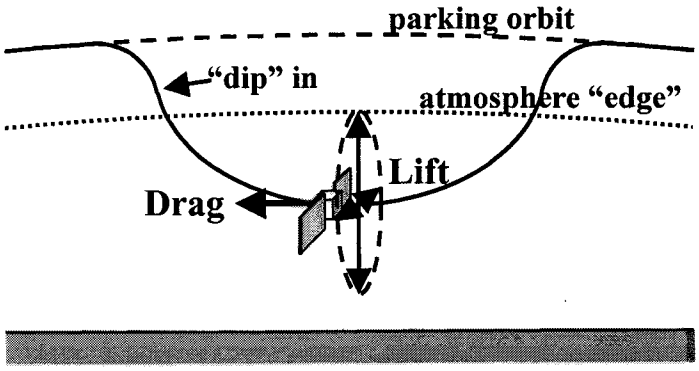


Figure 6. The Aeroassist Maneuver

Cowell's formulation of the equations of motion can describe the flight of the satellite both in orbit and through the atmosphere.[1] In system state format, the equations of motion in an Earth-centered inertial frame are shown in Equation 2.

During orbital flight, the accelerations of a_I, a_J, a_K are from perturbations due to the J2 effect, other Earth effects, and other heavenly bodies. For this paper's analysis of the aeroassist maneuver, they will be assumed to be zero. During the atmospheric portion of the flight, the accelerations will be quite large and will depend on the atmospheric forces on the satellite through some appropriate coordinate transformation.

In the body frame of the satellite, the accelerations due to lift and drag are found from Equation 3. There does not exist a complete analytical solution to the above combination of equations.

$$X = \begin{bmatrix} r_I \\ r_J \\ r_K \\ v_I \\ v_J \\ v_K \end{bmatrix}, \quad \dot{X} = \begin{bmatrix} v_I \\ v_J \\ v_K \\ -\frac{\mu r_I}{r^3} + a_I \\ -\frac{\mu r_J}{r^3} + a_J \\ -\frac{\mu r_K}{r^3} + a_K \end{bmatrix} \qquad (2)$$

However, several authors have found approximate solutions based on making several assumptions.[4,5] In order to not be constrained within the boundaries of those assumptions for this paper, the ΔV costs will be determined through numerical integration of the equations of motion.

$$a = \frac{F}{m}$$
$$a_{lift} = \frac{1}{2}\rho V^2 \frac{C_L S}{m} \qquad (3)$$
$$a_{drag} = \frac{1}{2}\rho V^2 \frac{C_D S}{m}$$

Alignment of the Lift Vector

Although the drag vector is always opposite the direction of flight (Figure 6), the lift vector can be in any direction perpendicular to the drag vector. To specify the direction of the lift vector, the angle, ϕ, will be defined as the angle from the radial R vector in the R-W plane as shown in Figure 7.

To maximize the amount of Ω or inclination change during an atmospheric pass, it is best to let ϕ be perpendicular to the orbital plane (ϕ = +/- 90 deg). This can be seen by looking at the perturbation equations (Equations 4) in the RSW satellite frame of reference.

In this frame of reference, F_S = - D, F_R = L x cos(ϕ), and F_W = L x sin(ϕ) from simple trigonometry. Any ϕ other than +/- 90 deg will lessen the Ω or inclination change and cause a change in the semi-major axis and eccentricity.

Another point to note while looking at the perturbation equations, is the effect of ω on ΔΩ. The argument of periapsis, ω, shows up through the di/dt and dΩ/dt equations via the u term (u = ω + ν). An atmospheric pass will be approximately symmetric with the perigee, thus ν will take on values that are symmetric to ν = 0 deg. For example, if, say, the satellite enters the atmosphere when ν = -10 deg, then it will leave the atmosphere when ν = 10 deg. If the di/dt and dΩ/dt equations are integrated across this range of ν's, different results will be produced depending on the value of ω. If ω = 0 or 180 deg, then dΩ/dt will be zero and di/dt will be some value. If ω = -90 or 90 deg, then di/dt will be zero and dΩ/dt will be some value. Therefore, to maximize the Ω change and not change the inclination, the atmospheric pass needs to be conducted when the orbit has an ω of -90 or 90 deg. If, for example, the apogee of the orbit is desired to be overlooking the North Pole, then ω should be chosen to be -90 deg. More discussions about the ω parameter will be detailed in a later section.

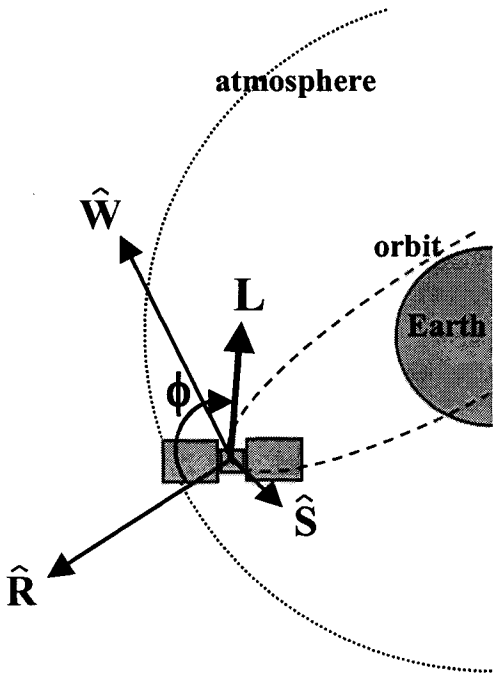


Figure 7. Alignment of the Lift Vector

$$\frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}}\left\{e\sin(\upsilon)F_R + \frac{p}{r}F_S\right\}$$
$$\frac{de}{dt} = \frac{\sqrt{1-e^2}}{na}\left\{\sin(\upsilon)F_R + \left(\cos(\upsilon) + \frac{e+\cos(\upsilon)}{1+e\cos(\upsilon)}\right)F_S\right\} \qquad (4)$$
$$\frac{di}{dt} = \frac{r\cos(u)}{na^2\sqrt{1-e^2}}F_W$$
$$\frac{d\Omega}{dt} = \frac{r\sin(u)}{na^2\sqrt{1-e^2}\sin(i)}F_W$$
$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae}\left\{-\cos(\upsilon)F_R + \sin(\upsilon)\left(1+\frac{r}{p}\right)F_S\right\} - \frac{r\cot(i)\sin(u)}{h}F_W$$
$$\frac{dM_o}{dt} = \frac{1}{na^2e}\{(p\cos(\upsilon) - 2er)F_R - (p+r)\sin(\upsilon)F_S\}$$

RESULTS

Simulation of Aeroassist

A computer program was written in MATLAB to analyze the change in orbital parameters during an atmospheric pass. The program solves Cowell's equations for a given set of initial conditions. To help focus the discussion, a generic orbit was chosen from which to make comparisons. The parameters of the orbit and satellite are summarized in Table 1.

Orbit Parameter	Value
Initial apogee	500 km
Initial perigee	300 km
Inclination, i	30 deg
Argument of periapsis, ω,	-90 deg
Directed perigee of atmospheric pass	97 km
Number of atmospheric passes	1
Satellite Parameters	**Value**
Ballistic coefficient (BC)	125 kg/m^2
Lift-to-drag (L/D) ratio	1
Lift pointing vector, ϕ,	-90 deg

Table 1. Summary of Orbital and Satellite Parameters

The orbit chosen starts with the 500km x 300km orbit. An impulsive ΔV propulsive burn is done at apogee to send it into an initial 500 km x 97 km orbit. This orbit will take the satellite into the atmosphere, and the aerodynamic forces will actually end up dipping the satellite lower than the initial 97 km perigee. At the lowest point of its atmospheric dip, the satellite will conduct an impulsive ΔV burn in the direction of the velocity vector that will send it back out of the atmosphere and back to its original apogee.

The graphs in Figure 8 show the results of the simulation for one orbit. The graphs show how the orbital parameters change during the flight. The simulation starts at apogee, and then progresses near perigee where the satellite enters the atmosphere. After a propulsive impulse, it soon leaves the atmosphere and returns back up to apogee. Of particular concern is how much Ω changes during the atmospheric pass.

As expected, the orbital parameters remain constant at the start of the simulation since there are no atmospheric forces applied to the satellite at this point. When the satellite enters the atmosphere, however, the orbital parameters will change. The impulsive boost at perigee will also produce changes in some variables. The orbital parameters will stop changing after the satellite leaves the atmosphere. By design, some orbital parameters will return, effectively, to their original value (semi-major axis and eccentricity). Other parameters, however, will have a net change.

Results from Simulation

Table 2 summarizes the results from the simulation for one atmospheric pass. As seen, a change in Ω is produced. The change in i, as also desired, is negligible. Secondarily, it should also be noted that there is a substantial change in ω.

Item	Value
Ω change	-1.6066 deg
ω change	3.7711 deg
i change	0.0498 deg
ΔV cost for boost	113.52 m/s
Minimum height	94.1618 km

Table 2. Summary of Simulation Results

Required ΔV for Maneuver

From the simulation, the ΔV cost of the boost can be determined, but there are also other ΔV's to consider. The ΔV cost of getting from the 500km x 300km orbit to the 500km x 97km orbit for the atmospheric pass and to return back to the 500km x 300km orbit needs to be included. This is most efficiently done at apogee (a Hohmann-like transfer)[6] and would cost 59.38 m/s to exit and 59.38 m/s to enter.

A very minor ΔV required is worth some mention. This ΔV is due to slight changes in the semi-major axis and eccentricity. If the boost ΔV is adjusted to achieve a zero net semi-major axis change, the eccentricity will be slightly different. If the boost ΔV is adjusted to achieve a zero net eccentricity change,

the semi-major axis will be slightly different. These differences can be adjusted most efficiently at apogee and the ΔV's are very small. In the example simulation, an additional 0.846 m/s would have to be included

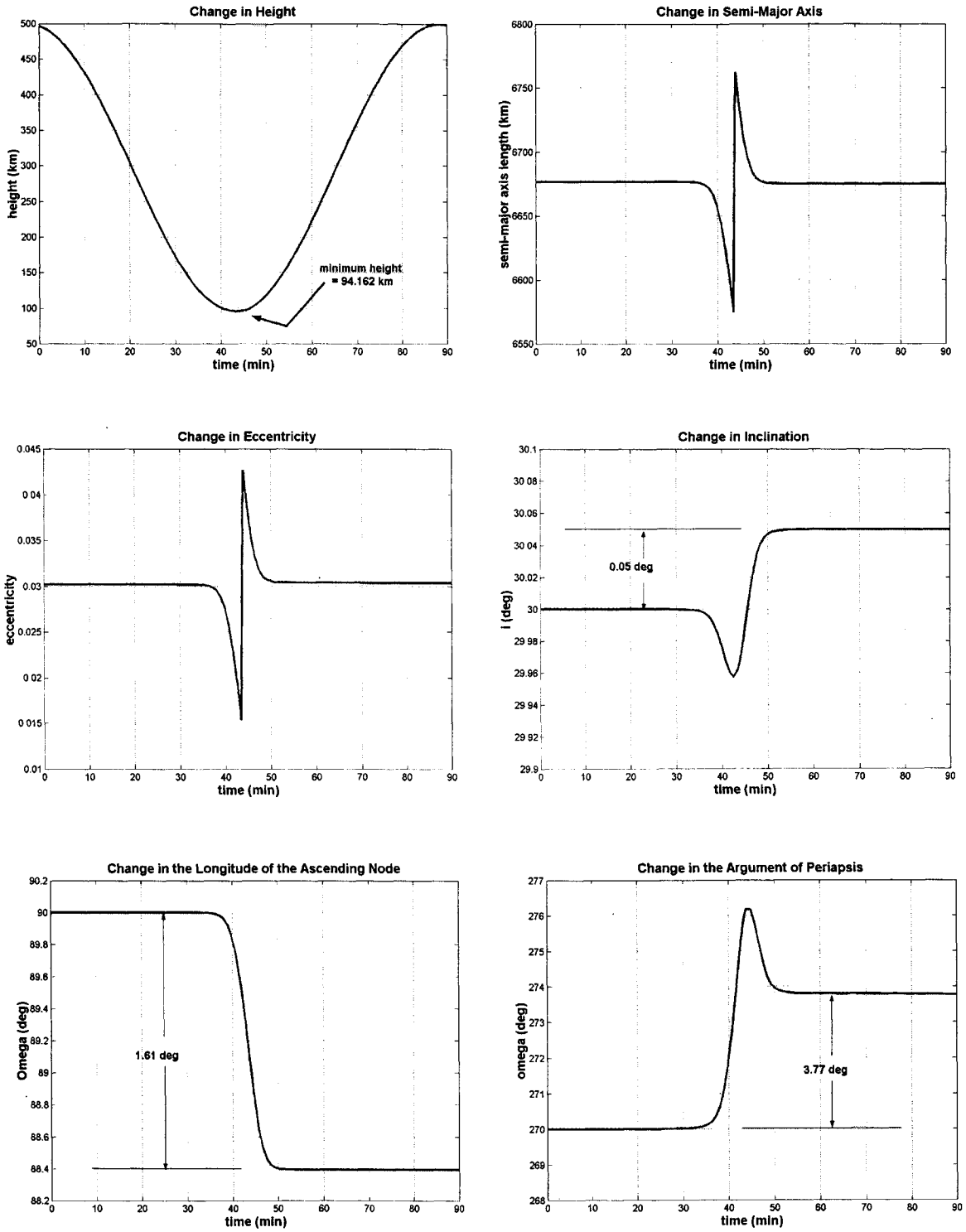


Figure 8a-f. Simulation of an Atmospheric Pass

for the adjustment in order to return both the semi-major axis and eccentricity to their original values. This adjustment will be called the "apogee adjustment" cost. The total cost for one atmospheric pass is calculated in Table 3.

This atmospheric pass changes Ω by 1.6066 deg. This ratio of ΔV to $\Delta\Omega$ will be called the "efficiency" of the Ω change. For this example, the efficiency is $\Delta V/\Delta\Omega$ = 145.11 m/s per deg of change.

ΔV to enter atmospheric pass orbit	59.38 m/s
ΔV for perigee boost	113.52 m/s
ΔV for apogee adjustment	0.85 m/s
ΔV to enter back to original orbit heights	59.38 m/s
Total ΔV	233.13 m/s

Table 3. Cost for One Atmospheric Pass

This ΔV required for the $\Delta\Omega$ produced is quite high. However, consecutive passes, one right after another, can make an improvement. For each pass, there will still be the boost cost and apogee adjustment cost, but the exit and entering costs only have to be paid once. For example, by using this multiple pass method for a total $\Delta\Omega$ of 180 deg, 112 atmospheric passes are needed at 1.6066 deg per pass. The total cost would be calculated as in Table 4.

The cost is 12.93 km/s for 180 deg of change. The efficiency greatly improves to 71.82 m/s per deg of change.

ΔV to enter atmospheric pass orbit	59 m/s
ΔV for 112 perigee boosts (113.52 m/s x 112)	12,714 m/s
ΔV for 112 apogee adjustment (0.85 m/s x 112)	95 m/s
ΔV to enter back to original orbit heights	59 m/s
Total ΔV	12,927 m/s

Table 4. Cost for Multiple Atmospheric Passes

Unfortunately, it turns out that this efficiency ratio cannot be achieved with the given orbital and satellite values. The assumption that a $\Delta\Omega$ of 1.6066 deg per pass is obtained will not always hold true. The change in ω that also occurs during the maneuver will sometimes make the $\Delta\Omega$ per pass less as the next section explains.

Effect of Changing ω

As discussed and seen from the perturbation equations, ω plays a large role in how much $\Delta\Omega$ can be achieved. The graphs in Figure 9a-c display from simulation how Ω, ω, and i change with an initial orbit's ω for various inclinations for the 300km x 500km orbit.

Note that in the graphs, the value of ϕ is -90 deg. If the value of ϕ is +90 deg, then the sign of the values will be the opposite. There is positive nodal regression change for $0 < \omega < 90$ deg (with ϕ = -90 deg). If positive nodal regression change is desired when $180 < \omega < 360$ deg, then ϕ should be changed to +90 deg.

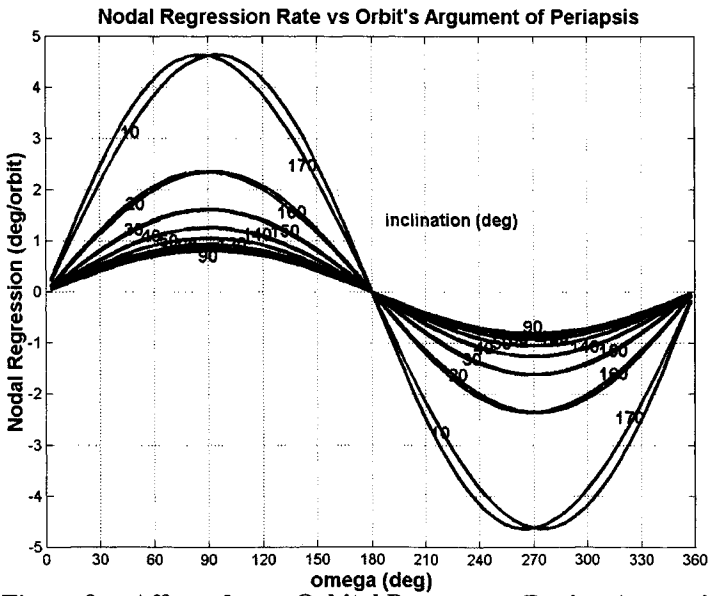


Figure 9a. Affect of ω on Orbital Parameters During Aeroassist

The relationships are of a sinusoidal nature. By looking at the perturbation equations, it is easy to see why this should be so.

The aeroassist results assumed that atmospheric passes would be made with ω always equal to +/- 90 deg. But as seen by the above graphs, with each atmospheric pass, the value of ω will change as well. If back-to-back atmospheric passes are done without doing

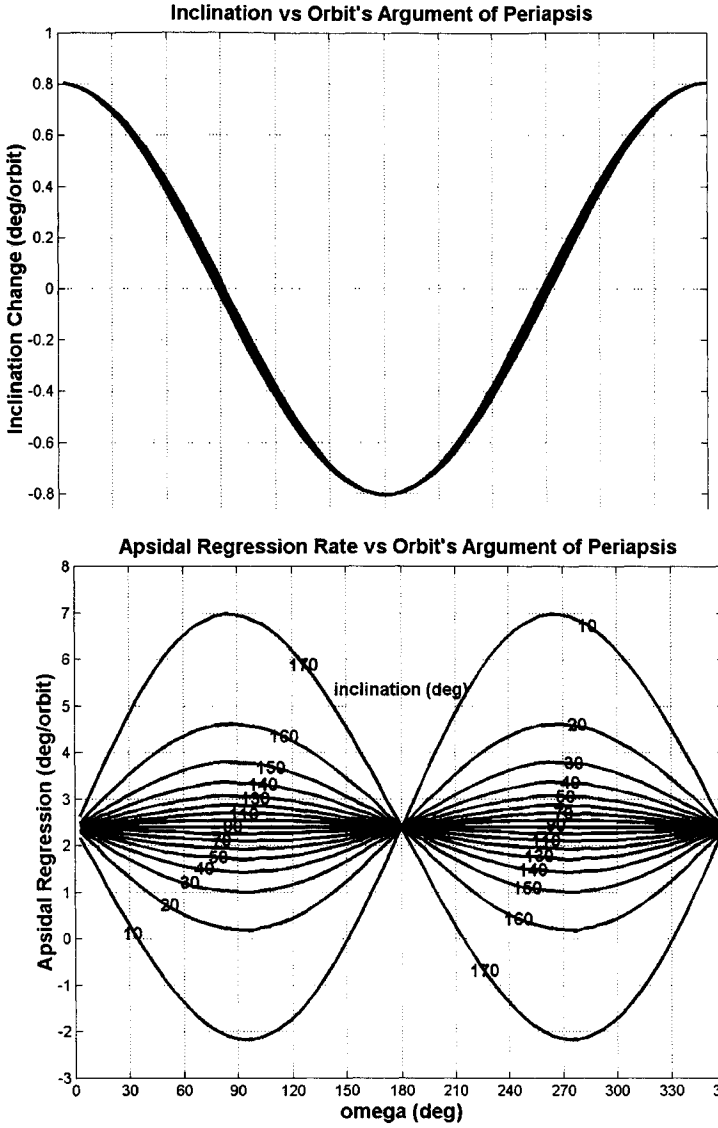


Figure 9b-c. Affect of ω on Orbital Parameters During Aeroassist

anything else to change the orbital parameters, most of the passes will produce sub-optimal $\Delta\Omega$'s because ω will not always be at the +/- 90 deg required for the maximum $\Delta\Omega$ for a given set of parameters.

As an example, with an original inclination of 30 deg and ω = 90 deg, the first pass will produce a $\Delta\Omega$ of about 1.6 deg. At the same time, though, ω will change by 1 deg. After, say, 30 atmospheric passes, ω will have changed to be about 120 deg. At this point the passes will produce a $\Delta\Omega$ of only1.4 deg per pass and ω will start changing by about 1.2 deg per pass. Eventually, ω will change to a point where $\Delta\Omega$ doesn't change much at all. These essentially would be wasted trips where ΔV is spent but no $\Delta\Omega$ is accomplished.

With a very large number of atmospheric passes, the average amount these parameters change can be determined and this can be used as a guide for the amount of Ω change that can effectively be produced. By looking at the graphs, ω changes, on average, about 2.4 deg per orbit. The change in Ω will vary in magnitude from 0 to 1.6 deg per orbit. The average, however, will be zero. If it is assumed the sign of Ω is always positive, by integrating over the sine curve, the average amount Ω will vary is 0.7071 x 1.6066 deg/orbit = 1.13 deg/orbit.

Note that if positive Ω change is required, the sign of ϕ needs to be switched appropriately when 180 deg < ω < 360 deg. Doing so will also affect how ω changes. Instead of varying from 1 deg to 3.8 deg, ω will vary from 1 deg to 2.4 deg with the average being ((2.4 deg - 1 deg) x 0.7071) + 1 deg = 1.99 deg/orbit.

With a very large number of atmospheric passes, it really doesn't matter what the original starting value of ω is since ω will sweep across the entire range of values anyway. Starting at a particular ω won't help the average Ω change.

What about a smaller number of atmospheric passes? For the most overall $\Delta\Omega$ change, from Figure 9a it is clear that a near equal number of passes be made from the time ω is a bit less than and from the time ω is a bit more than the optimal ω of +/- 90 deg. In this way, ω will sweep through the range of values that produce the most $\Delta\Omega$. During a small number of passes, the average $\Delta\Omega$ will be higher than

the large number average of 1.13 deg per pass. It will be much nearer the optimal $\Delta\Omega$ change of 1.6066 deg per pass.

Note that no mention has been made yet about the change of inclination. That certainly is a factor, but it is very minor. Where it is typically desired to change Ω (near ω = +/- 90 deg), the change in i is zero. On one side of this ω, Δi is positive, and on the other side, Δi is negative. So if the passes are made symmetrically around ω = +/- 90 deg, there will be no net inclination change. Also, the values of Δi are, in general, quite small in magnitude. Therefore, the perturbations about the original i will be quite small as well. For the strategies discussed, a good assumption is that the original i remains constant for analysis purposes.

The original aeroassist analysis thus needs to be modified to take into account the changing ω. For a total $\Delta\Omega$ change of 180 deg, about 159 atmospheric passes would be required (using the newly previously calculated $\Delta\Omega$ average of 1.13 deg per orbit). This definitely qualifies as a large number of passes. For this many passes, the total cost would be calculated as in Table 5.

The 18.30 km/s is much more than the 12.93 km/s originally estimated. The efficiency now averages to 101.68 m/s per degree of Ω change.

ΔV to enter atmospheric pass orbit	59 m/s
ΔV for 159 perigee boosts (113.52 m/s x 159)	18,050 m/s
ΔV for 159 apogee adjustment (0.85 m/s x 159)	135 m/s
ΔV to enter back to original orbit heights	59 m/s
Total ΔV	18,303 m/s

Table 5. Cost of Multiple Aeroassist with Changing ω

The only other alternative to dealing with this sub-optimal $\Delta\Omega$ average, is to find a way to always have ω = +/- 90 deg. This would require a modification to the orbit before each pass. One way to do this would be to circularize the orbit when the satellite gets up to its apogee, then when the satellite is at an appropriate point, de-orbit so that the resulting dipping orbit will again have ω = +/- 90 deg. Doing so will require the originally calculated 112 atmospheric passes at 1.6066 deg per pass, but adding in the cost of continually re-circularizing the orbit, the total cost would be as in Table 6.

A total $\Delta\Omega$ of 180 deg would cost 26.02 km/s. In this case, the efficiency is 144.58 m/s per deg. Clearly this is not the method for the best efficiency. It is more efficient to live with conducting sub-optimal $\Delta\Omega$ changes.

ΔV to enter atmospheric pass orbit (59 m/s x 112)	6,608 m/s
ΔV for 112 perigee boosts (113.52 m/s x 112)	12,714 m/s
ΔV for 112 apogee adjustment (0.85 m/s x 112)	95 m/s
ΔV to enter back to original orbit heights (59 m/s x 112)	6,608 m/s
Total ΔV	26,025 m/s

Table 6. Cost for Sub-Optimal Aeroassist

Are there any other ways to modify the orbit so the atmospheric passes are always made with ω = +/- 90 deg for each pass? Possibly, but within the scope of this paper, this will not be investigated.

Improvements with Increasing L/D Ratio

So far, the aeroassist does not look very promising as compared to the propulsive impulse maneuver. However, one remaining yet important variable has yet to be investigated: the L/D ratio.

The computer simulation was run with different L/D ratios to see how it affects the changes in Ω, ω, and i. The other orbital and satellite parameters were not changed from Table 1. Figure 10 shows the results.

It is seen that the nodal regression varies almost linearly with the L/D ratio for these ranges and orbit parameters. As the L/D ratio increases, larger values of change are produced. Although not graphed, the ΔV for each of the runs remains the same as before. Therefore for the same ΔV cost, the $\Delta\Omega/\Delta$V

efficiency improves as the L/D ratio increases. For example, for an L/D ratio of 5, the $\Delta\Omega$ per atmospheric pass is 7.9 deg. For multiple passes, the average pass would change $\Delta\Omega$ by 0.7071 x 7.9 deg = 5.59 deg. For a total change of 180 deg, approximately 32 passes would be required. The overall cost would be as in Table 7.

ΔV to enter atmospheric pass orbit	59 m/s
ΔV for 32 perigee boosts (113.52 m/s x 32)	3,633 m/s
ΔV for 32 apogee adjustment (0.85 m/s x 32)	27 m/s
ΔV to enter back to original orbit heights	59 m/s
Total ΔV	3,778 m/s

Table 7. Cost for Improved L/D Ratio

The efficiency for this change is 20.99 m/s per deg. This is much improved from the 101.68 m/s per deg efficiency for an L/D ratio of 1. The total ΔV of 3.78 km/s is even much better than the 7.75 km/s for the propulsive impulse case.

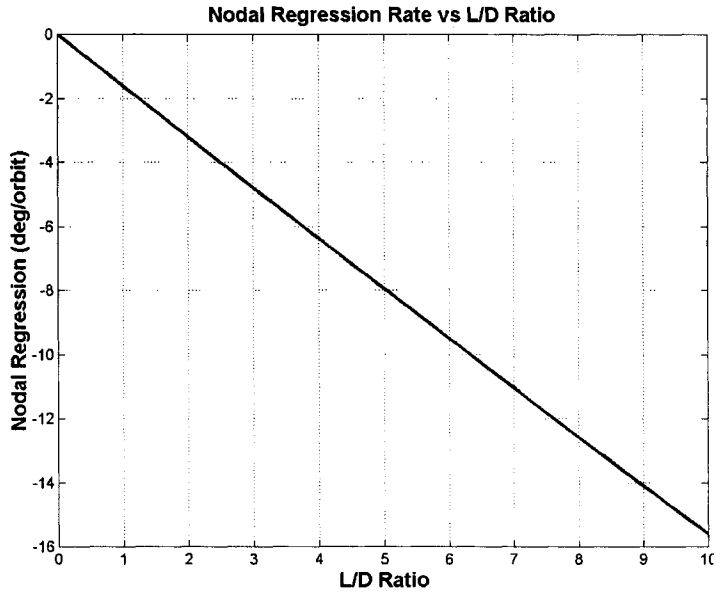


Figure 10. Effect of L/D Ratio on Nodal Regression Rates

It will be seen that even a modest L/D ratio increase from 1 will be more efficient than the propulsive impulse case for some orbits.

Effect of Ballistic Coefficient on ΔV Cost

The computer simulation was run for a variety of ballistic coefficients at different dipping depths (initial orbit perigees). The initial perigee height was varied from 95km to 105km. The ballistic coefficients were varied from 30 kg/m^2 to 200 kg/m^2. The L/D ratio was kept at 1. For these ballistic coefficients and heights, the nodal regression rate was plotted vs. the required boost velocity as seen in Figure 11.

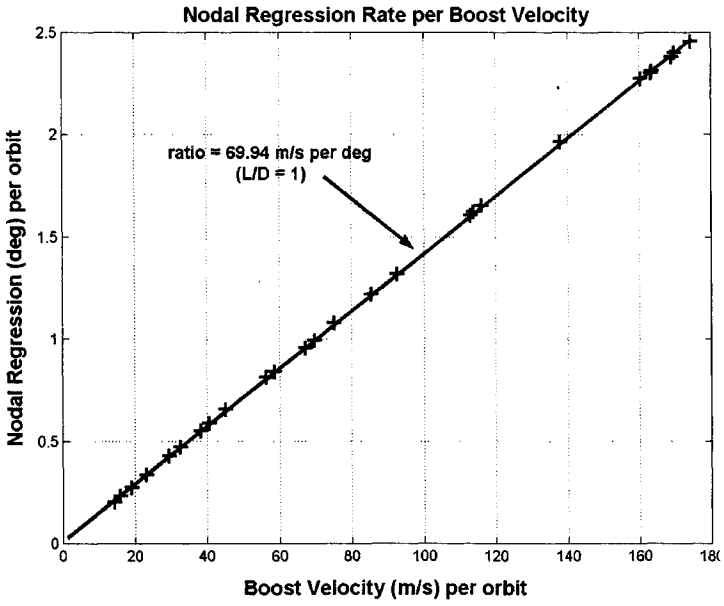


Figure 11. Efficiency for Various Heights and Ballistic Coefficients

Note that all of the points lie on the same line. This shows that the efficiency of the nodal regression is independent of both the ballistic coefficient and the depth of the atmospheric pass. The efficiency will only be dependent on the L/D ratio. It is easy to see why this should be so. The boost velocity required is directly related to the amount of drag the satellite experiences during the atmospheric pass because the boost velocity is required to make up for the velocity losses due to drag. As seen on a previous graph, the amount of $\Delta\Omega$ change is directly related to the L/D ratio. Therefore, commutatively, $\Delta\Omega$ is directly related to the boost

velocity by whatever factor the L/D ratio is. The ballistic coefficient and dipping depth do not play a role in this relationship.

The only effect that the ballistic coefficient and dipping depth will have is on the number of atmospheric passes required to complete a total $\Delta\Omega$ orbit change, and on the heating rate experienced on the satellite. The ballistic coefficient and dipping depth should be chosen to balance the heating rate and time concerns. This particular optimization will not be addressed in this paper.

Note that the ratio and costs described are only based on the boost velocity and not on the delta velocities required to exit and enter into the orbit or the adjustment at apogee.

Comparison of Aeroassist to ΔV Impulse Maneuver

From the previous discussions, it is known that the ΔV costs for the propulsive impulse maneuver depend on the amount of $\Delta\Omega$ desired. The efficiency of the aeroassist maneuver depends on $\Delta\Omega$, too, but it also depends on the L/D ratio. Figure 12 shows how the propulsive impulse case and aeroassist case compare in efficiency.

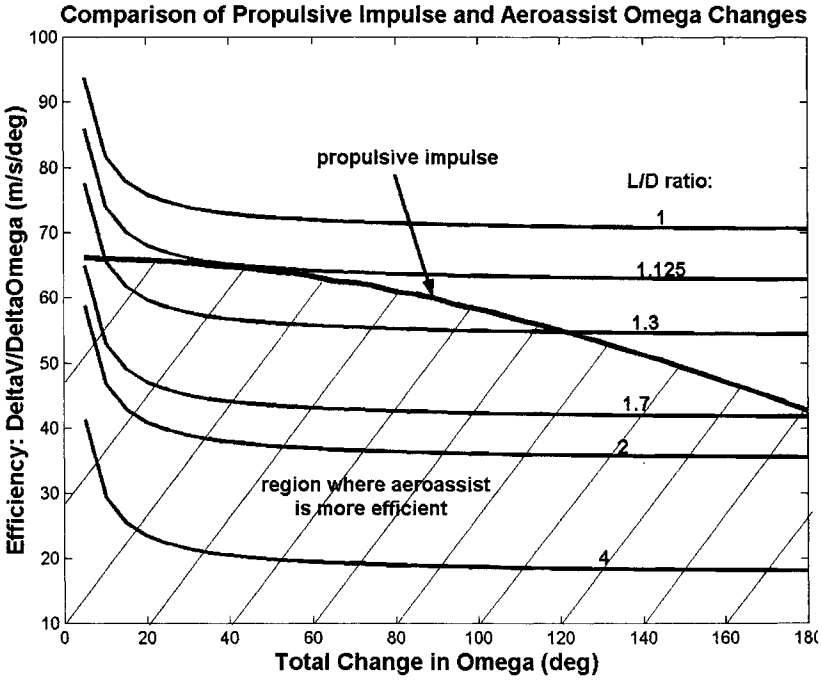


Figure 12. Efficiency of the Propulsive Impulse and Aeroassisted Cases

Note that the various L/D lines are not straight across. If the ΔV's to exit and enter the orbit for the dipping trajectory were not considered, then the L/D lines would be straight across. Since a set exit and enter ΔV amount for a dipping trajectory is spent, no matter what $\Delta\Omega$ is accomplished, it is more efficient to spread this cost over many orbits instead of just a few.

It is seen that for this orbit, if the L/D ratio is less than about 1.125, then the propulsive impulse maneuver is always more efficient than the aeroassist maneuver. If the L/D ratio is larger than 1.7, then the aeroassist maneuver is always more efficient than the propulsive impulse maneuver.

Although this paper has only focused on one sample orbit, similar trends are seen for other orbits as well. An L/D range can always be found where the aeroassist maneuver is better than the propulsive impulse one for certain $\Delta\Omega$ changes.

CONCLUSIONS

Advantages and Disadvantages of Each Maneuver

Table 8 summarizes, subjectively, the 3 different methods to produce a $\Delta\Omega$ change.

Costs	J2	Propulsive Impulse	Aeroassist
Time (operational costs)	High	Low	Low
Fuel (weight - launch costs)	Low	High	Medium
Design (weight - launch costs) (development costs)	Low	Low	High

Table 8. Summary of Costs for Each Method

While the J2 method is generally the lowest in fuel and design costs, it may take 6 months or more to move satellites into their desired constellation. During this time, the desired science objectives may not be able to take place. If there is a substantial cost in operating the satellites while in orbit, the large amount of "dead" time may end up costing more than if the science could be conducted immediately.

The propulsive impulse method is the quickest and allows the science to be started almost immediately. However the fuel costs will be large.

The aeroassist method, as long as the satellite can be designed with a large enough L/D ratio, is a nice compromise of fuel and time savings. With a suitable L/D ratio, the aeroassist maneuver will cost less fuel than the propulsive impulse maneuver, and can be completed on the order of a week. The costly part of this method is in the design of such a satellite. Designing a satellite to handle aerospace maneuvers (increased heating, increased forces, complex control requirements, etc.) may be too expensive and the satellite may end up with more weight.

The costs described are very subjective in nature. It will be left for future analysis to quantitatively determine the actual costs and determine when one method should be chosen over another.

Aeroassist Maneuver Conclusions

This paper introduces the concept of using an aeroassisted maneuver to change the Ω of a satellite's orbit. The paper outlines how to calculate the efficiency of this maneuver and compares it with the J2 and propulsive impulse methods. It is found that if a satellite is designed with a certain minimum L/D ratio, the aeroassist maneuver will be the most efficient.

It was also learned that the efficiency of an aeroassist maneuver is independent of the ballistic coefficient and the dipping depth. The efficiency only depends on the L/D ratio and the fuel required to exit into and return from a dipping trajectory.

The paper also addresses several practical concerns when conducting an aeroassist maneuver. The change of ω with each atmospheric pass prevents one from changing Ω continuously at the largest $\Delta\Omega$ rate. However, overall, it is cheaper to change Ω at a sub-optimal rate than the cost to continually adjust the orbit to change at the optimal rate for a given set of orbital parameters.

REFERENCES

1. Vallado, David A., Fundamentals of Astrodynamics and Applications, McGraw-Hill, New York, NY, 1997.
2. Gobetz, F. W., Doll, J. R., "A Survey of Impulsive Trajectories", AIAA Journal, Vol. 7, No. 5, May 1969, pp. 801-834.
3. Eckel, K. G., "Numerical Solutions of Non-coaxial Optimum Transfer Problems", The Journal of the Astronautical Sciences, Vol. 10, No. 3, Fall 1963, pp. 82-92.
4. London, H. S., "Change of Satellite Orbit Plance by Aerodynamic Maneuvering", Journal of the Aerospace Sciences, Vol. 29, No. 2, 1962, pp.323-332.
5. Vinh, N. X., Optimal Trajectories in Atmospheric Flight, Elsevier Scientific Publishing Co., New York, NY, 1981.
6. Bate, Roger R., Mueller, Donald D., White, Jerry E., Fundamentals of Astrodynamics, Dover Publications Inc, New York, NY, 1971.

Session 3: Attitude Determination, Sensor Calibration and Computational Techniques

CALIBRATION OF GYROS WITH TEMPERATURE DEPENDENT SCALE FACTORS[⊥]

Sheela V. Belur*
Computer Sciences Corporation, 7700 Hubble Drive, Lanham-Seabrook, MD 20706

Richard Harman
Flight Dynamics Branch, Code 572, NASA Goddard Space Flight Center, Greenbelt, MD 20771

ABSTRACT

The general problem of gyro calibration can be stated as the estimation of the scale factors, misalignments and drift-rate biases of the gyro using the on-orbit sensor measurements. These gyro parameters have been traditionally treated as temperature-independent in the operational flight dynamics ground systems at NASA Goddard Space Flight Center (GSFC), a scenario which has been successfully applied in the gyro calibration of a large number of missions. A significant departure from this is the Microwave Anisotropy Probe (MAP) mission where, due to the high thermal variations expected during the mission phase, it is necessary to model the scale factors as functions of temperature.

This paper addresses the issue of gyro calibration for the MAP gyro model using a manufacturer-supplied model of the variation of scale factors with temperature. The problem is formulated as a least squares problem and solved using the Levenberg-Marquardt algorithm in the MATLAB© library function 'NLSQ'. The algorithm was tested on simulated data with Gaussian noise for the quaternions as well as the gyro rates and was found to consistently converge close to the true values. Significant improvement in accuracy was noticed due to the estimation of the temperature-dependent scale factors as against constant scale factors.

INTRODUCTION

The general problem of gyro calibration can be stated as the estimation of the scale factors, misalignments and the biases of the gyro using the on-orbit sensor measurements

---

[⊥] NASA GSFC Contract No: GS-35F-4381-G, Task order No: S-43411-G

* Email: sbelur@csc.com, Phone: (301) 794-2317

(ref.1). Usually, the scale factors of the gyros are deemed to be constants which will be estimated a priori and only small, possibly slowly varying corrections to these manufacturer-supplied values will be estimated during the mission (ref. 2-5). But in some missions such as the Microwave Anisotropy Probe (MAP) (ref. 6), due to the high thermal variations expected during the mission, the scale factors are themselves functions of temperature. In most of these cases, the functional form of the variation of the gyro scale factors will be known a priori from the manufacturer and certain associated parameters for their actual values on-orbit are to be estimated as part of the gyro calibration exercise.

The existing gyro calibration utilities used at the NASA Goddard Space Flight Center (GSFC) assume that the scale factors are constants. Therefore the objective of the work reported in this document is to formulate this temperature-dependent gyro calibration and provide a methodology and a tool to solve it.

PROBLEM FORMULATION

The problem of temperature-dependent gyro calibration is formulated in this section. With a view to applying this methodology to the MAP gyro calibration scenario, the gyro model conforming to MAP (ref.7) was used here. The 3x1 vector of gyro rates is given by:

$$\omega(t) = \boldsymbol{M}\begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix}\begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} - \boldsymbol{b} \qquad (1)$$

where

$$S_i = a_{io} + a_{i1}v + a_{i2}v^2 + a_{i3}v^3, \quad i=1,2,3 \qquad (2)$$

are the temperature dependent scale factors, v is the voltage, N_i are the gyro telemetry counts, $\boldsymbol{b}$ is the 3x1 bias vector to be estimated, and $\boldsymbol{M}$ is the 3x3 misalignment matrix. The scale factors are actually functions of temperature but are provided by the manufacturer in terms of gyro thermistor voltage variations.

The gyro calibration problem now reduces to estimating the 28x1 state vector $\boldsymbol{X}$

$$\boldsymbol{X} = [\{a_{ij}\},\{ M_{kl}\}, \boldsymbol{b}^T, \boldsymbol{q}_0{}^T]^T \qquad (3)$$

where $\{a_{ij}\}$ are the 12 coefficients of the scale factors defined in Equation (2), $\{M_{kl}\}$ are the 9 elements of $\boldsymbol{M}$, and $\boldsymbol{q}_0$ is the 4x1 epoch inertial-to-body quaternion . It is convenient to define the 3x4 matrix $\boldsymbol{A}$ whose elements are $\{a_{ij}\}$.

It is possible to reduce the number of parameters in the state vector by 3 by redefining the alignment matrix to incorporate the linear part of the scale factor corrections by

normalization. We are considering applying this reduction in a future version of the algorithm.

The kinematic equation relating the attitude to the gyro rates is (ref. 8)

$$\dot{\boldsymbol{q}} = \tfrac{1}{2}\boldsymbol{\Omega}\boldsymbol{q} \tag{4}$$

where $\boldsymbol{q}$ is the attitude quaternion and $\boldsymbol{\Omega}$ is given by

$$\boldsymbol{\Omega} = \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix} \tag{5}$$

Here $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)$ is the spacecraft angular velocity measured by the gyros with components along the body axes.

X is estimated by minimizing the error between the attitude quaternions computed from sensor measurements and the attitude obtained by integrating the kinematic equation above. The closed form solution to the equation is given by (ref. 8):

$$\boldsymbol{q}(t_{i+1}) = \left[\cos\left(\frac{\omega \Delta t}{2}\right) \mathbf{I} + \frac{1}{\omega} \sin\left(\frac{\omega \Delta t}{2}\right) \boldsymbol{\Omega}_i \right] \boldsymbol{q}(t_i) \tag{6}$$

where

$$\omega = \left[\omega_x^2(t_i) + \omega_y^2(t_i) + \omega_z^2(t_i) \right]^{1/2}$$

$$\Delta t = t_{i+1} - t_i$$

$$\boldsymbol{\Omega}_i = \boldsymbol{\Omega}(\boldsymbol{\omega}(t_i))$$

$\boldsymbol{q}(t_i)$ is the attitude quaternion at time t_i, and $\boldsymbol{q}(t_{i+1})$ is the propagated attitude quaternion at time t_{i+1} . This closed form solution is used in the present formulation to get the propagated attitude quaternions $\boldsymbol{q}_p(t)$ for evaluating the objective function matrix for minimization using the Levenberg-Marquardt algorithm. The quaternion residuals $f(X)$ are now defined by

$$f_i(\boldsymbol{X}) = \boldsymbol{q}_p(t_i) - \boldsymbol{q}_t(t_i) \tag{7}$$

where $\boldsymbol{q}_t(t)$ is the true quaternion at time t obtained using the attitude sensor measurements. The problem will then be to find the state vector that minimizes the cost function

$$L(\boldsymbol{X}) = \Sigma_i \, (f_i(\boldsymbol{X}))^2 \tag{8}$$

with $(f_i(X))^2$ defined as the sum of the squares of the individual components of $f_i(X)$. As will be seen below, the above cost function formulation works well. Note that the residuals defined in Equation (7) are different from the traditional attitude residuals represented by three small rotations about the true attitude (see for example, ref. 3). This traditional approach to the current problem is being investigated. Also, by suitably introducing a weight matrix in the cost function, non-uniform weighting can be allowed. This is not done in the current formulation and will be attempted in the future version.

SOLUTION METHODOLOGY

The Levenberg-Marquardt Method

The general nonlinear least squares problem can be stated as:

Find X such that L(X), defined in equation (8), is a minimum.

The Levenberg-Marquardt (L-M) (ref. 9) method uses a trust-region approach to shrink the step sizes to minimize the cost function at each iteration and the state update is given by

$$X_{new} = X_{old} + \text{Inv}\,(J^T J + D).J^T F$$

where D is the diagonal matrix given by

$$D = \lambda.\ \text{Diag}\,(J^T J\,),$$

λ is a multiplication factor, J is the matrix of first partial derivatives $\partial F/\partial X$, and F is the residual vector .

The general procedure for this method is to set λ to 1.0 for the first iteration. If the first attempt at an iteration reduces the cost function then λ is reduced for the next iteration by a factor of 10. If the first step increases the cost function, then λ is increased by a factor of 5 until the cost function is reduced. The final value of λ for the iteration is used for the next iteration. Rather than compute the sum of squares of $f_i(X)$, 'NLSQ' requires the user-defined function to compute the matrix-valued function

$$F(X) = [f_1(X)\, f_2(X)\, f_3(X) \ldots f_m(X)]^T$$

Here each $f_i(X)$ could be a vector in which case the objective function will be matrix-valued.

The MATLAB Function 'NLSQ'

The MATLAB function 'NLSQ' is a general function to solve the non-linear least squares problem defined in the previous section. The default algorithm used in this function is that of L-M algorithm. The calling parameters for this function are

[X,OPTIONS,F,J] = NLSQ('FUN',X0,OPTIONS,'GRADFUN',P1,P2,..)

This function starts at the state vector $\boldsymbol{X_0}$ and finds a minimum to the sum of squares of the functions described in *FUN*. *FUN* is usually an M-file which returns a vector of objective functions: $\boldsymbol{F}$=*FUN*($\boldsymbol{X}$). *FUN* should return $\boldsymbol{F}(\boldsymbol{X})$ and not the sum-of-squares since the cost function is computed implicitly in the algorithm.

OPTIONS is a vector of optional parameters to be defined. GRADFUN is an optional function which returns the partial derivatives of the functions at a given $\boldsymbol{X}$. If *GRADFUN* is not supplied, numerical derivatives are computed within the function.

APPLICATION TO SIMULATED DATA

General Simulation Procedure

- Assign true values for the misalignment matrix, $\boldsymbol{M}$, the scale factor coefficients matrix, $\boldsymbol{A}$, the bias vector, $\boldsymbol{b}$, and the epoch quaternion $\boldsymbol{q}_0$.
- Assign start and end times t_i and t_f.
- Assign m, the number of time tags.
- Assign a temperature/voltage profile for the simulation time interval $[t_i,t_f]$.
- Assume an angular velocity profile for the spacecraft in $[t_i,t_f]$.
- Calculate the scale factor vector $\boldsymbol{S}$ =$[\boldsymbol{A}]$ $\boldsymbol{V}$, where $\boldsymbol{A}$ is the 3x4 matrix of scale factor coefficients, $\boldsymbol{V}$=$[\,1\ v\ v^2\ v^3]^T$, v being the voltage as a known function of time t.
- Calculate the telemetry counts after adding the bias using $[\boldsymbol{S}]^{-1}\,\boldsymbol{M}^{-1}\,(\boldsymbol{N}(t)+\boldsymbol{b})$ where $[\boldsymbol{S}]$ is the diagonal matrix of the scale factor components of $\boldsymbol{S}$.
- Starting from $\boldsymbol{q}_0$, get the quaternion history using the recursive relation given in Equation (6) above, normalize it and save it as "true" attitudes $\boldsymbol{q}_t$(t).

Specific Simulation Scenario

The simulated data spanned 1 hour 15 minutes and consisted of inertial pointing periods interspersed with three slews of rate 0.1 deg/sec rotating +/-45 deg about the x, y axes and +/-90 deg about z axis. The slews lasted 1 hour totally with a gap of 7 minutes between slews. The temperature variation in the gyros in the early launch phase was taken to be between 5-10 degrees which corresponds to about 1 to 2 volts variation in voltage as per discussion with the MAP Attitude Control System (ACS) engineers. Based on this, a sinusoidal voltage variation was assigned with an amplitude of 2 volts and period of 1 hour, i. e.,

$v = 2\sin 2\pi t/P$

where t is the time and P, the period, is 1 hour.

The following set of true gyro parameters and epoch quaternion, denoted by the subscript "t", were used:

$$A_t = \begin{bmatrix} 0.4945513578201 & .00036598584787 & .00001662910926 & -.00000087266463 \\ 0.4945513578201 & .00036598584787 & .00001662910926 & -.00000087266463 \\ 0.4945513578201 & .00036598584787 & .00001662910926 & -.00000087266463 \end{bmatrix} \text{X} 10^{-5}$$

$\boldsymbol{b}_t = [0.5, 0.5, 0.5]^T$ deg/hr

$$\boldsymbol{M}_t = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$\boldsymbol{q_0} = [0.5\ 0.5\ 0.5\ 0.5]^T$

The quaternion measurements were corrupted with noise equivalent to 40 arc seconds root-sum-squared (RSS) in Euler angles to correspond to the MAP star tracker accuracy and the gyro measurements were corrupted with Gaussian noise of 0.016 deg/hr (1-σ).

Estimation Steps

The core of the estimation procedure is the evaluation of the matrix of residuals $\boldsymbol{F(X)}$:

- Offset $\boldsymbol{M}, \boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{q_0}$ from the true values to get the initial state vector $\boldsymbol{X}_0$
- Calculate the gyro rates using $\boldsymbol{S=AV}$, $\omega = \boldsymbol{M\,S\,N - b}$ where $\boldsymbol{N}$ is the vector of telemetry counts
- Starting from $\boldsymbol{q_0}$, use Equation (6) above to get the propagated attitude quaternions, $\boldsymbol{q_p(t)}$, at the sampling intervals and normalize it at every time instant.
- Compute the residual vectors via Equation (7) at times $t_1, \ldots t_m$
- Update the 28x1 state vector $\boldsymbol{X}$ using the L-M algorithm via the MATLAB function 'NLSQ'.

This procedure is repeated till the convergence criterion of the L-M method is achieved. In order to scale the independent variables and make their range of variation uniform across the state vector, the components of the bias vector and the coefficients of the scale factors are multiplied by 10^{-6} before using them in the cost function evaluation. It was found that without this, the problem becomes ill-conditioned and the procedure terminates.

DISCUSSION OF RESULTS

Numerous estimation runs were made for widely ranging values of the a priori state vector X_0. Convergence close to the true value was obtained consistently. The following is a typical result:

$$\boldsymbol{b} = [\ 0.51543 \quad 0.5052 \quad 0.4954\ ] \text{deg/hr}$$

$$\boldsymbol{M} = \begin{bmatrix} -1.00000000000000 & 0.00000000131127 & -0.00000000043832 \\ 0.00000000160429 & -1.00000000000000 & -0.00000000916929 \\ 0.00000000030366 & -0.00000001133398 & -1.00000000000000 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} 0.44509621842136 & 0.00032942235090 & 0.00001491460493 & -0.00000080505695 \\ 0.44509603800549 & 0.00032951201204 & 0.00001499181614 & -0.00000080575904 \\ 0.44509511127285 & 0.00033038158483 & 0.00001508609330 & -0.00000097751931 \end{bmatrix} 10^{-5}$$

$$\boldsymbol{q_0} = [\ 0.500001\ 0.500008\ 0.49994\ 0.49995]^T$$

The biases agree to an accuracy of .015 deg/hr and the products of misalignment and scale factor matrices match to an accuracy of 10^{-10}, which corresponds to 0.017% error. Figure 1 gives the plot of the errors between estimated and true Euler angles and Figure 2 gives the deviations plot of the components of the products of the M and S matrices defined in equation (1) as a function of time (i.e., as temperature varies).

It was found that the solutions were independent of the initial guesses. Some of the initial guesses tried were: 0.9 for all diagonal elements of $\boldsymbol{M}$, 0.1 for all diagonal elements of $\boldsymbol{M}$, zero for all components of $\boldsymbol{b}$, 2 deg/hr for all components of $\boldsymbol{b}$, zero for all scale factor coefficients and 1.0 for all scale factor coefficients and different combinations thereof. Results differing only in the 12^{th} or 13^{th} significant figure were obtained with these initial guesses, which is essentially due to the limitation of the machine precision. Runs were also made not including measurements during maneuver times and the resulting accuracies did not vary significantly.

Further, estimation was carried out with only constant scale factors in the state vector and setting all other coefficients to zero as in the conventional gyro calibration methods. As seen in Figure 3, which gives the errors in the Euler angles for this case, as compared to those of the temperature-dependent gyro calibration case above, the improvement in accuracy is significant due to the enhanced state vector. Note that the asterisks in this figure represent the same data as in Figure 1, but on a different scale.

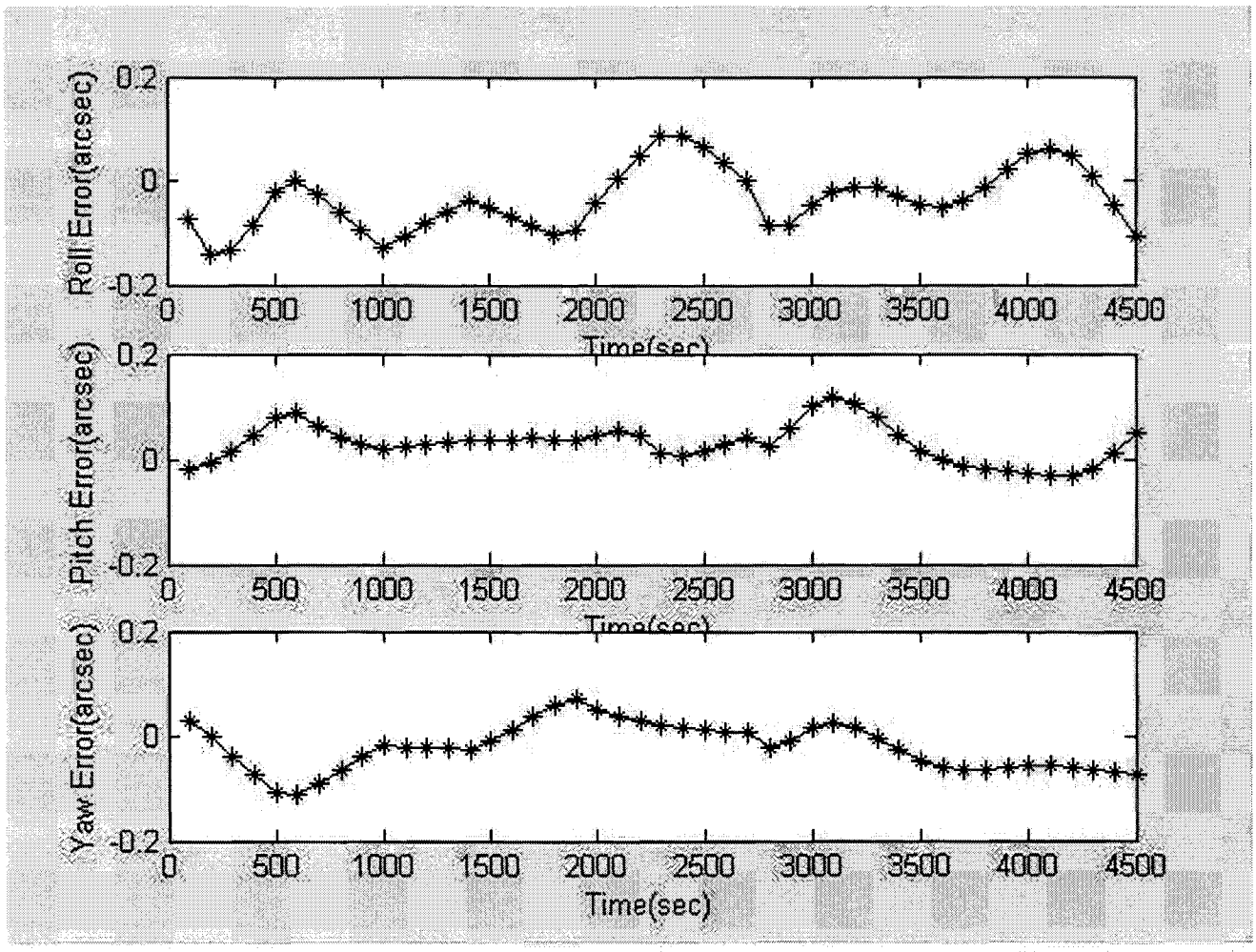


Fig 1. Euler Angle Error Plot After Gyro Calibration

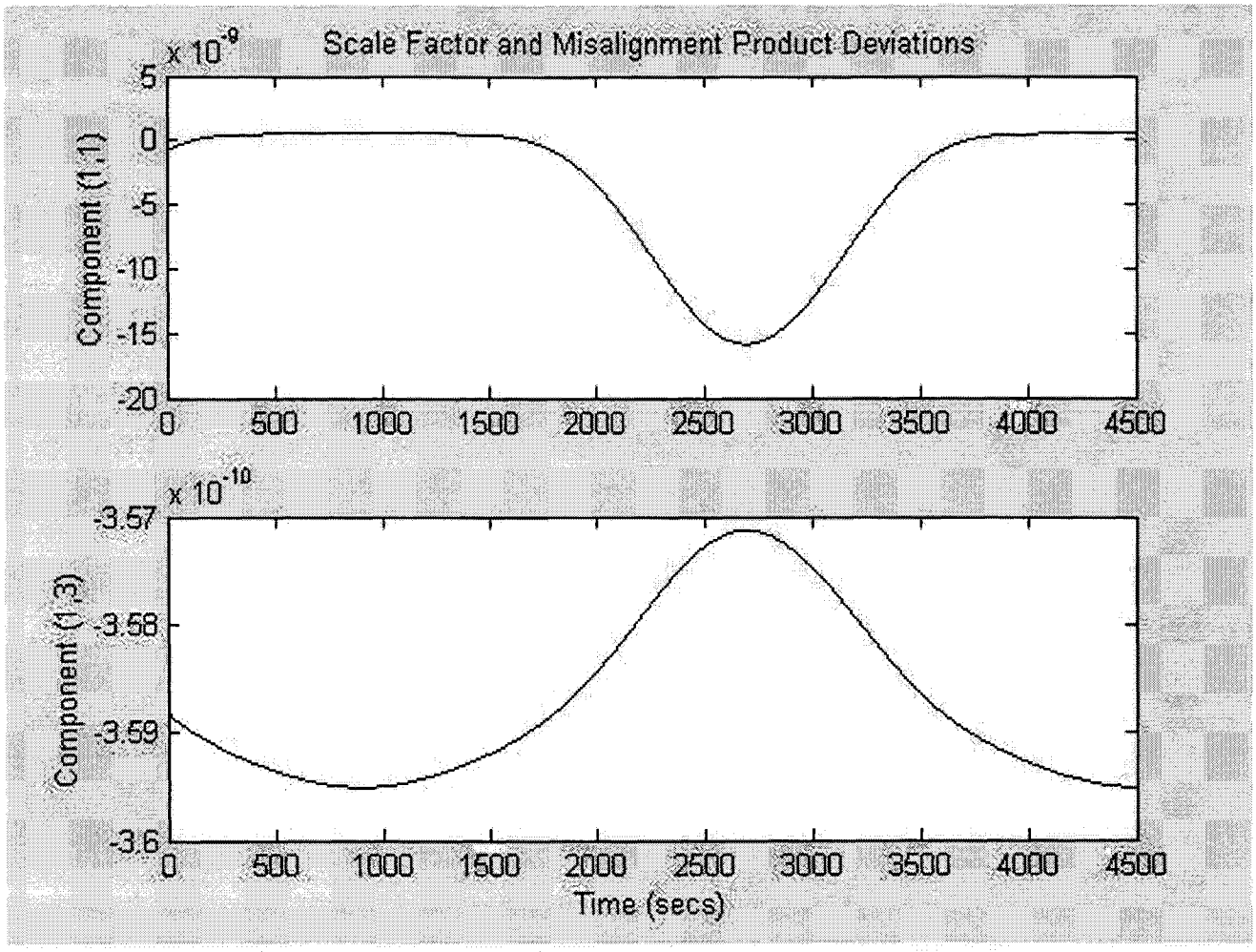


Fig. 2. Deviations Plot of the Components of the Products of Misalignment and Scale Factors

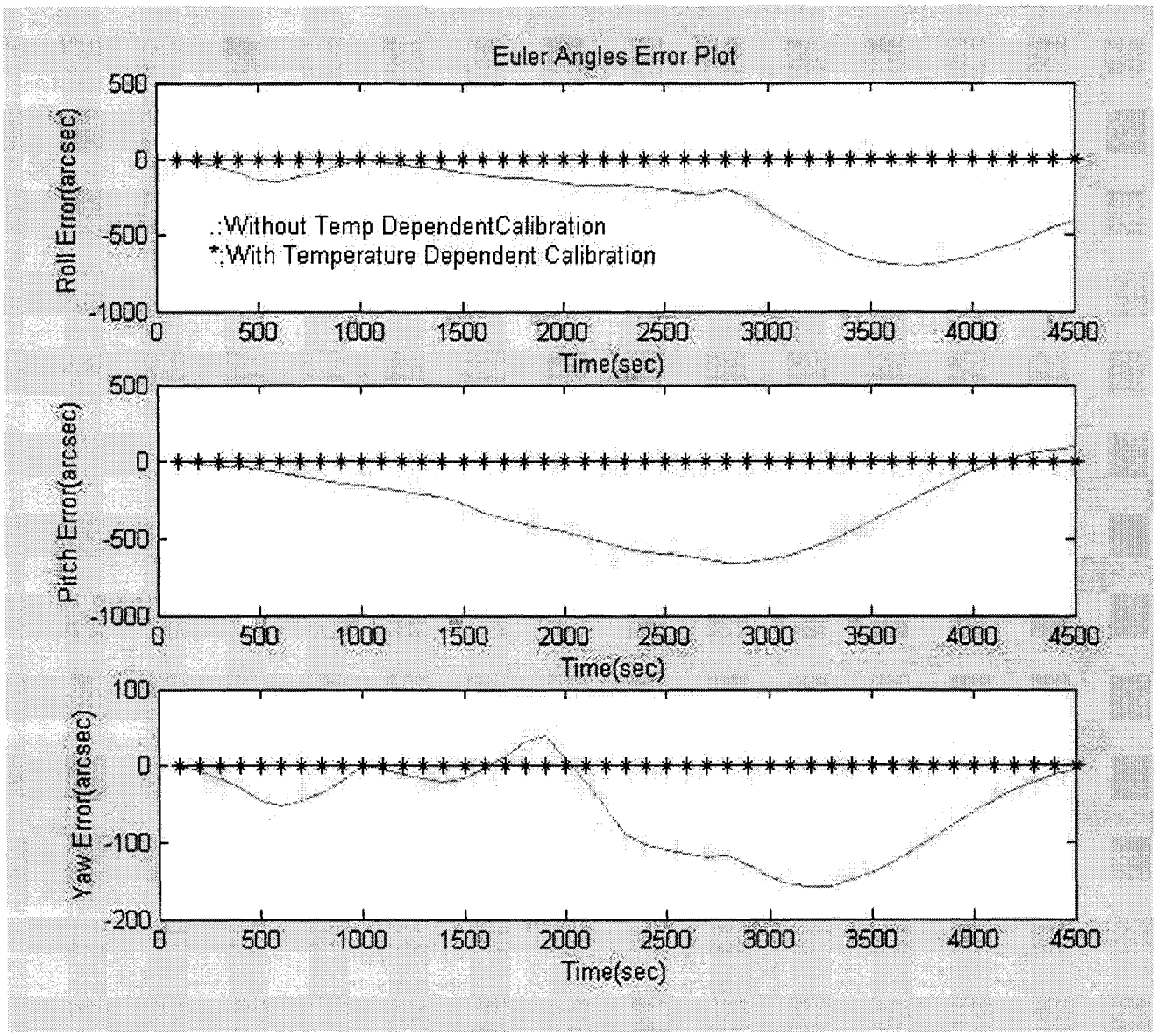


Fig 3. Euler Angle Errors Improvement Due to Temperature-Dependent Gyro Calibration

CONCLUSIONS

A gyro calibration problem with temperature dependent scale factors was formulated. The Levenberg-Marquardt algorithm in the MATLAB function NLSQ was used to solve the resulting least squares problem. Encouraging results using simulated data, conforming to the MAP mission, show the feasibility of applying it for recovering the spacecraft gyro scale factors. The advantage of this approach is that no partial derivatives of the cost function with respect to the state vector are needed. This helps to augment or remove

components from the state vector very easily. Also, since the entire state vector is estimated using a single span of data, no operator intervention is needed.

SYMBOLS

ω	Vector of gyro rates
M	Misalignment matrix
S	Temperature-dependent scale factors
N	Gyro telemetry counts
b	Bias vector
X	State vector
a_{ij}	Scale factor coefficients
q_0	Epoch quaternion
Ω	Matrix of gyro rate components
$f(X)$	Vector of residuals
q_t	True quaternion
q_p	Propagated quaternion
δq	Quaternion error
Δt	Time interval

REFERENCES

1. J. Keat, "*Gyro Calibration Analysis for the High Energy Astronomy Observatory-A (HEAO-A)*", Computer Sciences Corporation, CSC/TM-77/6082, prepared for NASA-GSFC, 1977.
2. P. Davenport and G. Welter, "*Algorithm for In-flight Gyroscope Calibration*," Proceedings of the Flight Mechanics/Estimation Theory Symposium, NASA CP-3011, NASA-GSFC, Greenbelt, MD 20771, May 10-11,1988
3. G.A.Natanson, "*A Transition Matrix Approach to the Davenport Gyro Calibration Scheme*", Paper No. AAS 98-334, Proceedings of the American Astronautical Society (AAS)/GSFC International Symposium on Space Flight Dynamics, Vol. 100 Part I of Advances in Astronautical Sciences, AAS, San Diego, CA, 1998
4. G. Welter et. al., "*Variations on the Davenport Gyroscope Calibration Algorithm*", Proceedings of the Flight Mechanics/Estimation Symposium, NASA-GSFC, NASA CP-3333, Greenbelt, MD 20771, May 14-16, 1996
5. Hashmall, J.A., M.Radomski, and J. Sedlak, " *On-Orbit Calibration of Satellite Gyroscopes*", Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, Denver, August 2000
6. MAP Home Page is at: *http://map.gsfc.nasa.gov/*
7. MAP Flight Software Document, *http://sfandrews.gsfc.nasa.gov/mapdoc/sdp.html*
8. Wertz., J. R.,(Editor), "*Spacecraft Attitude Determination and Control*", D. Riedel Publishing Co., Boston, 1978
9. More, J.J., "*The Levenberg-Marquardt Algorithm: Implementation and Theory*," Numerical Analysis, ed. G.A. Watson, Lecture Notes in Mathematica 630, Springer Verlag, pp. 105-116, 1977.

On-Orbit Calibration of Redundant Spacecraft Gyros by Optimal Reduction to Three Axes *

M. S. Radomski
Computer Sciences Corporation
Lanham-Seabrook, Maryland, USA 20706

Abstract

The Aqua spacecraft will carry four single-axis gyros configured with three orthogonal axes and one skew axis. This redundancy presents a challenge for batch methods of on-orbit gyro calibration that use a spacecraft rotation model deterministically related to gyro data, in that sensor data can respond to at most three angular velocity components. When the number of gyros, N, is greater than 3, the $3\times N$ matrix, **G**, that reduces the N gyro measurements to three body-frame angular-velocity components cannot be fully determined by such methods; there are many such matrices that produce essentially the same angular-velocity history. In such a case, spacecraft operators require information about the $N\times 3$ gyro linear response matrix, **R**, that relates gyro outputs to the body-frame angular velocities causing them. This matrix provides sufficient information to determine multiple reduced-dimension G-matrices for use in case of failure or degradation of one or more gyros, as well as to determine an optimal $3\times N$ **G** for the fully-functional configuration.

A viable proposal is to apply a $3\times N$ pre-filter matrix, **F**, to the N gyro outputs before carrying out a conventional gyro calibration procedure. The angular-velocity history emerging from conventional calibration may then be used as input data, together with the same gyro data that generated it, to fit the alignment, scale-factor, and bias parameters of each gyro axis in turn. A difficulty of such a proposal is the arbitrariness in the choice of **F**. Due to gyro noise, different pre-filter matrices produce different calibrations. This paper presents a method of choosing **F** that is based on optimizing gyro consistency in the limit of infinite weight on gyro data, as compared to sensor data. The choice of **F** is independent of *a priori* alignment and is based on the gyro data alone. The method is applicable to any N of three or more, but reduces to conventional batch-estimation methodologies when N=3. Results of computational comparison among calibration simulations using various choices of **F** will be presented for the Aqua gyro configuration with N=4.

INTRODUCTION

When the Aqua mission is launched later this year or early the next, it will carry with it one more gyroscope sensitive axis than is needed to fully determine the angular velocity. The purpose of the fourth gyro axis (for brevity, "fourth gyro") is to provide redundancy in case of the failure of one of the others; current understanding is that it will be turned off awaiting need, in normal operation. Aqua can, however, be run with all four gyros turned on, connected to the control system, and reporting their measurements to the ground via telemetry. This presents the opportunity of using any of the gyros to check the accuracy of the other three. But it presents, as well, the challenge of resolving the discrepancies, due to finite measurement accuracy, that will inevitably arise. It will not, in general, be just one of the gyros that is responsible for the discrepancy, and there is no way uniquely to identify how the errors on the gyros have combined to produce an observed discrepancy.

This would be all opportunity and no problem if there were accurate knowledge of the on-orbit angular velocities so that gyro calibration could be a process of fitting the calibration-curve slopes and intercepts to the observations of

* This work was supported by the National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center (GSFC), Greenbelt, MD, Contract GS-35F-4381G, Task Order no. S-43411-G.

NASA/GSFC, Guidance, Navigation and Control Center, *Flight Mechanics Symposium*, Greenbelt, MD USA, June 2001.

gyro readouts and angular velocities, like calibrating a thermometer against a standard. Instead, the sensors against which one must compare the gyro readouts are attitude, not rate, sensors and relatively less accurate than the gyros (unless long time-intervals are used). The gyro calibration methods which have successful operational experience (Reference 1) at Goddard Space Flight Center (GSFC) all require that one be able to model the estimated angular velocity as a function of the gyro readouts and calibration parameters via an equation equivalent, in some form, to

$$\vec{\omega}^* = \mathbf{G}\bar{g} - \vec{b} \tag{1}$$

where $\vec{\omega}^*$ is the estimate of the angular velocity vector in body-fixed coordinates, $\bar{g}$ is a column-vector of gyro readouts, **G** is a matrix of the appropriate size, and $\vec{b}$ is a bias vector (which may or may not have some slight programmed time dependence). This expression is integrated to generate a family of models of some attitude observables (such as attitude sensor data or slew offsets) depending on **G**, $\vec{b}$ and other model parameters such as initial attitudes. The calibration parameters can then be adjusted to fit the observables. At first thought, it appears necessary only to expand **G** from 3 columns to N. The problem is that, however many components $\bar{g}$ may have, the manifold of $\bar{g}$, to the extent that the gyros respond only to rotation rates, is a space of, at most, three-dimensions mapped out by the variations of the true angular velocity. All G-matrices that produce equivalent results on that space are forever indistinguishable.

Equation (1), in reducing the dimensions from N to three, implicitly resolves any internal disagreements in the gyros. The better the accuracy of the gyros, the less is the observable difference among competing models for doing that and the more ambiguous and subtle the choice among them becomes. In the end we find that the choice cannot be left purely to the retrospective comparison of attitude trajectories propagated from model angular velocities and compared to attitude sensor data. The extra dimensions of information in the gyro data, in other words, must not be discarded at the start by merely modeling the three-dimensional angular velocities. Yet we shall show that the existing three-axis calibration methods are still usable with the addition of a little superstructure to resolve the retrospective observation ambiguity by an imposed requirement of predictive optimization.

MATHEMATICAL STATEMENT OF THE PROBLEM

The N outputs at n times g_{ij} (i=1...N, j=1...n) of a well behaved gyro assembly can be stacked into column N-vectors and put side by side in an $N\times n$ matrix. The time sequence does not need to consist of continuous gyro readouts, but can include any number of separate time spans. The linear response of each gyro to the stimulus of angular velocity (as represented by n angular velocity 3-vectors lined up in the $3\times n$ matrix ω) is described by the equation

$$\mathbf{g} = \mathbf{R}\omega + \mathbf{B}\mathbf{U} + \mathbf{v} \tag{2}$$

where **R**, the gyro linear-response matrix, is $N\times 3$, **B**, the bias matrix, is $N\times m$, **U** is $m\times n$, and **v** is $N\times n$ gyro error, to be modeled as noise. [The form **BU** we have written for the biasing function requires some explanation. It is constructed to allow different constant biases for any of the separate gyro time spans. Each column of **B** is one of m independent bias N-vectors, each applicable to one or more of a set of m or more disjoint time intervals. The matrix **U** indicates which of the m biases corresponds to each of the n times. Matrix **U** is taken to be made up of n columns, each of m-1 zeros and a single 1. In each row of **U**, the 1's are grouped in clumps corresponding to the time intervals. More general forms of **U** might be used, but this corresponds to the type of time-dependent biasing function supported by the Batch Least-Squares IRU Calibration method (BICAL) (Reference (1)) and is convenient for this analysis.] The angular velocities, ω, are written in an appropriately defined orthonormal body-fixed coordinate system, so that **R** and **B** may be approximately constant for strap-down gyros.

One would then wish to solve this equation for an estimate, ω^*, of the body angular velocities in the form

$$\omega^* = \mathbf{G}\mathbf{g} - \mathbf{b}\mathbf{U} \tag{3}$$

The angular velocity estimation error, $\omega^* - \omega$, will be devoid of systematic variation with ω, if and only if

$$\mathbf{GR} = \mathbf{I}_3 \tag{4}$$

$$\mathbf{GB} = \mathbf{b} \tag{5}$$

When N=3, these equations are sufficient to determine the gyro calibration parameters **G** and **b** for given gyro physics represented by **R** and **B**. When N>3, Equation (4) is insufficient to determine **G** for given **R** (or **R** for given **G**). **R** has many left inverses (and no right inverses). The existence of extra gyros is, however, a boon not a bane, and one can use them to reduce the effect of noise on ω^* by a factor something like $\sqrt{(3/N)}$. For equally noisy gyros, the optimal noise-reducing **G** (if **R** is known) is the well-known pseudoinverse of **R**

$$\mathbf{G} = \mathrm{pinv}(\mathbf{R}) \equiv (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T \tag{6}$$

which is one of the many matrices satisfying Equation (4). For sensors considered noisy relative to the gyros, however, one should not expect that any one of the left inverses of **R** would provide unambiguously best agreement with the sensor data. The residual disagreement will be dominated by the sensor error, whereas the differences provided by various left-inverses of **R** are of the order of gyro noise.

With this background, we introduce a loss function for gyro calibration consisting of the sum of gyro and attitude-sensor terms:

$$J = J_{\mathrm{gyro}}(\mathbf{R},\mathbf{B},\boldsymbol{\omega}) + J_{\mathrm{sensor}}(\boldsymbol{\omega},a) \tag{7}$$

J_{sensor} measures, in a least-squares sense, the agreement of some complement of attitude-sensitive measurements or attitude estimates with the results of propagating the epoch attitudes a using the angular velocities ω. It is the usual BICAL or Davenport-algorithm loss function, which we don't need to specify further at this stage. The gyro loss function is

$$J_{\mathrm{gyro}} = \frac{1}{2}\mathrm{Tr}\left\{[\mathbf{R}\boldsymbol{\omega} + \mathbf{BU} - \mathbf{g}]^T\,\mathbf{M}[\mathbf{R}\boldsymbol{\omega} + \mathbf{BU} - \mathbf{g}]\mathbf{W}\right\} \tag{8}$$

where **W** and **M** are symmetric positive semidefinite square matrices of dimension $n\times n$ and $N\times N$ giving the observation weights by time and by gyro-number, respectively. Although more general forms may be used, we shall take them to be diagonal matrices, corresponding to the assumption of lack of correlation among gyro measurements. Specifically,

$$\begin{aligned} \mathbf{M} &= \mathbf{S}^T\mathbf{S} \\ \mathbf{S} &= \mathrm{diag}\left([1/\sigma_1,\ldots,1/\sigma_N]\right) \\ \mathbf{W} &= \mathrm{diag}\left([\Delta t_1,\ldots,\Delta t_n]\right) \end{aligned} \tag{9}$$

with the sigma's being gyro noise specifications in radians/sec$^{1/2}$. This is a general least-square loss function for gyros, except for the assumptions that gyro observation noise has a factorable dependence on time and gyro instrument number, and that the time dependence is pure white rate-noise.

This loss function depends on $N(3+m)+3n+3n_a$ independent variables (where n_a is the number of attitudes a), which is very large when the number of gyro measurements, n, is large (20,000 or so in the Terra calibration effort). It might be doubted whether the minimum in such a large number of variables can be either well-defined or practically calculable. At any fixed **R** and **B**, however, the optimization with respect to each column of ω is well-defined, even with only J_{gyro} in play, and J_{sensor} cannot make it less well-defined. Much of the information in J_{gyro} will be consumed in that conditional minimization, but all of the information in J_{sensor} remains available to constrain the remaining $N(3+m)+3n_a$ degrees of freedom of **R**, **B**, and a. We shall show proof, below, that the minimum of Equation (7) is well-defined in the limit of infinite weight on J_{gyro} relative to J_{sensor}. So the minimum cannot be degenerate for all relative sensor/gyro weightings. As to the computability of the minimum, it can be stated, without proof here, that a change of variables can give the normal equations of the problem a sparsity structure permitting their solution by sequential methods (bearing a close resemblance to a sequential filter-smoother) with computational effort that grows only with the first power of n, for large n. Numerical stability and ultimate practicability of that solution are unproven at this time, and it is not the subject of this paper.

In this paper, we shall accept the more modest goal of establishing contact with previous successful on-orbit gyro calibration methods by using the limit of infinite gyro weight to establish a form for a low-dimensionality family

$\{\omega^*\}$ of approximate solutions for ω. This limit of infinite gyro weight in the loss function must not be confused with the limit of zero gyro noise. It is distinct in that

- it does require the loss-function to become non-optimal, of course, but this disadvantage is trumped by the fact that
- it is realizable with real gyro data, which do not satisfy Equation (2) with zero noise.

This limit corresponds to the assumptions made in the successful gyro calibration methods of Reference 1 for N=3. In that case, Equation (3) is the obvious solution to the infinite-weight limit, but J_{gyro} drops out of the problem by achieving a minimum at exactly zero (taking **R** as the inverse of **G**) and so only J_{sensor} requires explicit minimization. For N>3, that limit does not trivialize J_{gyro} (unless the gyro data are noise-free simulated data).

Our plan for the approximate minimization of the loss function of Equation (7), carried out in the succeeding sections, will then be as follows:

- For N=3, minimization of J_{gyro} with respect to ω implies:
  - Equation (3) with **G** and **b** determined completely by Equations (4) and (5).
  - **G** and **b** have the same information content as **R** and **B**, and may replace them as the independent variables of the problem, as in traditional treatments such as Reference 1.
  - The information content of J_{gyro} is completely consumed in the minimization with respect to ω, the minimum of J_{gyro} is exactly 0, and it disappears from the problem. This explains why J_{gyro} need not appear explicitly when N=3, as in Reference 1.
- For N>3, unconditional minimization of J_{gyro} with respect to ω will be shown to require
  - Equation (3) with **G** and **b** satisfying, but not completely determined by, Equations (4) and (5).
  - A generalization of Equation (6) (see Equation (11) below) will be shown to be a sufficient (but only weakly necessary) condition for the minimization.
  - **R** and **B** must be maintained as the independent variables and may not be replaced by **G** and **b**.
  - Some of the information content of J_{gyro} survives the partial minimization and its retention is crucial to the solution of the problem.
- The useful result that shall be established is the reduction of Equation (3), with N-column inputs **g**, to Equation (39), with 3-column inputs obtained linearly from **g** (Equation (40)) by multiplication with a 3×N "filter" matrix **F**. The condition of minimizing the gyro loss function, subject to Equation (11), provides the value of **F**.
- Any method applicable to the calibration of a 3-gyro complement (Reference 1) may then be used to adjust the parameters $\mathbf{G}_3$ and **b** of Equation (39) to those that minimize the sensor loss-function, J_{sensor}. That process provides the G-matrix as the product $\mathbf{G}_3\mathbf{F}$.
- If the gyros have sufficiently little noise, filter matrices **F** other than the one that minimizes J_{gryo} can give almost equivalent angular-velocity sequences in Equation (39). G-matrices from such calibrations will not be optimal noise-reducing ones. Decent values of **R** and **B** will then result from a constrained minimization of J_{gryo} with ω replaced by these approximate angular velocities, since they differ little from the optimal ones. More nearly optimal **G** and **b** may be derived from these **R** and **B** determinations via Equation (11).
- This process could be iterated and would be expected to converge to the results that would have been obtained immediately if the optimal **F** had been used from the beginning.

SOLUTION IN THE LIMIT

In the limit where unlimited faith is reposed in the gyros compared to the sensors, J_{sensor} can only be minimized subject to the constraint that J_{gyro} is already at its absolute minimum. (That minimum must be degenerate for J_{sensor} to have any effect on the problem, but it is—see below). In particular, that assumption allows the angular velocities to be solved for in terms of **R** and **B** by using $0 = \partial J_{gyro} / \partial \boldsymbol{\omega} \implies$

$$\mathbf{R}^T\mathbf{M}(\mathbf{R}\boldsymbol{\omega}^* + \mathbf{B}\mathbf{U} - \mathbf{g}) = \mathbf{0} \tag{10}$$

The angular velocities solving this equation are given by Equation (3), provided that **b** is defined by Equation (5) and **G** by

$$\mathbf{G} = \mathbf{G}_{pinv}(\mathbf{R}) \equiv (\mathbf{R}^T\mathbf{M}\mathbf{R})^{-1}\mathbf{R}^T\mathbf{M} \tag{11}$$

This slightly generalized (i.e. **M**-weighted) pseudoinverse form of **G** has the properties

$$\mathbf{R} = \mathbf{M}^{-1}\mathbf{G}^T(\mathbf{G}\mathbf{M}^{-1}\mathbf{G}^T)^{-1} \tag{12}$$

$$(\mathbf{RG})^T\mathbf{MRG} = \mathbf{MRG} = (\mathbf{MRG})^T \tag{13}$$

It should be emphasized that Equation (11) comes from a solution for ω^*, not from a solution for **G**. This solution is time-independent in that it can be established separately for each individual column of ω, independently of the time-dependence and of the relative weighting (**W**). The G-matrix is only a surrogate for a family of ω-solutions. In particular, for (nearly) noise-free gyro data when $N>3$, as previously discussed, many G-matrices (all those solving Equation (4)) will produce (nearly) the same ω^*. Nevertheless, this need not discourage one from using Equation (11)—it will always be able to produce the desired ω^* even when that ability is not unique. Even when there might be other G-matrices that could produce marginally better fits to the retrospective data in the loss-function J, if one thinks one knows something about **R**, one is better off using the weighted pseudoinverse for **G** to reduce gyro data in the future. Perhaps Equation (11) should be viewed as a constraint on the problem, rather than something that can be determined from the data.

Even apart from the limit of infinite-gyro-weight, Equation (3) may be introduced as an additional assumption to reduce the number of degrees of freedom in the loss function (by removing all the thousands of angular velocity components as independent variables). Substituting Equation (3) into (8) and (7):

$$J(\mathbf{R},\mathbf{G},\mathbf{B}) = J_{\text{gyro}}(\mathbf{R},\mathbf{G},\mathbf{B}) + J_{\text{sensor}}(\mathbf{G},\mathbf{b},a) \tag{14}$$

$$J_{\text{gyro}}(\mathbf{R},\mathbf{G},\mathbf{B}) = \tfrac{1}{2}\text{Tr}\left\{[(\mathbf{I}_N - \mathbf{RG})(\mathbf{g} - \mathbf{BU})]^T\mathbf{M}[(\mathbf{I}_N - \mathbf{RG})(\mathbf{g} - \mathbf{BU})]\mathbf{W}\right\} \tag{15}$$

Equation (11) is sufficient to guarantee zero for the gradients of J_{gyro} with respect to **G** (and necessary, as well, unless the gyros are actually noiseless).

Let us now minimize the loss function with respect to the m N-dimensional biases **B**:

$0 = \partial J_{\text{gyro}}/\partial\mathbf{B} \quad \Rightarrow$

$$(\mathbf{I}_N - \mathbf{RG})^T\mathbf{M}(\mathbf{I}_N - \mathbf{RG})(\mathbf{g} - \mathbf{BU})\mathbf{W}\mathbf{U}^T = 0 \tag{16}$$

$$(\mathbf{I}_N - \mathbf{RG})\mathbf{B} = (\mathbf{I}_N - \mathbf{RG})\langle\mathbf{g}\rangle \tag{17}$$

where

$$\langle\mathbf{g}\rangle \equiv \mathbf{g}\mathbf{W}\mathbf{U}^T(\mathbf{U}\mathbf{W}\mathbf{U}^T)^{-1} \tag{18}$$

is a matrix (of the same size as **B**) of the gyro measurements averaged over each biasing subinterval. Equation (17) supplies the information required in addition to Equation (5) to determine a **B** that minimizes the total loss-function from a **b** that minimizes only J_{sensor}. When that latter equation is substituted for **GB** in the left-hand side of the former equation we have

$$\mathbf{B} = (\mathbf{I}_N - \mathbf{RG})\langle\mathbf{g}\rangle + \mathbf{Rb} \tag{19}$$

which satisfies Equation (5) thanks to Equation (4). Furthermore, substituting Equation (17) into Equation (15)

$$J_{\text{gyro}}(\mathbf{R},\mathbf{G}) = \frac{1}{2}\text{Tr}\left\{\mathbf{M}[\mathbf{I}_N - \mathbf{RG}]\mathbf{C}_g[\mathbf{I}_N - \mathbf{RG}]^T\right\} \tag{20}$$

where

$$\mathbf{C}_g = \delta\mathbf{g}\mathbf{W}\delta\mathbf{g}^T \tag{21}$$

$$\delta\mathbf{g} = \mathbf{g} - \langle\mathbf{g}\rangle\mathbf{U} \tag{22}$$

Patently, J_{gyro} is invariant under the transformation $\mathbf{G}\rightarrow\mathbf{AG}$ together with $\mathbf{R}\rightarrow\mathbf{R}\mathbf{A}^{-1}$ (where A is an arbitrary nonsingular 3×3 matrix). This transformation preserves the mutual pseudoinverse relationship of **R** and **G**. Minimization of J_{gyro} is degenerate over at least these nine degrees of freedom. Therefore the constraints provided by infinitely weighting J_{gyro} relative to J_{sensor} still allow us to minimize the latter using

$$\mathbf{G} = \mathbf{G}_3\mathbf{F} \tag{23}$$

where $\mathbf{G}_3$ is a 3×3 matrix which may be arbitrarily varied along with **b** to optimize the agreement with sensor observations, while the 3×N matrix **F** is responsible for maintaining the optimization of the gyro loss function. Note that **Fg** is a set of 3-dimensional gyro-derived data. The minimization of J_{sensor} with respect to $\mathbf{G}_3$ and **b** is precisely the same problem faced in the calibration of spacecraft with 3 gyros and can be treated with any of the existing methods, for example, BICAL. The new problem is now to find the **F** that optimizes J_{gyro}.

As a preliminary to solving that problem, let us digress upon the meaning and properties of the matrix $\mathbf{C}_g$. It is proportional to the sample covariance matrix (over the sample of gyro observations) of **g**, with means subtracted on a subinterval-dependent basis. Since the **g** are constrained to obey an equation of the form of Equation (3), with angular velocities of three dimensions only, one should expect this matrix to have only 3 eigenvalues substantially differing from zero. Consider a set of gyro data from a truth model given by Equation (2) with $\mathbf{R}=\mathbf{R}_t$ and $\omega=\omega_t$. Then

$$\begin{aligned}\mathbf{S}\mathbf{C}_g\mathbf{S}^T/n &= \mathbf{S}\,(\mathbf{R}_t\delta\omega_t + \delta\nu)\mathbf{W}(\mathbf{R}_t\delta\omega_t + \delta\nu)^T\mathbf{S}^T/n \\ &\approx \mathbf{S}\,\mathbf{R}_t(\delta\omega_t\mathbf{W}\delta\omega_t/n)\mathbf{R}_t{}^T\mathbf{S}^T + \mathbf{I}_N\end{aligned} \tag{24}$$

Assuming that the set of angular velocities in the calibration dataset is adequately three-dimensional, the first matrix in the expression above is of rank three having N-3 null eigenvalues and 3 of order $(\delta\omega/[\sigma\Delta t^{-1/2}])^2$. We assume, as we must if we are to expect any gyro calibration to yield good results, that the spacecraft has been maneuvered about all axes at statistically significant rates, so that the three nonzero eigenvalues are large compared with 1, the eigenvalue of the noise term approximating $\mathbf{I}_N$. The addition of the second (identity) term ensures that $\mathbf{S}\mathbf{C}_g\mathbf{S}^T$ is nonsingular with 3 large eigenvalues and N-3 much smaller.

Varying J_{gyro} (Equation (20)) with respect to **G** at fixed **R** (not yet assuming Equation (11))

$0 = \partial J_{\text{gyro}}/\partial\mathbf{G}\,|_{\mathbf{R}} \quad \Rightarrow$

$$\mathbf{R}^T\mathbf{M}\,(\mathbf{I}_N - \mathbf{R}\mathbf{G})\,\mathbf{C}_g = \mathbf{0} \tag{25}$$

Equation (11) is implied by this only to the extent that $\mathbf{C}_g$ is nonsingular. As we have seen, $\mathbf{C}_g$ is nearly singular in N-3 dimensions, and can in fact be singular if there is no gyro noise, such as with error-free simulated gyro data. This simply reminds us that the choice of a particular left-inverse of **R** for **G** is of secondary importance to the necessity for **GR** to be $\mathbf{I}_3$. The choice of $\mathbf{G}_{\text{pinv}}(\mathbf{R})$ is motivated by noise reduction and is not important if the noise is actually absent. Nevertheless, even when $\mathbf{C}_g$ is singular, this choice of **G** still does satisfy Equation (25) and minimize the loss function—it simply no longer does so *uniquely*. Of course, it only does so exactly in the limit of infinite gyro weight. Since some of the gradients of J_{gyro}, i.e. those in the direction of the small-eigenvalue eigenvectors of $\mathbf{C}_g$, would be small even for other **G**s, the influence of the gradients of J_{sensor} cannot necessarily be neglected. We will have gained little if we consider that we have added $J_{\text{gyro}}(\mathbf{R},\mathbf{G},\mathbf{B})$ to $J_{\text{sensor}}(\mathbf{G},\mathbf{b},a)$ along with the "new" variables **R** and **B**; the power of J_{gyro} is largely used up in defining **R** and **B** as functions of the old variables **G** and **b**—a useful thing, since **R** can imply G-matrices to use for different weightings, **M**, and for degraded 3-gyro configurations, but it won't make the minimum much more unique. We have to remind ourselves, however, that we now view **G** and **b** as but proxies for ω^*, and that it is quite reasonable to impose constraints on **G** so that it does not import nearly irrelevant new destabilizing degrees of freedom into the problem.

Now let us finally work out the constraints on **R** corresponding to minimizing J_{gyro} with respect to **R** at the corresponding minima of **G** and **B**. Substituting Equation (13) into (20).

$$J_{\text{gyro}}(\mathbf{R}) = \frac{1}{2}\text{Tr}\{\mathbf{M}\,[\mathbf{I}_N - \mathbf{R}\mathbf{G}]\mathbf{C}_g\} \quad \textit{for} \quad \mathbf{G} = \mathbf{G}_{\text{pinv}}(\mathbf{R}) \tag{26}$$

Minimizing this is equivalent to maximizing the benefit function

$$B_{\text{gyro}} = \text{Tr}\{\mathbf{M}\,\mathbf{R}\mathbf{G}\,\mathbf{C}_g\} \tag{27}$$

In order to find the constrained (by Equation (11)) maximum of B_{gryo}, we now consider the "economy-size" singular value decomposition (SVD) of **SR**:

$$\mathbf{SR} = \mathbf{V_R D_R U_R^T} \tag{28}$$

where

$$\begin{aligned} &\mathbf{U_R} = 3\times 3 \text{ unitary} \\ &\mathbf{D_R} = 3\times 3 \text{ diagonal, nonsingular} \\ &\mathbf{V_R} = N\times 3, 3 \text{ columns of an } N\times N \text{ unitary matrix, } \mathbf{V_R}^T\mathbf{V_R} = \mathbf{I}_3 \quad \mathbf{V_R V_R}^T \neq \mathbf{I}_N \end{aligned} \tag{29}$$

Using this SVD we find, after some algebra,

$$\mathbf{G} = \mathbf{U_R D_R}^{-1}\mathbf{V_R^T S} \tag{30}$$

$$\begin{aligned} \mathbf{MRG} &= \mathbf{S^T V_R D_R U_R^T U_R D_R}^{-1}\mathbf{V_R^T S} \\ &= \mathbf{S^T V_R V_R^T\ S} \\ &= \mathbf{S}^T\left(\sum_{i=1}^{3}\mathbf{v}_{\mathrm{R}i}\mathbf{v}_{\mathrm{R}i}^T\right)\mathbf{S} \end{aligned} \tag{31}$$

where $\mathbf{v}_{\mathrm{R}i}$ are the orthonormal columns of $\mathbf{V}_\mathrm{R}$. Substituting this into (27)

$$B_{\mathrm{gyro}} = \sum_{i=1}^{3}\mathbf{v}_{\mathrm{R}i}^T\mathbf{S C}_\mathrm{g}\mathbf{S}^T\mathbf{v}_{\mathrm{R}i} \tag{32}$$

The maximization of this, subject to the orthonormality constraints on the columns of $\mathbf{V}_\mathrm{R}$, has the general solution

$$\mathbf{V_R} = \mathbf{V_C U_3}^T \tag{33}$$

where $\mathbf{U}_3$ is an arbitrary 3×3 unitary matrix and

$$\mathbf{V_C} = [\mathbf{v}_{\mathrm{C1}} \quad \mathbf{v}_{\mathrm{C2}} \quad \mathbf{v}_{\mathrm{C3}}] \tag{34}$$

is composed of columns which are the orthonormal eigenvectors corresponding to the largest three eigenvalues of $\mathbf{SC}_\mathrm{g}\mathbf{S}^T$:

$$\mathbf{SC}_\mathrm{g}\mathbf{S}^T\mathbf{V_C} = \mathbf{V_C D_C} \tag{35}$$

$\mathbf{D}_\mathrm{C}$ is a 3×3 diagonal matrix with entries >> n. It is not necessary that eigenvalues be nondegenerate; these eigenvectors do not need to be cleanly distinguished from each other. They do need to be clearly distinguished from the eigenvectors of the N-3 smaller eigenvalues. What is important is the subspace spanned by the three eigenvectors; any mixture of the top three eigenvectors may be absorbed into the unitary matrix $\mathbf{U}_3$.

Then the **G**s that maximize B_{gyro} are of the form

$$\begin{aligned} \mathbf{G} &= \mathbf{U_R D_R}^{-1}\mathbf{U_3 V_C}^T\mathbf{S} \\ &= \mathbf{G_3 F} \end{aligned} \tag{36}$$

where

$$\mathbf{G_3} = \mathbf{U_R D_R}^{-1}\mathbf{U_3 A}^{-1} \tag{37}$$

$$\mathbf{F} = \mathbf{A V_C}^T\mathbf{S} \tag{38}$$

for an arbitrary nonsingular 3×3 matrix **A** (which will be used, below, to match *a priori* conditions). The expression for $\mathbf{G_3 A}$, above, is the SVD of a perfectly general 3×3 matrix, there being no constraints on $\mathbf{U_R}, \mathbf{D_R}, \mathbf{U_3}$ other than those defining the SVD. $\mathbf{G}_3$ is not constrained or restricted by the minimization of the gyro loss function. It is the 3×N matrix **F** that bears the full responsibility for that minimization.

We are therefore free to use $\mathbf{G}_3$ and **b** to minimize the sensor loss function with the equation

$$\boldsymbol{\omega}^* = \mathbf{G_3}\boldsymbol{\omega_R} - \mathbf{bU} \tag{39}$$

where

$$\boldsymbol{\omega}_R = \mathbf{F}\mathbf{g} \tag{40}$$

is a three-dimensional reduction of the gyro measurements. It is just this problem that the existing 3-dimensional BICAL algorithm solves, as well as other conventional 3-dimensional gyro calibration methods. The resulting ω^* could be substituted into Equation (2) to provide data for a linear fit of that equation to determine the m degrees of freedom of **B** and 3(N-3) of **R** and that are undetermined by Equations (4) and (5). But since this linear fit is already implicit in the gyro loss function, we get equivalent results by minimizing Equation (20) with respect to **R.** Since we satisfy Equation (25) with Equation (11), we do not need to worry that **G** varies implicitly with **R** in now finding the global minimum of J_{gyro} by imposing $0 = dJ_{\text{gyro}}/d\,\mathbf{R} = \partial J_{\text{gyro}}/\partial\,\mathbf{R}\,|_{\mathbf{G}} \;\Rightarrow$

$$\mathbf{S}(\mathbf{I}_N - \mathbf{R}\mathbf{G})\,\mathbf{C}_g\mathbf{G}^T = \mathbf{0} \tag{41}$$

with the solution

$$\mathbf{R} = \mathbf{C}_g\mathbf{G}^T(\mathbf{G}\mathbf{C}_g\mathbf{G}^T)^{-1} = \tilde{\mathbf{R}}\mathbf{G}_3^{-1} \tag{42}$$

$$\tilde{\mathbf{R}} \equiv \mathbf{C}_g\mathbf{F}^T(\mathbf{F}\mathbf{C}_g\mathbf{F}^T)^{-1} \tag{43}$$

The usefulness of the above equation is not restricted to the infinite-gyro-weight limit, nor to **F** being optimal, i.e., as given by Equation (38). One might, for instance, define a set of conservative calibration strategies by assuming particular fixed values for **F**. For example, if **F** were taken to be the first 3 rows of $\mathbf{I}_4$, the sensor loss-function fit would be a conventional calibration of the first 3 gyros. This would provide the first 3 rows of the R-matrix as $\mathbf{G}_3^{-1}$, and the first three rows of **B** as $\mathbf{G}_3^{-1}\mathbf{b}$, but the fourth rows would have to be found by comparing the fourth gyro's measurements to the angular velocities ω^*. Equation (42), with the help of Equation (19), performs that task, and it is easy to show, with those equations, that **FR** (in this example, the first 3 rows of **R**) is $\mathbf{G}_3^{-1}$ and **FB** is $\mathbf{G}_3^{-1}\mathbf{b}$.

We now have another way to derive the optimality condition for **F** (Equation (38)) as being the condition of consistency between Equation (11) (the requirement of best noise reduction by **G** for known **R**) and Equation (42) (the requirement of best fit of the gyro data to ω^* for known **G**). For, if both these conditions are true, then the partials of J_{gyro} with respect to both **G** (at fixed **R**) and **R** (at fixed **G**) are zero, as shown above, J_{gyro} must be at a minimum and Equation (38) follows. And indeed, Equation (42) can be shown consistent with Equation (11), and the proof uses Equation (35) essentially.

If **F** is optimal and thus of the form given by Equation (38) we have, substituting that equation into Equation (43),

$$\begin{aligned}\mathbf{S}\tilde{\mathbf{R}} &= \mathbf{S}\mathbf{C}_g\mathbf{S}^T\mathbf{V}_C\mathbf{A}^T(\mathbf{A}\mathbf{V}_C{}^T\mathbf{S}\mathbf{C}_g\mathbf{S}^T\mathbf{V}_C\mathbf{A}^T)^{-1} \\ &= \mathbf{V}_C\mathbf{D}_C(\mathbf{V}_C{}^T\mathbf{V}_C\mathbf{D}_C)^{-1}\mathbf{A}^{-1} \\ &= \mathbf{V}_C\mathbf{A}^{-1}\end{aligned} \tag{44}$$

And then

$$\begin{aligned}\mathbf{G}_{\text{pinv}}(\tilde{\mathbf{R}}) &\equiv \left(\left(\mathbf{S}\tilde{\mathbf{R}}\right)^T\mathbf{S}\tilde{\mathbf{R}}\right)^{-1}\left(\mathbf{S}\tilde{\mathbf{R}}\right)^T\mathbf{S} \\ &= \left(\left(\mathbf{V}_C\mathbf{A}^{-1}\right)^T\mathbf{V}_C\mathbf{A}^{-1}\right)^{-1}\left(\mathbf{V}_C\mathbf{A}^{-1}\right)^T\mathbf{S} \\ &= \mathbf{A}\mathbf{V}_C{}^T\mathbf{S} = \mathbf{F}\end{aligned} \tag{45}$$

$$\mathbf{G}_{\text{pinv}}(\mathbf{R}) = \mathbf{G}_{\text{pinv}}(\tilde{\mathbf{R}}\mathbf{G}_3^{-1}) = \mathbf{G}_3\mathbf{G}_{\text{pinv}}(\tilde{\mathbf{R}}) = \mathbf{G}_3\mathbf{F} = \mathbf{G}$$

The set of conservative calibration procedures that pre-defines **F**, not taking advantage of optimal noise reduction in the 3-dimensional J_{sensor} calibration problem, might proceed as follows. Minimize J_{sensor} as a function of $\mathbf{G}_3$ and **b**. Then fit **R** and **B** to the estimated ω^* implied by Equation (39) by using Equations (42) and (19). With this **R** in hand, one would want to support the mission's future gyro data using optimal noise reduction, so one would define

$$\mathbf{G}_{\text{new}} = \mathbf{G}_3\mathbf{G}_{\text{pinv}}(\tilde{\mathbf{R}}) = \mathbf{G}_3\mathbf{F}_{\text{new}} \tag{46}$$

for this purpose. One might think that there could be some gain in iterating this procedure further, each time replacing **F** according to the above and re-minimizing J_{sensor} with respect to $\mathbf{G}_3$ and **b**. This would be burdensome,

since the minimization of J_{sensor} is by far the most computationally intensive part of the process. Such a procedure would, however, probably converge, and, if it did, it would have to be to an optimal **F** according to Equation (38). Using the optimal **F**, determined in advance from eigen-analysis of the gyro-only data matrix $\mathbf{C}_g$, permits an equivalent answer to be achieved with a single J_{sensor} minimization.

An *a priori* value for $\mathbf{G}_3$ is needed to initialize the 3-dimensional calibration (by a method such as BICAL or the Davenport algorithm) using the sensor loss-function. This may be obtained conveniently using Equation (4) if one has an indication of a nominal R-matrix, $\mathbf{R}^0$. An *a priori* $\mathbf{G}_3$ of $\mathbf{G}_3{}^0=\mathbf{I}_3$ can be used if one chooses

$$\mathbf{A} = (\mathbf{V}_C{}^T \mathbf{S} \mathbf{R}^0)^{-1} \tag{47}$$

(or use this expression for $\mathbf{G}_3{}^0$ with $A=\mathbf{I}_3$).

The *N*-row matrices **R** and **B** provide full information on all the gyros' response to spacecraft rotation. Knowledge of them permits us to make recommendations for versions of the G-matrix (and three-dimensional biases **b**) to be used when any one of the gyros must be down-weighted because of increased noise or even totally excluded from use. One simply uses Equations (11) and (5) with the same (calibrated) **R** and **B** but with a new weight matrix $\mathbf{M}^{new}$. If a row and column of $\mathbf{M}^{new}$ is set to be zero, the corresponding gyro will have a null column in the resulting weighted pseudoinverse, so that that gyro's data are removed from further use in calculating angular velocities. A re-calibration is not necessary (unless, of course, a partially degraded gyro is thought to have changed its biases or sensitivities).

RELATION TO THE SOLUTION FOR OPTIMAL WEIGHTING

The limit of infinite weight on J_{gyro} versus J_{sensor} has been used in the above section. Strong arguments can be made that the minimum of Equation (14) is little dependent upon the relative weighting of the two terms. (Little, however, can be said about the use of the infinite weight limit to get from Equation (7) to that point.) J_{sensor} has only small sensitivity, vanishing with gyro noise, to variations d**G** of **G** for which $(d\mathbf{G})\mathbf{R}_t=\mathbf{0}$. J_{gyro}, on the other hand, has full sensitivity to variation in those directions proportional to the inverse square of the gyro noise parameter. Furthermore, the 12-dimensionally degenerate minimum of J_{gyro} can easily be attained using only such variations at little cost to J_{sensor}. Under such circumstances, the relative weighting of the two terms has marginal effect on the solution. A fully mathematical demonstration of this is possible, but requires a large expansion of the notation and will not be presented.

SIMULATED NUMERICAL EVALUATION

We have tried these methods on simulated data corresponding to the Aqua spacecraft, with all four gyros operating and reporting data, while undergoing a calibration maneuver sequence like that executed by the Terra mission on April 13, 2000 (Reference 1). The eight calibration slews started with a positive (x-axis) 10-degree roll followed by a roll back to nominal. This sequence was repeated in the negative roll direction. There were similar 30-degree positive and negative yaw (z-axis) offset maneuvers. Terra, like Aqua, is an Earth-oriented spacecraft and at all times it maintained its 1 rotation per orbit attitude motion about orbit normal. Because of this constant rotation, the offset attitudes resulted in changes of the pitch rates by the cosine of the offset angle and therefore provided rather weak, but realistic, pitch (y-axis) scale-factor observability. This sequence was approximated by simulated rate data with the nominal body y-rate of –0.06 deg/sec. Slews were performed at 0.1 deg/sec, offset attitudes were held for 1000 sec with intermediate nominal holds for 5000 sec, and the whole sequence was begun and ended with a 1000-sec hold at nominal rate. The entire sequence required 22600 sec. The angular velocities corresponding to this scenario were recorded every second (and each second had constant rates).

This angular-velocity history was integrated exactly to provide a truth model of the attitude history. Attitude quaternion "measurements" every 8 seconds stood in for sensor data, with pseudorandom Gaussian noise standard deviation of 10 arcsec per axis. Gyro data were formed from Equation (2) with a truth R-matrix, $\mathbf{R}_t$ corresponding to the nominal for the Aqua mission:

$$\mathbf{R}_t = \begin{pmatrix} 0 & -\sqrt{2/3} & -\sqrt{1/3} \\ \sqrt{1/2} & \sqrt{1/6} & -\sqrt{1/3} \\ -\sqrt{1/2} & \sqrt{1/6} & -\sqrt{1/3} \\ 0 & 0 & 1 \end{pmatrix} \tag{48}$$

This is an orthonormal triple of gyros supplemented by a fourth gyro sensitive to the relative [-1 -1 -1]-direction, with the whole assembly rotated to put the fourth axis in the z (yaw) direction. The angle between gyro axis 4 and the others is 125.26 deg, rather than the 109.47 deg that it would be for a regular tetrahedral configuration. Pseudorandom Gaussian noise impressed on the gyro data included white rate noise at three different levels (4, 57, and 229 $\times 10^{-6}$ deg/sec$^{1/2}$) and bias random walk (6×10^{-9} deg/sec$^{3/2}$). The lower level of gyro noise realistically corresponds to the specification for a NASA standard gyro of the UARS/EUVE era, the Dry-Rotor Inertial Reference Unit, Model II. The other two levels are moderately to highly pessimistic, and were included in order to exacerbate the differences among different calibration methods for illustrative purposes. All three simulations used the same random-number samples, scaled differently.

For each gyro noise level, four different calibration methods were exercised. All used BICAL to minimize J_{sensor} with respect to $\mathbf{G}_3$ and $\mathbf{b}$.

- The "optimal" method allowed $\mathbf{F}$ to vary and find its optimum (Equation (38)) with $\mathbf{G}=\mathbf{G}_{\text{pinv}}(\mathbf{R})$ in the infinite-gyro-weight limit.
- The "suboptimal" method fixed $\mathbf{F}$ at the pseudoinverse of a nominal $\mathbf{R}^0$ close to $\mathbf{R}_t$. (This $\mathbf{R}^0$ takes the place of pre-launch calibration assumptions and has to differ from $\mathbf{R}_t$. Gaussian pseudorandom numbers of standard deviation 0.002 were added to $\mathbf{R}_t$ to form $\mathbf{R}^0$ with scale-factor errors ranging from -1.6 to +5.1 parts per thousand and alignment errors ranging from 2.1 to 3.9 milliradians (0.12 to 0.23 degrees).) Calibrated $\mathbf{R}$-matrices were then calculated, in all cases, with Equation (42) (linear fit to the estimated angular velocities using $\mathbf{G}=\mathbf{G}_3\mathbf{F}$).
- The third calibration method (called "no-4") set $\mathbf{F}$ at the left-inverse of $\mathbf{R}^0$ that ignores the fourth gyro axis, doing a three-axis calibration of the remaining orthonormal triple of gyro axes, before fitting the fourth gyro to its data.
- The last method (called "no-3") used a fixed $\mathbf{F}$ analogous to the "no-4" method, but with gyro axis 3 rather than 4 initially discarded.

The differences among these methods would be nil except for gyro error.

Table 1 shows the results for the maximum R-matrix element error in each simulation.

Table 1. Summary of Simulated Calibration R-Matrix Errors

R-Matrix Errors				
Rate Noise (microdeg/s$^{1/2}$)	Calibration Method			
	Optimal	Suboptimal	no - 4	no - 3
4	1.4e-4	1.4e-4	1.6e-4	1.2e-4
57	4.6e-4	4.6e-4	5.1e-4	2.0e-3
229	1.5e-3	1.5e-3	1.6e-3	2.6e-2

The optimal and suboptimal methods offer little difference; their R-matrices never differ by more than 12 parts per million. This is not unexpected, as the difference between these two is proportional to the product of gyro noise and *a priori* alignment error. The method that initially discards gyro axis number 4 also differs little from these two; that is perhaps more surprising because this foregoes the possibility of a 50% gyro noise variance reduction in the y-axis during the BICAL fit. The calibration method discarding gyro axis 3 performs poorly compared to the others, except at the realistic, low gyro noise level. All methods based on this paper perform quite well (0.01%) at realistic levels of noise.

Gyro rate noise dominates the errors at the two higher levels of noise, and, in view of that the poor performance of the "no-3" method is not so surprising. Consider the angular-velocity noise error 1-sigma ellipsoids in three dimensions, assuming unit uncorrelated rate noise in each of the four gyros. For the nominal Aqua configuration, this is a spheroid, oblate in the z-axis, with semiminor axis of 0.71 ($1/\sqrt{2}$) and unit semimajor axes. If axis four is excluded, the figure becomes a unit sphere. If one of the three orthogonal axes is excluded, on the other hand, the resulting canted triaxial ellipsoid has semiminor axis 0.74 and semimajor axis 2.3. This difference depending on which gyro is discarded is a bit surprising if one has been thinking that the difference between the Aqua configuration and a regular tetrahedral configuration is minor. For a regular tetrahedral configuration, the nominal error surface is a sphere of radius 0.87. Loss of any gyro only stretches that surface by a factor of 2 along the lost axis, only, to semimajor axis 1.7 ($\sqrt{3}$).

CONCLUSIONS

Supernumerary gyro axes present a challenge for batch methods of on-orbit gyro calibration that use a spacecraft rotation model deterministically related to gyro data, but one that they can meet without fundamental changes. The $3\times N$ matrix, **G**, that reduces N measurements to three body-frame angular-velocity components cannot be directly determined by such methods because the needed information is filtered out from the beginning by such an approach. But, by re-inserting the full suite of gyro measurements as observations in a least-squares loss function, the $N\times 3$ gyro linear response matrix, R, and a full N biases, can be determined from the results of 3x3 calibration.

A variable element in this approach is the $3\times N$ pre-filter matrix, **F**, applied to the N gyro outputs to give a 3-axis data stream to which ordinary gyro calibration is applied. The results of our simulations with N=4 show that, for reasonable levels of gyro error, any well-founded choice performs adequately. The method of letting the gyro data determine an optimal **F**, developed in this work, certainly does well, but does not perform noticeably better than the use of the *a priori* nominally optimal **F**. Even without taking any advantage at all of the noise-reduction possible using the extra axes, such as by performing the 3x3 calibration with only 3 of the gyros, good results are usually achieved if the gyro to be discarded is chosen intelligently. This is not altogether surprising since, with only one extra axis, the amount of noise reduction is at most a factor of $\sqrt{0.75}$. Use of the optimal choice of **F** would probably be of greater consequence for larger amounts of redundancy, although even for N=6 the potential for gyro noise reduction is only 30%.

Certain facets of the method remain to be determined. Covariance analysis of the method is incomplete and will not be presented at this time. Since the infinite-gyro-weight limit for which exact results have been derived is a non-optimal smoother, there are some formal complications. A related issue is that of how to properly insert *a priori* information into the loss functions. Such information complicates the solution since it induces correlations between the 3×3 ($\mathbf{G}_3$) and $(N-3)\times 3$ (**F**) degrees of freedom that are absent from the on-orbit information, at least when the time-dependence (**W**) and gyro-axis dependence (**M**) of the gyro noise amplitude are factorable as we have assumed herein.

REFERENCES

1. Joseph A. Hashmall, Mark Radomski, and Joseph Sedlak, "On-Orbit Calibration of Satellite Gyroscopes", **AIAA/AAS Astrodynamics Specialist Conference**, AIAA-2000-4244, Denver, Colorado, August 14-17, 2000.

ESTIMATION OF WHEEL AND CMG ALIGNMENTS FROM ON-ORBIT TELEMETRY

Mason A. Peck
University of California, Los Angeles

ABSTRACT

Using an accurate estimate of the alignments of a spacecraft's actuators helps ensure efficient and precise maneuvers, particularly for missions that implement open-loop or feedforward input sequences. Often actuator alignments are estimated as small biases identified through Kalman filtering. This study takes a different approach. First, we acknowledge that alignment estimation is dual to attitude determination, the problem of finding the representation (or attitude) of one set of basis vectors relative to another. Existing attitude-determination techniques can be incorporated into the actuator-identification problem through an iterative solution. Dynamic parameters of the system and small-angle representations of the alignments are identified from measurement data via least-squares estimation. Then these parameters, along with the measurements, are used as vector observations for optimal attitude-determination for each alignment matrix of interest. The magnitude of the actuator torque applied at the instant of measurement serves as a weighting factor on the vector observation. If further iterations are necessary, the optimal alignments are used as inputs for a new iteration of the least-squares system identification step and new optimal alignments found. The iteration can be considered complete when the attitude-determination step returns identity rotations within some tolerance. This general method is applied explicitly to a class of problems: finding the alignments of control-moment gyros for a spinning spacecraft with unknown mass properties. The solution is demonstrated through simulations.

INTRODUCTION

Common sense tells us that high-performance systems under active control benefit from accurate knowledge of plant dynamics and the effects of actuators. This principle has inspired decades of work in the field of system identification[1]. The performance of systems that implement feedforward or open-loop control sequences also depends heavily on the accuracy of such knowledge. For spacecraft that must slew quickly and precisely, it is the system mass properties and the actuator alignments that must be known. This paper presents a method for estimating both the inertias and large-angle actuator (mis)alignments for a virtually rigid spacecraft from flight data. Rather than treating the alignments as small biases to be estimated in a Kalman filter[2], this method consists of a batch process that identifies the mass properties and large-angle misalignments with the help of attitude-determination techniques.

The term *attitude* refers to the orientation of one coordinate system (or basis vectors in a reference frame) relative to another. Elegant techniques are available for finding the optimal attitude parameterization from vector observations, such as the q method[3] and its successors: QUEST[4], the SVD method[5], FOAM[6], and others. These algorithms were originally devised to provide optimal estimates of a spacecraft's attitude given, for example, many measurements of stars in spacecraft coordinates and given the representation of these vectors in some reference coordinates. There have also been some successful applications of these techniques in other areas: for example, determining the alignment of attitude sensors in spacecraft coordinates[7]. Kalman filters have also been devised that incorporate such algorithms explicitly for this purpose[8]. In this paper attitude-determination algorithms are applied to the problem of determining the orientation, or alignment, of torque and momentum actuators in spacecraft coordinates.

In developing this method, we will keep in mind the environment in which it is likely to be used: an operational setting, where possibly intermittent telemetry is recorded for later processing. Accordingly, this method is based on measurements that are directly available from typical spacecraft: attitude estimates, angular-velocity measurements, and estimated actuator outputs. Angular acceleration is rarely measured, and there is no simple sensor that directly

measures the angular momentum of a spacecraft, desirable as that might be. Furthermore, this method must be able to accommodate sporadic data, for which the sampling rate is not constant and which may contain large periods of unusable data. Certainly there exists some motivation to develop a recursive algorithm or a real-time filter, but this exercise is left for future research.

ANGULAR MOMENTUM OF A SPACECRAFT WITH CONTROL-MOMENT GYROS

The estimation method is based on conservation of angular momentum--not directly on the equations of motion, but on integrals of the motion. This approach circumvents having to extract angular-velocity and momentum derivatives from noisy data. It will also turn out to offer a direct way to incorporate attitude estimates from such devices as star trackers. The derivation begins with a definition of the angular momentum of the system of interest, a rigid body with control-moment gyros (CMGs). This expression is ultimately rewritten as a Wahba-style loss function to be minimized.

A CMG is a device that influences the rigid-body angular momentum $\boldsymbol{I}\cdot\boldsymbol{\omega}^{B/N}$ through torques produced when any number of gimbals precess the spin axis of its spinning rotor. Although the multiple-gimbal case is the general one from a theoretical perspective, single-gimbal CMGs are preferred in practice because the torque produced by precessing the rotor is reacted by structural components rather than the motors on other gimbal axes, saving weight and power. Acknowledging the common practice of incorporating single-axis CMGs, this derivation assumes that $\boldsymbol{h}$ comprises any number of internal components with at most one gimbal axis each. A momentum wheel or reaction wheel is a special case of such a device: one with no gimbal axis.

Consider a spinning rigid spacecraft whose angular-momentum vector ($\boldsymbol{H}$) is given by

$$\boldsymbol{H} = \boldsymbol{I}\cdot\boldsymbol{\omega}^{B/N} + \boldsymbol{h} \tag{1}$$

where

$\boldsymbol{I}$ is the inertia dyadic of the rigid spacecraft.
B is a reference frame fixed in the spacecraft body.
N is an inertial, or Newtonian, reference frame.
$\boldsymbol{\omega}^{B/N}$ is the angular velocity of B in N.
$\boldsymbol{h}$ is the sum of the angular-momentum of internal components, such as momentum wheels or control-moment gyros (CMGs), that rotate about axes through their respective mass centers.

The assumption that the internal components are statically balanced about their angular-velocity vectors implies that the rigid body's mass center does not accelerate--that equations for the rotational dynamics completely describe the motion of the system.

Given these assumptions we define some parameters that describe the i^{th} CMG:

$\boldsymbol{g}_i$ is the gimbal axis of the i^{th} CMG. $\boldsymbol{g}_i$ is fixed in B.
$\boldsymbol{\omega}_i^{G/B} = \Omega_{gi}\boldsymbol{g}_i$ is the angular velocity in B of a frame G_i that is fixed in the i^{th} CMG in such a way that it rotates with the gimbal. Ω_{gi} is the gimbal rate.
$\boldsymbol{\omega}_i^{G/N} = \boldsymbol{\omega}_i^{G/B} + \boldsymbol{\omega}^{B/N} = \Omega_{gi}\boldsymbol{g}_i + \boldsymbol{\omega}^{B/N}$ is the angular velocity of G_i in N.
$\boldsymbol{w}_i$ is the spin axis of the i^{th} CMG's rotor. $\boldsymbol{g}_i$ is fixed in B.
$\boldsymbol{\omega}_i^{W/G} = \Omega_{wi}\boldsymbol{w}_i$ is the angular velocity in N of a frame W_i that is fixed in the CMG in such a way that it rotates with the wheel. Ω_{wi} is the i^{th} rotor's spin rate.
$\boldsymbol{\omega}_i^{W/N} = \boldsymbol{\omega}_i^{W/G} + \boldsymbol{\omega}_i^{G/B} + \boldsymbol{\omega}^{B/N} = \Omega_{wi}\boldsymbol{w}_i + \Omega_{gi}\boldsymbol{g}_i + \boldsymbol{\omega}^{B/N}$ is the angular velocity of W_i in N.
$\boldsymbol{I}_{gi}$ is the inertia dyadic of the rigid i^{th} gimbal. $\boldsymbol{I}_{gi}$ is constant in G_i, but it is not constant in B unless the device under consideration is a momentum wheel or a reaction wheel.

$\boldsymbol{I}_{gi}$ is the inertia dyadic of the rigid i^{th} gimbal. $\boldsymbol{I}_{wi}$ is constant in W_i and in G_i.

Based on these definitions, the internal angular momentum of all k CMGs is

$$\begin{aligned}\boldsymbol{h} &= \sum_{i=1}^{k} \boldsymbol{h}_i \\ &= \sum_{i=1}^{k} \boldsymbol{I}_{gi} \cdot \boldsymbol{\omega}_i^{G/N} + \boldsymbol{I}_{wi} \cdot \boldsymbol{\omega}_i^{W/N} \\ &= \sum_{i=1}^{k} \boldsymbol{I}_{gi} \cdot \left(\Omega_{gi}\boldsymbol{g}_i + \boldsymbol{\omega}^{B/N}\right) + \boldsymbol{I}_{wi} \cdot \left(\Omega_{wi}\boldsymbol{w}_i + \Omega_{gi}\boldsymbol{g}_i + \boldsymbol{\omega}^{B/N}\right)\end{aligned} \tag{2}$$

ATTITUDE PARAMETERIZATION

The foregoing discussion is based on a vector-dyadic representation of the dynamics, which has its merits; but when these quantities are measured by hardware mounted on the satellite, those measurements are obviously scalar quantities associated with some basis vectors. For example, let $\boldsymbol{v}$ be any vector. Then in terms of axes fixed in some frame A,

$$\boldsymbol{v} = v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + v_3\boldsymbol{a}_3 = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix} {}^{A}v\,. \tag{3}$$

The 3×1 matrix ${}^{A}v$ contains the components of $\boldsymbol{v}$ in A axes. The attitude of frame A with respect to some other frame B is denoted ${}^{B}_{A}Q$, such that

$${}^{B}v = {}^{B}_{A}Q\,{}^{A}v\,. \tag{4}$$

The rotation ${}^{B}_{A}Q$ transforms the representation of $\boldsymbol{v}$ in A axes to its representation in B axes. ${}^{B}_{A}Q$ can be derived more explicitly as follows:

$$\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3\end{bmatrix} {}^{B}v = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix} {}^{A}v\,. \tag{5}$$

Taking the dot product with $\begin{bmatrix}\boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{b}_3\end{bmatrix}^T$ results in

$$\begin{bmatrix}\boldsymbol{b}_1\cdot\boldsymbol{b}_1 & \boldsymbol{b}_1\cdot\boldsymbol{b}_2 & \boldsymbol{b}_1\cdot\boldsymbol{b}_3 \\ \boldsymbol{b}_2\cdot\boldsymbol{b}_1 & \boldsymbol{b}_2\cdot\boldsymbol{b}_2 & \boldsymbol{b}_2\cdot\boldsymbol{b}_3 \\ \boldsymbol{b}_3\cdot\boldsymbol{b}_1 & \boldsymbol{b}_3\cdot\boldsymbol{b}_2 & \boldsymbol{b}_3\cdot\boldsymbol{b}_3\end{bmatrix} {}^{B}v = \begin{bmatrix}\boldsymbol{b}_1\cdot\boldsymbol{a}_1 & \boldsymbol{b}_1\cdot\boldsymbol{a}_2 & \boldsymbol{b}_1\cdot\boldsymbol{a}_3 \\ \boldsymbol{b}_2\cdot\boldsymbol{a}_1 & \boldsymbol{b}_2\cdot\boldsymbol{a}_2 & \boldsymbol{b}_2\cdot\boldsymbol{a}_3 \\ \boldsymbol{b}_3\cdot\boldsymbol{a}_1 & \boldsymbol{b}_3\cdot\boldsymbol{a}_2 & \boldsymbol{b}_3\cdot\boldsymbol{a}_3\end{bmatrix} {}^{A}v \tag{6}$$

given that the $\boldsymbol{b}_i$ are orthogonal unit vectors,

$$\begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{bmatrix} {}^{B}v = \begin{bmatrix}\boldsymbol{b}_1\cdot\boldsymbol{a}_1 & \boldsymbol{b}_1\cdot\boldsymbol{a}_2 & \boldsymbol{b}_1\cdot\boldsymbol{a}_3 \\ \boldsymbol{b}_2\cdot\boldsymbol{a}_1 & \boldsymbol{b}_2\cdot\boldsymbol{a}_2 & \boldsymbol{b}_2\cdot\boldsymbol{a}_3 \\ \boldsymbol{b}_3\cdot\boldsymbol{a}_1 & \boldsymbol{b}_3\cdot\boldsymbol{a}_2 & \boldsymbol{b}_3\cdot\boldsymbol{a}_3\end{bmatrix} {}^{A}v\,. \tag{7}$$

Recognizing that the matrix on the right-hand side is simply the direction-cosine matrix reduces equation (7) to the earlier form

$${}^{B}v = {}^{B}_{A}Q\,{}^{A}v\,. \tag{8}$$

The columns of ${}^{B}_{A}Q$ are now seen to be the projection of the A axes onto the B axes. For example, the first column is the representation of $\boldsymbol{a}_1$ in terms of $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$. It is also worth noting that ${}^{B}_{A}Q$ is an orthonormal, or "proper orthogonal" matrix[9]:

$$ {}^{B}_{A}Q^{-1} = {}^{B}_{A}Q^{T} = {}^{A}_{B}Q . \tag{9} $$

Having established the meaning of such a transformation helps define how N, B, G_i, and W_i are related:

${}^{N}_{B}Q$ is the attitude of B in N. This direction-cosine matrix is a parameterization of the spacecraft attitude. A numerically robust way to track ${}^{N}_{B}Q$ is by propagating Euler parameters (a quaternion), whose derivative is a simple function of $\boldsymbol{\omega}^{B/N}$. Alternatively, if a star-tracker or some other device can compute ${}^{N}_{B}Q$ from vector observations, the quaternions need not be propagated except for periods when, for example, no stars can be identified.

${}^{B}_{G_i}Q$ is the attitude, or alignment, of G_i in B. This quantity is of particular interest. Spacecraft that employ CMGs generally do so because they are required to slew quickly and accurately. Often, such slewing maneuvers incorporate a feed-forward torque: based on estimates of the rigid-body mass properties and the gimbal alignment, an open-loop torque profile can be constructed that comes close to an optimal closed-loop sequence. Naturally, closed-loop control is also incorporated to eliminate hang-off and inevitable imprecision in knowledge of the system parameters. The objective of this study is to estimate ${}^{B}_{G_i}Q$.

${}^{G_i}_{W}Q$ is the attitude of W in G_i. Let us assume that each rotor rotates about an axis of symmetry. I.e., there are no products of inertia with respect to $\boldsymbol{w}$, and the transverse moments of inertia (those associated with axes perpendicular to $\boldsymbol{w}$) are equal. Therefore, the wheel's inertia dyadic representation in G_I axes is constant. For some definition of W, its representation in G_i is equal to its representation in W. So, without loss of generality, ${}^{G_i}_{W}Q$ can be taken as the identity. A more rigorous way to account for the axisymmetry would be to define yet another frame between G_i and W, but this added rigor does not enhance the solution.

PARAMETERIZATION OF CMG ALIGNMENT

In order to parameterize the misalignment of the gimbals, we define a rotation ${}_{i}E$ that describes the error in the i^{th} alignment such that

$$ {}^{B}_{G_i}Q = {}_{i}E\,{}^{B}_{G_i}\bar{Q} , \tag{10} $$

where ${}^{B}_{G_i}\bar{Q}$ is the best guess of the alignment, or what the alignment is meant to be. Similarly, a best guess of ${}^{Gi}g$ can be defined:

$$ \begin{aligned} {}^{B}g_i &= {}^{B}_{G_i}Q\,{}^{G_i}g_i \\ &= {}_{i}E\,{}^{B}_{G_i}\bar{Q}\,{}^{G_i}g_i \\ &= {}_{i}E\,{}^{B}\bar{g}_i \end{aligned} \tag{11} $$

Since $\boldsymbol{w}$ is fixed in G_i,

$$ \begin{aligned} {}^{B}w &= {}^{B}_{G_i}Q\,{}^{G_i}w \\ &= {}_{i}E\,{}^{B}_{G_i}\bar{Q}\,{}^{G_i}w \end{aligned} \tag{12} $$

We seek a relation that allows ${}_iE$ to be estimated. Beginning with the matrix equation

$$ {}^{B}h_i = {}^{B}_{G_i}Q\,{}^{G_i}h_i , \tag{13} $$

substituting the above results, and collecting some terms ultimately leads to

$$ {}^{B}h_i = {}_iE\,{}^{B}_{G_i}\bar{Q}\left[\left({}^{G_i}I_{gi} + {}^{G_i}I_{wi}\right){}^{B}_{G_i}\bar{Q}^{T}\left(\Omega_{gi}\,{}^{B}\bar{g}_i + {}^{B}\omega^{B/N}\right) + {}^{G_i}I_{wi}\Omega_{wi}\,{}^{G_i}\bar{w}_i\right], \tag{14} $$

an expression in which the quantities to be estimated, the contents of ${}_iE$, have been factored out so that they appear only once. Since everything that postmultiplies ${}_iE$ is known, let this quantity be represented by

$$ A_i = {}^{B}_{G_i}\bar{Q}\left[\left({}^{G_i}I_{gi} + {}^{G_i}I_{wi}\right){}^{B}_{G_i}\bar{Q}^{T}\left(\Omega_{gi}\,{}^{B}\bar{g}_i + {}^{B}\omega^{B/N}\right) + {}^{G_i}I_{wi}\Omega_{wi}\,{}^{G_i}\bar{w}_i\right]. \tag{15} $$

for notational simplicity. Thus,

$$ {}^{B}h_i = {}_iEA_i . \tag{16} $$

MASS-PROPERTIES IDENTIFICATION WITH SMALL-ANGLE ALIGNMENTS

Although a useful intermediate step, this equation does not immediately provide a solution for ${}_iE$ because ${}^{B}h_i$ is still unknown. To estimate it, we turn to an expression for the total system angular momentum and recognize that the B axes are in general time-varying in N:

$$ {}^{N}H = {}^{N}_{B}Q(t)\,{}^{B}H(t) . \tag{17} $$

Often one can successfully assume that H is fixed in N. In other words, the torques acting on the spacecraft are generally negligible compared to the angular momentum in the wheels and/or in the body. It will be shown later that this assumption can be relaxed in some cases of practical interest. For now, taking H to be constant in N leads to a useful form of conservation of angular momentum across two times t_0 and t_n:

$$ {}^{N}_{B}Q(t_0)\,{}^{B}H(t_0) = {}^{N}_{B}Q(t)\,{}^{B}H(t) . \tag{18} $$

Equation (18) can be rewritten in a way that parameterizes the rigid-body inertia dyadic as a 6×1 matrix[10]:

$$ {}^{N}_{B}Q(t_0)\,{}^{B}\left[\tilde{\omega}(t_0)\tilde{I} + h(t_0)\right] = {}^{N}_{B}Q(t)\,{}^{B}\left[\tilde{\omega}(t)\tilde{I} + h(t)\right], \tag{19} $$

where

$$ \tilde{I} = \begin{bmatrix} I_{11} \\ I_{22} \\ I_{33} \\ I_{12} \\ I_{13} \\ I_{23} \end{bmatrix} \quad \text{and} \quad \tilde{\omega} = \begin{bmatrix} \omega_1 & 0 & 0 & \omega_2 & \omega_3 & 0 \\ 0 & \omega_2 & 0 & \omega_1 & 0 & \omega_3 \\ 0 & 0 & \omega_3 & 0 & \omega_1 & \omega_2 \end{bmatrix}. \tag{20} $$

Collecting on $\tilde{I}$ results in

$$\left[{}^{N}_{B}Q(t_0)\,{}^{B}\tilde{\omega}(t_0)-{}^{N}_{B}Q(t)\,{}^{B}\tilde{\omega}(t)\right]\tilde{I}={}^{N}_{B}Q(t)\,{}^{B}h(t)-{}^{N}_{B}Q(t_0)\,{}^{B}h(t_0). \tag{21}$$

The representation of the B axes in N changes from t_0 to t_n. This relative rotation is

$${}^{B(t)}_{B(t_0)}Q={}^{N}_{B}Q(t)^{T}\,{}^{N}_{B}Q(t_0). \tag{22}$$

Multiplying through by ${}^{N}_{B}Q(t_n)^{T}$ and using this identity yields

$$\left[{}^{B(t)}_{B(t_0)}Q\,{}^{B}\tilde{\omega}(t_0)-{}^{B}\tilde{\omega}(t)\right]\tilde{I}={}^{B}h(t)-{}^{B(t)}_{B(t_0)}Q\,{}^{B}h(t_0). \tag{23}$$

For notational simplicity let

$$W(t_0,t)=\left[{}^{B(t_n)}_{B(t_0)}Q\,{}^{B}\tilde{\omega}(t_0)-{}^{B}\tilde{\omega}(t)\right]. \tag{24}$$

Substituting the expression for CMG momentum in terms of the alignment error leads to

$$W(t_0,t)\,\tilde{I}=\sum_{i=1}^{k}{}_{ei}QA_i(t)-{}^{B(t)}_{B(t_0)}Q\,{}_{i}EA_i(t_0). \tag{25}$$

At this point let us assume that the alignment error is small so that the rotation ${}_{i}E$ is approximately equal to the first-order terms of the Taylor-series expansion of $e^{-\left[{}^{B}\theta\right]^{\times}}$ where

$${}^{B}\theta_i=\left[\theta_1\quad\theta_2\quad\theta_3\right]_i^{T} \tag{26}$$

are infinitesimal rotations about the B axes and

$$\left[{}^{B}\theta_i\right]^{\times}=\begin{bmatrix}0 & -\theta_3 & \theta_2\\ \theta_3 & 0 & -\theta_1\\ -\theta_2 & \theta_1 & 0\end{bmatrix}_i. \tag{27}$$

Then,

$${}_{i}E\approx 1-\left[{}^{B}\theta_i\right]^{\times}. \tag{28}$$

Implementing the identity

$$u^{\times}v=-v^{\times}u \tag{29}$$

leads to

$$W(t_0,t)\,\tilde{I}=\sum_{i=1}^{k}\left(A_i(t)-{}^{B(t)}_{B(t_0)}QA_i(t_0)-\left\{\left[A_i(t)\right]^{\times}-{}^{B(t)}_{B(t_0)}Q\left[A_i(t_0)\right]^{\times}\right\}{}^{B}\theta_i\right). \tag{30}$$

At long last we have an expression that is linear in column matrices of the unknown parameters, $\tilde{I}$ and ${}^{B}\theta_i$, from which follows

$$\left(W(t_0,t) \quad \left\{{}^{B(t_n)}_{B(t_0)}Q\left[A_1(t_0)\right]^{\times}-\left[A_1(t)\right]^{\times}\right\} \quad \cdots \quad \left\{{}^{B(t)}_{B(t_0)}Q\left[A_k(t_0)\right]^{\times}-\left[A_k(t)\right]^{\times}\right\}\right)\begin{bmatrix}\tilde{I}\\ {}^{B}\theta_1\\ \vdots\\ {}^{B}\theta_k\end{bmatrix}=\sum_{i=1}^{k}\left[A_i(t)-{}^{B(t)}_{B(t_0)}QA_i(t_0)\right]. \quad (31)$$

Again, for the sake of simplicity, we represent the matrices of known quantities as $X(0,n)$ and $Y(0,n)$ and the column matrix of parameters as Θ so that this unwieldy expression becomes

$$X(0,n)\Theta=Y(0,n). \quad (32)$$

LEAST-SQUARES SOLUTION

$X(0,n)$ is a 3×(3k+6) matrix, and $Y(0,n)$ is (3k+6)×1. For a series of m measurement events we can cast this problem as a least-squares one with a loss function

$$L(\Theta)=\frac{1}{2}\sum_{n=1}^{m}\left[X(0,n)\Theta-Y(0,n)\right]^{T}\left[X(0,n)\Theta-Y(0,n)\right], \quad (33)$$

and a corresponding least-squares solution

$$\Theta=\begin{bmatrix}X(0,1)\\ \vdots\\ X(0,m)\end{bmatrix}^{+}\begin{bmatrix}Y(0,1)\\ \vdots\\ Y(0,m)\end{bmatrix}. \quad (34)$$

The pseudoinverse shown here can be robustly computed by way of a singular-value decomposition or simply by

$$\begin{bmatrix}X(0,1)\\ \vdots\\ X(0,m)\end{bmatrix}^{+}=\left\{\begin{bmatrix}X(0,1)\\ \vdots\\ X(0,m)\end{bmatrix}^{T}\begin{bmatrix}X(0,1)\\ \vdots\\ X(0,m)\end{bmatrix}\right\}^{-1}\begin{bmatrix}X(0,1)\\ \vdots\\ X(0,m)\end{bmatrix}^{T} \quad (35)$$

if the measurement matrix has nonzero singular values. The solution is a least-squares estimate of the inertia dyadic in B axes and a least-squares fit of a small-angle representation of the alignments. The equations of motion have not been linearized in any way. Only the alignments have been represented approximately.

LARGE-ANGLE ALIGNMENTS

A natural objection to this formulation is that it accommodates only small misalignments. Furthermore, even those have some error associated with the Taylor-series approximation of the misalignment matrix. To remedy this situation we turn to a venerable body of work on the subject of estimating orthonormal matrices like ${}_{i}E$: optimal attitude determination.

There are very successful solutions to the problem, first posed by Wahba, of how to find the components of the orthonormal matrix Q that minimize the loss function

$$L(Q)=\frac{1}{2}\sum_{j=1}^{n}a_j\left|{}^{B}v_j-Q^{A}v_j\right|^2, \quad (36)$$

where the v_j are known vector components in the two frames A and B whose relationship (the attitude Q) is to be found. The a_j are weighting factors often taken to be unity.

The solutions are legion, and Markley has summarized them quite effectively[3]. This paper does not go into them in detail. Using an attitude-determination method with the CMG model described above requires that everything but the alignments be represented as pairs of vectors. For this, we return to the expression for the total angular momentum for two discrete measurement events and isolate the p^{th} CMG, whose alignments are to be refined:

$$ {}^{N}H = {}^{N}_{B}Q(t)\left({}^{B}\tilde{\omega}(t)\tilde{I} + \sum_{i=1,i\neq p}^{k} {}^{B}h_i(t) + {}^{B}h_p(t) \right). \tag{37} $$

Using some of the earlier definitions yields

$$ {}^{N}_{B}Q(t)^{T}\,{}^{N}H - {}^{B}\tilde{\omega}(t)\tilde{I} - \sum_{i=1,i\neq p}^{k} {}^{B}h_i(t) = {}_{p}EA_p(t). \tag{38} $$

Although perhaps difficult to recognize, this expression is in the form required for attitude determination:

$$ {}^{B}v_j = Q\,{}^{A}v_j, \tag{39} $$

where the formulation of the CMG alignment problem maps to the attitude-determination one as

$$ \begin{aligned} &\frac{{}^{N}_{B}Q(t)^{T}\,{}^{N}H - {}^{B}\tilde{\omega}(t)\tilde{I} - \sum_{i=1,i\neq p}^{k} {}^{B}h_i(t)}{\left|A_p(t)\right|} \rightarrow {}^{B}v \\ &{}_{i}E \rightarrow Q \\ &\frac{A_p(t)}{\left|A_p(t)\right|} \rightarrow {}^{A}v \\ &\left|A_p(t)\right|^2 \rightarrow a \end{aligned} \tag{40} $$

It practice there is no need to divide by $\left|A_p(t)\right|$ because the weighting factor a re-scales the vector measurements anyway. It makes sense that the wheel-momentum measurements with the largest magnitude be weighted the most heavily—presumably the noise-to-signal ratio is smaller for such measurements and they are that much more trustworthy. In the absence of more explicit statistics, we use the (unconditional) mean of all n measurements to compute

$$ {}^{N}H = \frac{1}{n}\sum_{j=1}^{n} {}^{N}_{B}Q(t_j)\left({}^{B}\tilde{\omega}(t_j)\tilde{I} + \sum_{i=1}^{k} {}^{B}h_i(t_j) \right). \tag{41} $$

Using the measurements and the previously estimated I and ${}_{i}E$, $i \neq p$ to assemble H and A creates a set of n vector pairs for use with the q method or whatever attitude-determination routine is preferred. The attitude-determination step is repeated for each of the k CMGs. It is worth noting that the transformation from B to N may not be known. However, all that really matters is that the measurements be referenced to some common orientation of B, such as the 0^{th} one, in whatever reference frame is used to express ${}^{B(t_n)}_{B(t_0)}Q$.

Equation (41) assumes that the absolute attitude of B in N is available for each measurement so that a smoothed ${}^{N}H$ can be used. If that is not the case, one can incorporate the measurement at each time step, sacrificing the benefits of the smoothed ${}^{N}H$:

$$ {}^{B(t)}_{B(t_0)}Q^T\left({}^{B}\tilde{\omega}(t_0)\tilde{I} + \sum_{i=1}^{k} {}^{B}h_i(t_0)\right) - {}^{B}\tilde{\omega}(t)\tilde{I} - \sum_{i=1,i\neq p}^{k} {}^{B}h_i(t) = {}_{p}EA_p(t). \tag{42} $$

If further refinement is required, the following substitutions are made:

$$ {}_{i}E\,{}^{B}_{G_i}Q \rightarrow {}^{B}_{G_i}\overline{Q} \tag{43} $$

$$ {}_{i}E\,{}^{B}g_i \rightarrow {}^{B}\overline{g}_i. \tag{44} $$

Then the two-step sequence—linear estimation and nonlinear attitude estimation—is repeated until ${}_{i}E$ converges to the identity within some tolerance. A useful check for such convergence is tr(${}_{i}E$), which approaches 3 as the rotation approaches identity.

Known applied torques can certainly be implemented, but often models of these are inaccurate and the torques themselves tend to be very small (and likely negligible) for practical spacecraft. An important exception for some Earth satellites is the gravity-gradient torque. Conveniently, this torque is accurately modeled as a linear function of the inertia matrix[11], suggesting that this torque can play a useful role in the linear estimation problem. The torque ${}^{B}\tau$ is given in vector-dyadic form by

$$ \boldsymbol{\tau} = 3\left(\frac{\mu}{R_c}\right)^3 \hat{\boldsymbol{o}} \times \boldsymbol{I} \cdot \hat{\boldsymbol{o}}, \tag{45} $$

where $\left(\frac{\mu}{R_c}\right)$ is related to the radial distance of the satellite from the barycenter and $\hat{\boldsymbol{o}}$ is the nadir direction (the vector from the spacecraft mass center to the Earth). Converting equation (45) to matrix form and using some of the operations in equation (20) leads to

$$ {}^{N}_{B}Q(t)\,{}^{B}\tau = {}^{N}_{B}Q(t)3\left(\frac{\mu}{R_c}\right)^3 \left(\hat{o}^{\times}\tilde{\hat{o}}\right)\tilde{I}. \tag{46} $$

Generally everything on the RHS can vary with time except $\tilde{I}$. Integration provides the change in angular momentum in N from one measurement to another:

$$ {}^{N}H(t) - {}^{N}H(t_0) = \left[\int_{t_0}^{t} {}^{N}_{B}Q(\sigma)3\left(\frac{\mu}{R_c}\right)^3 \left(\hat{o}^{\times}\tilde{\hat{o}}\right)d\sigma\right]\tilde{I}. \tag{47} $$

This effect can be included with $W(t_0,t)$ if the added precision warrants it:

$$ \left\{W(t_0,t) + 3{}^{B}_{N}Q(t)\left[\int_{t_0}^{t} {}^{N}_{B}Q(\sigma)\left(\frac{\mu}{R_c}\right)^3 \left(\hat{o}^{\times}\tilde{\hat{o}}\right)d\sigma\right]\right\}\tilde{I} = \sum_{i=1}^{k}\left(A_i(t) - {}^{B(t)}_{B(t_0)}QA_i(t_0) - \left\{\left[A_i(t)\right]^{\times} - {}^{B(t)}_{B(t_0)}Q\left[A_i(t_0)\right]^{\times}\right\}{}^{B}\theta_i\right). \tag{48} $$

Then the least-squares solution and subsequent iterations proceed as before.

EXAMPLES

We use two examples to illustrate the effectiveness of this algorithm. The first case is that of a spacecraft with two CMGs with large misalignments, using measurements with no sensor noise. The misalignment is about 2 radians, in arbitrarily chosen directions. Since there are three attitude parameters for each of the CMGS and six inertia parameters for the system inertia dyadic, we expect the algorithm to yield an approximately correct estimate after 10 three-component vector measurements. Furthermore, we expect comparatively few iterations of the attitude-determination step. The second example is a more stressing case. The spacecraft includes four CMGs, each of which is misaligned by about 2 radians in arbitrarily directions. The noise is also conservatively large, with a variance of about 5%. Convergence here is expected to require more data and more attitude iterations.

In both cases, Markley's SVD method of attitude determination is used. In brief, this method constructs the optimal direction-cosine matrix ${}^a_bQ$ from a matrix B such that

$$B = \sum_{j=1}^{n} a_j \, {}^b v \left({}^a v \right)^T \tag{49}$$

and

$${}^a_bQ = VU^T \tag{50}$$

where V and U are taken from the singular-value decomposition (svd) of B:

$$USV^T = B \tag{51}$$

The method is only slightly more complicated when B is not full rank. This method was chosen for its numerical robustness and simplicity of implementation.

For both cases the rigid-body inertia matrix is

$$I = \begin{bmatrix} 5000 & 50 & -40 \\ 50 & 4000 & 30 \\ -40 & 30 & 3000 \end{bmatrix} kg \cdot m^2$$

The momentum in the rotor of each CMG relative to its gimbal frame is 380 Nms. The total momentum in the spacecraft at any instant is constant in N, and its magnitude is 760 Nms. Each spacecraft is assumed to undergo a large-angle maneuver that distributes the angular-velocity vector and the CMG momentum uniformly throughout a 4π sphere in body coordinates. This data, consisting of 1000 points, is extremely rich. The problem of accuracy and what is observable for poorer data is left as a possible direction for further work.

Case 1. Small Misalignments, No Noise

The direction-cosine matrix for each of the two CMGs is

CMG_1:

```
   0.21274626549828  -0.97035773459249   0.11465118155980
  -0.48767135935684  -0.20712573641758  -0.84810115822143
   0.84670872900874   0.12451825661324  -0.51728090240256
```

CMG_2:

```
 0.65572255717022   -0.35950983599517   -0.66391310112158
-0.58789549910923    0.30861032766778   -0.74775567385604
 0.47371595935250    0.88063178654845   -0.00899146123961
```

Figure 1 shows the convergence of the I_{11} moment of inertia. For brevity, and because the other moments and products of inertia exhibit the same behavior, they are not shown. Figure 2 shows the convergence of the misalignment estimate for each CMG in terms of the total rotation magnitude. This magnitude is the angle of rotation about the eigenaxis, one of many ways to parameterize the direction-cosine matrix[9]. It is clear that only about seven iterations of the attitude-determination algorithm drive the error to 0 within machine precision for this noiseless case. This trend is also apparent in the total rotation magnitude

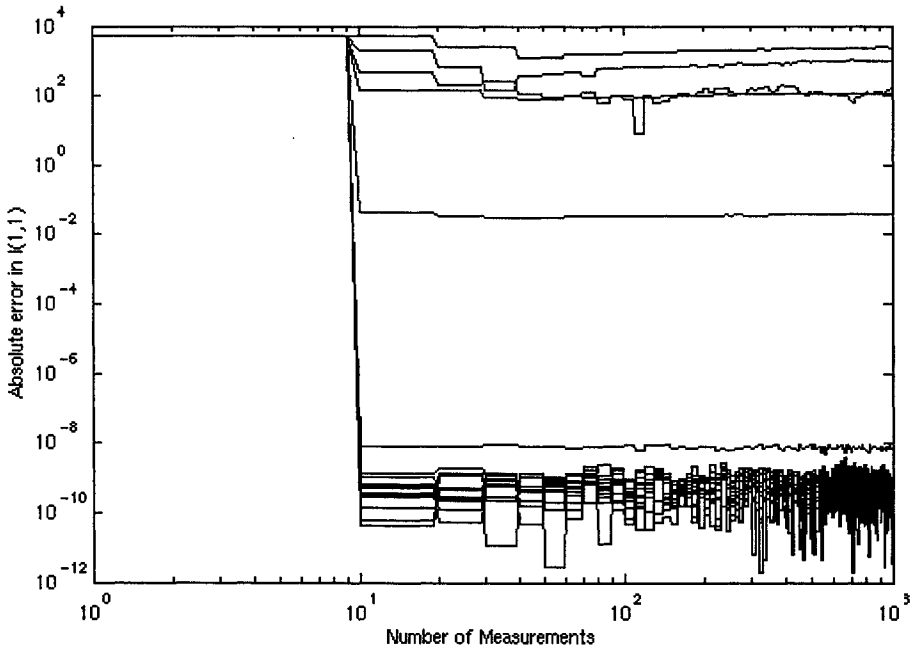


Figure 1. Convergence of I_{11} (kg-m^2) estimate for Case 1.

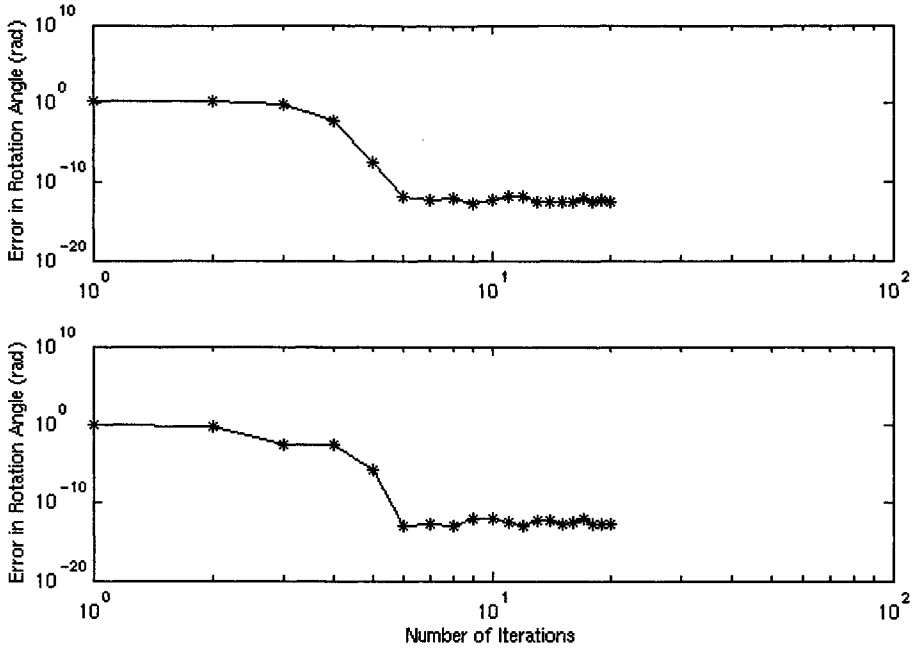


Figure 2. Convergence of Alignment Estimate for Case 1 Through Multiple SVD Iterations.

Case 2. Large Misalignments, 5% Noise Covariance

The direction-cosine matrix for each of the four CMGs is

CMG_1:

```
 0.92101881794847  -0.34579322358611  -0.17930806871621
-0.08104899278043   0.28012912142771  -0.95653475425487
 0.38099264790031   0.89551924705730   0.22997799980834
```

CMG_2:

```
 0.09369581387249  -0.31366664313298   0.94489911178308
 0.29799923806021  -0.89672746205582  -0.32722517157220
 0.94995660352371   0.31223884412515   0.00945281127940
```

CMG_3:

```
 0.34575826210188  -0.07402382057905  -0.93539921860945
-0.93782065120759   0.00537397214515  -0.34707858849539
 0.03071889250702   0.99724199390633  -0.06756297235082
```

CMG_4:

```
 0.55755488967896   0.78200975433685   0.27855571995044
-0.68043895365647   0.23829514987036   0.69297781486506
 0.47553693373068  -0.57591333173515   0.66497252498712
```

Figure 3 shows the convergence of the I_{11} moment of inertia. For brevity, and because the other moments and products of inertia exhibit the same behavior, they are not shown. Figure 4 shows the convergence of the misalignment estimate for each CMG in terms of the total rotation magnitude. Iterations clearly improve the result in this case as well. The downward trend in I_{11} suggests more precision might be achieved with more measurements.

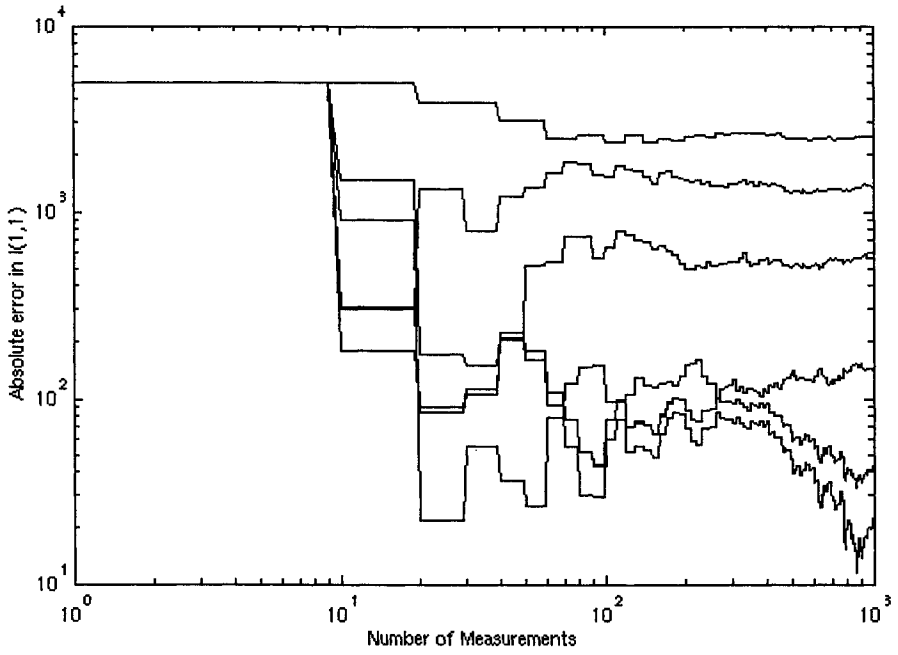


Figure 3. Convergence of I_{11} (kg-m^2) estimate for Case 2.

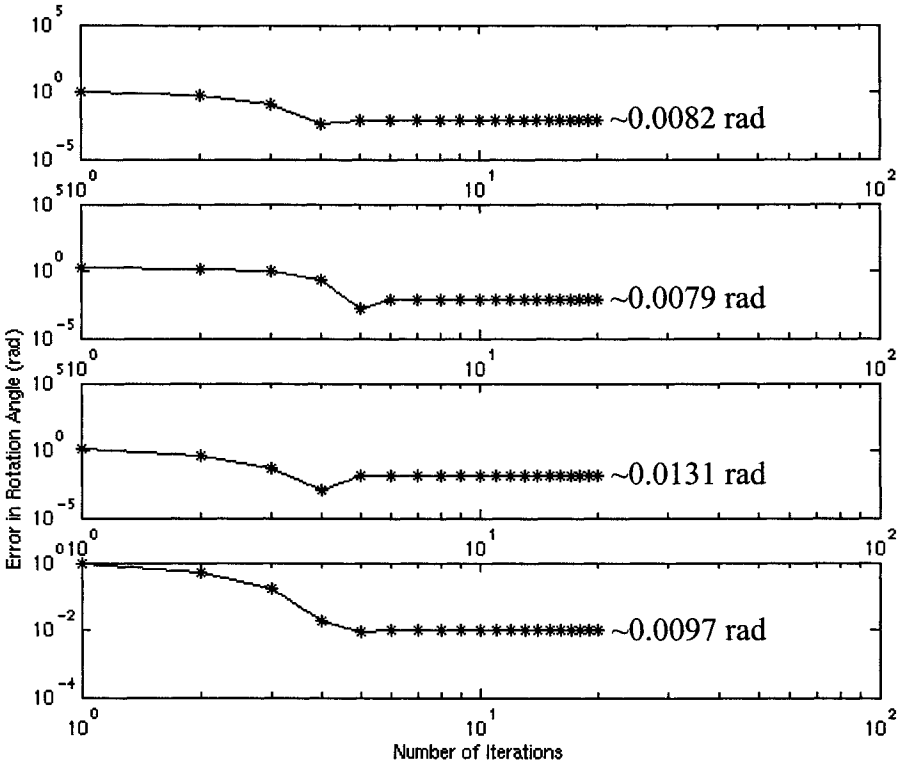


Figure 4. Convergence of Alignment Estimate for Case 2 Through Multiple SVD Iterations.

CONCLUSIONS

This algorithm has several appealing features. It uses measured quantities directly, rather than requiring numerical derivatives (which would be the case if the equations of motion were used in a traditional filtering application). Also, continuous measurements are unnecessary if attitude determination is based on some means other than propagating the quaternion. In this case, ${}^{B(t_n)}_{B(t_0)}Q$ might be available from a device such as a star tracker, whose errors may be completely uncorrelated with those of the gyro measurements of ω.

Simulations show that this algorithm can converge for very large-angle misalignments and in the presence of significant sensor noise. An aspect of this convergence that is not discussed at length here, but deserves further attention, is that actuators misaligned by 180° can lead to mass-properties estimates that are correct except that they are opposite in sign. The resulting inertia matrix is the negative of the true one. To rectify this problem, it is likely sufficient to constrain the moments of inertia to positive numbers, using only their magnitude, in the various iterations of the algorithm.

REFERENCES

1. Astrom and P. Eykhoff, "System Identification, a Survey", *Automatica*, V. 7. pp. 123-162.

2. Jaswinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1979.

3. Markley, F. L., "How to Estimate Attitude from Vector Observations" AAS Paper 99-427, AAS/AIAA Astrodynamics Specialist Conference, Girdwood, AK, August 16-19, 1999.

4. Shuster, M. D., "In Quest of Better Attitudes" AAS Paper 99-427, Plenary Lecture, AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, CA, February 11-14, 2001.

5. Markley, F. L., "Attitude Determination Using Vector Observations and the Singular Value Decomposition", *Journal of the Astronautical Sciences*, Vol. 36, No. 3, pp. 245-258, July-September, 1988.

6. Markley, F. L., "Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm", *Journal of the Astronautical Sciences*, Vol. 41, No. 2, pp. 261-280, April-June, 1993.

7. Shuster, M. D., Chitre, D. K., and Niebur, D. P., "Inflight Estimation of Spacecraft Attitude-Sensor Accuracy and Alignments," *Journal of Guidance, Control and Dynamics*, Vol. 5, No. 4, pp. 339-343, July-August 1982.

8. Shuster, M. D., "Kalman Filtering of Spacecraft Attitude and the Quest Model", *Journal of the Astronautical Sciences*, Vol. 38, No. 3, pp. 377-393, July-September, 1990.

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

10. Peck, Mason A., "Attitude Determination for Gyrostats in Non-Equilibrium Spins from Infrequent Vector Observations," AAS/AIAA Paper 2000-3946, AAS/AIAA Astrodynamics Specialist Conference, Denver, CO, August 2000.

11. Hughes, P.: *Spacecraft Attitude Dynamics*, John Wiley & Sons, Toronto, Canada, 1986.

VALIDATING GOES INSTRUMENT THERMAL DEFORMATIONS[1]

Peter Harter and Benny Ghaffarian
ITT Industries
Fort Wayne, Indiana

Ray Ng and Brett Pugh
Boeing Satellite Systems
El Segundo, California

Paul Wilkin, Chetan Sayal and Don Chu
Swales Aerospace
Beltsville, Maryland

ABSTRACT

Comparison of GOES instrument thermal model predictions with on-orbit data shows that the models capture the observed temperature and misalignment trends. Lack of precise knowledge as to spacecraft pointing precludes such comparison with instrument pointing predictions. Based on the models, thermally induced instrument attitude variation will dominate GOES N-Q Image Motion Compensation (IMC). Errors due to day-to-day changes in the attitude profiles are predicted to be under 10 microradians except for rapid scans where disturbances may reach 30 microradians.

THE VALIDATION PROBLEM

The ITT Industries Imager and Sounder are scanning multi-channel imaging instruments flown on the Geostationary Operational Environmental Satellites (GOES). To ensure that the line-of-sight points in the desired direction, nominal scan mirror orientation is adjusted to compensate for predicted instrument pointing and gimbal misalignment errors. This is called Image Motion Compensation (IMC) and ideally avoids the need for image adjustments on the ground. When there is a problem, it usually comes about because the pointing and misalignment predictions used to compute IMC are not correct.

On the current GOES I-M momentum-bias spacecraft built by Space Systems Loral (SSL), pointing errors come not only from instrument thermal deformations but also from the Earth sensors used to control spacecraft attitude. Upcoming GOES N-Q three-axis-stabilized spacecraft being built by Boeing Satellite Systems (BSS) control spacecraft attitude using star trackers and gyros. This approach is expected to leave instrument thermal deformation as the primary source of pointing error.

The archive of GOES IMC sets provides instrument pointing and misalignment profiles for every day of the year. Although instrument misalignment is spacecraft-independent, instrument pointing includes spacecraft attitude. Obtaining pure instrument pointing from the IMC set requires precise knowledge of spacecraft attitude. Unfortunately, it is the instruments themselves that provide the most accurate observations of spacecraft attitude, and no way has been found to distinguish between instrument pointing and spacecraft attitude effects.

[1] This work was supported under NASA contract NAS5-01090.

If there were a way to propagate spacecraft pointing either kinematically or dynamically, it would be possible to remove spacecraft pointing from the IMC pointing. Propagation accuracy, however, would have to be on the order of one-tenth the maximum IMC or about 100 microradians over 24 hours. Unfortunately GOES I-M gyro propagation is not this accurate. GOES I-M does provide relatively accurate actuator telemetry, but 100 microradians out of the daily 360° pitch rotation is only one part in fifty thousand, and knowledge of the spacecraft inertia is not that accurate.

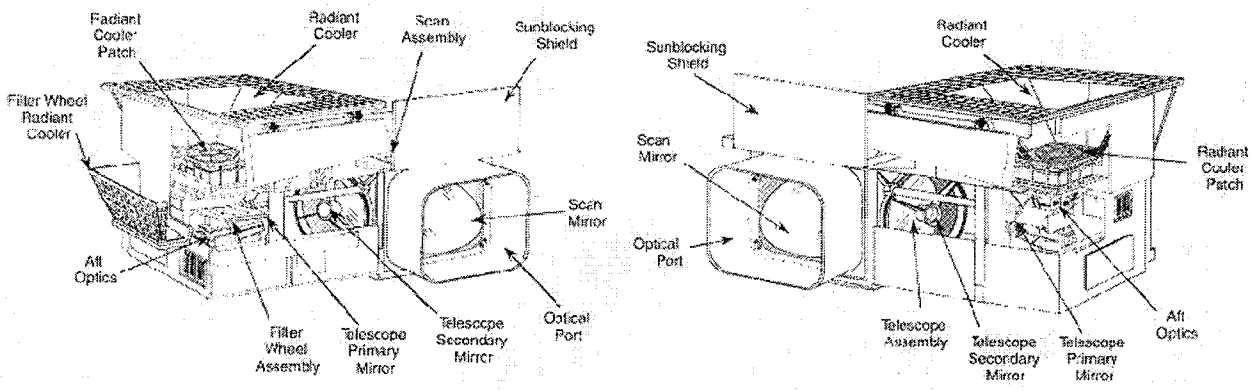
To determine instrument pointing and misalignment, ITT, SSL and BSS developed detailed thermal, structural and optical models for the instruments. The thermal model simulates electronic and solar heating plus reradiation and then computes temperatures at hundreds of instrument points. From these temperatures, the expansion of instrument parts is computed. The new dimensions are fed into a structural model that computes the movement of thousands of instrument points. Deflections and rotations of four critical optical points are then extracted and transformed to pointing and misalignment errors.

INSTRUMENT MODELS

The Imager and Sounder are similar in construction. The optical components consist of a flat scanning mirror, a Cassegrain telescope made up of parabolic primary and secondary mirrors plus a detector array. These are housed in an aluminum box having an optical port for incoming light, cooling louvers above the scan mirror and radiant coolers above the detector array. Light from the Earth enters the optical port, is reflected at the scan mirror and enters the telescope and detector. The baseplate at the back of the instruments is attached to the spacecraft and is heated to keep its temperature above 12° C. Figure 1 shows the structure of the two instruments.

The instruments are insulated everywhere except over the optical port, radiant cooler and louver. At midnight (local solar time), the optical points toward the Sun. This is the time of greatest thermal loading and deformation. For much of the year, the Sun crosses the instrument field-of-view and around the equinoxes is eclipsed by the Earth. This intense heating interrupted by sudden cooling causes the most rapid thermal pointing and misalignment disturbances of all. Because Sun shining into the coolers reduces their effectiveness, 180° yaw maneuvers are planned for equinox to keep the coolers pointing away from the Sun.

Figure 1. Imager and Sounder Structure



For purposes of Image Navigation and Registration (INR), the instruments are characterized by three pointing and two internal misalignment angles which are collectively called the instrument attitude. The roll (ϕ), pitch (θ) and yaw (ψ) pointing angles are defined with respect to the orbital coordinate system x-, y- and z-axes respectively. In the upright orientation, orbital coordinates coincide with nominal instrument

coordinates. In the inverted orientation, the instrument x- and y-axes coincide with the orbital minus x- and minus y-axes. This means that the same deformations in instrument coordinates imply opposite roll and pitch pointing in the upright and inverted orientations.

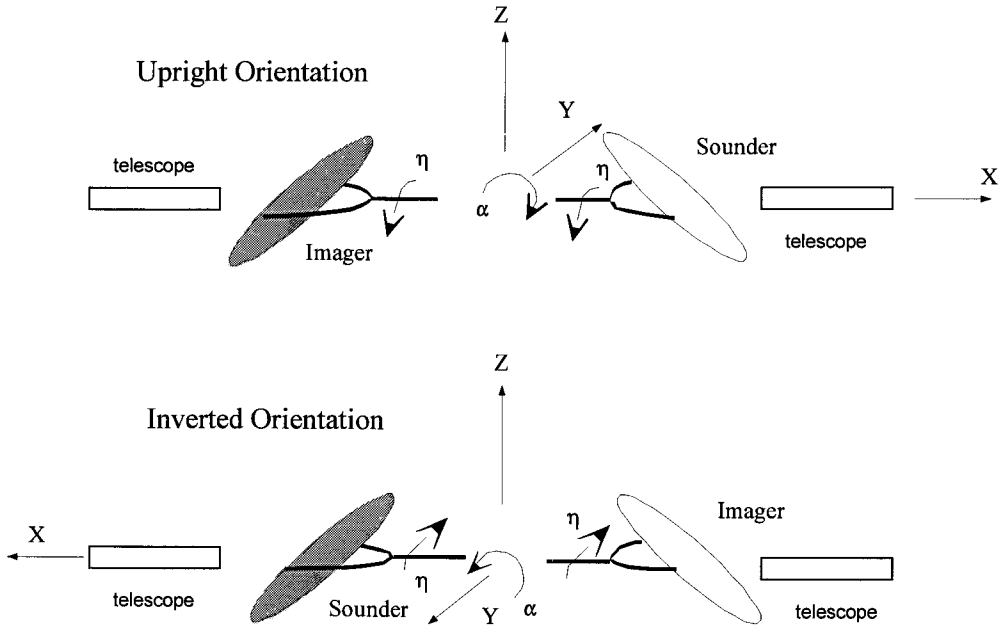
Misalignments are those of the outer scan mirror gimbal axis with respect to the Cassegrain telescope axis. Roll misalignment (ϕ_{ma})and pitch misalignment (θ_{ma}) are not tied to body coordinates as might be expected but are yaw-dependent. This is done to make the misalignments correspond to roll and pitch in either yaw orientation. Neither are the polarities the same for Imager and Sounder. This is done to make the upright Imager look like an inverted Sounder. The directions of positive misalignment rotation are indicated in Table 1.

Table 1. Misalignment Sign Conventions

		Upright	Inverted
Imager	ϕ_{ma}	-z_axis	z_axis
	θ_{ma}	-y_axis	y_axis
Sounder	ϕ_{ma}	z_axis	-z_axis
	θ_{ma}	-y_axis	y_axis

Instrument observations are east-west (E) and north-south (N) scan angles derived from the scan mirror inner and outer gimbal angles α and η. The zero position for α is such that the mirror normal makes a 45°. angle with the outer gimbal axis. The zero position for η is such that the mirror normal lies in the x-z plane. Unlike the misalignment angles, these angles are fixed in instrument coordinates along the positive body y– and x-axes respectively. This is Figure 2 shows this schematic INR model for the two instruments.

Figure 2. INR Instrument Model



The magnitude of the η angle corresponds approximately to the north-south scan angle (N), but the magnitude of the α angle is approximately half the east-west scan angle (E). Because the gimbal angles are defined with respect to instrument coordinates, however, the signs change with yaw orientation. If *yf* equal to +1 indicates the upright orientation and –1 indicates the inverted orientation and *is* equal to +1 indicates the Imager while *is* equal to –1 indicates the Sounder, the dependence of scan angle errors on pointing and misalignment for all four cases can be represented by the following pair of equations

$$\Delta E \approx \begin{pmatrix} 0 & -C_N & -S_N & -is \cdot yf \cdot S_N & 0 \end{pmatrix} \cdot \begin{pmatrix} \phi_n & \theta_n & \psi & \phi_{ma} & \theta_{ma} \end{pmatrix}^T \quad (1)$$

$$\Delta N \approx \begin{pmatrix} -1 & -S_N T_E & C_N T_E & 1 - \dfrac{C_N}{C_E} & \dfrac{S_N}{C_E}(S_E + is \cdot yf) \end{pmatrix} \cdot \begin{pmatrix} \phi_n & \theta_n & \psi & \phi_{ma} & \theta_{ma} \end{pmatrix}^T \quad (2)$$

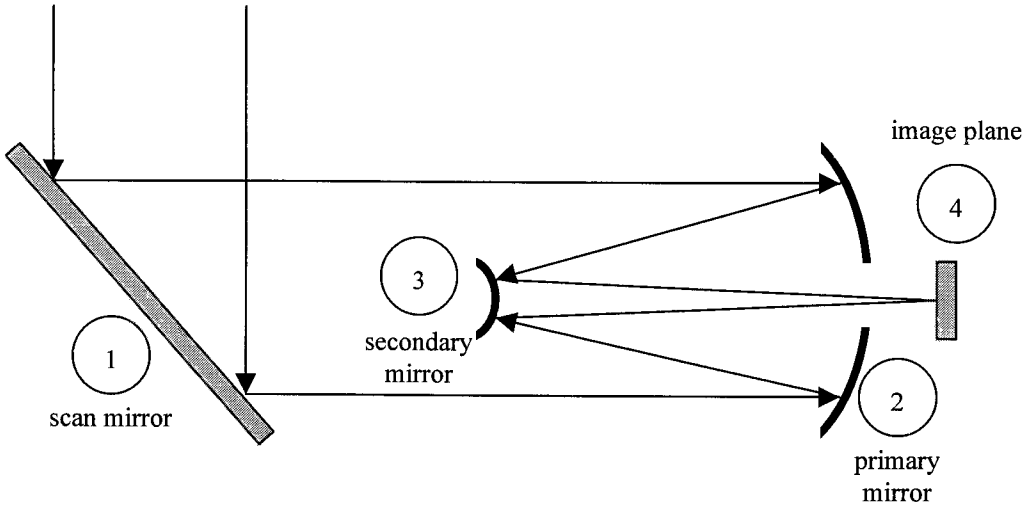
Here, *C* and *S* denote the cosine and sine of the subscript angles, and ϕ_n and θ_n are the modified roll and pitch used in the GOES Orbit and Attitude Tracking System (OATS)

$$\phi_n \equiv \phi + \phi_{ma} \quad (3)$$

$$\theta_n \equiv \theta + \theta_{ma} \quad (4)$$

Although the thermal and structural models predict deformations for thousands of instrument points, the optical model requires the three displacements and three rotations of only four points [1, 2, 3, 4, 5]. As shown in Figure 3, these four points are the centers of the scan mirror, primary mirror, secondary mirror and detector. Their twenty-four coordinates are multiplied by an Optical Sensitivity Matrix (OSM) that is specific for each instrument in each of the two possible yaw orientations. No deformations of the optical components are considered.

Figure 3. Four-Point Optical Model



GOES I-M MODEL VALIDATION

Over the years since 1987 when the ITT instrument model was first developed, it has been repeatedly compared to ground test or on-orbit data and against general purpose modeling software. Temperature predictions have been checked to test the thermal model. Natural vibration frequency predictions have

been checked to test the structural model. Pointing and misalignment predictions have been checked to test the optical model. When necessary, the model has been corrected or enhanced.

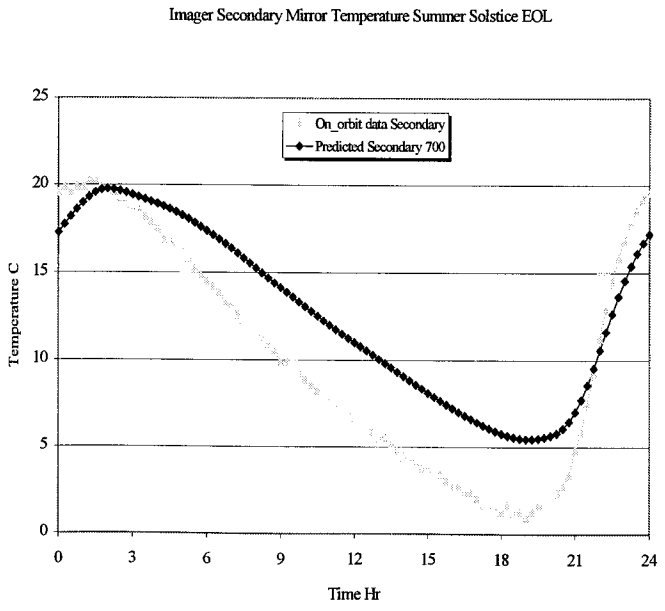
Thermal modeling is the first step in the simulation process, and predicted temperatures have been compared to both ground test and on-orbit telemetry data. In 1994, Harter showed that temperature predictions matched test data and predicted INR performance [6]. In 1995, Zurmehly showed that predicted temperatures matched GOES-8 on-orbit temperatures [7]. In 1996, Ghaffarian and Sprunger predicted that secondary mirror temperatures would exceed operating limits and verified their predictions with GOES-8 thermistor data [8].

Structural models relate temperatures to deformations. In 1997, Harter validated the structural model by successfully predicting instrument natural vibration frequencies [9]. He also showed that uniform temperature gives minute pointing errors as expected. In 1998, he identified the contributions of various instrument sub-assemblies by setting coefficients of thermal expansion to zero for all but the components under consideration [10]. In this way, Harter showed that roughly three-fifths of INR errors came from deformations of the instrument housing and one third came from deformations of the scan assembly.

The optical model transforms the three translations and three rotations for each of the four optical points into three pointing and two misalignment angles. Predictions have been compared with on-orbit IMC sets by Harter in 1991, Walker in 1996, Hampton in 1997, and Harter in 1997, but agreement was weak due to the overriding effect of Earth sensor errors [11, 12, 13, 14]. From the 1997 study, Harter found that structural translations and rotations corresponding to an instrument rigid body rotation wrongly produced optical internal misalignments. He also discovered that the OSM predicted results of the wrong sign for the Sounder. In 1999, Harter revised the Imager OSM and created a distinct OSM for the Sounder [15]. In 2000, Harter and Wickholm showed that the corrected OSM matched results obtained with the Code V optics modeling program [16].

GOES instrument thermal models are important for day-to-day operations as well as INR prediction. The instrument is susceptible to overheating, and operators avoid scanning close to the Sun rather than risk damage. The cost of this caution is lost images. So, there is motivation to predict temperatures as accurately as possible. Current models capture trends and actual temperatures within several degrees. Figure 5 shows typical agreement between predicted and observed Imager secondary mirror temperatures for GOES-10 at summer solstice.

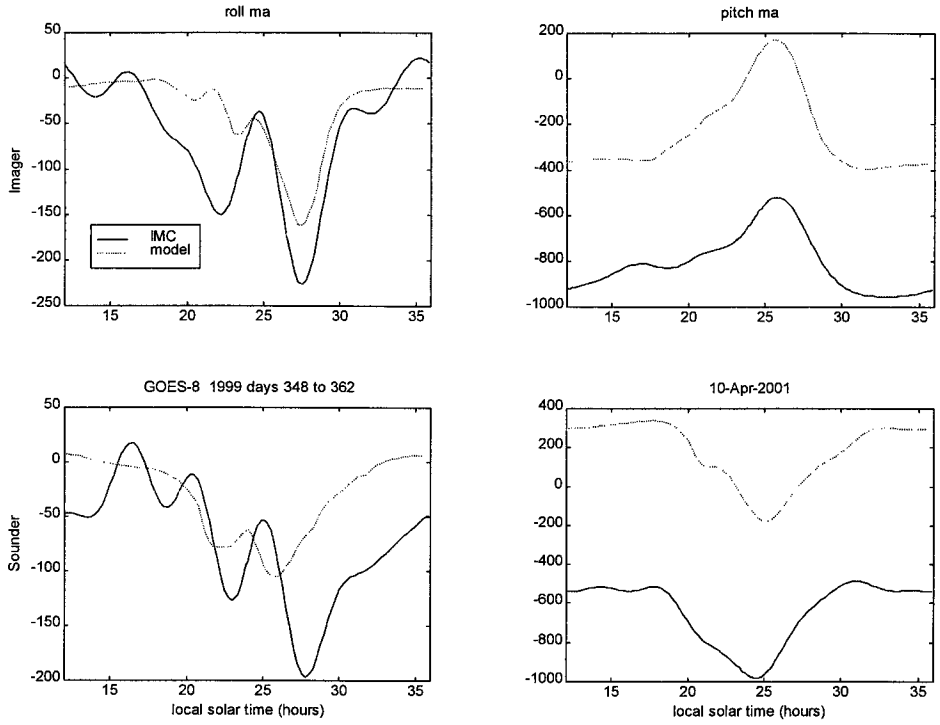
Figure 5. Predicted and Observed Mirror Temperatures



Unfortunately, there are no strain gauges or other devices to measure deformations on-board. So, structural predictions cannot be checked. The next level of on-orbit instrument validation possible is that of instrument attitude. As mentioned earlier, attitude is available from the GOES I-M IMC sets but is not ideal. Pointing includes Earth sensor errors, and misalignment observability is often poor. To minimize the effect of day-to-day variability, the IMC values in the following plots were averaged over fifteen days.

Figure 5 shows predicted and observed misalignments for winter solstice. The pointing curves show little agreement. When the shapes of the predicted and observed misalignment curves are compared, they are almost identical. Roll misalignment values are within the uncertainty of the IMC profiles. The larger pitch misalignment bias may be due to unmodeled effects either in the thermal or INR models. Lower bounds on the pitch misalignment estimate standard deviation are 10 microradians, and the misalignments are not among the most highly correlated of the solved-for state variables.

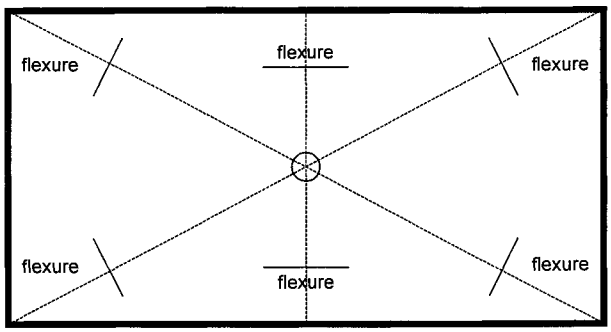
Figure 5. Predicted and Observed Misalignment



GOES N-Q MODEL VALIDATION

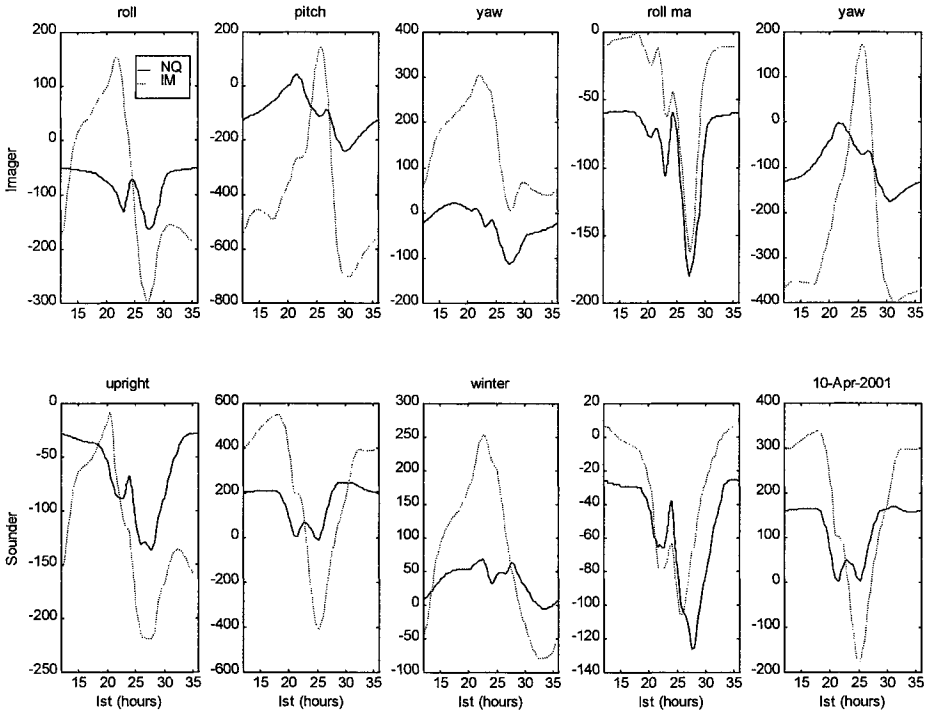
The preceding comparisons with on-orbit data were made against the GOES I-M model, but it is primarily the GOES N-Q models that are of interest now. Without GOES N-Q temperatures or IMC sets for comparison, the N-Q models can still be checked for reasonableness. The N-Q models are qualitatively similar to those for GOES I-M. The primary differences are the attitude stability of the spacecraft and the instrument mounting. Thin metal strips called *flexures* attach the instrument to the bench. This holds the instrument in place but allows it more freedom to expand and contract. As shown in Figure 7, the instrument is mounted using six flexures whose normal vectors intersect at the center of the instrument footprint.

Figure 7. Flexure Mounting



Harter identified flexures in 1998 as a way to reduce pointing and misalignment errors without redesigning the instrument [11]. As shown in Figure 8, the GOES N-Q predictions are generally smaller than those for GOES I-M. Pointing improvement may be due in part to the GOES N-Q bus attitude stability, but misalignment improvement is due to the flexures. Pitch misalignment is greatly reduced while roll misalignment variation is only slightly reduced. The results for summer and equinox are similar.

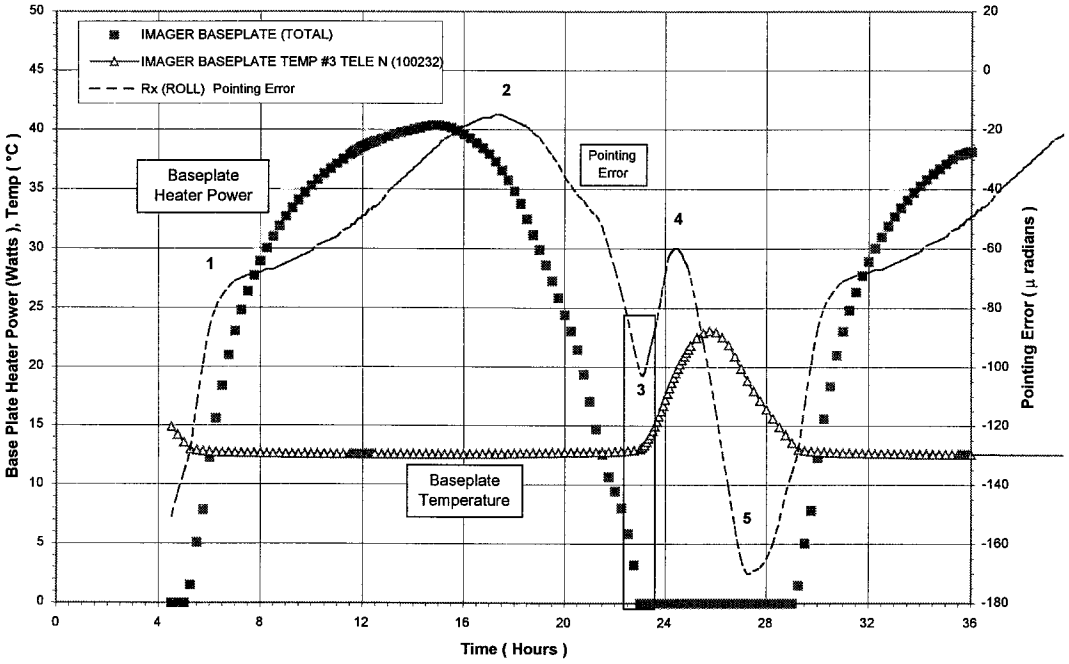
Figure 8. GOES I-M / N-Q Model Comparison



The strange shape of some of the GOES N-Q predicted curves, particularly the double peaks, raised concern that the GOES N-Q model might not be consistent with real instrument behavior. To resolve that question, GOES I-M IMC sets were checked to see if double peaks had been seen in operations and temperature predictions were checked for anything that could cause the double peaks. . In addition, special simulations were run holding the spacecraft at a uniform temperature and holding the instrument at a uniform temperature in order to isolate the contributions of the spacecraft alone and the instrument alone.

Checking IMC sets showed some days with and some without double peaks. That day-to-day variability may have been due to Earth sensor and estimation errors. So, other verification was still necessary. Given the success of the thermal model in capturing thermal variations, baseplate heater power and instrument temperatures were examined for features coincident with instrument attitude variations. The instrument temperatures checked were those of the baseplate, optical port sunshade, scan mirror gussets, north panel, primary and scan mirror. What was found were that the jogs in the INR profiles did correspond to instrument thermal events and that reradiation or *backloading* from the optical port sunshade was a significant source of heating. Figure 9 shows one such plot of roll overplotted on top of baseplate heater power [17].

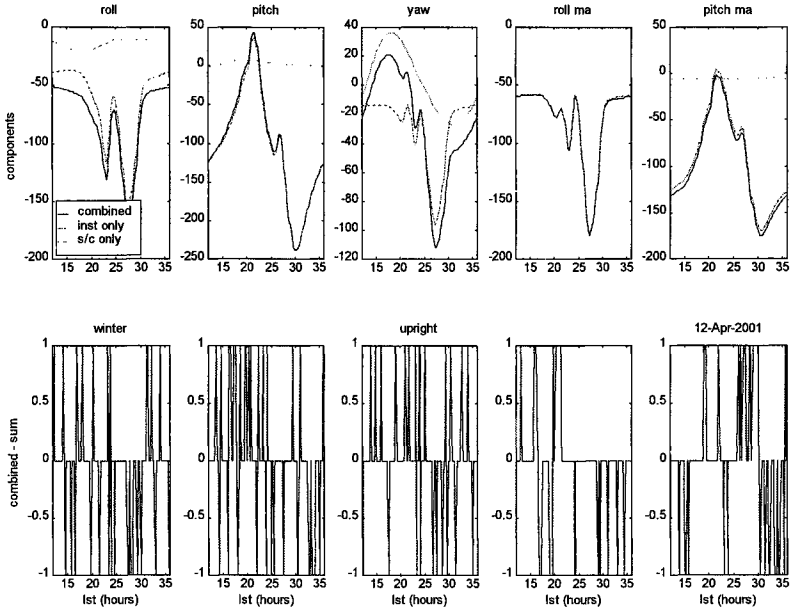
Figure 9. Roll and Baseplate Heater Power



The special simulations to separate instrument and spacecraft effects also suggested that the double peaks were due to the instrument. Figure 10 shows the original combined attitude profiles plus those for the instrument alone and spacecraft alone. The prediction for the instrument alone closely follows that for the spacecraft and instrument combination. Due to the spacecraft stability, the spacecraft-only profile is a small fraction of the combination. Also as expected, spacecraft thermal variation contributes very little to instrument misalignment.

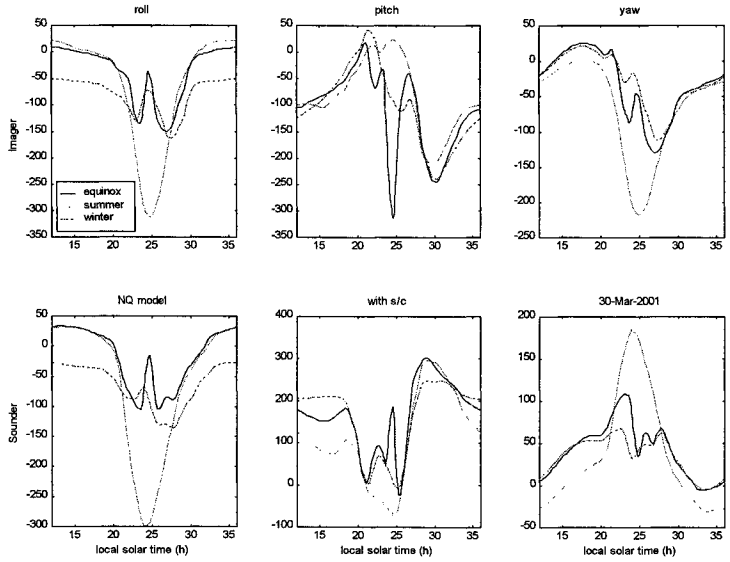
Underlying the interpretation of these simulations as the effect due to spacecraft alone and that due to the instrument alone is the assumption that the combined profile equals the sum of the individual profiles, i.e. that the spacecraft and instrument effects are independent of each other. Conceivably, there could be interactions between components of the spacecraft and instrument attitudes that would cause the individual profiles not to add up. When summed together, however, the spacecraft-only plus instrument-only profiles do match the combined profiles within one microradian. This agreement lends credence to the interpretation of the results as being spacecraft-only and instrument-only effects.

Figure 10. Spacecraft versus Instrument Effects



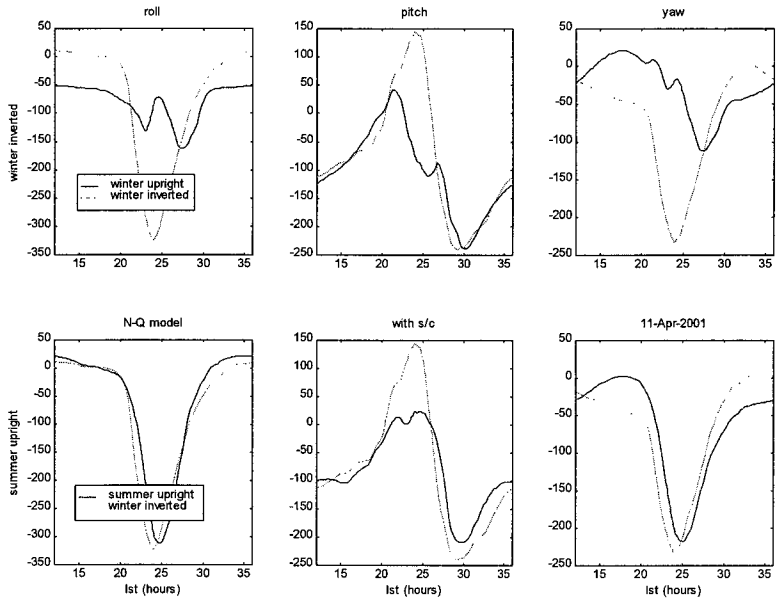
The variation of pointing and misalignment with season also provides insight into the thermal behavior of the instruments. Figure 11 shows predicted pointing for winter, spring and summer in the upright yaw orientation. At midnight, the summer Sun illuminates the north face of the instruments, shines into the louvers and heats the scan cavity. This causes the large long-lasting excursions that dominate the roll and yaw profiles. In contrast, the winter Sun illuminates the south face of the instruments and does not shine into the louvers. It does not cause the same large deformations at midnight. As expected, the equinox case is intermediate between the winter and summer cases for most of the day. At midnight, however, the Earth blocks the Sun, and the instruments cool down rapidly pushing the equinox profile in the winter direction.

Figure 11. Predicted Pointing for Different Seasons



The dominant effect of Sun coming into the louvers suggests that summer in the upright orientation should look more like winter in the inverted orientation than summer in the upright orientation. This is borne out in Figure 12 where the curves in the second row of plots are more similar than those in the first row. The differences between the summer upright and winter inverted curves may be due to the fact that the Sun travels in different directions with respect to the instrument and also to the greater distance from the Sun during summer.

Figure 12. Summer Upright ~ Winter Inverted



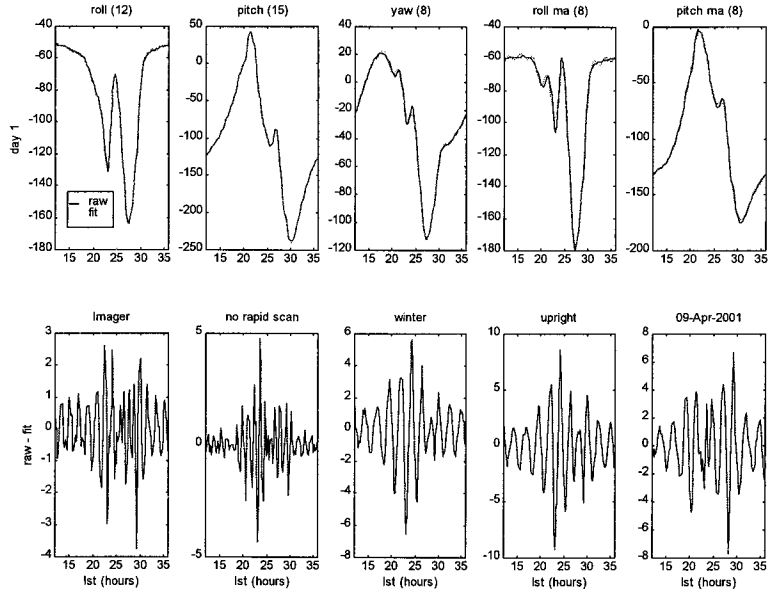
INR PREDICTIONS

To ensure that GOES N-Q meets its INR requirements, day-to-day variations in the thermal profiles are specified to be under 10 microradians. The thermal model provides a means of predicting whether or not this requirement will be met. By fitting the day 1 profile to a Fourier series and comparing that fit to the raw profile for day 2, one can predict the day-to-day INR error. This error depends on season because the thermal profiles depend on season. Vernal and autumnal equinox are expected to be the worst cases, but special short span IMC sets will be used over the eclipse periods. The only case considered here is that of winter solstice which is in the normal season for upright yaw.

Rather than fit points at the same time as the next day's "observations", day 1 points were first interpolated to uniform 15 minute intervals staggered 7.5 minutes from the original points. Then the interpolated values were fit to Fourier series with the recommended [12 15 8 8 8] fit orders for the roll, pitch, yaw, roll misalignment and pitch misalignment. This was considered a better simulation of the random landmark observation spacing encountered in operations. As shown in Figure 13, fit errors range from 3 microradians for roll to 10 microradians for more jagged roll misalignment.

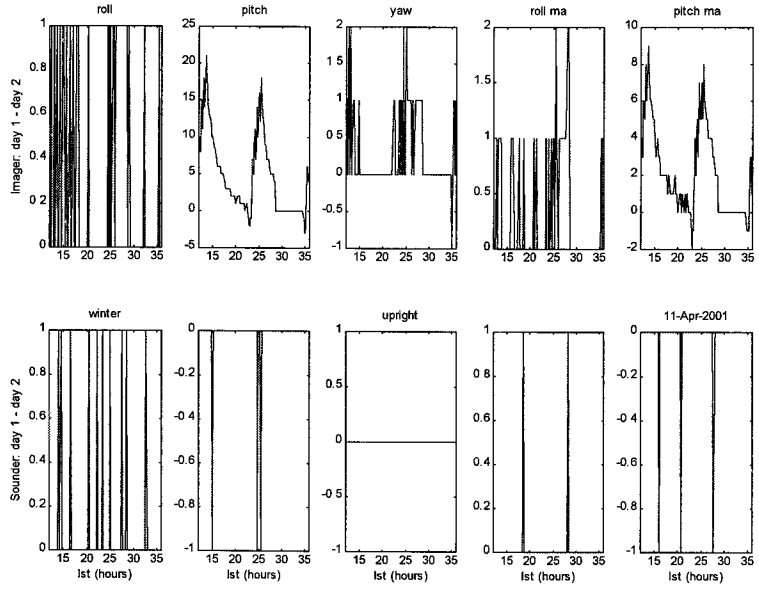
In addition to solar heating, another source of nonrepeatability is the rapid scan mode of operation used to image severe storms. Rapid scanning may heat and deform the instrument on one day but not necessarily the next day. To assess the impact on repeatability, rapid scans were simulated for the Imager on day 2 but not for the Sounder.

Figure 13. Curve Fit Error



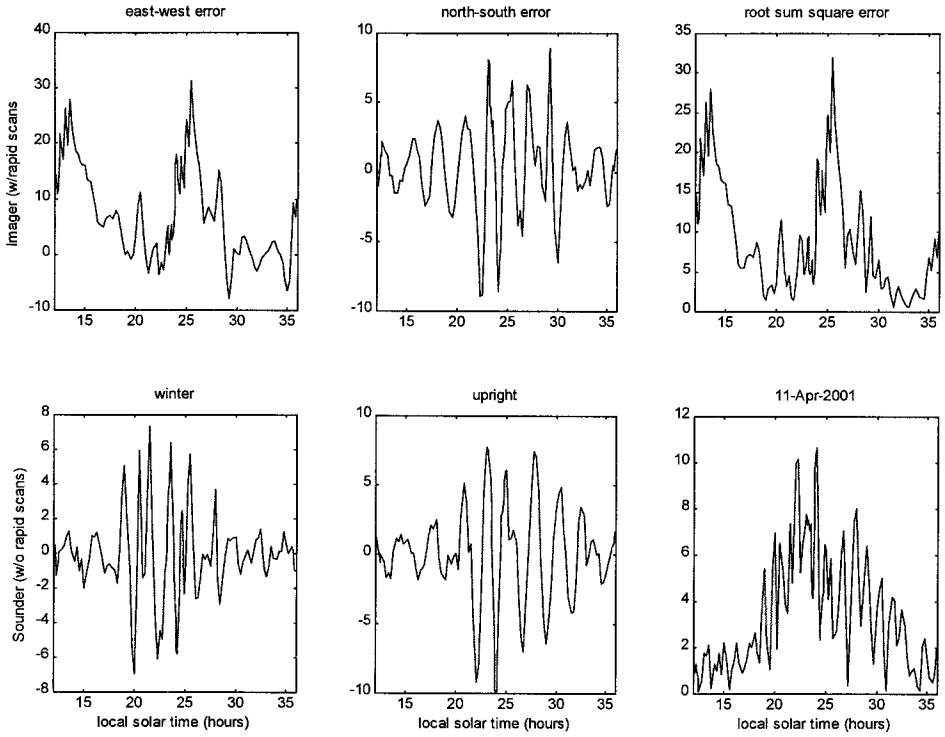
During rapid scanning, the servo motor generates more heat than usual. This alters the thermal deformation profiles and causes errors in the profiles predicted from the previous day. Figure 14 shows the error for the Imager which performs two rapid scans on day 2 and for the Sounder which does none. The error is greater at noon when the scan cavity is otherwise cool than at midnight when sunlight enters the optical port. Without rapid scans, repeatability differences are under one microradian (Sounder). With rapid scans (Imager), differences reach 20 microradians in pitch and 8 microradians in pitch misalignment. This exceeds the specification for day-to-day variation and requires special attention.

Figure 14. Repeatability



Although instrument pointing and misalignment have their own requirements, navigation error is the bottom line for INR. Scan angle errors can be computed using the equations (1-4) given earlier from instrument attitude and the scan angles themselves. Roll, pitch, roll misalignment and pitch misalignment effects depend weakly on scan angle, but yaw effects increase from zero at nadir to a maximum at the edge of the field of view. For a point on the Earth limb 8.3° to the north and east of nadir, the east-west, north-south angle and rss errors are as shown in Figure 15. Overall navigation error may be computed as the root sum square of the east-west and north-south scan angle errors. With rapid scans, navigation error reaches 30 microradians. Without rapid scans, navigation error is only 10 microradians.

Figure 15. Navigation Errors



CONCLUSIONS

The GOES I-M Imager and Sounder thermal, structural and optical models have been shown to agree with on-orbit data. The GOES N-Q instrument models are derived from those for GOES I-M but predict smaller pointing and misalignment errors due to improved spacecraft attitude stability and stress-relieving instrument flexure mounts. In the absence of on-orbit GOES N-Q data for comparison, the instrument models have been shown to be reasonable and self-consistent.

The GOES N-Q models predict that the instrument itself will be the primary source of pointing and misalignment errors. In the absence of rapid scans, day-to-day pointing and misalignment repeatability are predicted to be 1 microradian at winter solstice. Curve fitting these profiles with the planned [12 15 8 8 8] order Fourier series introduces additional error on the order of 8 microradians. Rapid scanning heats the instruments and causes deformations that add 20 microradians to the nonrepeatable error. The resulting root sum square of the east-west and north-south errors is 30 microradians which is of concern.

REFERENCES

1. Roy, N. and W. Haile, "Optical Ray Tracing in Finite Element Models", Advanced Technology Optical Telescopes II, SPIE Proc. Vol. 444, 1983
2. "Thermal/Structural Analysis for GOES Image Navigation/Registration", ITT Rept MD-1053, September 14, 1988, by J. Monirian and P. Harter
3. "Thermal/Structural Analysis for GOES Image Navigation/Registration Instrument-Instrument Cross-over Effect", ITT Rept MD-1127, February 20, 1989, by P. Harter and J.Monirian
4. Memo, "GOES One-Model Report", September 19, 1994, E. Zurmehly to C. Young
5. "GOES INR Prediction Evaluation", S-GOES-NSH-99-0140, September 28, 1999, by N. Hodgman
6. "Analysis of GOES Image Navigation/Registration Performance with Test-Correlated Temperatures", ITT Rept MD-1546, February 21, 1994, by P. Harter
7. Memo, "Correlation of the GOES Instrument Thermal Model to GOES-8 Flight Data", 26 April 1995, E. Zurmehly to C. Young
8. "Solar Intrusion Thermal Analysis", SPIE Vol. 2812, pp. 251-259, 1996, by Benny Ghaffarian and Kent Sprunger
9. "GOES Imager Model Correlation with Overall Damping", S-GOES-PKH-97-0044, May 9, 1997, by P. Harter
10. "Effect of Composite Structural Components on GOES-NQ INR Performance", February 6, 1998, S-NQ-PKH-98-0011, by P. Harter
11. "Preliminary Analysis of GOES Image Navigation/Registration Performance in Response to CCN #116, ITT Rept MD-1355, August 16, 1991, by P. Harter
12. "NAS5-96090, Systems Engineering Analysis #4, Development and Correlation of Thermal Distortion Model", GSFC #214.2-G03670, December 10, 1996, Clelia Walker (GSFC) Correspondence to C. Cierniak (ITT)
13. "Sounder_18 Correlation", June 12, 1997, J. Hampton (Swales) to M. Clark (GSFC)
14. "Analysis of GOES INR Performance w/ Second Generation Instrument Models", S-GOES-PKH-97-0082, August 15, 1997, by P. Harter
15. "Revised Optical Sensitivity Matrix for GOES Sounder INR Error Analysis", S-GOES-PKH-99-0146, October 4, 1999, by P. Harter
16. "Verification of GOES Optical Sensitivity Matrix using Code V Program", S-GOES-PKH-00-0031, March 1, 2000, by P. Harter and D. Wickholm
17. "Investigation of Integrated Thermal Model Effect on INR pointing Error", ITT MEA report01-0010, NOPQDCN-274, 20 Feb 2001, by B. Ghaffarian

A Long-Term Characterization of GOES I-M Attitude Errors

Joe Harris and Jim Carr
Carr Astronautics
Washington, DC

Don Chu
Swales Aerospace
Beltsville, MD

ABSTRACT

A year of GOES-8 and GOES-10 in-flight data is analyzed to characterize long-term GOES attitude errors. GOES Imager payload star sensing is the primary means of observing attitude. Attitude is estimated for each star window and the diurnally repeatable and nonrepeatable components of the attitude error are determined. The diurnally repeatable component is due to thermal distortions of the spacecraft and instruments. It evolves slowly with the seasons. The nonrepeatable component is due to the response of the Earth sensor, used by the GOES attitude control system, to IR radiance gradients and cold clouds near the Earth limb.

INTRODUCTION

The GOES I-M satellites are the first geostationary three-axis stabilized series of weather satellites. Prior satellites in the GOES series were spinners, providing a highly stable platform for imaging, but a low duty cycle for viewing the Earth since the Earth subtends less than 20° of arc from a geostationary orbit. The three-axis attitude control of GOES I-M permits continuous viewing of the Earth; however, the problem of stabilizing the images in the presence of orbit and attitude motion becomes much more challenging. This challenge is addressed by the GOES I-M Image Navigation and Registration (INR) system.

A significant amount of in-flight operational experience has been accumulated for the GOES I-M series. In this paper, we examine the in-flight attitude errors for GOES-8 (East satellite) and GOES-10 (West satellite) over the course of an entire year (CY 1999). Our focus is on long-term errors, for which thermal deformations (driven by solar heating) and Earth sensor random walk (driven by the response of the Earth sensor to IR radiance gradients and high clouds) are the primary contributors. This work should be of interest for future three-axis stabilized weather satellite programs (*e.g.*, GOES N-Q,

GOES R+, and MTSAT) and commercial geostationary communications satellite programs where Earth sensors are routinely used for attitude control.

Image Navigation and Registration (INR) System Description

The INR system block diagram for GOES I-M is shown in Figure 1. The GOES I-M series Attitude and Orbit Control System (AOCS) is a momentum-biased three-axis stabilized control system using an Earth Sensor (ES) to steer the spacecraft towards the center of the Earth. The Imager and Sounder meteorological instruments each employs a two-axis servo-controlled scan mechanism to scan their visible and infrared detectors across the Earth. During scanning, the AOCS can generate real-time error compensation signals for the instrument servos, derived from an Image Motion Compensation (IMC) set, to compensate for orbit and attitude motion.

The Imager and Sounder radiometric data stream is sent to the ground as wideband telemetry, where it is calibrated, formatted and rebroadcast to the user community via the GOES VARiable (GVAR) data link. A single-station two-way range measurement is made every 15 minutes using the GVAR broadcast and received at the GOES Command and Data Acquisition (CDA) station.

Star sensing is performed by slewing the Imager and Sounder lines-of-sight to view space and dwelling near a star. The inertial pitch rate of the spacecraft causes the stellar image to be dragged across the visible detectors. Nominally, 4 stars are observed every half-hour by the Imager and 3 stars are observed every half-hour by the Sounder.

Navigation landmark observations are made during normal scanning of the Earth by the Imager. Operationally, navigation landmarks are derived from the Imager visible (1 km resolution at nadir) and infrared (4 km resolution at nadir) images. A typical landmark rate is about 100 landmarks per day. Sounder landmarks are not routinely used operationally because of the coarser Sounder resolution (10 km at nadir).

Orbit and attitude determination is performed by the Orbit and Attitude Tracking System (OATS) every 24 hours using the range, star, and landmark measurements. The orbit is observable from the ranges and landmarks. Instrument attitude (separate solution for Imager and Sounder) is observable from the landmarks and stars. The OATS propagates the orbit and attitude solutions into the future and packages them into an IMC set, which is transferred to the spacecraft once per day. An IMC set is used to compensate the images for orbit and attitude motion for 24 hours; however, unpredictable attitude errors often require Short Span Attitude Adjustments (SSAA's) to be performed. An SSAA is triggered when the star measurement residuals are observed to deviate significantly from the attitude predictions.

Long-Term Attitude Error Sources

Thermal Distortions. The sun rises and sets over the Earth face of the GOES spacecraft much like it does over a point on the Earth. The solar heating geometry causes thermal deformations of the instruments and spacecraft that are nearly repeatable from one day to

the next. These deformations can be parameterized by five angles for each satellite – roll, pitch, yaw, roll misalignment and pitch misalignment. The latter two angles describe the instrument internal misalignment between the telescope and scan mechanism, resulting in off-axis pointing errors which depend on scan elevation and azimuth angles. The IMC set represents the thermal distortions by Fourier series because of their nearly periodic behavior. Thermal deformation errors generally will evolve slowly in time as the declination of the sun changes with the seasons; however, rapid variations will occur during eclipse season. Little is known from INR operations in eclipse or just after eclipse because imaging ceases before eclipse entry and does not resume until well after eclipse exit.

Earth Sensor (ES) Random Walk. The GOES ES is used by the AOCS to observe satellite roll and pitch attitude. Yaw is controlled by the quarter-orbit roll-yaw coupling of a momentum biased attitude control system. The ES scans a pair of IR detectors across the Earth disk to measure four semi-chord lengths (NE horizon-to-midscan, midscan-to-NW horizon, SE horizon-to-midscan and midscan-to-SW horizon). The horizon crossings for the Northern and Southern chords occur in the mid-latitudes. The measured locations for the NE, NW, SE, and SW horizons are susceptible to seasonal IR radiance gradients, the presence of cold clouds near the limb, and by stratospheric waves. These influences cause the roll and pitch pointing of the GOES spacecraft to perform a slow random walk which frequently requires an SSAA to correct.

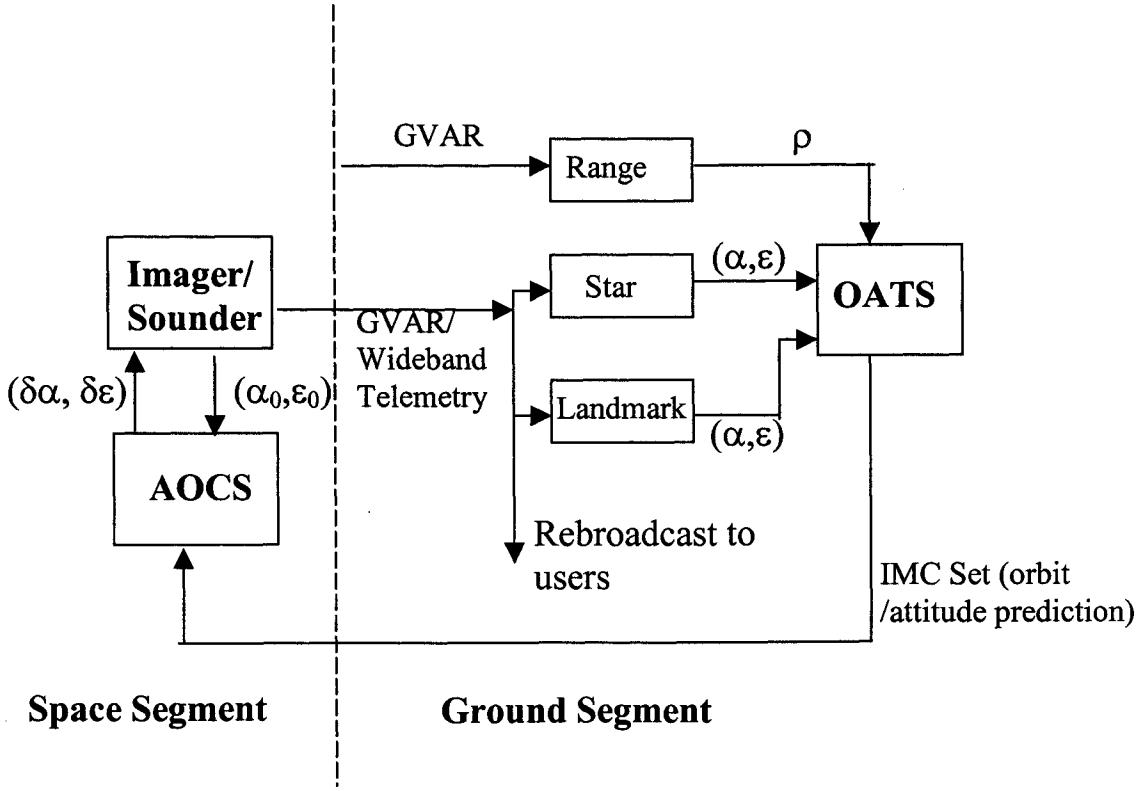


Figure 1. Ground and Space segments for GOES I-M INR System

CHARACTERIZATION OF INR ERROR SOURCES

To capture the errors mentioned in the previous section, operational observations from 1999 for GOES-8 and 10 are used. The process consists of the following steps:

- using ranges, landmarks, and stars, a daily estimation is performed that solves for orbit and Imager attitude. For stars, the quantity $\hat{y} = \begin{bmatrix} \alpha - \alpha_0 \\ \varepsilon - \varepsilon_0 \end{bmatrix}$ is computed for each star measurement, where (α,ε) are the measured scan angles and (α_0,ε_0) are the ideal scan angles based on the true orbit plane. The vector y represents the total attitude signature.
- The vector y is then recast into instrument attitude, which is described in next section.
- The attitude time series are then described by a Fourier series; there is a unique Fourier series for each day of the data span. The Fourier series are averaged together to reduce the effects of non-repeatable errors.

Attitude and Instrument Misalignment Estimation

Imager star measurements consist of a pair of scan angles. The attitude is estimated when there are at least 3 stars within a 2.25 minute window, with the assumption that the attitude does not change significantly during the window. The observation model for stars can be expressed in terms of perturbations from the ideal scan angles by three euler angles (ϕ, θ, ψ) and the two internal misalignments (ϕ_{ma}, θ_{ma}).

$$\alpha = \alpha_0 + \delta\alpha(\alpha_0,\varepsilon_0,\phi,\theta,\psi) + \Delta\alpha_{ma}(\alpha_0,\varepsilon_0,\phi_{ma},\theta_{ma}) \qquad (1)$$

$$\varepsilon = \varepsilon_0 + \delta\varepsilon(\alpha_0,\varepsilon_0,\phi,\theta,\psi) + \Delta\varepsilon_{ma}(\alpha_0,\varepsilon_0,\phi_{ma},\theta_{ma}) \qquad (2)$$

The ideal scan angles, at time t_i , depend on the spacecraft orbit and the star's inertial direction vector.

$$\hat{s}_{ocs} = \begin{bmatrix} \sin\alpha_0^i \\ -\cos\alpha_0^i \sin\varepsilon_0^i \\ \cos\alpha_0^i \cos\varepsilon_0^i \end{bmatrix} = M_{ocs}\hat{s}_{eci} = \begin{bmatrix} (\vec{R}_i x \vec{V}_i) x \vec{R}_i / |(\vec{R}_i x \vec{V}_i) x \vec{R}_i| \\ \vec{R}_i x \vec{V}_i / |\vec{R}_i x \vec{V}_i| \\ -\vec{R}_i / |\vec{R}_i| \end{bmatrix} \hat{s}_{eci} \qquad (3)$$

where the Orbital Coordinate System (OCS) is defined by the spacecraft's position and velocity.

The sensitivity matrix, h_i, with respect to the five attitude angles can be written as:

$$h_i = \begin{bmatrix} \frac{\partial}{\partial \hat{x}} \alpha_i \\ \frac{\partial}{\partial \hat{x}} \varepsilon_i \end{bmatrix} = \begin{bmatrix} 0 & -\cos\varepsilon_0^i & -\sin\varepsilon_0^i & -a*\sin\varepsilon_0^i & 0 \\ -1 & -\sin\varepsilon_0^i \tan\alpha_0^i & \cos\varepsilon_0^i \tan\alpha_0^i \left(1 - \frac{\cos\varepsilon_0^i}{\cos\alpha_0^i}\right) & a\frac{\sin\varepsilon_0^i}{\cos\alpha_0^i}(1 + a\sin\alpha_0^i) \end{bmatrix} \qquad (4)$$

with $\hat{x} = [\phi \quad \theta \quad \psi \quad \phi_{ma} \quad \theta_{ma}]^T$

and a = 1 for upright spacecraft and –1 for inverted. For a given star window, the instrument attitude is obtained by applying a linear estimation method to the following:

$$\hat{x} = (H^T \cdot H)^{-1} \bullet H^T \cdot \hat{y} \qquad (5)$$

with

$$H = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \quad \text{and} \quad \hat{y} = \begin{bmatrix} \alpha_1 - \alpha_0^1 \\ \varepsilon_1 - \varepsilon_0^1 \\ \\ \vdots \\ \\ \vdots \\ \alpha_n - \alpha_0^n \\ \varepsilon_n - \varepsilon_0^n \\ \\ \end{bmatrix}$$

where n is the number of stars in a window and must be at least 3.

Repeatable Attitude Profiles

The attitude time series were separated into repeatable and nonrepeatable components. The repeatable portion was determined by using a symmetric averaging technique that was applied to a 14-day sliding window of attitude solutions. The daily attitude solutions were represented typically by 15^{th} order Fourier series for ϕ, θ, and ψ outside of days with eclipse and 8^{th} order inside of eclipse; in both cases a 4^{th} order fit was used for the misalignments.

GOES-East and West repeatable profiles are shown in Figures 2 and 3 for the Winter and Summer seasons, respectively. The average profiles are over-plotted from noon to noon spacecraft local time. The GOES-West spacecraft is inverted, with the instrument coolers oriented towards the earth's south pole. An inverse relationship can be observed in the roll, pitch, and misalignments. The yaw profiles are related by a time shift of about 10 hours. The profiles presented in Figures 2 and 3 are in agreement with those found by an independent method that averaged operational IMC sets over 14-day periods[2].

Non-repeatable Attitude Characterization

The non-repeatable attitude can be obtained by removing the repeatable attitude from the original time series. This method is illustrated in Figure 4. A 3-day snapshot of the time series for roll is shown for GOES-East and West at the Winter and Summer Solstices. The average roll profile for the period is also shown and the non-repeatable roll. The roll for GOES-East is well determined at Winter solstice since the scan mirror and coolers are shielded from the sun by the spacecraft bus. Although there are still discontinuities present at the Winter solstice, they are not as pronounced as at Summer solstice. The discontinuities are due to a lack of star measurements around spacecraft midnight; in some cases there are sufficient star windows around spacecraft midnight but not containing the requisite 3 stars to solve for the five attitude angles. An interesting feature of the GOES-West roll profile is the smoothness and lack of discontinuities. In part this is due to smaller peak-peak excursions around midnight.

An attempt is made to establish non-repeatable attitude correlations between GOES-East and West. A data arc of 40 days for roll and pitch is shown in Figure 5. While a visual correlation can be detected for the first 20 days, this trend did not continue for other periods analyzed in 1999. The intent of Figure 5 is to show the structure and random walk of the non-repeatable roll and pitch.

The distributions of the non-repeatable attitude are shown in Figure 6 and over-plotted with a normal distribution. Yaw for both spacecraft (and also Goes-East pitch) correspond more closely to the normal distribution. Yaw appears to be driven more so by the white noise processes of the spacecraft control system and influenced very little by the ES random walk; this is also shown in Figure 7. The statistics shown for GOES-East are in agreement with those presented in reference [1].

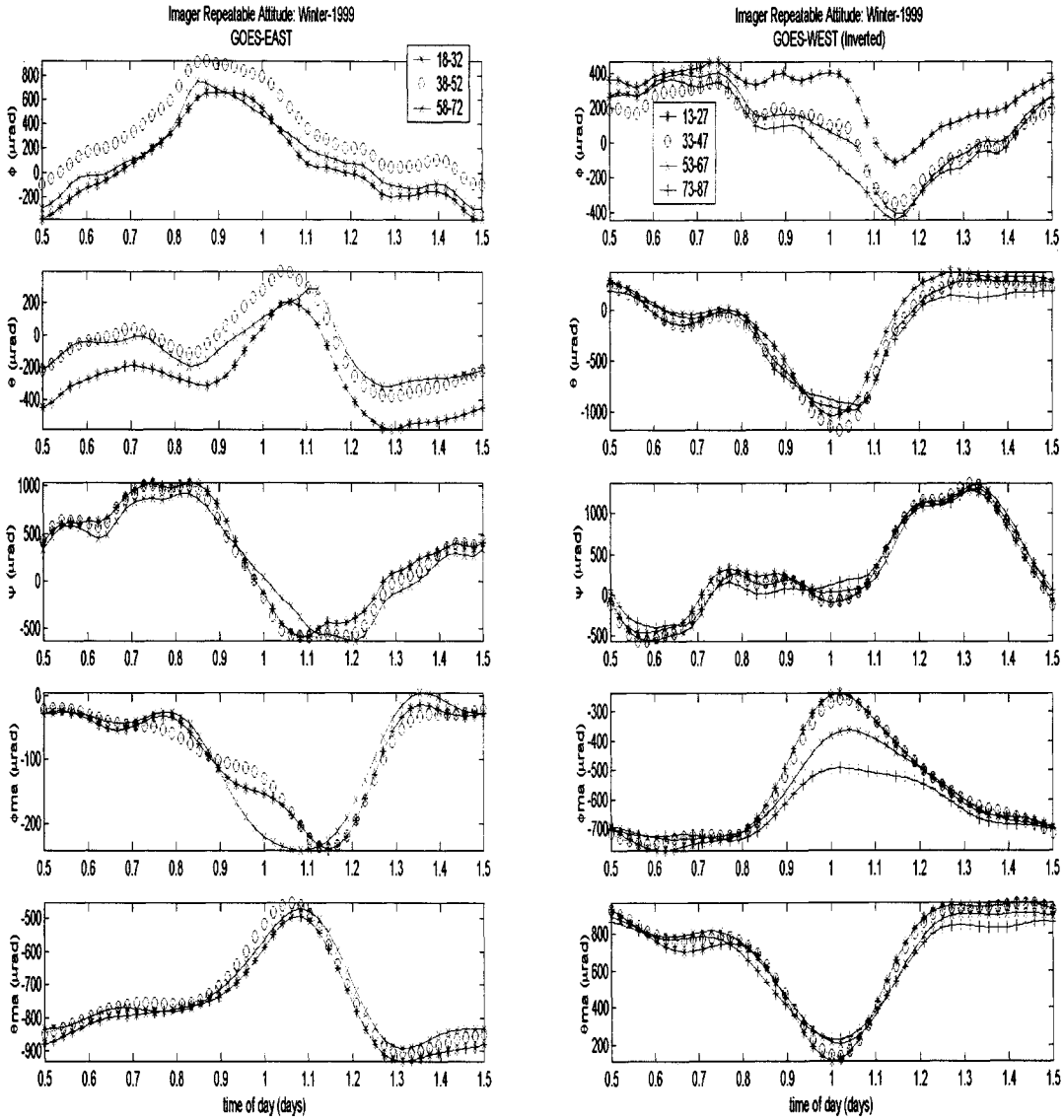


Figure 2. Winter Repeatable Attitude Profiles

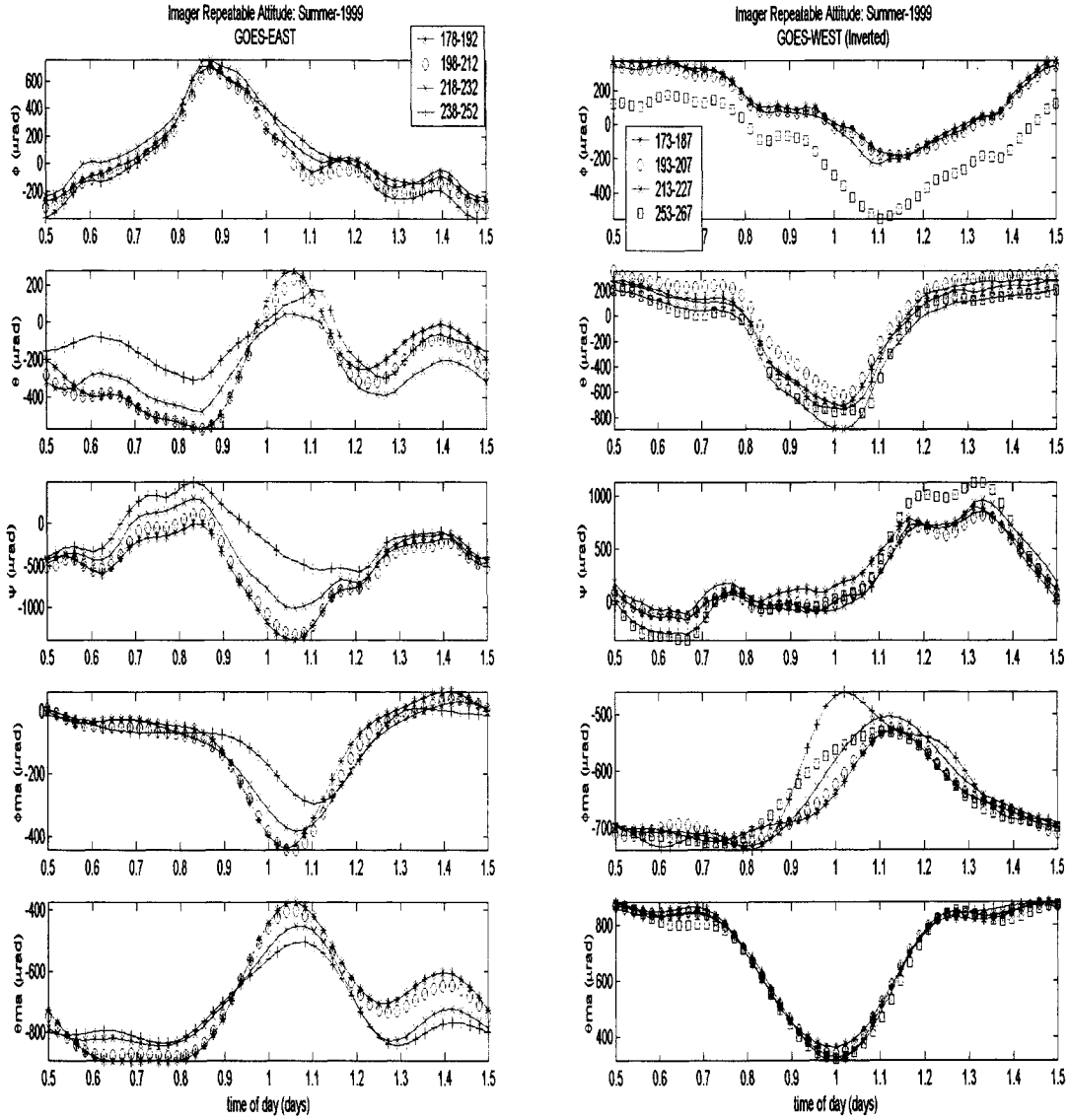


Figure 3. Summer Repeatable Attitude Profiles

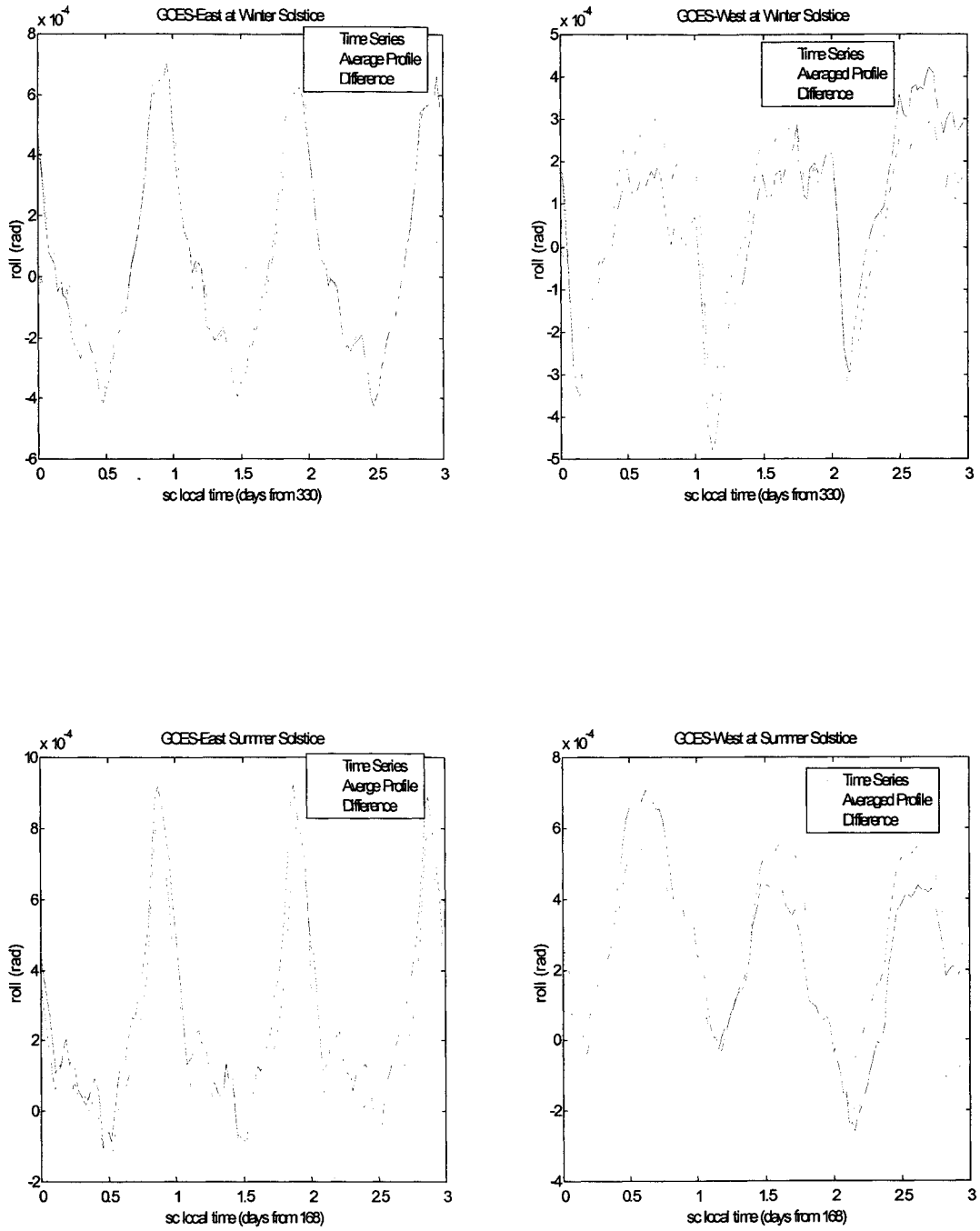


Figure 4. Average roll comparisons with actual roll time series

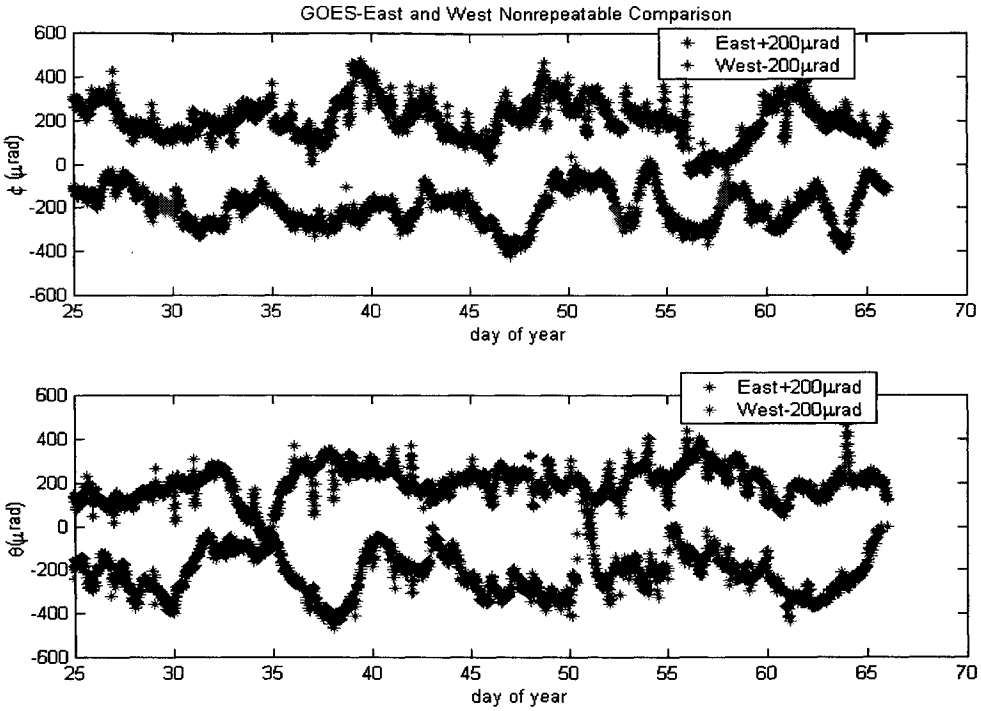


Figure 5. Roll and pitch non-repeatable time series comparisons

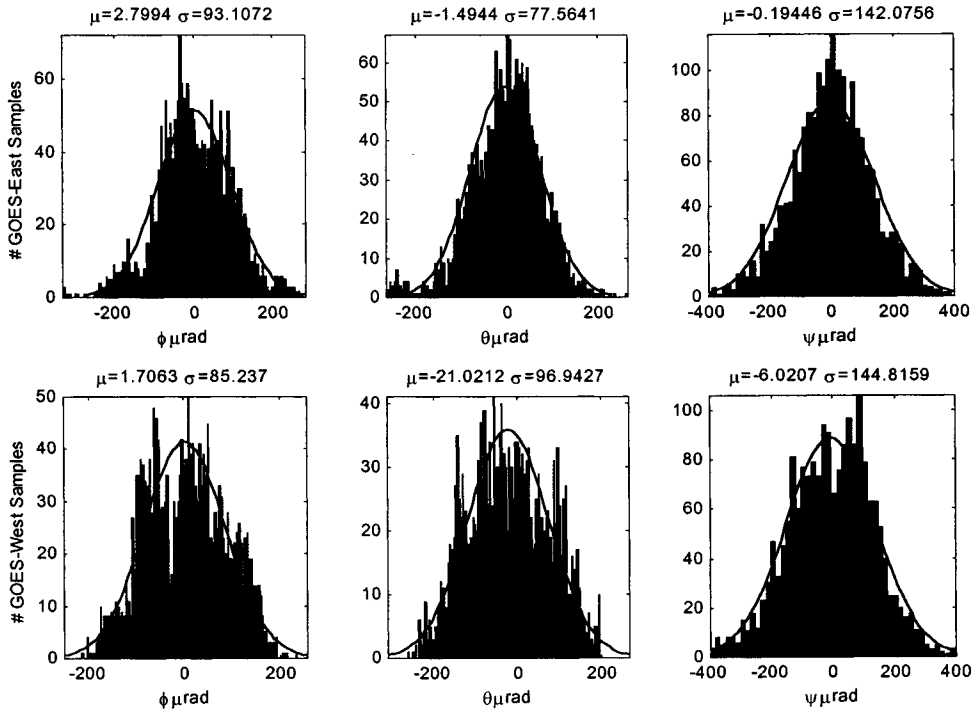


Figure 6. Histograms of time series compared with normal distribution

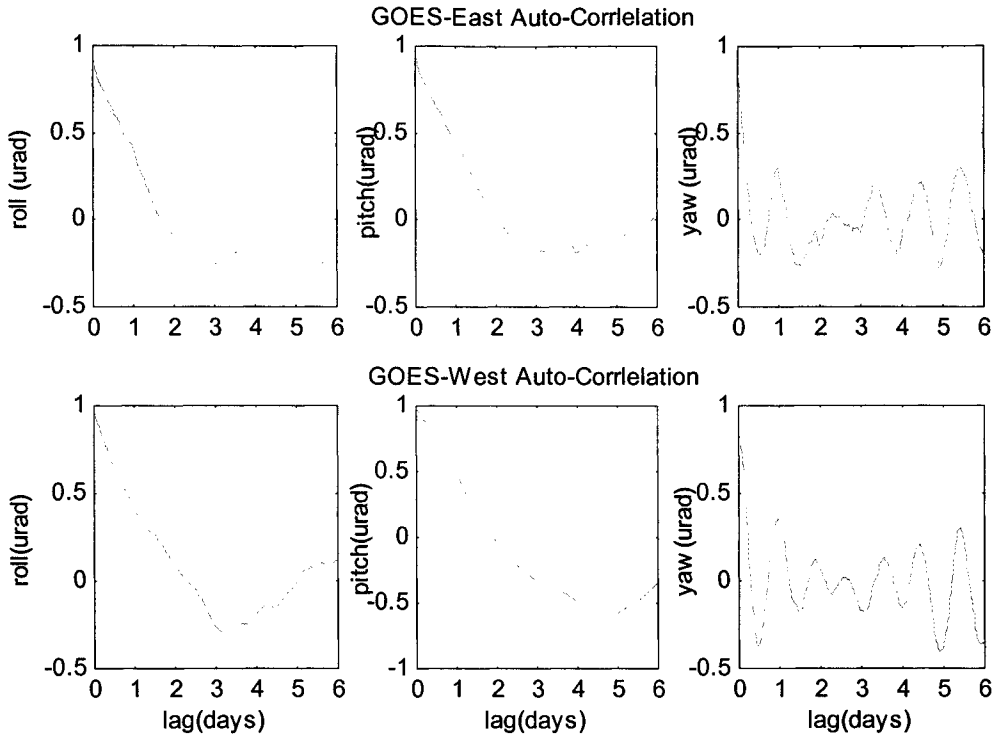


Figure 7. Correlations for non-repeatable attitude

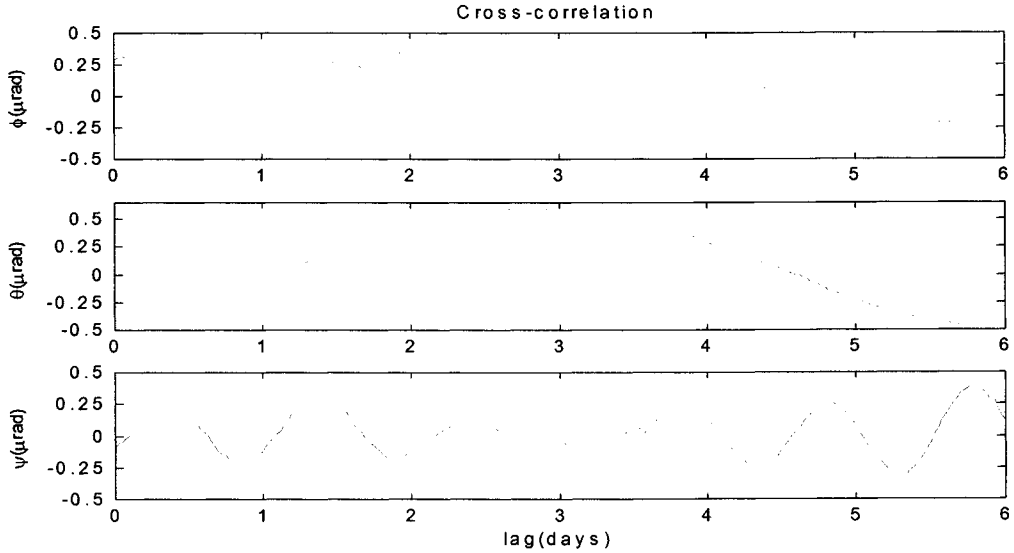


Figure 8. GOES-East/West cross-correlations

Auto-correlation coefficients are shown in Figure 7. The roll and pitch auto-correlation in Figure 7 can be modeled very well by a low pass filter for lags of a day or less, which is the period of interest for INR. For GOES-East, the filter's time constants for roll and pitch are 1.12 and 1.45 days, respectively. For GOES-West, the roll and pitch time constants are 1.15 and 1.35 days, respectively.

The cross-correlation coefficients are shown in Figure 7. A 25% correlation is observed between GOES-East and West. A correlation of this size is not insignificant since the longitude separation is 60° and the orbital inclination difference is 0.25° for most of 1999. But two of the four ES chords are spatially co-located and the errors can be influenced by radiance gradients and cold clouds.

CONCLUSIONS

The analysis presented in this paper has been a follow-on to the work presented in reference [1]. The objective of this work has been to develop accurate attitude error model characteristics that can be used in a high fidelity simulation of the INR system described in Figure 1. Such a simulation will be of interest to the GOES I-M flight operation teams as an important tool in helping to characterize the performance of operational spacecraft.

The repeatable and non-repeatable models developed here can also be used in simulations of the MTSAT-2 system, which will employ an Imager and a type-II ES similar to GOES I-M.

REFERENCES

[1]Carr, J. , "Long Term Stability of GOES-8 and 9 Attitude Control", International Symposium on Optical Science, Engineering, and Instrumentation, SPIE, Denver, Co, August 1996.

[2]Harter, P. et al, "Validating GOES Instrument Thermal Deformations", to be published in the 2001 Flight Mechanics Symposium Proceedings.

IN-SPACE CALIBRATION OF A GYRO QUADRUPLET

Itzhack Y. Bar-Itzhack[1] and Richard R. Harman[2]
Flight Dynamics Analysis Branch, Code 572
Guidance Navigation and Control Center
NASA-Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

This work presents a new approach to gyro calibration where, in addition to being used for computing attitude that is needed in the calibration process, the gyro outputs are also used as measurements in a Kalman filter. This work also presents an algorithm for calibrating a quadruplet rather than the customary triad gyro set. In particular, a new misalignment error model is derived for this case. The new calibration algorithm is applied to the EOS-AQUA satellite gyros. The effectiveness of the new algorithm is demonstrated through simulations.

INTRODUCTION

Gyro calibration as well as calibration of other instruments includes two stages. In the first stage the instrument error parameters are estimated. During the second stage those errors are continuously removed from the gyro readings. In the classical approach to gyro calibration, the gyro outputs are used to maintain or compute body orientation rather than being used as measurements in the context of filtering. In inertial navigation, for example (ref. 1), gyro errors cause erroneous computation of velocity and position, and then when the latter are compared to measured velocity and position, a great portion of the computed velocity and position errors can be determined. The latter errors are then fed into a Kalman filter (KF) that uses the Inertial Navigation System (INS) error model to infer on the gyro errors. Similarly, when applying the classical approach to spacecraft (SC) attitude determination, the gyro outputs are used to compute the attitude and then attitude measurements (refs. 2, 3) are used to determine the attitude errors, which again using a KF, indicates what the gyro errors are.

In the approach adopted in this work, the gyro outputs are used as angular rate measurements and are compared to estimated angular rate measurements. However, this approach requires the knowledge of the angular rate. In the past (ref. 4), the estimated angular rate was computed in a rather simplistic way assuming basically that the rate was constant. In the present work, the estimated angular rate is derived using a KF whose input can be any kind of attitude measurement; therefore, the angular rate experienced by the SC can be continuously changing, and yet a good estimate of the rate, necessary for calibration, can be obtained.

The calibration algorithm presented in this work was derived for a set of quadruplet gyros. This required the derivation of a new error model, particularly for the gyro misalignments. The new calibration algorithm was applied to the gyro package of the EOS-AQUA satellite. The latter consists of four gyros, which are given the task of measuring the three components of the SC angular velocity vector resolved in the body Cartesian coordinates.

In the next section the gyro error model is derived. The section that follows presents an algorithm for computing the calibration parameters when the rate is known, and then in the section that follows we

---

[1] Sophie and William Shamban Professor of Aerospace Engineering.
On sabbatical leave from the Faculty of Aerospace Engineering,
Technion-Israel Institute of Technology. Member Technion Asher
Space Research Institute. NRC NASA Resident Research Associate.
Tel: (301) 286-9216, Fax: (301) 286-036, Email: ibaritz@pop500.gsfc.nasa.gov

[2] Aerospace Engineer. Tel: (301) 286-5125, Fax: (301) 286-036,
Email: richard.r.harman.1@gsfc.nasa.gov

present the same when the rates are unknown. In the following section we present the compensation procedure that needs to take place to complete the calibration process, and in the subsequent section we present simulation results. Finally, in the last section, the conclusions are presented.

GYRO ERROR MODEL

The gyro errors that are considered in this work are: misalignment, scale factor error, and bias (constant drift rate). The gyro error model is basically a linear model, which associates small error sources to the gyro outputs. *Due to the linearity of the model we can compute the contribution of each error source independently, and then sum up all the contributions into one linear model.*

We start the description of the error model, by deriving the expression for the gyro misalignments.

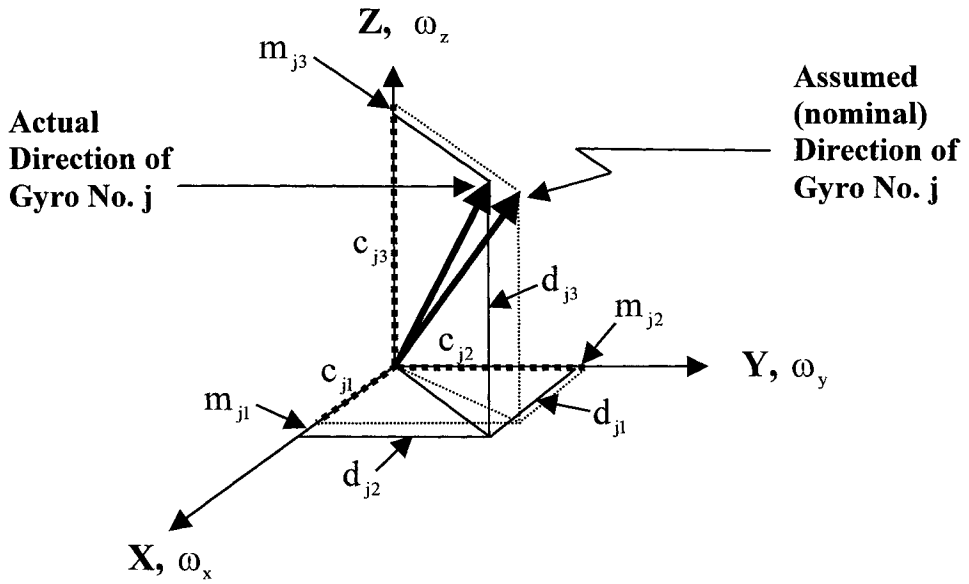


Fig. 1: The geometry of the Assumed and the Actual Direction of the Gyro Input Axis

Misalignment Model

The assumed direction of the sensitive axis of gyro j, which is one of the four gyros, is presented in Fig. 1 where the body coordinate axes are also presented and are denoted by **X**, **Y**, and **Z**. The orientation of this gyro is expressed by a vector of unit length in the direction of the gyro sensitive axis. The direction of this unit vector in the body coordinates is expressed by its three direction cosines, which are identical to its components when the unit vector is resolved in the body coordinates. These components are $c_{j1}, c_{j2}, \text{and } c_{j3}$. Being direction cosines, or equivalently, components of a unit vector, the sum of their squares adds up to 1; that is,

$$c_{j1}^2 + c_{j2}^2 + c_{j3}^2 = 1 \qquad (1)$$

The rate that this gyro reads is the projection of the angular velocity vector on this unit vector. If we express the angular velocity vector in the body coordinates, where its components are $\omega_x, \omega_y, \text{and } \omega_z$, then this projection is given by

$$\mathbf{1}_j \cdot \boldsymbol{\omega} = c_{j1}\omega_x + c_{j2}\omega_y + c_{j3}\omega_z \qquad (2.a)$$

where $\mathbf{1}_j$ is the unit vector along the $j^{\underline{th}}$ gyro sensitive axis, and $\boldsymbol{\omega}$ is the angular rate vector. The nominal (error-less) reading of this gyro is then

$$G_{jn} = c_{j1}\omega_x + c_{j2}\omega_y + c_{j3}\omega_z = \begin{bmatrix} c_{j1} & c_{j2} & c_{j3} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{2.b}$$

where the subscript *n* denotes the nominal or design value. Combining all four gyros we obtain

$$\begin{bmatrix} G_{1n} \\ G_{2n} \\ G_{3n} \\ G_{4n} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \\ c_{41} & c_{42} & c_{43} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{3.a}$$

Define

$$\mathbf{G}_n^T = \begin{bmatrix} G_{1n} & G_{2n} & G_{3n} & G_{4n} \end{bmatrix} \tag{3.b}$$

and

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \\ c_{41} & c_{42} & c_{43} \end{bmatrix} \tag{3.c}$$

Eq. (3.a) can be written as

$$\mathbf{G}_n = C\boldsymbol{\omega} \tag{4}$$

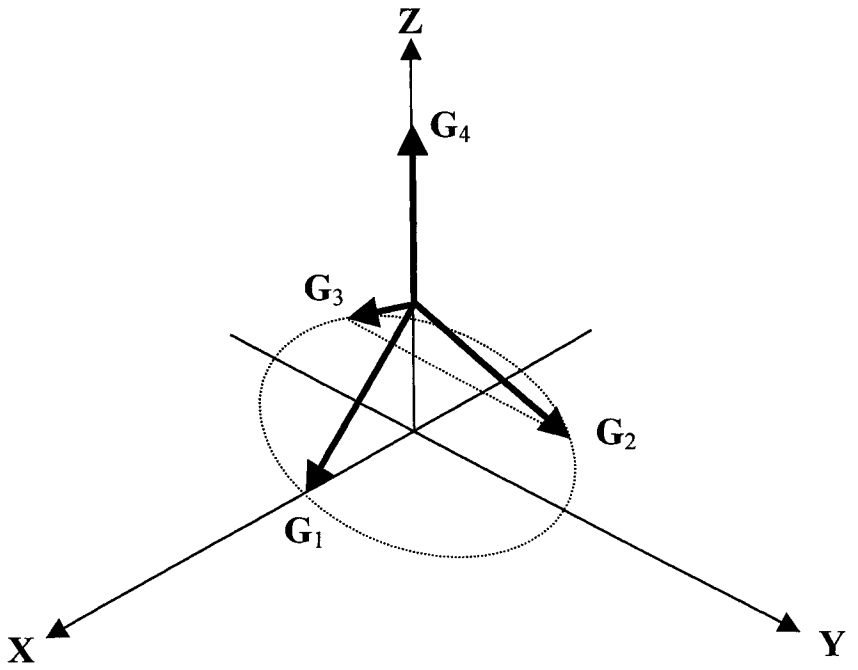


Fig. 2: The Gyro Configuration in the EOS-AQUA Satellite.

In the EOS-AQUA satellite the gyro configuration is as shown in Fig. 2 where, as mentioned before, **X**, **Y**, and **Z** are the axes of the body frame. The C matrix in this case is

$$C = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 & -\sqrt{\frac{1}{3}} \\ -\sqrt{\frac{1}{6}} & \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{3}} \\ -\sqrt{\frac{1}{6}} & -\sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{3}} \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Note that the vector described by each row is of unit length like it should be (see Eq. 1).

Due to misalignment, the sensitive axis of each gyro may actually point at a slightly different direction than the assumed one. This is illustrated in Fig. 1 where the components of this direction (which is still a unit vector) are $d_{j1}, d_{j2}, \text{and } d_{j3}$ respectively. Following the steps that led to the development of the nominal gyro reading presented in Eq. (2.b), the actual gyro reading is found to be

$$G_{ja} = d_{j1}\omega_x + d_{j2}\omega_y + d_{j3}\omega_z = \begin{bmatrix} d_{j1} & d_{j2} & d_{j3} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{6}$$

where the subscript *a* denotes an actual value. The difference, ΔG_j, between the reading of the $j^{\underline{th}}$ gyro and its assumed nominal reading is computable using Eqs. (2.b) and (6), as follows;

$$\Delta G_j^m = G_{ja} - G_{jn} = \begin{bmatrix} d_{j1} - c_{j1} & d_{j2} - c_{j2} & d_{j3} - c_{j3} \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{7}$$

where the superscript *m* denotes the fact that the error is due to misalignment. We denote by m_{ji} the differences $d_{j1} - c_{j1}$, $d_{j2} - c_{j2}$, and $d_{j3} - c_{j3}$ as follows

$$m_{j1} = d_{j1} - c_{j1} \quad (8.a) \qquad m_{j2} = d_{j2} - c_{j2} \quad (8.b) \qquad m_{j3} = d_{j3} - c_{j3} \quad (8.c)$$

Using Eqs. (8) we can write Eq. (7), as follows;

$$\Delta G_j^m = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix} \begin{bmatrix} m_{j1} \\ m_{j2} \\ m_{j3} \end{bmatrix} \tag{9}$$

The m_{ji} differences are shown in Fig. 1. Actually only two of the m_{ji} of each gyro are independent. This results from the fact that the nominal as well as the actual directions of the gyros are given by vectors of unit length. For a reason that will be clear later, let us choose to present the third component of m_j by the

first two; that is, we express m_{j3}, the misalignment along the sensitive axis, by m_{j1} and m_{j2}. Since similarly to Eq. (1), it is also true that

$$d_{j1}^2 + d_{j2}^2 + d_{j3}^2 = 1 \qquad (10)$$

then using this relation and Eq. (8.c) we can write

$$m_{j3} = \sqrt{1 - d_{j1}^2 - d_{j2}^2} - \sqrt{1 - c_{j1}^2 - c_{j2}^2} \qquad (11)$$

For the case described by the fourth gyro (G_4 in fig. 3) the nominal direction of the gyro sensitive axis is along the Z axis; therefore, $c_{41} = c_{42} = 0$, then from Eqs. (8.a) and (8.b)

$$m_{41} = d_{41} \qquad (12.a) \qquad\qquad m_{42} = d_{42} \qquad (12.b)$$

and from Eq. (11)

$$m_{43} = \sqrt{1 - d_{41}^2 - d_{42}^2} - 1 \qquad (13)$$

In the case where the misalignments are small, d_{41}^2 and d_{42}^2 are small too. Therefore we can expand the square root function of Eq. (11) in a Taylor series, as follows;

$$\sqrt{1 - d_{41}^2 - d_{42}^2} = 1 - \tfrac{1}{2} d_{41}^2 - \tfrac{1}{2} d_{42}^2 \qquad (14)$$

(Note that the linear term of the series vanishes). Substituting of the last equation into Eq. (13) yields

$$m_{43} = -\tfrac{1}{2} d_{41}^2 - \tfrac{1}{2} d_{42}^2 = -\tfrac{1}{2} m_{41}^2 - \tfrac{1}{2} m_{42}^2 \qquad (15)$$

When d_{41} and d_{42} are indeed small, such as this case, then m_{43} is negligible with respect to m_{41} and m_{42}. Then using Eqs. (12) we can write

$$\begin{bmatrix} m_{41} \\ m_{42} \\ m_{43} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d_{41} \\ d_{42} \end{bmatrix} \qquad (16)$$

It is the choice to express the component of m_j along the gyro sensitive axis (in this case m_{43}) that enables its elimination.

For gyros whose sensitive axes are not aligned along one of the body axes the computation is more elaborate. Consider for example G_2, the second gyro of the EOS-AQUA satellite. In order to define its misalignment in the body coordinates let us define a coordinate system in which the gyro sensitive axis is *nominally* aligned along one of its axes. Such a system (**X"**, **Y"**, **Z"**) is presented in Fig. 3, where the sensitive axis of the G_2 gyro is aligned along the system **Y"** axis. Following the preceding development for the G_4 gyro we conclude that

$$m''_{21} = d''_{21} \qquad (16.a) \qquad\qquad m''_{23} = d''_{23} \qquad (17.b)$$

$$m''_{22} = -\tfrac{1}{2}d''^{\,2}_{21} - \tfrac{1}{2}d''^{\,2}_{23} = -\tfrac{1}{2}m''^{\,2}_{21} - \tfrac{1}{2}m''^{\,2}_{23} \qquad (17.c)$$

(The " sign denotes the fact that the values are expressed in the **X"**, **Y"**, **Z"** coordinate system.) Here too, the misalignment along the sensitive axis, m''_{22}, is normally negligible. In order to compute the misalignment error in the gyro reading we have to use Eq. (9) where the angular rate vector is transformed to the double prime coordinate system and the misalignment parameters are those given in Eqs. (16.a and b). As shown in Fig. 3 the transformation from the body to the double prime coordinates is performed by

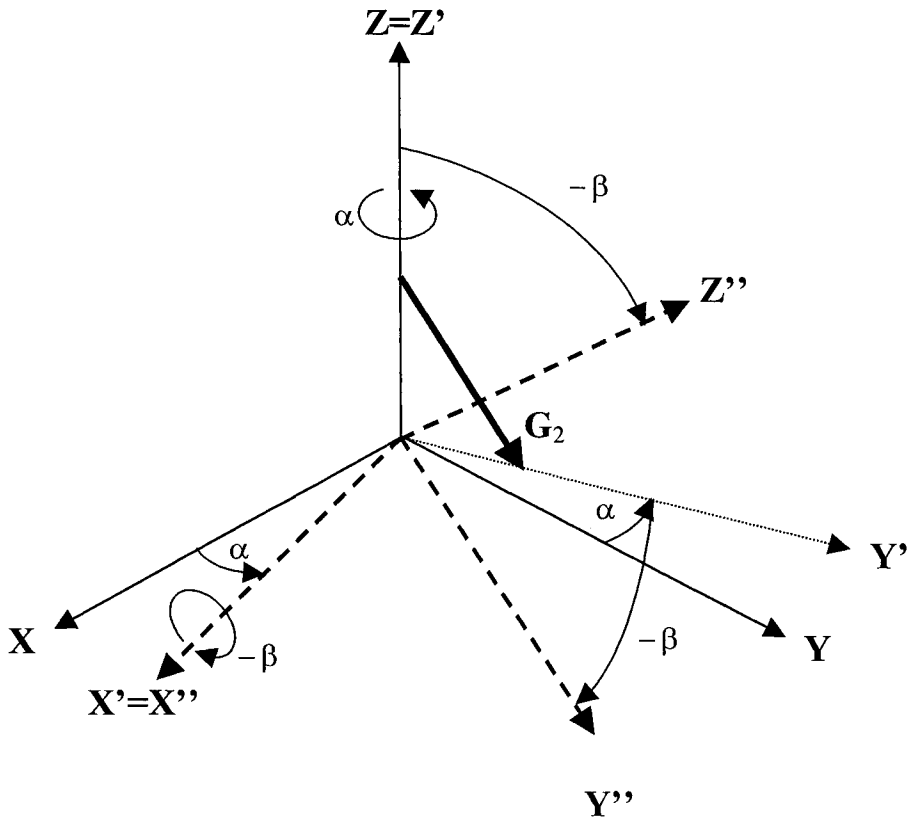


Fig. 3: The Transformation from the G_2 Body to Gyro Coordinates.

two rotations. The first rotation is by an angle α about the **Z** axis, and the second is by an angle $-\beta$ about the **X'** axis. The resulting transformation matrix from the body to the G_2 coordinates is therefore

$$R_2^b = \begin{bmatrix} c\alpha & s\alpha & 0 \\ -s\alpha \cdot c\beta & c\alpha \cdot c\beta & -s\beta \\ -s\alpha \cdot s\beta & c\alpha \cdot s\beta & c\beta \end{bmatrix} \qquad (18)$$

and

$$\boldsymbol{\omega}_2 = R_2^b \boldsymbol{\omega}_b \qquad (19)$$

then following Eq. (9)

$$\Delta G_2^m = [R_2^b \boldsymbol{\omega}_b]^T \mathbf{d}_{2*}'' = \boldsymbol{\omega}_b^T R_b^2 \mathbf{d}_{2*}'' \tag{20}$$

where

$$[\mathbf{d}_{2*}'']^T = \begin{bmatrix} d_{21}'' & 0 & d_{23}'' \end{bmatrix} \tag{21}$$

It is easy to see that

$$R_b^2 \mathbf{d}_{2*}'' = \begin{bmatrix} c\alpha & -s\alpha \cdot s\beta \\ s\alpha & c\alpha \cdot s\beta \\ 0 & c\beta \end{bmatrix} \begin{bmatrix} d_{21}'' \\ d_{23}'' \end{bmatrix} \tag{22}$$

Define

$$E_2 = \begin{bmatrix} c\alpha & -s\alpha \cdot s\beta \\ s\alpha & c\alpha \cdot s\beta \\ 0 & c\beta \end{bmatrix} \quad (23.a) \qquad \mathbf{d}_2 = \begin{bmatrix} d_{21}'' \\ d_{23}'' \end{bmatrix} \quad (23.b)$$

then Eq. (20) can be written as

$$\Delta G_2^m = \boldsymbol{\omega}_b^T E_2 \mathbf{d}_2 \tag{24}$$

To evaluate E_2, we need to compute the angles α and β. For this we turn to Fig. 4 where these angles

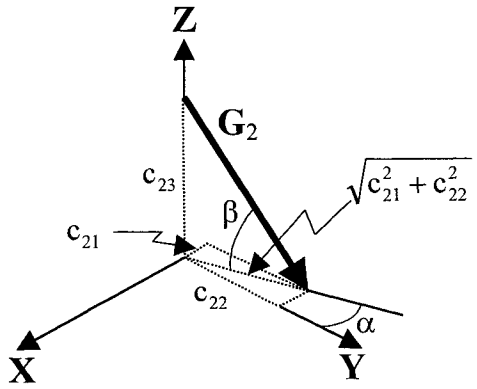


Fig. 4: Definition of the Rotation Angles α and β.

are defined in the projection of the G_2 direction on the body axes. From this figure we conclude that

$$\alpha = \tan^{-1}\left(\left|\frac{c_{21}}{c_{22}}\right|\right) \tag{25.a}$$

$$\beta = \cos^{-1}\left(\frac{\sqrt{c_{21}^2 + c_{22}^2}}{1}\right) = \cos^{-1}\left(\sqrt{c_{21}^2 + c_{22}^2}\right) \tag{25.b}$$

Using the EOS-AQUA satellite values (see Eqs. (4) and (5)) we obtain

$$\alpha = \tan^{-1}\left(\left|\frac{c_{21}}{c_{22}}\right|\right) = \tan^{-1}\left(\frac{\sqrt{\frac{1}{6}}}{\sqrt{\frac{1}{2}}}\right) = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = 30^\circ \qquad (25.c)$$

$$\beta = \cos^{-1}\left(\sqrt{c_{21}^2 + c_{22}^2}\right) = \cos^{-1}\left(\sqrt{\frac{1}{6} + \frac{1}{2}}\right) = \cos^{-1}\left(\sqrt{\frac{2}{3}}\right) = 35.26^\circ \qquad (25.d)$$

Due to the symmetry between the positioning of the G_2 and the G_3 gyros, it is easy to see that when considering the G_3, the first rotation is about the new **Z** axis by the angle $\pi - \alpha$ and then about the **X'** axis by the angle $-\beta$. Therefore

$$R_3^b = \begin{bmatrix} -c\alpha & s\alpha & 0 \\ -s\alpha \cdot c\beta & -c\alpha \cdot c\beta & -s\beta \\ -s\alpha \cdot s\beta & -c\alpha \cdot s\beta & c\beta \end{bmatrix} \qquad (26)$$

and similarly to Eq. (19) for this transformation we obtain

$$\boldsymbol{\omega}_3 = R_3^b \boldsymbol{\omega}_b \qquad (27)$$

then following Eq. (20)

$$\Delta G_3^m = [R_3^b \boldsymbol{\omega}_b]^T \mathbf{d}_{3*}'' = \boldsymbol{\omega}_b^T R_b^3 \mathbf{d}_{3*}'' \qquad (28)$$

where

$$[\mathbf{d}_{3*}'']^T = \begin{bmatrix} d_{31}'' & 0 & d_{33}'' \end{bmatrix} \qquad (29)$$

It is easy to see that

$$R_b^3 \mathbf{d}_{3*}'' = \begin{bmatrix} -c\alpha & -s\alpha \cdot s\beta \\ s\alpha & -c\alpha \cdot s\beta \\ 0 & c\beta \end{bmatrix} \begin{bmatrix} d_{31}'' \\ d_{33}'' \end{bmatrix} \qquad (30)$$

Define

$$E_3 = \begin{bmatrix} -c\alpha & -s\alpha \cdot s\beta \\ s\alpha & -c\alpha \cdot s\beta \\ 0 & c\beta \end{bmatrix} \quad (31.a) \qquad \mathbf{d}_3 = \begin{bmatrix} d_{31}'' \\ d_{33}'' \end{bmatrix} \quad (31.b)$$

then Eq. (28) can be written as

$$\Delta G_3^m = \boldsymbol{\omega}_b^T E_3 \mathbf{d}_3 \qquad (32)$$

For the G_1 gyro we have only one transformation, which brings the body **X**-axis into coincidence with the G_1 gyro sensitive axis (see Fig. 2). It is about the **Y**-axis by an angle which we denote by γ. For this gyro we have then

$$R_1^b = \begin{bmatrix} c\gamma & 0 & -s\gamma \\ 0 & 1 & 0 \\ s\gamma & 0 & c\gamma \end{bmatrix} \qquad (33)$$

and the angular rate in coordinate system 1 is then

$$\boldsymbol{\omega}_1 = R_1^b \boldsymbol{\omega}_b \tag{34}$$

and

$$\Delta G_1^m = [R_1^b \boldsymbol{\omega}_b]^T \mathbf{d}_{1*}'' = \boldsymbol{\omega}_b^T R_b^1 \mathbf{d}_{1*}' \tag{35}$$

where

$$[\mathbf{d}_{1*}']^T = \begin{bmatrix} 0 & d_{12}' & d_{23}' \end{bmatrix} \tag{36}$$

We denote the misalignment parameters of this gyro by a single prime because it takes only one rotation (to a single prime system) to align the coordinate axis with the sensitive axis of the G_1 gyro. Note that here the misalignments that are not negligible are d_{12}' and d_{13}'. It is easy to see that

$$R_b^1 \mathbf{d}_{1*}'' = \begin{bmatrix} 0 & s\gamma \\ 1 & 0 \\ 0 & c\gamma \end{bmatrix} \begin{bmatrix} d_{12}' \\ d_{13}' \end{bmatrix} \tag{37}$$

Define

$$E_1 = \begin{bmatrix} 0 & s\gamma \\ 1 & 0 \\ 0 & c\gamma \end{bmatrix} \quad (38.a) \qquad \mathbf{d}_1 = \begin{bmatrix} d_{12}'' \\ d_{13}'' \end{bmatrix} \quad (38.b)$$

then Eq. (35) can be written as

$$\Delta G_1^m = \boldsymbol{\omega}_b^T E_1 \mathbf{d}_1 \tag{39}$$

From fig. 2 it is easy to see that the rotation angle, γ, is computable, as follows;

$$\gamma = \sin^{-1}\left(\frac{|c_{13}|}{1}\right) = \sin^{-1}\left(\sqrt{\tfrac{1}{3}}\right) = 35.26^\circ \tag{40}$$

Similarly to the computations carried out for the misalignment errors for gyros 1, 2, and 3, we can write for gyro 4

$$\Delta G_4^m = \boldsymbol{\omega}_b^T E_4 \mathbf{d}_4 \tag{41}$$

where, based on Eq. (16),

$$E_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (42.a) \qquad \mathbf{d}_4 = \begin{bmatrix} d_{41} \\ d_{42} \end{bmatrix} \quad (42.b)$$

Let

$$[\Delta \mathbf{G}^m]^T = \begin{bmatrix} \Delta G_1^m & \Delta G_2^m & \Delta G_3^m & \Delta G_4^m \end{bmatrix} \tag{43.a}$$

$$\Omega^m = \begin{bmatrix} \omega_x & \omega_y & \omega_z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_x & \omega_y & \omega_z & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \omega_x & \omega_y & \omega_z & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \omega_x & \omega_y & \omega_z \end{bmatrix} \tag{43.b}$$

$$E = \begin{bmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 \\ 0 & 0 & E_3 & 0 \\ 0 & 0 & 0 & E_4 \end{bmatrix} \quad (43.c) \qquad \mathbf{d}^T = \begin{bmatrix} \mathbf{d}_1^T & \mathbf{d}_2^T & \mathbf{d}_3^T & \mathbf{d}_4^T \end{bmatrix} \quad (43.d)$$

Then Eqs. (39), (23), (32) and (41) can be unified into the following single equation

$$\Delta \mathbf{G}^m = \Omega^m E \, \mathbf{d} \quad (44)$$

Scale Factor Error Model

As mentioned, another error source that causes the difference between the correct value of the rates and their measurements are the scale factor errors. The error model for the scale factor error is simply

$$\Delta \mathbf{G}^k = \begin{bmatrix} \omega_{G1} k_1 \\ \omega_{G2} k_2 \\ \omega_{G3} k_3 \\ \omega_{G4} k_4 \end{bmatrix} \quad (45)$$

where the subscript *k* denotes the fact that this error is caused by gyro scale factor error, ω_{Gi}, $i = 1 - 4$ is the angular velocity measured by gyro number *i*, and k_i is the scale factor error of that gyro. The actual components ω_{Gi} are obtained by transforming the angular velocity expressed in body coordinates to the actual misaligned gyro sensitive axes using the matrix D; however, because D is unknown to us we use instead the matrix C that transforms the body rate to the nominal gyro axes, and is close enough to D. Thus,

$$\begin{bmatrix} \omega_{G1} \\ \omega_{G2} \\ \omega_{G3} \\ \omega_{G4} \end{bmatrix} = C \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (46)$$

where in the case of EOS-AQUA, C is as given in Eq. (5). Equation (45) can be written as follows

$$\Delta \mathbf{G}^k = \begin{bmatrix} \omega_{G1} & 0 & 0 & 0 \\ 0 & \omega_{G2} & 0 & 0 \\ 0 & 0 & \omega_{G3} & 0 \\ 0 & 0 & 0 & \omega_{G4} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} \quad (47)$$

Define

$$\Omega^k = \begin{bmatrix} \omega_{G1} & 0 & 0 & 0 \\ 0 & \omega_{G2} & 0 & 0 \\ 0 & 0 & \omega_{G3} & 0 \\ 0 & 0 & 0 & \omega_{G4} \end{bmatrix} \quad (48.a)$$

and

$$\mathbf{k}^T = \begin{bmatrix} k_1 & k_2 & k_3 & k_4 \end{bmatrix} \tag{48.b}$$

then Eq. (47) can be written as

$$\Delta\mathbf{G}^k = \Omega^k \mathbf{k} \tag{48.c}$$

Bias Model

The bias error model is quite simple and is given by

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{49}$$

where b_i is the bias of gyro number *i*.

The Augmented Gyro Error Model

The total gyro error is the sum of all the errors discussed before; namely, bias, scale factor and misalignment errors; that is

$$\Delta\mathbf{G} = \Delta\mathbf{G}^m + \Delta\mathbf{G}^k + \mathbf{b} \tag{50.a}$$

or using Eqs. (44) and (48.c)

$$\Delta\mathbf{G} = \Omega^m E\,\mathbf{d} + \Omega^k \mathbf{k} + \mathbf{b} \tag{50.b}$$

The last equation can be written in the following form

$$\mathbf{G}_a - C\boldsymbol{\omega}_r = \begin{bmatrix} \Omega^m E & \Omega^k & I_4 \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{k} \\ \mathbf{b} \end{bmatrix} \tag{50.c}$$

where $\boldsymbol{\omega}_r$ is the reference angular velocity vector. It is the angular velocity, which the SC experiences in reality. As mentioned before, $\mathbf{G}_n$ is the nominal angular velocity measured by the four gyros. The vector of the left hand side of the last equation as well as the matrix on the right hand side are functions of the body angular rate, $\boldsymbol{\omega}_r$. We denote them as follows

$$\mathbf{y}(\boldsymbol{\omega}_r) = \mathbf{G}_a - C\boldsymbol{\omega}_r \tag{51.a}$$

$$H(\boldsymbol{\omega}_r) = \begin{bmatrix} \Omega^m E & \Omega^k & I_4 \end{bmatrix} \tag{51.b}$$

also let

$$\mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{k} \\ \mathbf{b} \end{bmatrix} \tag{51.c}$$

then Eq. (50.c) can be written as

$$\mathbf{y}(\boldsymbol{\omega}_r) = H(\boldsymbol{\omega}_r)\mathbf{x} \tag{51.d}$$

CALIBRATION-PARAMETERS ESTIMATION FOR KNOWN RATE

Our goal now is to estimate $\mathbf{x}$, and for that we need to know the angular rate, which the gyros are set to measure. We distinguish between two major cases. One case is that where $\boldsymbol{\omega}_r$, the reference SC angular velocity, is known, and the other case is that where the rate is not known and has to be evaluated simultaneously with the estimate of $\mathbf{x}$. In the first case we also distinguish between the deterministic and stochastic cases. All these cases are discussed next.

Deterministic Case

When the SC rotates at a certain angular rate and a one-time measurement of the four gyro readings is taken at that time, which we denote by t_k, we obtain one matrix equation, as follows:

$$\mathbf{y}(\boldsymbol{\omega}_{r,k}) = \mathrm{H}(\boldsymbol{\omega}_{r,k})\mathbf{x} \tag{52}$$

where $\boldsymbol{\omega}_{r,k}$ denotes the angular rate at time t_k. This yields four equations for the 16 unknowns of $\mathbf{x}$. If the rate does not change, then more measurements do not change the equations. A change in the angular rate of the SC is needed to generate more equations. It should be noted that even if we have 16 equations, it does not mean that they are all independent and that we can solve for $\mathbf{x}$. We have to design the profile of $\boldsymbol{\omega}_r$ and the times when measurements are taken in such a way that we will be able to find 16 independent equations. Let us denote the 16 independent equations by one matrix equation, as follows:

$$\tilde{\mathbf{y}} = \tilde{\mathrm{H}}\mathbf{x} \tag{53}$$

Because we have 16 independent equations, $\tilde{\mathrm{H}}$ has an inverse; therefore, we can solve for $\mathbf{x}$ using

$$\mathbf{x} = \tilde{\mathrm{H}}^{-1}\tilde{\mathbf{y}} \tag{54}$$

Stochastic Case

In this case we assume that the measurements are contaminated by noise, which is the most likely case. Therefore, the matrix equation that describes this case at time t_k is

$$\mathbf{y}(\boldsymbol{\omega}_{r,k}) = \mathrm{H}(\boldsymbol{\omega}_{r,k})\mathbf{x} + \mathbf{v}_k \tag{55}$$

Even if we find 16 independent equations from measurements done at different time points we still want to use all available measurements and obtain $\mathbf{x}$ as a least squares estimate. We have

$$\begin{bmatrix} \mathbf{y}(\boldsymbol{\omega}_{r,1}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,2}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,3}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,4}) \end{bmatrix} = \begin{bmatrix} \mathrm{H}(\boldsymbol{\omega}_{r,1}) \\ \mathrm{H}(\boldsymbol{\omega}_{r,2}) \\ \mathrm{H}(\boldsymbol{\omega}_{r,3}) \\ \mathrm{H}(\boldsymbol{\omega}_{r,4}) \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \end{bmatrix} \tag{56}$$

Let

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}(\boldsymbol{\omega}_{r,1}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,2}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,3}) \\ \mathbf{y}(\boldsymbol{\omega}_{r,4}) \end{bmatrix} \quad (57.a) \qquad \mathbf{H} = \begin{bmatrix} H(\boldsymbol{\omega}_{r,1}) \\ H(\boldsymbol{\omega}_{r,2}) \\ H(\boldsymbol{\omega}_{r,3}) \\ H(\boldsymbol{\omega}_{r,4}) \end{bmatrix} \quad (57.b)$$

then $\hat{\mathbf{x}}$, the least squares fit to $\mathbf{x}$, is as follows (ref. 5)

$$\hat{\mathbf{x}} = (H^T H)^{-1} H^T \mathbf{Y} \quad (58)$$

The profile of $\boldsymbol{\omega}_r$ has to be chosen in a careful way as to enhance the observability of $\mathbf{x}$.

CALIBRATION-PARAMETERS ESTIMATION FOR UNKNOWN RATE

In this case we have to find the angular rate vector while estimating the calibration parameters. The information that we have is attitude information and gyro measurements. We need the attitude information in order to estimate the angular rate, and we need the gyro measurements, as well as the estimated angular rate, for the calibration process. The attitude information can be supplied in various ways; namely, we may have it in the form of raw vector measurements or we can have it in an already processed form as attitude quaternion for example. The angular rate behaves according to the following SC angular dynamics equation

$$\dot{\boldsymbol{\omega}} = I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times]\boldsymbol{\omega} + I^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \quad (59.a)$$

where I is the SC inertia tensor, $[(I\boldsymbol{\omega} + \mathbf{h})\times]$ is the cross product matrix of the vector $(I\boldsymbol{\omega} + \mathbf{h})$, $\mathbf{h}$ is the angular momentum of the momentum wheels, and $\mathbf{T}$ is the external torque operating on the SC. Because $\mathbf{x}$ is a constant vector, it obeys the following differential equation

$$\dot{\mathbf{x}} = 0 \quad (59.b)$$

We are tempted to combine the last two equations into one dynamics equation, as follows;

$$\begin{bmatrix} \dot{\boldsymbol{\omega}} \\ \dot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times] & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \end{bmatrix} + \begin{bmatrix} I^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \\ 0 \end{bmatrix} \quad (59.c)$$

This dynamics model calls for the use of a Kalman Filter (KF). In fact the most appropriate filter is the Pseudo-Linear Kalman (PSELIKA, ref. 6) filter. To find the suitable measurement equation, we turn to Eqs. (50.c) and (51.b) from which it is obvious that

$$\mathbf{G}_a = \begin{bmatrix} C & H(\boldsymbol{\omega}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \end{bmatrix} \quad (60)$$

In order to apply the filter algorithm we need to add some process noise to Eq. (59.c) and some measurement noise to Eq. (60). However, it is easy to see though that the last system is unobservable. Theoretically, if the actual noise values are small, and if we know the initial angular velocity, then we can compute the angular velocity separately. Once we know the angular rate and we command the SC to execute a suitable angular rate profile, we should be able to compute the calibration parameters. In reality though the computation of the angular rate is not accurate enough because it is done in an open loop manner. Therefore, we have to add attitude measurements in order to check the divergence of the computed angular rate. In this case we can indeed combine all the dynamics equations into one augmented matrix

equation and estimate the augmented state vector. This is so because it is possible to find an angular rate profile that will render the system observable.

As mentioned, attitude can be given in several ways; namely, it can be given in a raw form as vector measurements or in processed attitude parameters like a quaternion or direction cosine matrix (DCM). Let us consider two cases, one where attitude is represented by a processed quaternion, and the other case when we have raw vector measurements. The case where attitude is given in the form of a DCM can be inferred from the development presented in Ref. 7 and the way we handle quaternion representation of attitude.

Estimation When Attitude is Presented by the Attitude Quaternion

Let us assume first that the attitude is given in a form of a quaternion (ref. 8). In this case the filter dynamics is as follows (ref. 7)

$$\begin{bmatrix} \dot{\boldsymbol{\omega}} \\ \dot{\mathbf{x}} \\ \dot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} I^{-1}[(I\boldsymbol{\omega}+\mathbf{h})\times] & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{2}Q & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \mathbf{q} \end{bmatrix} + \begin{bmatrix} I^{-1}(\mathbf{T}-\dot{\mathbf{h}}) \\ 0 \\ 0 \end{bmatrix} \tag{61}$$

where

$$Q = \begin{bmatrix} q_4 & -q_3 & q_2 \\ q_3 & q_4 & -q_1 \\ -q_2 & q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix} \tag{62}$$

and the corresponding measurement equation is

$$\mathbf{q}_m = \begin{bmatrix} 0_{4\times3} & 0_{4\times16} & I_{4\times4} \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \mathbf{q} \end{bmatrix} \tag{63}$$

The matrix $I_{4\times4}$ is a fourth order identity matrix. The combined measurement equation consists of Eqs. (60) and (63); that is

$$\begin{bmatrix} \mathbf{G}_a \\ \mathbf{q}_m \end{bmatrix} = \begin{bmatrix} C & H(\boldsymbol{\omega}) & 0_{4\times4} \\ 0_{4\times3} & 0_{4\times16} & I_{4\times4} \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \mathbf{q} \end{bmatrix} \tag{64}$$

Estimation When Attitude is Given by Vector Observations

Normally, in space missions attitude is determined from vector observations. These observations can be used directly to check the divergence of the angular velocity estimates (ref .7). This is shown next. Suppose that we have N vector measurements at a certain time point. Let $\mathbf{r}_i$ denote some abstract i^{th} vector as expressed in the reference coordinate system, and let $\mathbf{b}_i$ denote the same vector when expressed in the body coordinates. From the laws of dynamics it is known that

$$D\dot{\mathbf{r}}_i = \dot{\mathbf{b}}_i + \boldsymbol{\omega} \times \mathbf{b}_i \tag{65}$$

where $\dot{\mathbf{r}}_i$ is the time derivative of $\mathbf{r}_i$ as seen by an observer in the reference coordinates, D is the matrix that transforms vectors from the reference to body coordinates, and $\dot{\mathbf{b}}_i$ is the time derivative of $\mathbf{b}_i$ as seen by an observer in body coordinates. The vector $\mathbf{b}_i$ is a measured vector and $\dot{\mathbf{b}}_i$ is its time derivative. We can write Eq. (65), as follows;

$$\dot{\mathbf{b}}_i = [\mathbf{b}_i\times]\boldsymbol{\omega} + \mathrm{D}\dot{\mathbf{r}}_i \tag{66}$$

Note that $\dot{\mathbf{r}}_i$ is computable since $\mathbf{r}_i$ is usually known because, generally, the vector is a direction to a certain known planet whose location is given in an Almanac or, like with magnetometer measurements, the vector can be computed using a model. (It should be noted that quite often the rate of change of $\mathbf{r}_i$ is so small that $\dot{\mathbf{r}}_i$ is negligible). Define

$$\dot{\boldsymbol{\beta}} = \begin{bmatrix} \dot{\mathbf{b}}_1 \\ \cdot \\ \cdot \\ \dot{\mathbf{b}}_N \end{bmatrix} \quad (67.a) \qquad \mathrm{B} = \begin{bmatrix} [\mathbf{b}_1\times] \\ \cdot \\ \cdot \\ [\mathbf{b}_N\times] \end{bmatrix} \quad (67.b) \quad \text{and} \quad \mathbf{u} = \begin{bmatrix} \mathrm{D}\dot{\mathbf{r}}_1 \\ \cdot \\ \cdot \\ \mathrm{D}\dot{\mathbf{r}}_N \end{bmatrix} \quad (67.c)$$

then we can augment all the N equations of Eq. (66) into one matrix equation, as follows;

$$\dot{\boldsymbol{\beta}} = \mathrm{B}\boldsymbol{\omega} + \mathbf{u} \tag{68}$$

therefore, instead of Eq. (61), we obtain in this case of vector measurement the augmented equation

$$\begin{bmatrix} \dot{\boldsymbol{\omega}} \\ \dot{\mathbf{x}} \\ \dot{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} \mathrm{I}^{-1}[(\mathrm{I}\boldsymbol{\omega}+\mathbf{h})\times] & 0 & 0 \\ 0 & 0 & 0 \\ \mathrm{B} & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \mathrm{I}^{-1}(\mathbf{T}-\dot{\mathbf{h}}) \\ 0 \\ \mathbf{u} \end{bmatrix} \tag{69}$$

and the corresponding measurement equation is

$$\boldsymbol{\beta}_m = \begin{bmatrix} 0_{3\times3} & 0_{3\times16} & \mathrm{I}_{3\times3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix} \tag{70}$$

whereas the augmented measurement equation is

$$\begin{bmatrix} \mathbf{G}_a \\ \boldsymbol{\beta}_m \end{bmatrix} = \begin{bmatrix} \mathrm{C} & \mathrm{H}(\boldsymbol{\omega}) & 0_{4\times4} \\ 0_{3\times3} & 0_{3\times16} & \mathrm{I}_{3\times3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix} \tag{71}$$

COMPENSATION

To complete the calibration process we need to perform its second stage; namely, compensation where we eliminate the estimated errors from the gyro readings. From Eq. (50.c) we obtain

$$C\boldsymbol{\omega}_r = \mathbf{G}_a - \begin{bmatrix}\Omega^m E & \Omega^k & I_4\end{bmatrix}\begin{bmatrix}\mathbf{d}\\ \mathbf{k}\\ \mathbf{b}\end{bmatrix} \tag{72.a}$$

As mentioned before, $\mathbf{G}_a$ is a vector of the gyro readings and $\boldsymbol{\omega}_r$ is the correct angular velocity vector. However, we do not have $\boldsymbol{\omega}_r$ which is what we are trying to measure; therefore, to compute Ω^m and Ω^k, which have to be computed using $\boldsymbol{\omega}_r$, we use the measured uncompensated angular rate vector. This vector is derived from the uncompensated gyro measurements which we called *actual* and denoted by *a*. Also, we do not have the actual values of **d**, **k** or **b**, but rather their estimate; therefore, using the values on hand Eq. (72.a) becomes

$$C\hat{\boldsymbol{\omega}}_r = \mathbf{G}_a - \begin{bmatrix}\Omega_a^m E & \Omega_a^k & I_4\end{bmatrix}\begin{bmatrix}\hat{\mathbf{d}}\\ \hat{\mathbf{k}}\\ \hat{\mathbf{b}}\end{bmatrix} \tag{72.b}$$

where *a* denotes the *actual* values and ^ denotes estimated vectors. To obtain the compensated measurements of the angular rate vector define the A matrix, as follows:

$$A = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{1}{2}\end{bmatrix} \tag{73}$$

It is easy to verify that

$$AC^TC = I \tag{74}$$

therefore pre multiplying Eq. (72.b) by AC^T yields

$$\hat{\boldsymbol{\omega}}_r = AC^T\mathbf{G}_a - AC^T\begin{bmatrix}\Omega_a^m E & \Omega_a^k & I_4\end{bmatrix}\begin{bmatrix}\hat{\mathbf{d}}\\ \hat{\mathbf{k}}\\ \hat{\mathbf{b}}\end{bmatrix} \tag{75}$$

SIMULATION RESULTS

In lieu of actual SC data, a simulator was developed to produce the gyro, reaction wheel, and star tracker data. Much care was devoted to the simulation since any dynamics errors would be perceived by the KF as a state error. First, the maneuver strategy was developed. An inertial period before any maneuver would facilitate the estimation of the gyro bias. The scale factors of each gyro could be estimated by a maneuver about that gyro axis. The misalignments could be estimated by the same scale factor maneuvers. An additional two maneuvers about the SC **X** and **Y** body axes, respectively, were added to assist in the alignment estimation of gyro 4 which senses rate about the body **Z** axis. Second, the maneuvers were modeled as a ramp-coast-ramp where the linear ramp time was 5 seconds. The rate profile was then obtained and $\dot{\boldsymbol{\omega}}$ was derived by simple subtraction. Third, the dynamics Eq. (59.a) was re-written in terms of $\dot{\mathbf{h}}$ and the ordinary differential equation (ODE) was solved using the $\boldsymbol{\omega}$ and $\dot{\boldsymbol{\omega}}$ generated above. Fourth, using the newly generated system momentum profile $\mathbf{h}$, the rate, $\boldsymbol{\omega}$, was determined from

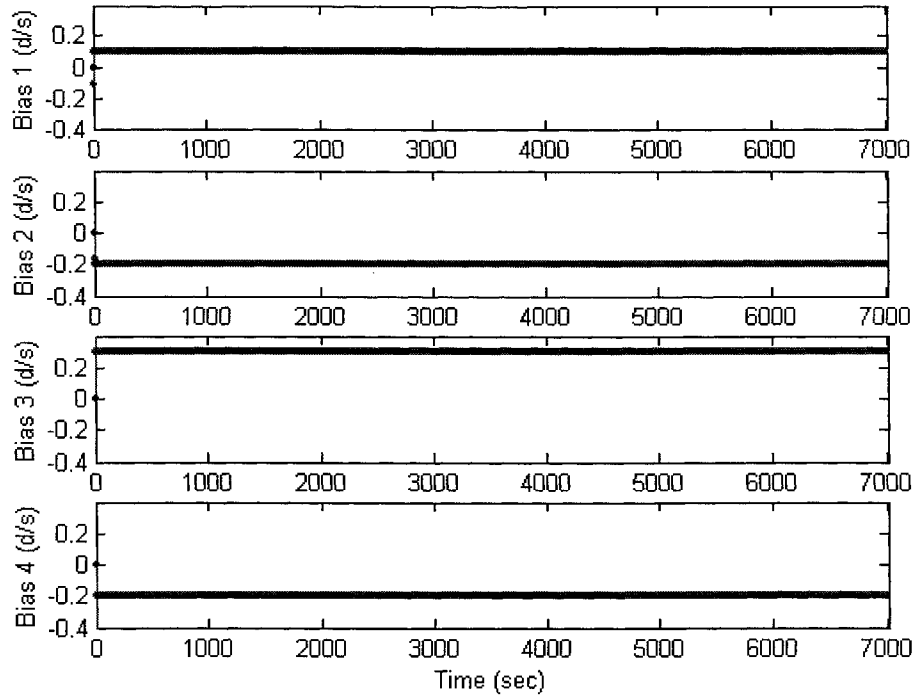


Fig. 5: Gyro Bias Estimate (bold) versus Truth (thin)

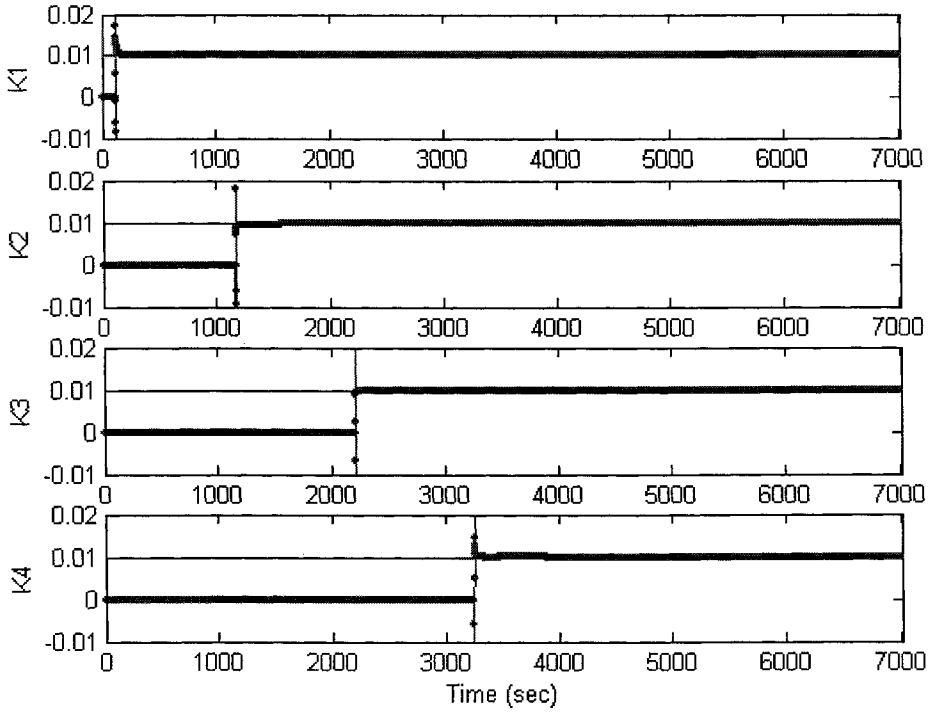


Fig. 6: Gyro Scale Factor Estimate (bold) versus Truth (thin)

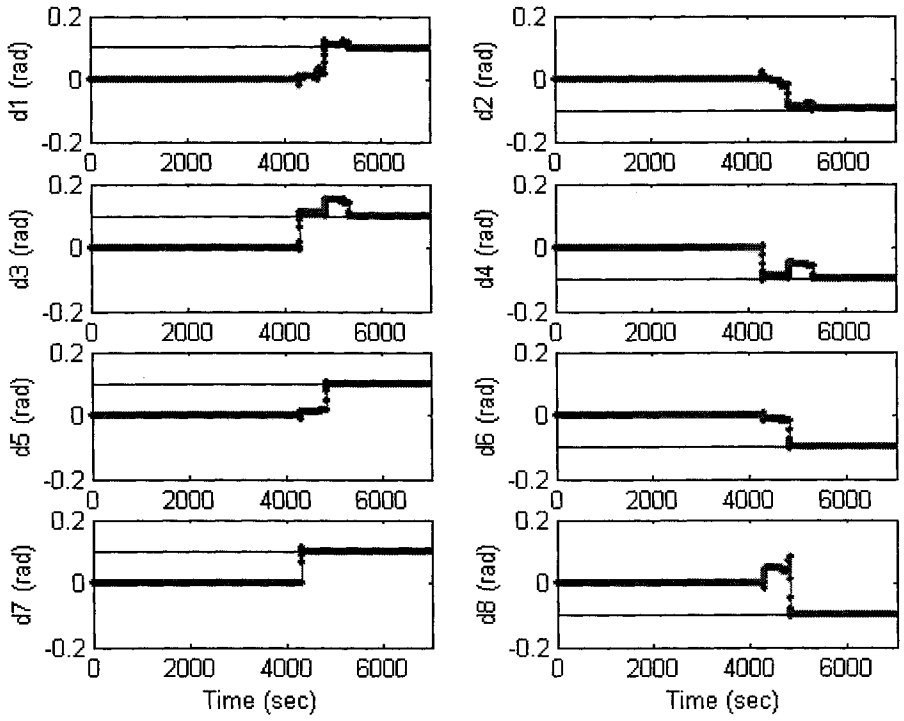


Fig. 7: Gyro Misalignment Estimate (bold) versus Truth (thin)

the ODE in Eq. (59.a). Fifth, using this $\boldsymbol{\omega}$, the quaternion was estimated using the $\dot{\mathbf{q}}$ portion of Eq. (61). Lastly, the body rate was resolved in the AQUA gyro frame by use of Eq. (50.c).

The KF was then executed on the simulated data. The biases were estimated using the initial inertial period. Also, to facilitate the bias estimation, the scale factor and misalignment parameter estimation was terminated by zeroing out their respective rows and columns in the state covariances and process noise. The scale factors for each gyro were then estimated using the respective maneuver about that axis and by zeroing out the influences of the other scale factors, the misalignments, and biases as described with the bias estimation. Lastly, the gyro misalignment slews, which were a repeat of the scale factor slews with the addition of **X** and **Y**-axis maneuvers, were executed with the biases and scale factors zeroed out. All states were estimated with less than a 1% deviation from truth. The KF bias estimate can be seen in Fig. 5, followed by the scale factor estimate in Fig. 6, and lastly the misalignment estimate in Fig. 7.

CONCLUSIONS

In this paper we presented a new method of gyro calibration. Normally, we have to calibrate a cluster of three gyros whose sensitive axes are along the body axes. Here, the rate is read by four gyros only one of which is aligned along the body coordinate axes. Therefore, a new algorithm was devised for calibrating a quadruplet rather than the customary triad gyro set. In particular, a new model had to be developed for the gyro misalignment errors. Normally, the gyro outputs are used to supply data for a differential equation, which is solved in order to compute attitude. According to the new method the gyro outputs are also used as measurements, which are fed into a Kalman filter that estimates the gyro misalignment, scale factors, and biases. The new calibration algorithm was developed in particular for the

calibration of the EOS-AQUA satellite gyros. The effectiveness of the new algorithm was demonstrated through simulations with error of each estimated parameter being less than 1%.

REFERENCES

1. Chatfield, A.B., *Fundamentals of High Accuracy Inertial Navigation*, Vol. 174 Progress in Astronautics and Aeronautics, AIAA, 1997, pp. 93-106.

2. Bar-Itzhack, I.Y., and Harman, R.R., "True Covariance Simulation of the EUVE Update Filter," *Proceedings of the 1989 Flight Mechanics Estimation Theory Symposium*. NASA Conference Publication 3050, pp. 223-236.

3. Deutschmann, J.K., and Bar-Itzhack, I.Y., "Evaluation of Attitude and Orbit Estimation Using Actual Earth Magnetic Field Data," to be published in the *Journal of Guidance, Control, and Dynamics*.

4. NASA-Goddard Space Flight Center, Multimission Three-Axis Stabilized Spacecraft (MTASS), 553-FDD-93/032R0UD0, 1933, pp. 3.3.2-1 – 3.3.2-10.

5. Gelb, A. (Ed.), *Applied Optimal Estimation*, MIT Press, Cambridge MA, 1974, p. 103.

6. Harman, R.R., and Bar-Itzhack, I.Y., "Pseudo-Linear and State Dependent Riccati Equation Filters for Angular Rate Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 22, No. 5, Sept.-Oct. 1999, pp. 723-725. (Engineering Note).

7. Bar-Itzhack, I.Y., "Classification of Algorithms for Velocity Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 2, March-April 2001, pp. 214-218.

8. Azor, R., Bar-Itzhack, I.Y., Deutschmann, J.R. and Harman, R.R., "Angular-Rate Estimation Using Delayed Quaternion Measurements," AIAA Guidance, Navigation, and Control conference, Portland, OR, August 9-11, 1999.

A NONLINEAR SPACECRAFT ATTITUDE CONTROLLER AND OBSERVER WITH AN UNKNOWN CONSTANT GYRO BIAS AND GYRO NOISE

Julie Deutschmann
Flight Dynamics Analysis Branch
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

Robert M. Sanner
University of Maryland
Aerospace Engineering Department
College Park, Maryland 20742

ABSTRACT

A nonlinear control scheme for attitude control of a spacecraft is combined with a nonlinear gyro bias observer for the case of constant gyro bias, in the presence of gyro noise. The observer bias estimates converge exponentially to a mean square bound determined by the standard deviation of the gyro noise. The resulting coupled, closed loop dynamics are proven to be globally stable, with asymptotic tracking which is also mean square bounded. A simulation of the proposed observer-controller design is given for a rigid spacecraft tracking a specified, time-varying attitude sequence to illustrate the theoretical claims.

INTRODUCTION

Combined observer-controller designs for the attitude control of rigid flight vehicles are a subject of active research[1,2]. Successful design of such architectures is complicated by the fact that there is, in general, no separation principle for nonlinear systems. In contrast to linear systems, "certainty equivalence" substitution of the states from even an exponentially converging observer into a nominally stabilizing state feedback control law does not necessarily guarantee stable closed-loop operation for the coupled systems[3,4].

However, in reference 2, a separation principle is found to exist for the problem of forcing the attitude of a rigid vehicle to asymptotically track a (time-varying) reference attitude using feedback from sensors with persistent nonzero bias errors. A persistency of excitation argument demonstrated that the bias estimates provided by the observer are exponentially convergent to the true bias values. Reference 2 also proves that the certainty equivalence use of the observer bias estimates in the nonlinear feedback control algorithm proposed in reference 5 resulted in a stable closed-loop operation, with asymptotically perfect tracking.

Here we extend the analysis of reference 2 to include noise in the gyro reading. The Converse Lyaponov Theorem demonstrates that in this case the bias estimates provided by the observer converge exponentially to a mean square bound proportional to the variance of the noise. We then consider the certainty equivalence use of these observer estimates in the nonlinear feedback control algorithm and show that, as in reference 2, the perturbation introduced by this strategy into the closed-loop dynamics can be represented as a bounded function of the vehicle states multiplying the observer transients and the noise. We demonstrate that the stability properties of the controller are, in fact, maintained in the face of the perturbations, with asymptotic tracking to a mean square bound.

The paper is organized as follows. Section II contains definitions of the terms used in the controller and observer. In Section III the nonlinear observer for the constant gyro bias with added noise is developed and the convergence is proven. Section IV presents the nonlinear controller design and the proof of stability of the closed loop system and the convergence of the tracking errors. Section V presents simulation results, followed by conclusions in Section VI.

DEFINITIONS

The attitude of a spacecraft can be represented by a four component quaternion, consisting of a rotation angle and unit rotation vector, known as the Euler axis

$$\mathbf{q} = \begin{bmatrix} \mathbf{e}\sin\frac{\phi}{2} \\ \cos\frac{\phi}{2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}$$

where ϕ is the rotation angle, $\mathbf{e}$ is the Euler axis, $\boldsymbol{\varepsilon}$ and η are the vector and scalar portions of the quaternion, respectively. Note that $\|\mathbf{q}\| = 1$ by definition. The quaternion represents the rotation from an inertial coordinate system to the spacecraft body coordinate system. A rotation matrix can be computed from the quaternion as[6]

$$R(\mathbf{q}) = (\eta^2 - \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})I + 2\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^T - 2\eta S(\boldsymbol{\varepsilon})$$

where $S(\boldsymbol{\varepsilon})$ is a cross product matrix formed from the vector $\boldsymbol{\varepsilon}$.

$$S(\boldsymbol{\varepsilon}) = \begin{bmatrix} 0 & -\varepsilon_z & \varepsilon_y \\ \varepsilon_z & 0 & -\varepsilon_x \\ -\varepsilon_y & \varepsilon_x & 0 \end{bmatrix}$$

A desired target attitude is represented by the quaternion, $\mathbf{q}_d^T = [\boldsymbol{\varepsilon}_d^T, \eta_d]$. The attitude error used in the controller is defined as a rotation from the desired body frame to the actual body frame and is computed according to[7]

$$\tilde{\mathbf{q}}_c = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_c \\ \tilde{\eta}_c \end{bmatrix} = \mathbf{q} \otimes \mathbf{q}_d^{-1} = \begin{bmatrix} \eta_d I - S(\boldsymbol{\varepsilon}_d) & -\boldsymbol{\varepsilon}_d \\ \boldsymbol{\varepsilon}_d^T & \eta_d \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix}$$

Similarly, in the observer, the attitude error is defined as the rotation from the estimated body frame to the actual body frame as

$$\tilde{\mathbf{q}}_o = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_o \\ \tilde{\eta}_o \end{bmatrix} = \mathbf{q} \otimes \hat{\mathbf{q}}^{-1} = \begin{bmatrix} \hat{\eta} I - S(\hat{\boldsymbol{\varepsilon}}) & -\hat{\boldsymbol{\varepsilon}} \\ \hat{\boldsymbol{\varepsilon}}^T & \hat{\eta} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon} \\ \eta \end{bmatrix} \tag{1}$$

where $\hat{\mathbf{q}}$ represents the attitude state of the observer. Note that $\tilde{\boldsymbol{\varepsilon}}_c = 0$, $\tilde{\eta}_c = \pm 1$ indicates that the spacecraft is aligned with the desired attitude and, similarly, $\tilde{\boldsymbol{\varepsilon}}_o = 0$, $\tilde{\eta}_o = \pm 1$ indicates that the attitude estimate is aligned with the actual attitude.

The kinematics equation for the quaternion is given as

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{\boldsymbol{\varepsilon}} \\ \dot{\eta} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \eta I + S(\boldsymbol{\varepsilon}) \\ -\boldsymbol{\varepsilon}^T \end{bmatrix} \boldsymbol{\omega} = \frac{1}{2} Q(\mathbf{q})\boldsymbol{\omega}$$

where $\boldsymbol{\omega}$ is the spacecraft angular velocity. The angular velocity is typically measured by a gyro, which can be corrupted with both systematic and random errors. In the case of gyro bias and random noise, the gyro reading, $\boldsymbol{\omega}_g$, can be written as

$$\boldsymbol{\omega}_g = \boldsymbol{\omega} + \mathbf{b} + \mathbf{v}(t)$$

where $\boldsymbol{\omega}$ is the true angular velocity, $\mathbf{b}$ is the gyro bias (which in this work is treated without noise), and $\mathbf{v}(t)$ is an added noise. An estimate of the angular velocity is given as $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega}_g - \hat{\mathbf{b}}$. The bias error is defined as the difference between the true and estimated bias

$$\tilde{\mathbf{b}} = \mathbf{b} - \hat{\mathbf{b}} \tag{2}$$

Finally, a measure of the discrepancy between the actual and desired angular velocity in the controller is computed as[7]

$$\tilde{\boldsymbol{\omega}}_c = \boldsymbol{\omega} - R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d \tag{3}$$

which is defined such that $\dot{\tilde{\mathbf{q}}}_c = \frac{1}{2}Q(\tilde{\mathbf{q}}_c)\tilde{\boldsymbol{\omega}}_c$.

NONLINEAR OBSERVER FOR CONSTANT GYRO BIAS

Following the development of reference 1, a state observer for the bias can be defined as

$$\dot{\hat{\mathbf{q}}} = \frac{1}{2}\hat{Q}(\hat{\mathbf{q}})R^T(\tilde{\mathbf{q}}_o)(\boldsymbol{\omega}_g - \hat{\mathbf{b}} + k\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o)) \tag{4}$$

$$\dot{\hat{\mathbf{b}}} = -\tfrac{1}{2}\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o) \tag{5}$$

The gain, k, is chosen as a positive constant. The $R^T(\tilde{\mathbf{q}}_o)$ resolves the angular velocity terms in the observer frame.

Computing the derivatives of $\tilde{\mathbf{q}}_o$ in (1) and $\tilde{\mathbf{b}}$ in (2) results in the following differential error equations

$$\dot{\tilde{\mathbf{q}}}_o = \begin{bmatrix} \dot{\tilde{\boldsymbol{\varepsilon}}}_o \\ \dot{\tilde{\eta}}_o \end{bmatrix} = \tfrac{1}{2}\begin{bmatrix} \tilde{\eta}_o I + S(\tilde{\boldsymbol{\varepsilon}}_o) \\ -\tilde{\boldsymbol{\varepsilon}}_o^T \end{bmatrix}\left[-\tilde{\mathbf{b}} - \mathbf{v}(t) - k\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o)\right] \tag{6}$$

$$\dot{\tilde{\mathbf{b}}} = \tfrac{1}{2}\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o) \tag{7}$$

The above error equations are rewritten as

$$\begin{bmatrix} \dot{\tilde{\mathbf{q}}}_o \\ \dot{\tilde{\mathbf{b}}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\tilde{Q}(\tilde{\mathbf{q}}_o)(-\tilde{\mathbf{b}} - k\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o)) \\ \frac{1}{2}\tilde{\boldsymbol{\varepsilon}}_o \operatorname{sgn}(\tilde{\eta}_o) \end{bmatrix} + \begin{bmatrix} -\frac{1}{2}\tilde{Q}(\tilde{\mathbf{q}}_o)\mathbf{v}(t) \\ 0 \end{bmatrix} \tag{8}$$

The system in (8) is divided into the nominal system of reference 2 plus a perturbation

$$\dot{\mathbf{x}}(t) = f(t,\mathbf{x}) + D(t)$$

where $\mathbf{x}^T = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_o^T & \tilde{\mathbf{b}}^T \end{bmatrix}$ and $D(t)^T = \begin{bmatrix} (-\frac{1}{2}[\tilde{\eta}_o I + S(\tilde{\boldsymbol{\varepsilon}}_o)]\mathbf{v}(t))^T & 0 \end{bmatrix}$. Through a persistancy of excitation argument, the nominal system $\dot{\mathbf{x}}(t) = f(t,\mathbf{x})$ is proven to be exponentially stable[2]. Therefore, according to the Converse Lyapunov Theorem[3], a Lyapunov function and positive, finite constants c_1, c_2, c_3, and c_4 exist and satisfy the following

$$c_1\|\mathbf{x}\|^2 \le V_o \le c_2\|\mathbf{x}\|^2$$

$$\dot{V}_o \le -c_3\|\mathbf{x}\|^2$$
$$\left\|\frac{\partial V_o}{\partial \mathbf{x}}\right\| \le c_4\|\mathbf{x}\|$$

The perturbed Lyapunov function then satisfies

$$\dot{V}_o \le -c_3\|\mathbf{x}\|^2 + \left\|\frac{\partial V_o}{\partial \mathbf{x}}\right\|\|D\| \tag{9}$$

Since $\tilde{Q}^T\tilde{Q} = I$, $\|D\| = \frac{1}{2}\|\mathbf{v}\|$ and (9) becomes

$$\dot{V}_o \le -c_3\|\mathbf{x}\|^2 + \frac{c_4}{2}\|\mathbf{x}\|\|\mathbf{v}\| \tag{10}$$

If $\mathbf{v}(t)$ is uniformly bounded, the system is globally stable. The state $\mathbf{x}(t)$ converges exponentially to a ball determined by the bound on $\mathbf{v}(t)$, and then remains within that ball[3].

Consider the case that the noise $\mathbf{v}(t)$ is a bounded, zero mean, wide sense stationary (WSS) process with a mean square value of $\sigma^2 I$. Applying Young's inequality[4] to (10) results in

$$\dot{V}_o \le -\frac{c_3}{2}\|\mathbf{x}\|^2 + \frac{c_4}{2c_3}\|\mathbf{v}\|^2 \tag{11}$$

The time average of (11) is computed as

$$\frac{1}{T}\int_0^T\|\mathbf{x}\|^2 d\tau \le \frac{c_4}{c_3^2}\frac{1}{T}\int_0^T\|\mathbf{v}\|^2 d\tau + \frac{2}{c_3 T}[V_o(0) - V_o(T)]$$

Taking the limit as $T \to \infty$

$$\lim_{T\to\infty}\sup\frac{1}{T}\int_0^T\|\mathbf{x}\|^2 d\tau \le \lim_{T\to\infty}\frac{c_4}{c_3^2}\frac{1}{T}\int_0^T\|\mathbf{v}\|^2 d\tau = \frac{c_4}{c_3^2}3\sigma^2$$

The root mean square (RMS) bound is then given as

$$\|\tilde{\mathbf{b}}\|_{RMS} \le \|\mathbf{x}\|_{RMS} \le \sqrt{3c_4}\frac{\sigma}{c_3}$$

NONLINEAR CONTROLLER DESIGN

The complete attitude dynamics for a rigid spacecraft are given as

$$H\dot{\omega} - S(H\omega)\omega = \mathbf{u}$$
$$\dot{\mathbf{q}} = \frac{1}{2}Q(\mathbf{q})\omega$$

where H is a constant, symmetric inertia matrix and $\mathbf{u}$ is the applied external torque, for example, from attached rocket thrusters. The goal of the controller is for the actual, measured attitude $\mathbf{q}(t)$ to asymptotically track a

(generally) time-varying desired attitude $\mathbf{q}_d(t)$ and angular velocity $\boldsymbol{\omega}_d(t)$, related for consistency by $\dot{\mathbf{q}}_d = \frac{1}{2}Q(\mathbf{q}_d)\boldsymbol{\omega}_d$. It is assumed that $\boldsymbol{\omega}_d(t)$ is bounded and differentiable with $\dot{\boldsymbol{\omega}}_d(t)$ also bounded.

The passivity based controller of reference 5 utilizes the composite error metric

$$\mathbf{s} = \tilde{\boldsymbol{\omega}}_c + \lambda\tilde{\boldsymbol{\varepsilon}}_c = \boldsymbol{\omega} - \boldsymbol{\omega}_r \tag{12}$$

where from (3), $\boldsymbol{\omega}_r = R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d - \lambda\tilde{\boldsymbol{\varepsilon}}_c$, λ>o. Taking the derivative of (12) and multiplying by H results in

$$H\dot{\mathbf{s}} = H\dot{\boldsymbol{\omega}} - H\dot{\boldsymbol{\omega}}_r = \mathbf{u} + S(H\boldsymbol{\omega})\boldsymbol{\omega} - H\boldsymbol{\alpha}_r \tag{13}$$

where

$$\boldsymbol{\alpha}_r = \dot{\boldsymbol{\omega}}_r = R(\tilde{\mathbf{q}}_c)\dot{\boldsymbol{\omega}}_d - S(\tilde{\boldsymbol{\omega}}_c)R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d - \lambda Q_1(\tilde{\mathbf{q}}_c)\tilde{\boldsymbol{\omega}}_c$$

and $Q_1(\tilde{\mathbf{q}}_c) = \tilde{\eta}_c I + S(\tilde{\boldsymbol{\varepsilon}}_c)$ as defined above. With these definitions, the control law

$$\mathbf{u} = -K_D\mathbf{s} + H\boldsymbol{\alpha}_r - S(H\boldsymbol{\omega})\boldsymbol{\omega}_r \tag{14}$$

for any symmetric, positive definite K_D results in closed-loop dynamics

$$H\dot{\mathbf{s}} - S(H\boldsymbol{\omega})\mathbf{s} + K_D\mathbf{s} = 0$$

As shown in reference 5, these dynamics, together with the definition of the composite error, produces the desired stability and tracking properties.

In the current application, the control law (14) cannot be implemented because exact measurements of the angular velocity $\boldsymbol{\omega}$ are not available. Instead a certainty equivalence approach is employed using the estimates $\hat{\boldsymbol{\omega}}$ from above, resulting in

$$\mathbf{u} = -K_D\hat{\mathbf{s}} + H\hat{\boldsymbol{\alpha}}_r - S(H\hat{\boldsymbol{\omega}})\boldsymbol{\omega}_r \tag{15}$$

where $\hat{\mathbf{s}} = \hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_r$, $\hat{\tilde{\boldsymbol{\omega}}}_c = \hat{\boldsymbol{\omega}} - R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d$, and

$$\hat{\boldsymbol{\alpha}}_r = R(\tilde{\mathbf{q}}_c)\dot{\boldsymbol{\omega}}_d + S(R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d)\hat{\tilde{\boldsymbol{\omega}}}_c - \lambda Q_1(\tilde{\mathbf{q}}_c)\hat{\tilde{\boldsymbol{\omega}}}_c .$$

Substituting (15) into (13), along with (12), and noting that $\tilde{\mathbf{s}} = \mathbf{s} - \hat{\mathbf{s}} = -\tilde{\mathbf{b}} - \mathbf{v}(t)$, $\tilde{\boldsymbol{\alpha}}_r = [-S(R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d) + \lambda Q_1(\tilde{\mathbf{q}}_c)](\tilde{\mathbf{b}} + \mathbf{v}(t))$, and $\tilde{\boldsymbol{\omega}}_c - \hat{\tilde{\boldsymbol{\omega}}}_c = \boldsymbol{\omega} - \hat{\boldsymbol{\omega}} = -\tilde{\mathbf{b}} - \mathbf{v}(t)$ produces the closed-loop dynamics

$$H\dot{\mathbf{s}} - S(H\boldsymbol{\omega})\mathbf{s} + K_D\mathbf{s} = [S(\boldsymbol{\omega}_r)H + HS(R(\tilde{\mathbf{q}}_c)\boldsymbol{\omega}_d) - \lambda HQ_1(\tilde{\mathbf{q}}_c) - K_D](\tilde{\mathbf{b}} + \mathbf{v}(t)) \tag{16}$$

The terms on the right hand side of (16) can be rewritten as $\Delta(\tilde{\mathbf{q}}_c, \boldsymbol{\omega}_d)(\tilde{\mathbf{b}} + \mathbf{v}(t))$. Since $\|\tilde{\mathbf{q}}_c\| = 1$ by definition and $\|\boldsymbol{\omega}_d\| < \infty$ by assumption

$$\gamma \equiv \sup_{t \geq t_0} \sup_{\|\tilde{\mathbf{q}}_c\|=1} \|\Delta(\tilde{\mathbf{q}}_c, \boldsymbol{\omega}_d(t))\| < \infty$$

Using the Lyapunov function $V_c = \frac{1}{2}\mathbf{s}^T\mathbf{H}\mathbf{s}$, the derivative of V_c along closed-loop trajectories of (15) satisfies the inequality

$$\dot{V}_c = -\mathbf{s}^T K_D \mathbf{s} + \mathbf{s}^T \boldsymbol{\Delta}(\tilde{\mathbf{b}} + \mathbf{v}) \le -k_D\|\mathbf{s}\|^2 + \gamma\|\mathbf{s}\|(\|\tilde{\mathbf{b}}\| + \|\mathbf{v}\|)$$

where k_D is the smallest eigenvalue of K_D. Using Young's inequality[4] on the last term above, $\dot{V}_c$ is rewritten as

$$\dot{V}_c \le -\frac{k_D}{2}\|\mathbf{s}\|^2 + \frac{\gamma}{2k_D}(\|\tilde{\mathbf{b}}\|^2 + \|\mathbf{v}\|^2) \tag{17}$$

Thus, from the definition of V_c, and recalling from the observer analysis that $\|\tilde{\mathbf{b}}\|$ and $\|\mathbf{v}\|$ are bounded, $\mathbf{s}$ is also seen to be uniformly bounded. Similarly $\dot{\mathbf{s}}$ is uniformly bounded, since all the terms in (16) are bounded. Integrating (17)

$$\int_0^T\|\mathbf{s}\|^2 d\tau \le \frac{\gamma}{k_D^2}(\int_0^T\|\tilde{\mathbf{b}}\|^2 d\tau + \int_0^T\|\mathbf{v}\|^2 d\tau) + \frac{2}{k_D}(V_c(0) - V_c(T))$$

Substituting $\|\mathbf{s}\|^2 = \|\tilde{\boldsymbol{\omega}}_c\|^2 + 2\lambda\tilde{\boldsymbol{\omega}}_c^T\tilde{\boldsymbol{\varepsilon}}_c + \lambda^2\|\tilde{\boldsymbol{\varepsilon}}_c\|^2$ from (12)

$$\int_0^T\|\tilde{\boldsymbol{\varepsilon}}_c\|^2 d\tau \le \frac{\gamma}{k_D^2\lambda^2}(\int_0^T\|\tilde{\mathbf{b}}\|^2 d\tau + \int_0^T\|\mathbf{v}\|^2 d\tau) + \frac{2}{k_D\lambda^2}(V_c(0) - V_c(T)) - \frac{2}{\lambda}\int_0^T\tilde{\boldsymbol{\omega}}_c^T\tilde{\boldsymbol{\varepsilon}}_c d\tau \tag{18}$$

Noting that $\tilde{\boldsymbol{\omega}}_c^T\tilde{\boldsymbol{\varepsilon}}_c = 2\dot{\tilde{\eta}}_c$, (18) is then

$$\int_0^T\|\tilde{\boldsymbol{\varepsilon}}_c\|^2 d\tau \le \frac{\gamma}{k_D^2\lambda^2}(\int_0^T\|\tilde{\mathbf{b}}\|^2 d\tau + \int_0^T\|\mathbf{v}\|^2 d\tau) + \frac{2}{k_D\lambda^2}(V_c(0) - V_c(T)) - \frac{4}{\lambda}[\tilde{\eta}(T) - \tilde{\eta}(0)] \tag{19}$$

Computing the time average of (19) and taking the limit as $T \to \infty$

$$\lim_{T\to\infty}\sup\frac{1}{T}\int_0^T\|\tilde{\boldsymbol{\varepsilon}}_c\|^2 d\tau \le \lim_{T\to\infty}\frac{\gamma}{k_D^2\lambda^2}\frac{1}{T}(\int_0^T\|\tilde{\mathbf{b}}\|^2 d\tau + \int_0^T\|\mathbf{v}\|^2 d\tau) \le \frac{\gamma^2}{k_D^2\lambda^2}\left[\frac{c_4}{c_3^2}3\sigma^2 + 3\sigma^2\right]$$

The RMS limit of the tracking error is then

$$\|\tilde{\boldsymbol{\varepsilon}}_c\|_{RMS} \le \frac{\sqrt{3\gamma}\sigma}{k_D\lambda}\left[\frac{c_4}{c_3^2} + 1\right]^{\frac{1}{2}}$$

SIMULATION RESULTS

The spacecraft attitude controller/observer design is tested with a Matlab simulation. The inertia matrix is a diagonal matrix with principal moments of inertia of $[90, 100, 70]^T$ kg-m^2. Table I lists the initial conditions for the observer and controller, as well as the true gyro bias, true initial angular velocity, and desired angular velocity. The gains are chosen as k=1, $K_D=k_DI$ with k_D=10, and λ=3. The standard deviation of the gyro noise was first set to 0.57 deg/sec and then to 0.057 deg/sec.

Figures 1 and 2 show the observer bias errors with the two different standard deviations for the gyro noise, respectively. In each figure the top plot shows $\|\tilde{\mathbf{b}}(t)\|$ and the bottom plot shows the $\|\tilde{\mathbf{b}}(t)\|_{RMS}$. In Figure 1,

$\|\tilde{\mathbf{b}}(t)\|$ converges to less than 0.3 deg/sec, and in Figure 2 to less than 0.03 deg/sec. In both cases, the RMS errors are less than the standard deviation of the gyro noise.

Figures 3 and 4 show the attitude tracking error from the controller, again for the two different standard deviations of the gyro noise. In each figure the upper plot shows the angular error and the lower plot shows $\|\tilde{\varepsilon}_c(t)\|_{RMS}$. In Figure 3 the attitude error converges to less than 0.4 degrees and to less than 0.04 degrees in Figure 4. Note that the RMS error of the vector part of the quaternion is plotted with the standard deviation of the gyro noise (converted to rad/sec) for comparison of the magnitudes only.

Table I. Simulation Initial and True Values

Variable	Initial Value
$\hat{\mathbf{q}}$	$[0, 1, 0, 0]^T$
$\mathbf{q}$	$[0, 0, 1, 0]^T$
$\mathbf{q}_d$	$[0, 0, 0, 1]^T$
$\hat{\mathbf{b}}$	$[0, 0, 0]^T$ deg/sec
$\boldsymbol{\omega}$ - true	$[-5.7, 11.4, -22.9]^T$ deg/sec
$\mathbf{b}$ – true	$[2.9, -2.9, 1.9]^T$ deg/sec
$\boldsymbol{\omega}_d$	$[0, 6.3, 0]^T$ deg/sec

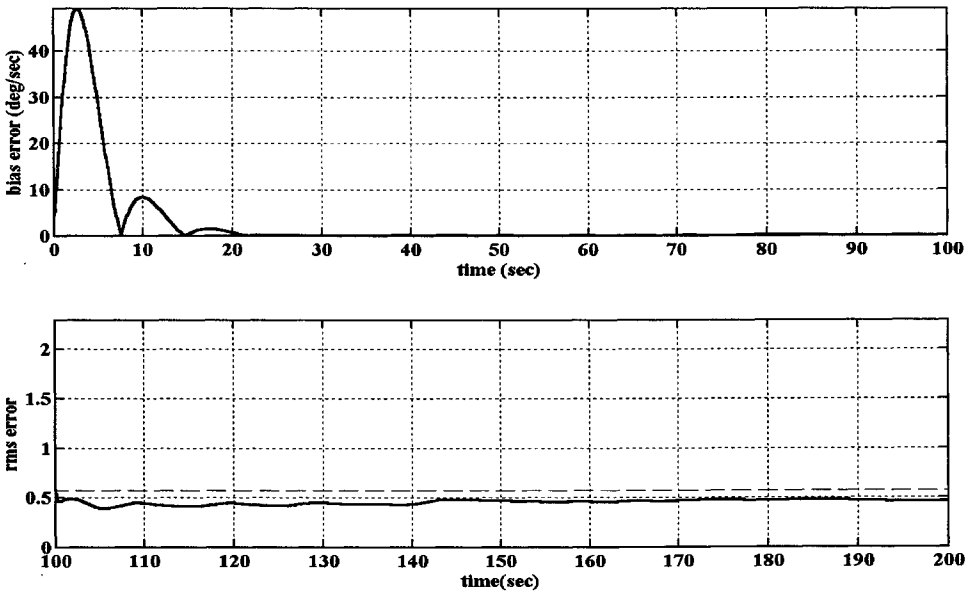


Figure 1. Bias Errors - $\|\tilde{\mathbf{b}}(t)\|$ and $\|\tilde{\mathbf{b}}(t)\|_{RMS}$ (solid lines), and σ=0.57 deg/sec (dashed line)

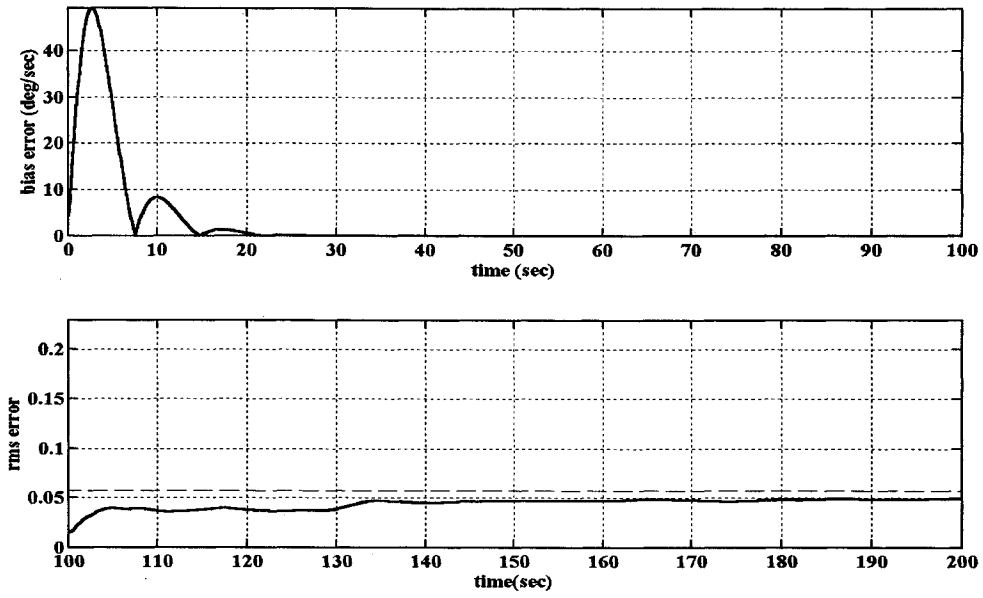


Figure 2. Bias Errors - $\|\tilde{\mathbf{b}}(t)\|$ and $\|\tilde{\mathbf{b}}(t)\|_{RMS}$ (solid lines), and σ=0.057 deg/sec (dashed line)

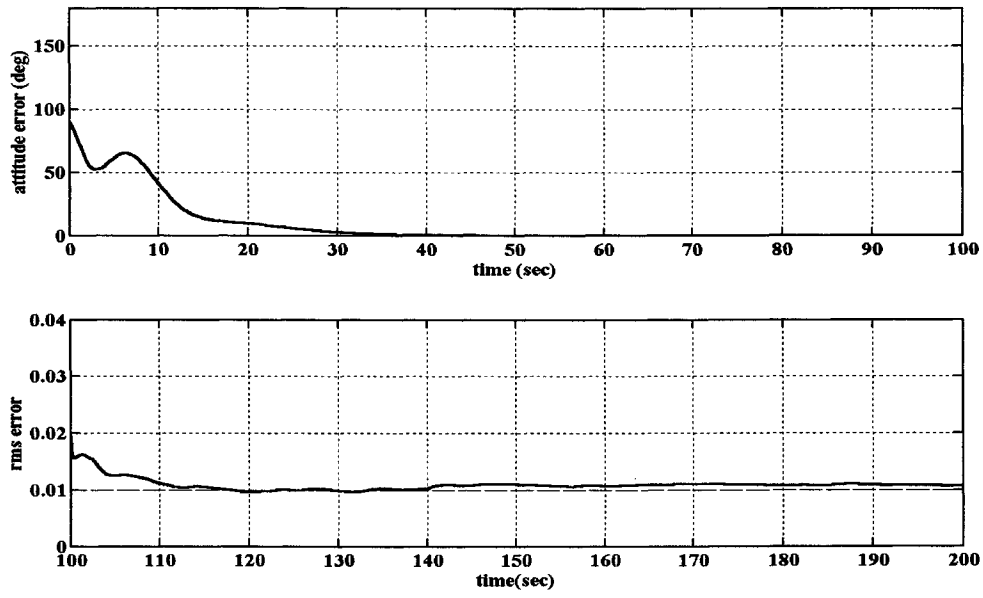


Figure 3. Attitude Tracking Error, $\|\tilde{\varepsilon}_c(t)\|_{RMS}$ (solid lines), and σ=0.01 rad/sec (dashed line)

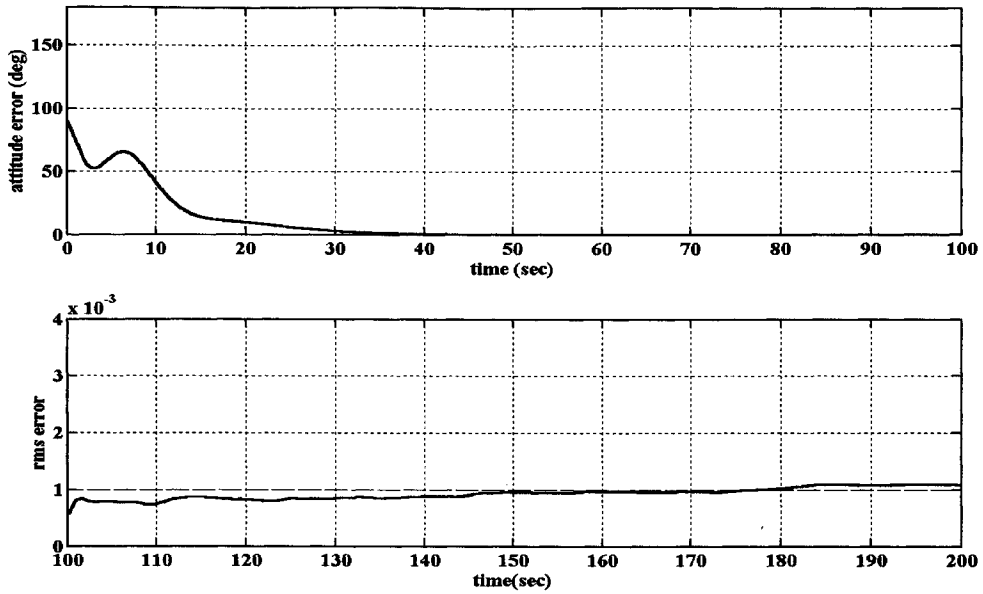


Figure 4. Attitude Tracking Error, $\|\tilde{\varepsilon}_c(t)\|_{RMS}$ (solid lines), and σ=0.001 rad/sec (dashed line)

CONCLUSIONS

A nonlinear controller/observer is presented for spacecraft attitude applications, given a constant gyro bias and gyro noise. The gyro bias estimates converge exponentially to a mean square bound determined by the standard deviation of the noise, relying on a persistency of excitation argument, which proves asymptotic identification of the gyro bias in the absence of noise. The nonlinear controller is a passivity based controller. The control input requires the use of the gyro bias estimate from the nonlinear observer. The closed loop stability properties of this nonlinear controller coupled with the nonlinear observer are analyzed and the system is found to be globally stable, leading to a separation principle for the nonlinear system, with asymptotic tracking to a mean square bound.

REFERENCES

1. H. Nijmeijer and T.I. Fossen, *Lecture Notes in Control and Information Sciences, New Directions in Nonlinear Observer Design*, Springer-Verlag, 1999, pp. 135-159.

2. J. Deutschmann and R.M. Sanner, "A Coupled Nonlinear Spacecraft Attitude Controller/Observer With an Unknown Constant Gyro Bias", submitted to the IEEE Conference on Decision and Control, paper no. CDC01-REG1379, Dec. 4-7, 2001, Orlando, FL.

3. H. Khalil, *Nonlinear Systems*, Prentice-Hall, Inc., 1996.

4. M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*, John Wiley & Sons, Inc., 1995.

5. O. Egland and J.M. Godhaven, "Passivity-Based Adaptive Attitude Control of a Rigid Spacecraft", *IEEE Transactions on Automatic Control*, Vol. 39, No. 4, April 1994, pp. 842-846.

6. B. Wie, Space Vehicle Dynamics and Control, AIAA Education Series, 1998.

7. R.M. Sanner, "Adaptive Attitude Control Using Fixed and Dynamically Structured Neural Networks", AIAA Guidance, Navigation and Control Conference, San Diego, CA, July 29-31, 1996.

STATE-DEPENDENT PSEUDO-LINEAR FILTER FOR SPACECRAFT ATTITUDE AND RATE ESTIMATION

Itzhack Y. Bar-Itzhack[1] and Richard R. Harman[2]
Flight Dynamics Analysis Branch, Code 572
Guidance Navigation and Control Center
NASA-Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

This paper presents the development and performance of a special algorithm for estimating the attitude and angular rate of a spacecraft. The algorithm is a pseudo-linear Kalman filter, which is an ordinary linear Kalman filter that operates on a linear model whose matrices are current state estimate dependent. The nonlinear rotational dynamics equation of the spacecraft is presented in the state space as a state-dependent linear system. Two types of measurements are considered. One type is a measurement of the quaternion of rotation, which is obtained from a newly introduced star tracker based apparatus. The other type of measurement is that of vectors, which permits the use of a variety of vector measuring sensors like sun sensors and magnetometers. While quaternion measurements are related linearly to the state vector, vector measurements constitute a nonlinear function of the state vector. Therefore, in this paper, a state-dependent linear measurement equation is developed for the vector measurement case. The state-dependent pseudo linear filter is applied to simulated spacecraft rotations and adequate estimates of the spacecraft attitude and rate are obtained for the case of quaternion measurements as well as of vector measurements.

INTRODUCTION

Precise angular-rate is required for spacecraft (SC) attitude determination, and a coarse rate is needed for attitude control damping. Classically, angular-rate information is obtained from gyros; however, these days, there is a tendency to build smaller, lighter and cheaper SC where rate accuracy can be compromised. Therefore, the inclination now is to do away with gyros and use other means to determine the SC angular-rate. In fact other means are needed even in gyro-equipped satellites when the angular-rate is out of range of the SC gyros.

There are several ways to obtain the angular-rate in a gyro-less SC. When the attitude is known, one can differentiate the attitude in whatever parameters it is given and use the kinematics equation that connects the derivative of the attitude with the satellite angular-rate in order to compute the latter (ref. 1, 2). However, the differentiation of the attitude introduces a considerable noise component in the computed angular-rate vector. To overcome this noise, one can use an active filter, like a Kalman filter (KF) (ref. 3, 4).

All these methods use the derivative of either the attitude parameters or of the measured directions, which normally determine the attitude parameters. Another approach is that of using the attitude parameters, or the measured directions themselves, as measurements in a KF. The dynamics model of that KF also includes the SC angular dynamics equation, which is a nonlinear first order vector differential equation. In this case, the kinematics equation that connects the attitude parameters, or the directions, with their derivatives are included in the dynamics model used by the filter thereby the need for differentiation is eliminated (ref. 5, 6). The KF dynamics model also includes the SC angular dynamics equation, which is a nonlinear first order vector differential equation.

---

[1] Sophie and William Shamban Professor of Aerospace Engineering.
On sabbatical leave from the Faculty of Aerospace Engineering,
Technion-Israel Institute of Technology. Member Technion Asher
Space Research Institute. NRC NASA Resident Research Associate.
Tel: (301) 286-9216, Fax: (301) 286-036, Email: ibaritz@pop500.gsfc.nasa.gov

[2] Aerospace Engineer. Tel: (301) 286-5125, Fax: (301) 286-0369,
Email: richard.r.harman.1@gsfc.nasa.gov

New sensor packages that uses star trackers and yield the SC attitude in terms of the attitude quaternion recently became available[1]. Therefore, it became possible to use the quaternion supplied by such sensors as measurements. This approach was indeed used in a recent work (ref. 7) where the nonlinear dynamics equation was converted to a *state-dependent pseudo-linear equation* but the measurement equation was linear. The pseudo-linear dynamics model together with he linear quaternion measurement model enabled the use of PSEudo LInear KAlman (PSELIKA) filter.

While the use of quaternion measurement yielded very satisfactory results, the algorithm was limited to one type of instrumentation package, consequently, the popular vector measuring devises like sun sensors, magnetometers, horizon sensors and others could not benefit from the PSELIKA estimator for estimating attitude and angular rate.

In the present work we add a development, which enables the use of the elementary vector observations as filter measurements rather than the quaternions derived from them. On one hand, this approach allows the use of a wide range of instruments to track the SC attitude, but on the other hand, it introduces nonlinear measurement equations. However, as shown in this paper, these measurement equations can be transformed into a state-dependent pseudo-linear equation thereby permitting the use of the PSELIKA filter.

In the next section we briefly describe the development of the filter dynamics equation and, in particular, the development of the state-dependent pseudo-linear SC angular equation. Then, in the following section we present the development of the pseudo-linear measurement equation associated with vector measurements. In the section that follows this one, we present the PSELIKA filter. Next, we present computer runs for vector measurements that, as mentioned, involve pseudo-linear dynamics as well as pseudo-linear measurement equations, and compare them with the results obtained when using quaternion measurements, which involve pseudo-linear dynamics but truly linear measurement equations. Finally, in the last section we present the conclusions derived from this work.

THE STATE-DEPENDENT DYNAMICS EQUATION

The angular dynamics of a constant mass SC is given in the following equation (ref. 8, p. 523)

$$\mathrm{I}\dot{\boldsymbol{\omega}} + \dot{\mathbf{h}} + \boldsymbol{\omega} \times (\mathrm{I}\boldsymbol{\omega} + \mathbf{h}) = \mathbf{T} \tag{1}$$

where I is the SC inertia matrix, $\boldsymbol{\omega}$ is the angular velocity vector, $\mathbf{h}$ is the angular momentum of the momentum wheels, and $\mathbf{T}$ is the external torque acting on the SC. Since the inertia matrix I is invertible we may write this equation as

$$\dot{\boldsymbol{\omega}} = \mathrm{I}^{-1}[(\mathrm{I}\boldsymbol{\omega} + \mathbf{h})\times]\boldsymbol{\omega} + \mathrm{I}^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \tag{2}$$

where $[\mathbf{a}\times]$ denotes the cross product matrix of the general vector $\mathbf{a}$. Defining

$$\mathrm{F}'(\boldsymbol{\omega}) = \mathrm{I}^{-1}[(\mathrm{I}\boldsymbol{\omega} + \mathbf{h})\times] \tag{3.a}$$

and

$$\mathbf{u}(\mathrm{t}) = \mathrm{I}^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \tag{3.b}$$

Eq. (2) can be written as

$$\dot{\boldsymbol{\omega}} = \mathrm{F}'(\boldsymbol{\omega})\boldsymbol{\omega} + \mathbf{u}(\mathrm{t}) \tag{3.c}$$

As was noted by Cloutier, D'Souza, and Mracek (refs. 9, 10) the decomposition of $\mathrm{I}\boldsymbol{\omega} \times \boldsymbol{\omega}$ into $[\mathrm{I}\boldsymbol{\omega}\times]\boldsymbol{\omega}$ is not unique. In fact there are eight possible ways to decompose the vector (ref. 11), consequently there are eight primary

[1] Van Bezooijen, R.W.H., "AST Capabilities," Lockheed Martin Advanced Technology Center, Palo Alto, CA 95304-1191. (Slide presentation).

dynamics matrices which express the angular dynamics of an SC; that is, we have $F_i'(\omega)\quad i=1,2,....,8$ for which $F_1'(\omega)\omega = F_2'(\omega)\omega = ... = F_8'(\omega)\omega$ but

$$F_1'(\omega) \neq F_2'(\omega) \neq ..... \neq F_8'(\omega) \tag{4}$$

Although there are only eight primary representations, one can generate infinite *secondary* dynamics matrices as linear combinations of the primary matrices by forming (refs. 9, 10)

$$E_i'(\omega) = \sum_{j=1}^{8} \alpha_{i,j} F_j'(\omega) \qquad \text{where} \qquad \sum_{j=1}^{8} \alpha_{i,j} = 1 \qquad i = 1,2,...,\rightarrow \infty \tag{5}$$

This can be easily proven when noting that

$$E_i'(\omega)\omega = \sum_{j=1}^{8} \alpha_{i,j} F_j'(\omega)\omega = \sum_{j=1}^{8} \alpha_{i,j}\mathbf{f}(\omega) = \mathbf{f}(\omega)\sum_{j=1}^{8} \alpha_{i,j} = \mathbf{f}(\omega) \tag{6}$$

where $\mathbf{f}(\boldsymbol{\omega}) = I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times]\boldsymbol{\omega}$. We conclude then that the representation of the SC dynamics equation in the form

$$\dot{\mathbf{x}} = F_\omega(\mathbf{x})\mathbf{x} + \mathbf{u}_\omega(t) \tag{7}$$

is not unique, and that there are exactly eight different ways to express the nonlinear SC dynamics equation by basic state-dependent linear equations, and infinite secondary such equations.

The attitude is best described by the quaternion of rotation. The quaternion dynamics equation is (ref. 8, p. 512)

$$\dot{\mathbf{q}} = \tfrac{1}{2} Q\omega \tag{8}$$

where

$$Q = \begin{bmatrix} q_4 & -q_3 & q_2 \\ q_3 & q_4 & -q_1 \\ -q_2 & q_1 & q_4 \\ -q_1 & -q_2 & -q_3 \end{bmatrix} \tag{9}$$

Grouping Eqs. (2) and (8) yields

$$\begin{bmatrix} \dot{\boldsymbol{\omega}} \\ \dot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times] & 0 \\ \tfrac{1}{2}Q & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \begin{bmatrix} I^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \\ 0 \end{bmatrix} \tag{10}$$

In order to use the last equation as a dynamics equation in a KF, we need to add to it white noise to account for modeling inaccuracies. This results in the following dynamics equation.

$$\begin{bmatrix} \dot{\boldsymbol{\omega}} \\ \dot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times] & 0 \\ \tfrac{1}{2}Q & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \begin{bmatrix} I^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{w}_\omega \\ \mathbf{w}_q \end{bmatrix} \tag{11}$$

where $\mathbf{w}_\omega$ accounts for the inaccuracies in the modeling of the SC angular dynamics and $\mathbf{w}_q$ accounts for modeling errors of the quaternion dynamics.

THE STATE-DEPENDENT MEASURMENT EQUATION

As mentioned earlier, a measurement system has been introduced recently[1], which yields the attitude directly in a form of a quaternion of rotation. In this case the measurement equation is linear; namely,

$$\mathbf{q}_m = \begin{bmatrix} 0_{4x3} & I_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \mathbf{v}_q \tag{12}$$

where $\mathbf{q}_m$ is the measured quaternion and $\mathbf{v}_q$ is the noise associated with this "measurement". The actual measurements are vector measurements obtained from star trackers out of which the "measured" quaternion is computed. Traditionally though, it is the vector measurements which are used directly to estimate the attitude. In this case the link between the measurements and the attitude expressed by the quaternion of rotation is a nonlinear one. However, this nonlinear relationship can too be expressed as a state-dependent pseudo-linear function. This is shown next. Let $\mathbf{r}$ denote the measured vector expressed in the reference coordinate system and let $\mathbf{b}$ denote the measurement of this vector as obtained in body coordinates. The relationship between the two vectors is expressed by

$$\mathbf{b} = D\mathbf{r} + \mathbf{v} \tag{13}$$

where D is the direction cosine matrix that transforms vectors from the reference to the body coordinates, and $\mathbf{v}$ denotes the zero mean white noise associated with this measurement. It is well known that D is the following function of the quaternion elements (ref. 8, p. 414)

$$D = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \tag{14}$$

when the last expression is substituted into Eq. (13) we obtain

$$\mathbf{b} = \begin{bmatrix} (q_1^2 - q_2^2 - q_3^2 + q_4^2)r_1 + 2(q_1q_2 + q_3q_4)r_2 + 2(q_1q_3 - q_2q_4)r_3 \\ 2(q_1q_2 - q_3q_4)r_1 + (-q_1^2 + q_2^2 - q_3^2 + q_4^2)r_2 + 2(q_2q_3 + q_1q_4)r_3 \\ 2(q_1q_3 + q_2q_4)r_1 + 2(q_2q_3 - q_1q_4)r_2 + (-q_1^2 - q_2^2 + q_3^2 + q_4^2)r_3 \end{bmatrix} + \mathbf{v} \tag{15}$$

An expansion and rearrangement of the terms in the last equation yields

$$\mathbf{b} = \begin{bmatrix} q_1^2r_1 + q_1q_2r_2 + q_1q_3r_3 - q_2^2r_1 + q_1q_2r_2 - q_2q_4r_3 - q_3^2r_1 + q_3q_4r_2 + q_1q_3r_3 + q_4^2r_1 + q_3q_4r_2 - q_2q_4r_3 \\ q_1q_2r_1 - q_1^2r_2 + q_1q_4r_3 + q_1q_2r_1 + q_2^2r_2 + q_2q_3r_3 - q_3q_4r_1 - q_3^2r_2 + q_2q_3r_3 - q_3q_4r_1 + q_4^2r_2 + q_1q_4r_3 \\ q_1q_3r_1 - q_1q_4r_2 - q_1^2r_3 + q_2q_4r_1 + q_2q_3r_2 - q_2^2r_3 + q_1q_3r_1 + q_2q_3r_2 + q_3^2r_3 + q_2q_4r_1 - q_1q_4r_2 + q_4^2r_3 \end{bmatrix} + \mathbf{v} \tag{16}$$

which can be written as

[1] Van Bezooijen, R.W.H., "AST Capabilities," Lockheed Martin Advanced Technology Center, Palo Alto, CA 95304-1191. (Slide presentation).

$$\mathbf{b} = \begin{bmatrix} q_1r_1 + q_2r_2 + q_3r_3 & -q_2r_1 + q_1r_2 - q_4r_3 & -q_3r_1 + q_4r_2 + q_1r_3 & q_4r_1 + q_3r_2 - q_2r_3 \\ q_2r_1 - q_1r_2 + q_4r_3 & q_1r_1 + q_2r_2 + q_3r_3 & -q_4r_1 - q_3r_2 + q_2r_3 & -q_3r_1 + q_4r_2 + q_1r_3 \\ q_3r_1 - q_4r_2 - q_1r_3 & q_4r_1 + q_3r_2 - q_2r_3 & q_1r_1 + q_2r_2 + q_3r_3 & q_2r_1 - q_1r_2 + q_4r_3 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} + \mathbf{v} \quad \ldots (17)$$

Let

$$H(\mathbf{r},\mathbf{q}) = \begin{bmatrix} q_1r_1 + q_2r_2 + q_3r_3 & -q_2r_1 + q_1r_2 - q_4r_3 & -q_3r_1 + q_4r_2 + q_1r_3 & q_4r_1 + q_3r_2 - q_2r_3 \\ q_2r_1 - q_1r_2 + q_4r_3 & q_1r_1 + q_2r_2 + q_3r_3 & -q_4r_1 - q_3r_2 + q_2r_3 & -q_3r_1 + q_4r_2 + q_1r_3 \\ q_3r_1 - q_4r_2 - q_1r_3 & q_4r_1 + q_3r_2 - q_2r_3 & q_1r_1 + q_2r_2 + q_3r_3 & q_2r_1 - q_1r_2 + q_4r_3 \end{bmatrix} \quad \ldots (18)$$

then Eq. (17) can be written as

$$\mathbf{b} = H(\mathbf{r},\mathbf{q})\mathbf{q} + \mathbf{v} \tag{19}$$

which can be written as

$$\mathbf{b} = \begin{bmatrix} 0 & H(\mathbf{r},\mathbf{q}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \mathbf{v} \tag{20}$$

We have thus succeeded in expressing the vector measurement equation as pseudo-linear equation.

THE PDEUDO-LINEAR KALMAN FILTER

The pseudo linear KF (PSELIKA) is an ordinary linear Kalman filter where the state-dependent coefficients are functions of the best available state estimate. Let

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} \tag{21.a}$$

$$F(\boldsymbol{\omega}) = \begin{bmatrix} I^{-1}[(I\boldsymbol{\omega} + \mathbf{h})\times] & 0 \\ \frac{1}{2}Q & 0 \end{bmatrix} \tag{21.b}$$

$$\mathbf{u}(t) = \begin{bmatrix} I^{-1}(\mathbf{T} - \dot{\mathbf{h}}) \\ 0 \end{bmatrix} \tag{21.c}$$

$$\mathbf{w}(t) = \begin{bmatrix} \mathbf{w}_\omega \\ \mathbf{w}_q \end{bmatrix} \tag{21.d}$$

then Eq. (11) can be written as

$$\dot{\mathbf{x}} = F(\boldsymbol{\omega})\mathbf{x} + \mathbf{u}(t) + \mathbf{w}(t) \tag{22}$$

From Eq. (12) we obtain the following corresponding *quaternion* measurement equation

$$\mathbf{q}_{m,k+1} = H^q_{k+1} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \mathbf{v}_{q,k+1} \tag{23.a}$$

where

$$H^q_{k+1} = \begin{bmatrix} 0_{4x3} & I_4 \end{bmatrix} \tag{23.b}$$

and from Eq. (20) we obtain the following corresponding *vector* measurement equation

$$\mathbf{b}_{m,k+1} = H^b_{k+1}(\mathbf{r},\mathbf{q}) \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \mathbf{v}_{b,k+1} \tag{23.c}$$

where

$$H^b_{k+1} = \begin{bmatrix} 0 & H(\mathbf{r},\mathbf{q}) \end{bmatrix} \tag{23.d}$$

Filter initialization:

Compute:

$$Q_k = E\{\mathbf{w}(t_k)\mathbf{w}(t_k)^T\} \tag{24.a}$$

$$R_{q,k+1} = E\{v_{q,k+1}v_{q,k+1}^T\} \tag{24.b}$$

$$R_{b,k+1} = E\{v_{b,k+1}v_{b,k+1}^T\} \tag{24.c}$$

and choose an approximate value for the initial estimate of the state vector. In the absence of such initial estimate, choose:

$$\mathbf{x}_0^T = \begin{bmatrix}\boldsymbol{\omega}_0^T & \mathbf{q}_0^T\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix} \tag{24.d}$$

Time propagation:

Solve simultaneously the differential equations

$$\dot{\hat{\mathbf{x}}} = F(\hat{\mathbf{x}})\hat{\mathbf{x}} + \mathbf{u}(t) \tag{25.a}$$

$$\dot{P} = F(\hat{\mathbf{x}})P + PF^T(\hat{\mathbf{x}}) + Q \tag{25.b}$$

between t_k and t_{k+1} using the definition $\hat{\mathbf{x}}(t_k) = \hat{\mathbf{x}}_{k/k}$, $P(t_k) = P_{k/k}$, $\hat{\mathbf{x}}_{k+1/k} = \hat{\mathbf{x}}(t_{k+1})$, and $P_{k+1/k} = P(t_{k+1})$. (The subscript q/p denotes the estimate at time t_q based on the measurements up to time t_p).

Measurement update:

For <u>quaternion</u> measurements:

$$K_{k+1} = P_{k+1/k}H_{k+1}^{qT}[H_{k+1}^q P_{k+1/k}H_{k+1}^{qT} + R_{q,k+1}]^{-1} \tag{26.a}$$

$$\hat{\mathbf{x}}_{k+1/k+1} = \hat{\mathbf{x}}_{k+1/k} + K_{k+1}[\mathbf{q}_{m,k+1} - H_{k+1}^q\hat{\mathbf{x}}_{k+1/k}] \tag{26.b}$$

$$P_{k+1/k+1} = [I - K_{k+1}H_{k+1}^q]P_{k+1/k}[I - K_{k+1}H_{k+1}^q]^T + K_{k+1}R_{q,k+1}K_{k+1}^T \tag{26.c}$$

For <u>vector</u> measurements:

$$K_{k+1} = P_{k+1/k}H_{k+1}^{bT}(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k})[H_{k+1}^b(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k})P_{k+1/k}H_{k+1}^{bT}(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k}) + R_{b,k+1}]^{-1} \tag{27.a}$$

$$\hat{\mathbf{x}}_{k+1/k+1} = \hat{\mathbf{x}}_{k+1/k} + K_{k+1}[\mathbf{b}_{m,k+1} - H_{k+1}^b(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k})\hat{\mathbf{x}}_{k+1/k}] \tag{27.b}$$

$$P_{k+1/k+1} = [I - K_{k+1}H_{k+1}^b(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k})]P_{k+1/k}[I - K_{k+1}H_{k+1}^b(\mathbf{r}_{k+1}, \hat{\mathbf{q}}_{k+1/k})]^T + K_{k+1}R_{b,k+1}K_{k+1}^T \tag{27.c}$$

SIMULATION RESULTS

In order to test the PSELIKA filter a simulation program was constructed. The simulation included a given profile of the wheel momentum, $\mathbf{h}(t)$, that generated the desired true angular velocity vector. Based on that data the corresponding true SC attitude was computed in terms of the true quaternion. White measurement noise was added to the true quaternion and the resultant noisy quaternion served as the measured signal for the case where the PSELIKA filter used measured quaternions to estimate the SC attitude and angular rate. The noise vector was zero mean and white. Its standard deviation was 0.00001 for each component. Fig. 1 presents the true and estimated angular velocity

components. The estimation error of the angular velocity for this case is presented in Fig. 2. Since the measured quantity was the quaternion itself, the attitude errors were on the level of the noise. They were, of course, very small.

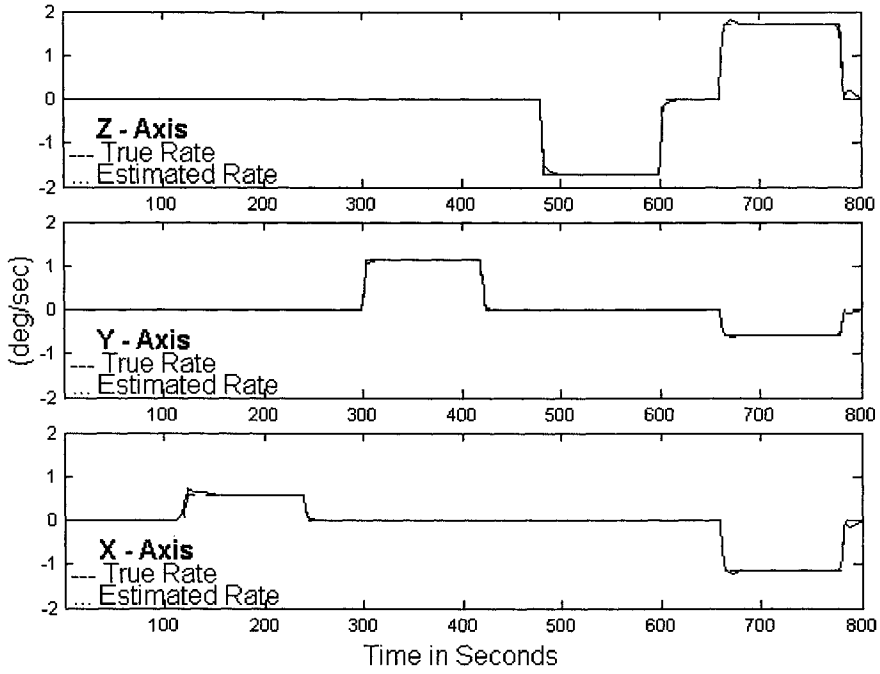


Fig. 1: True and Estimated SC Angular Rate Using Quaternion Measurements.

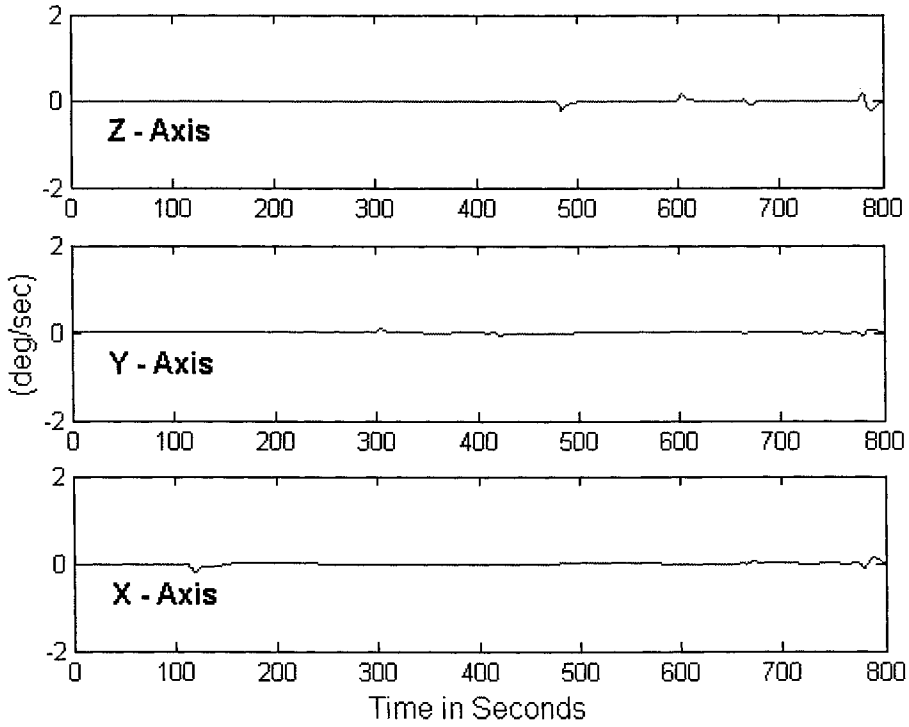


Fig. 2: Rate Estimation Errors Using Quaternion Measurements.

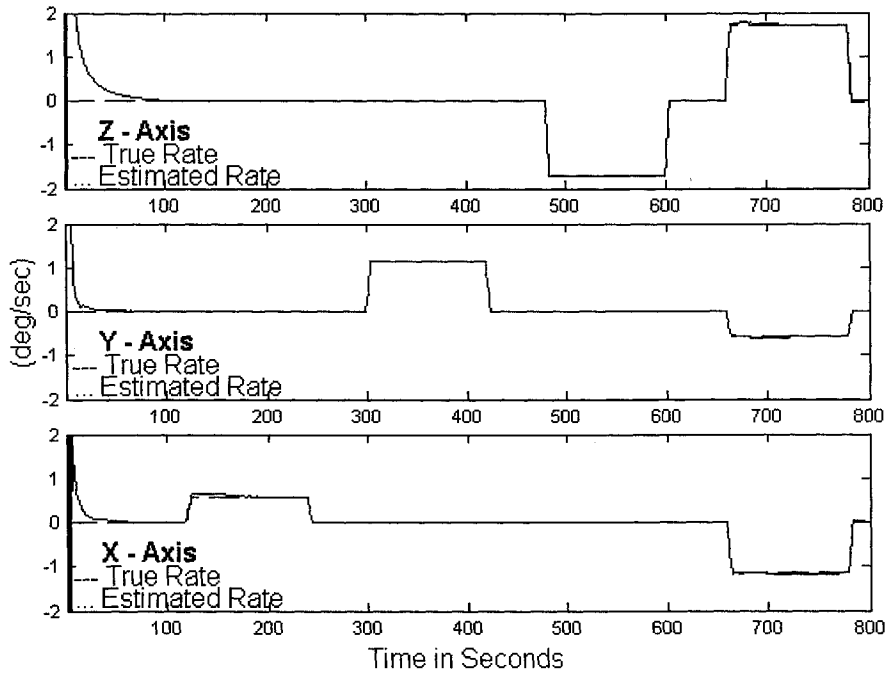


Fig. 3: True and Estimated SC Angular Rate Using Two Simultaneous Vector Measurements.

In order to simulate the case where the estimator used measured vectors, the following two reference vectors were assumed

$$\begin{aligned} \mathbf{r}_1^T &= [1 \quad 0 \quad 0] \\ \mathbf{r}_2^T &= [0 \quad 1 \quad 0] \end{aligned} \tag{28}$$

They represented sun sensor and magnetometer measurements, respectively. The corresponding vectors in body coordinates, $\mathbf{b}_1$ and $\mathbf{b}_2$, were computed by transforming $\mathbf{r}_1$ and $\mathbf{r}_2$ by the true attitude matrix. White measurement noise was added to the true $\mathbf{b}_1$ and $\mathbf{b}_2$ vector. The noise was zero mean and its standard deviation was 0.01 per component of $\mathbf{b}_1$ and 0.005 per component of $\mathbf{b}_2$. When the two measurements were used in a sequential manner at each update time point, the filter diverged. It was speculated that although in the truly linear case performing the two sequential measurement updates is identical to performing the two updates simultaneously, in the pseudo-linear case, where the matrices involved in the updates are functions of the estimate, there could be a difference between the two possible update methods. This is particularly true when estimating attitude using vectors because two vector measurements, determine attitude whereas one vector measurement is not sufficient for that. As a result of the foregoing discussion, the two measurements were stacked together to form the following measurement equation (see Eq. 23.c)

$$\begin{bmatrix} \mathbf{b}_{1,m,k+1} \\ \mathbf{b}_{2,m,k+1} \end{bmatrix} = \begin{bmatrix} H^b_{1,k+1}[\mathbf{r}, \hat{\mathbf{q}}_{k+1}(-)] \\ H^b_{2,k+1}[\mathbf{r}, \hat{\mathbf{q}}_{k+1}(-)] \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \mathbf{q} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{1,b,k+1} \\ \mathbf{v}_{2,b,k+1} \end{bmatrix} \tag{29}$$

Indeed, using the two vector measurements together in the filter updates, the filter converged and produced satisfactory results. Figure 3 presents the true and estimated rates obtained in this case, whereas Fig. 4 presents the angular rate estimation error for this case. Finally, in Fig. 5 the attitude estimation error is presented for this case of simultaneous vector measurements.

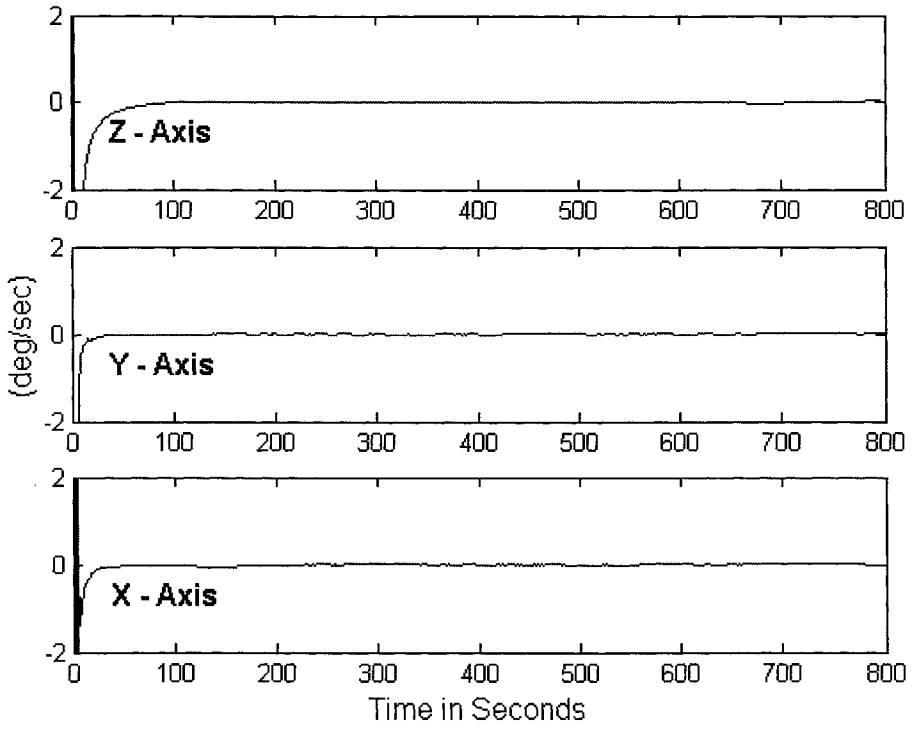


Fig. 4: Angular Rate Estimation Errors Using Two Simultaneous Vector Measurements.

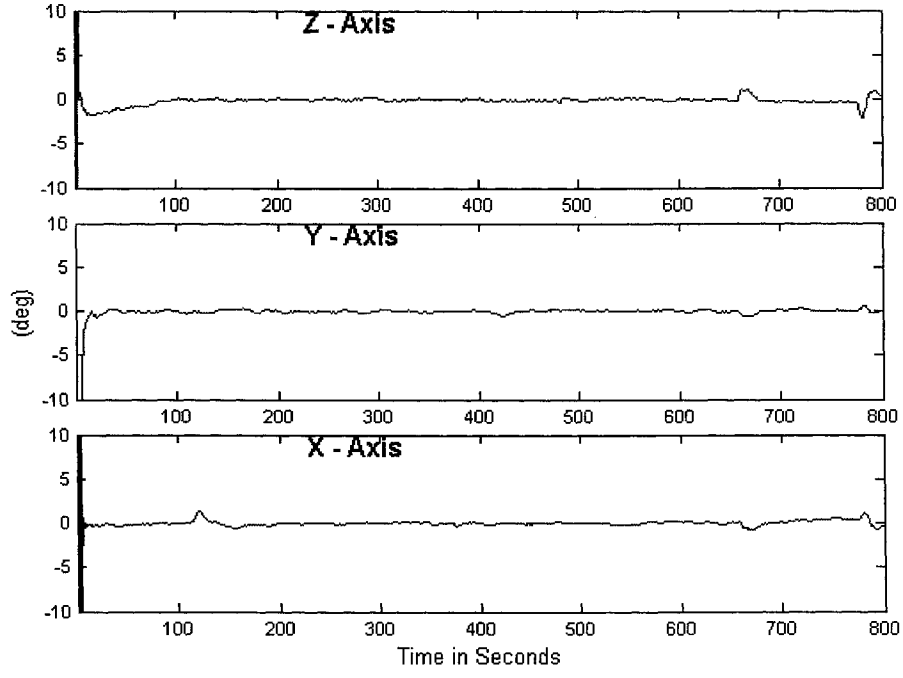


Fig. 5: Attitude Estimation Errors Using Two Simultaneous Vector Measurements.

CONCLUSIONS

In this work we presented two versions of a pseudo-linear Kalman filter for spacecraft attitude and rate estimation, which is based on the ability to express the nonlinear parts of the system model as state-dependent linear models. This formulation of the model enabled the evaluation of the state-dependent matrices in terms of the best available state estimate, and thus to express the system model as deterministic matrices multiplying random vectors.

In the first version of the filter, the measurement was that of the attitude quaternion, which is related linearly to the estimated state vector; therefore, only the nonlinear dynamics model of the spacecraft rotation had to be formulated in a state-dependent linear model. The second version treated a more general case where the measurements were vector measurements, in which case, the measurement model too was nonlinear and had to be formulated as a state dependent linear model.

Simulations were performed which showed that, as expected, the first filter whose quaternion measurements were based on star tracker measurements yielded very good rate estimates and exceptionally good attitude results. The second filter, though, was unable to produce good results when the vector measurements were performed sequentially. However, when the two vector measurements were performed simultaneously the filter performed satisfactorily. The results were, of course, less accurate than with the previous algorithm, but that was because the simulated measurements supplied by a sun sensor and magnetometers whose inherent accuracy is less than that of star trackers which supplied the quaternion measurement to the first algorithm.

REFERENCES

1. Flately, T.W., Forden, J.K., Henretty, D.A., Lightsey, E.G., and Markley, F.L., On-Board Attitude Determination and Control Algorithm or SAMPEX," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA Conference Publication 3102, NASA Goddard Space Flight Center, Greenbelt, MD,1990, pp. 379-398.

2. Harman, R.R., and Bar-Itzhack, I.Y., "Angular-rate Computation Using Attitude Differentiation," Faculty of Aerospace Engineering, Technion-Israel Institute of Technology, TAE-881, Haifa 32000, Israel, Dec. 97.

3. Harman, R.R., and Bar-Itzhack, I.Y., "Pseudo-Linear and State-Dependent Riccati Equation Filters for Angular Rate Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 22, No. 5, Sept.-Oct. 1999, pp. 723-725.

4. Azor, R., Bar-Itzhack, I. Y., and Harman, R. R., "Satellite Angular-rate Estimation from Vector Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 3, 1998, pp. 450-457.

5. Oshman, Y., and Markley, F. L., "Sequential Attitude and Attitude-Rate Estimation Using Integrated-Rate Parameters," *Journal of Guidance, Control, and Dynamics*, Vol. 22, No. 3, 1999, pp. 385-394.

6. Dellus, F., "Estimation of Satellite Angular Velocity Using Sequential Measurements of a Single Inertial Vector," M.Sc. Thesis, Faculty of Aerospace Engineering, Technion-Israel Institute of Technology, Haifa, Israel, June 1998 (In Hebrew).

7. Azor, R., Bar-Itzhack, I.Y., Deutschmann, J.K., and Harman, R. R., "Angular-rate Estimation Using Delayed Quaternion Measurements," *Journal of Guidance, Control, and Dynamics* (accepted for publication).

8. Wertz, J. R., (Ed.), *Spacecraft Attitude Dynamics and Control*, Reidel Publishing Co., Dordrecht, Holland, 1978.

9. Cloutier, J.R., D'Souza, C.N., and Mracek, C.P., "Nonlinear Regulation and Nonlinear H_∞ Control Via the State-Dependent Riccati Equation Technique: Part 1, Theory," The First International Conference on Nonlinear Problems in Aviation and Aerospace, Daytona Beach, FL, May 1996.

10. Cloutier, J.R., D'Souza, C.N., and Mracek, C.P., "Nonlinear Regulation and Nonlinear H_∞ Control Via the State-Dependent Riccati Equation Technique: Part 2, Examples," The First International Conference on Nonlinear Problems in Aviation and Aerospace, Daytona Beach, FL, May 1996.

11. Harman, R.R., and Bar-Itzhack, I.Y., "Pseudolinear and State-Dependent Riccati Equation Filters for Angular Rate Estimation," *J. of Guidance, Control, and Dynamics*, Vol. 22, No. 5, September-October, 1999, pp. 723-725.

Session 4: Attitude Control System Design and Simulation

Triana Safehold: A New Gyroless, Sun-Pointing Attitude Controller

J. (Roger) Chen (K&D Research), Wendy Morgenstern (GSFC), Joseph Garrick (GSFC)

ABSTRACT

Triana is a single-string spacecraft to be placed in a halo orbit about the sun-earth L1 Lagrangian point. The Attitude Control Subsystem (ACS) hardware includes four reaction wheels, ten thrusters, six coarse sun sensors, a star tracker and a three-axis Inertial Measuring Unit (IMU). The ACS Safehold design features a gyroless sun-pointing control scheme using only sun sensors and wheels. With this minimum hardware approach, Safehold increases mission reliability in the event of a gyroscope anomaly. In place of the gyroscope rate measurements, Triana Safehold uses wheel tachometers to help provide a scaled estimation of the spacecraft body rate about the sun vector. Since Triana nominally performs momentum management every three months, its accumulated system momentum can reach a significant fraction of the wheel capacity. It is therefore a requirement for Safehold to maintain a sun-pointing attitude even when the spacecraft system momentum is reasonably large.

The tachometer sun-line rate estimation enables the controller to bring the spacecraft close to its desired sun-pointing attitude even with reasonably high system momentum and wheel drags. This paper presents the design rationale behind this gyroless controller, stability analysis, and some time-domain simulation results showing performances with various initial conditions. Finally, suggestions for future improvements are briefly discussed.

INTRODUCTION

Triana is a mission dedicated to helping scientists construct more accurate models of the Earth's climate and to examining how solar radiation affects our climate. Triana is a single-string spacecraft destined for a Lissajous orbit about the Lagrangian point between the Earth and the Sun. This unique vantage point will provide Triana with a constant, sunlight view of the Earth.

Triana's on-board ACS software includes five different control modes: Science, Delta-V, Delta-H, Sun Acquisition and Safehold. During Science observations, the Triana ACS provides three axis stabilization with an accuracy of several arcminutes. Delta-V modifies the orbit trajectory, and Delta-H manages the spacecraft momentum. Sun Acquisition establishes a sun-pointing attitude. Safehold also establishes a thermal and power safe attitude, but it does so using a minimum hardware complement. This paper will focus on the design, implementation, and performance of Triana's gyroless Safehold control algorithm.

The Triana spacecraft, shown in Figure 1, uses a variety of sensors and actuators. Most of the hardware connects directly to the 1553 databus. A Honeywell Ring Laser Gyroscope (RLG) IMU provides rate information in all three axes. The star tracker is a Ball CT-633, which outputs an attitude quaternion. Triana includes two types of sun sensors. Six Adcole Coarse Sun Sensors (CSS) are arranged into three opposing pairs along three orthogonal axes for full sky coverage. Also on-board is an Adcole Digital Sun Senor (DSS), which provides partial sky coverage with a ±64 degree field of view and its boresight aligned along the sun-side solar array normal. The CSS and DSS are connected to the 1553 database through an interface known as the Utility Hub. Control actuation is provided by either reaction wheels or thrusters. Ten Kaiser-Marquardt one pound thrusters are arranged to provide both pure torque couples for attitude control and pure forces for changes in the orbit velocity. The thrusters are controlled by an Engine Valve Driver (EVD) card, which is also connected to the spacecraft databus via the Utility Hub. Four Integrated Reaction Wheel Assemblies (IRWA) are arranged in a pyramid configuration with the pyramid axis along the solar array normal.

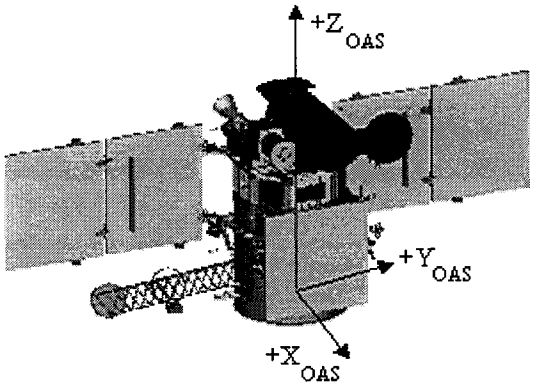


Figure 1: Fully Deployed Triana Spacecraft

Triana Safehold is a thermal and power safe sun pointing mode, based on a Proportional-Derative (PD) controller providing commands to the IRWAs, similar to the Safehold mode for the MAP spacecraft.[1] What makes the Triana Safehold unique is its derivation of rate information. Since Triana does not carry redundant gyros, it is desirable to

have a Safehold mode, capable of safing the spacecraft in the event of a gyro anomaly. CSS data provides the attitude errors for Safehold and, through their differentiation, body rates for the axes perpendicular to the sun-line. However, the rate about the sun-line is unobservable with the CSS. Instead, the wheel tach information is used to derive a rate about the commanded sun-line.

CONTROL LAWS

Two-Axis Control

A control law for commanding three or four wheels with only 2-axis attitude knowledge afforded by the sun sensors plays the primary role in Triana's Safehold control. This control law is

$$\mathbf{T_c} = \mathbf{I}\{\mathbf{k_p}\,\mathrm{Limit}(\theta_e) + \mathbf{k_v}[(\frac{\mathbf{d}}{\mathbf{dt}}\mathbf{s}) \times \mathbf{s}]\}$$

$$\theta_e = \mathbf{s} \times \mathbf{s_d}$$

$$\mathbf{T_{wc}} = \mathbf{W}^+\mathbf{T_c} + \mathbf{T_{mr}} + \mathbf{T_{dc}}$$

where the variables are

$\mathbf{T_c}$	: commanded net wheel torque vector
$\mathbf{T_{wc}}$	: column matrix (n by 1) consisting of individual wheel torque commands and n is the total number of wheels under control
$\mathbf{I}$	: spacecraft inertia matrix (3 by 3)
$\mathbf{s}$	: unit sun vector
$\mathbf{s_d}$	: desired unit sun vector (nominally along body -X axis for Triana)
$\mathbf{k_p}$,$\mathbf{k_v}$	: diagonal gain matrices (3 by 3)
$\mathbf{T_{mr}}$,$\mathbf{T_{dc}}$	: column matrices (n by 1) consisting of wheel momentum redistribution and drag compensation torque commands for individual wheels
$\mathbf{W}^+$	: pseudo-inverse (n by 3) of the wheel mounting matrix (3 by n), each column of which comprises the unit vectors along the positive spin axis of a wheel

Here and throughout this paper, a vector variable generally refers to a 3 by 1 matrix representation of a "basis-independent vector" in the unit vector basis fixed in the spacecraft's main body. Occasionally, a vector in a term will refer to a "basis-independent vector" itself and the term will be enclosed in quotation marks in this situation.

The "Limit" function in the above control law limits the absolute value of each element of the error angle vector θ_e to some predefined parameter value and preserves its sign. The error rate vector, being the cross product of the time derivative of $\mathbf{s}$ and $\mathbf{s}$, is approximately equal to ω, the angular velocity vector of the spacecraft in inertial frame, without the component in the direction of $\mathbf{s}$. This is the case because the unit "sun vector" is changing its direction very slowly in an inertial frame and therefore the derivative of $\mathbf{s}$ is approximately equal to $-\omega \times \mathbf{s}$. In this control law, momentum redistribution torque command matrix is used to minimize the maximum value of individual wheel momenta. Wheel drag compensation torque command matrix is meant to cancel a portion of the anticipated drag torque of each wheel so that the actual torque applied to each wheel is closer to what is commanded.

With this 2-axis control, a desired axis fixed in the spacecraft body represented by $\mathbf{s_d}$ can be reoriented toward the sun from any arbitrary starting orientation. In ideal conditions this control law will align $\mathbf{s}$ with $\mathbf{s_d}$ in the steady state. The ideal conditions are those in which either the wheel drag torque is zero, which is physically impossible, or the wheel drag is exactly compensated. Figure 2 shows some response curves of the sun angle for Triana under the 2-axis control in the ideal condition. Here and throughout this paper, sun angle refers to the angle between $\mathbf{s_d}$ and $\mathbf{s}$. Achieving a small steady-state sun angle from arbitrary initial orientation is thus the objective of a Safehold control. For Triana, the Safehold controller must reduce the sun angle to 15 degrees within 15 minutes. The sample responses seem to indicate the efficacy of the 2-axis control in achieving the desired goal of Safehold.

Results displayed in this section are all generated using a low-fidelity simulation with the control interval of 0.1 seconds and 1 cycle of computation delay consistent with Triana ACS implementation.

Even though this 2-axis control computes a 3-axis wheel torque command in each control cycle, only 2-axis measurements of the unit sun vector provide the feedback. Without the measurement in either angle or rate about the

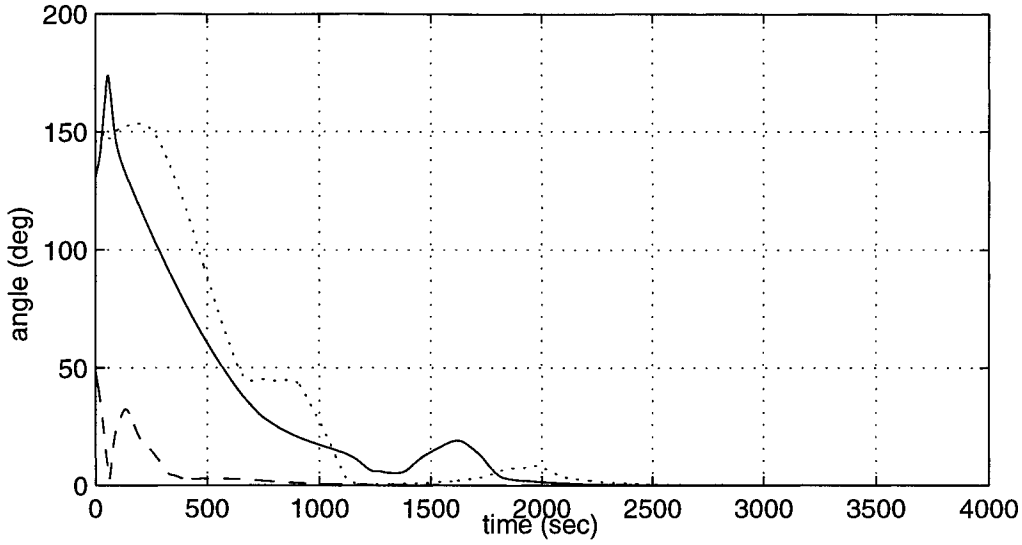


Figure 2: Sample Sun Angle Results Due to 2-Axis Control in Ideal Conditions

sun line, the system is not completely observable. This lack of feedback for one axis results in degraded sun pointing performance when the uncompensated wheel drag is included in the simulation. Figure 3 displays the simulation results of a case identical to the case yielding the solid line curve in Figure 2 except that the uncompensated wheel drags have some realistic non-zero values.

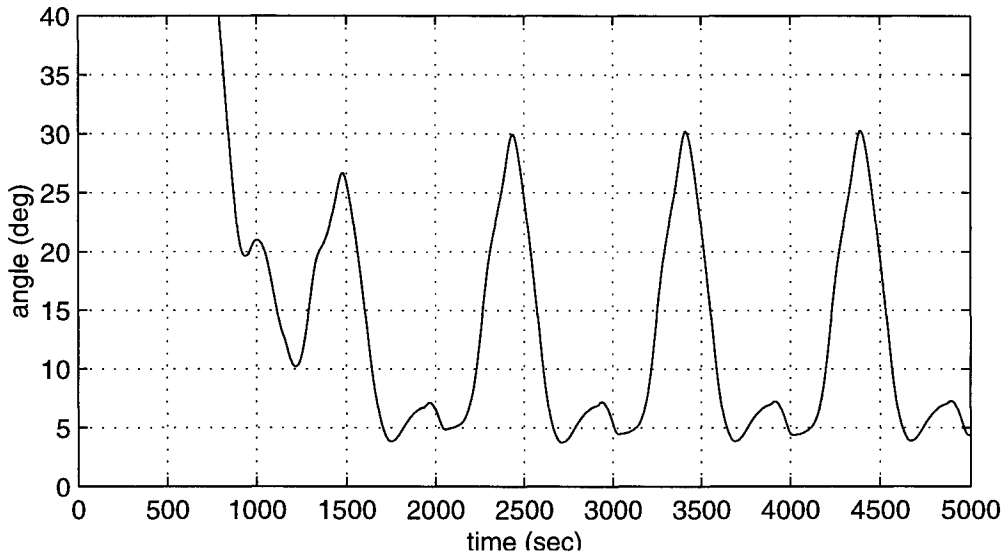


Figure 3: Sun Angle For Triana under 2-Axis Control in Non-ideal Condition

Not only does the steady-state sun angle deteriorate to a level beyond the 15 degree requirement, but also the wheel momenta and the wheel torque commands in Figures 4 and 5, respectively, show significant wheel activities in steady state. For Triana, reasonably high system momentum can occur due to the desire that the frequency of interruption caused by momentum management operations is minimized. Here system momentum refers to the magnitude of the system momentum vector, which is the sum of all inertial angular momentum of the spacecraft system with respect to the overall center of mass. Both high system momentum and uncompensated wheel drags can cause performance degradation; the greater their values are the worse the steady-state sun angle becomes. It is therefore desirable to have some enhancement of this 2-axis control that can provide improved steady-state sun angle performance.

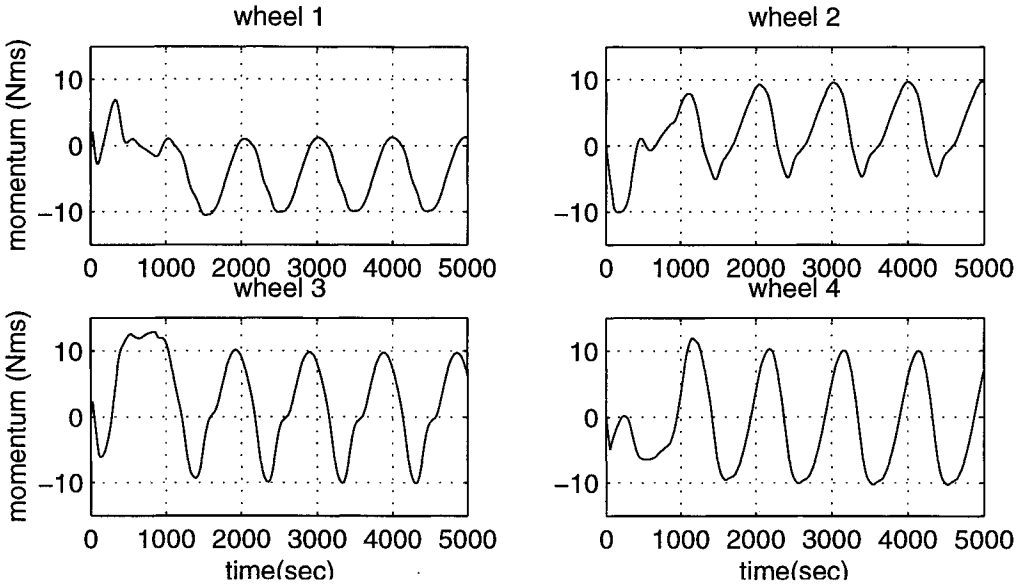


Figure 4: Wheel Momenta for Triana under 2-Axis control in Non-ideal Conditions

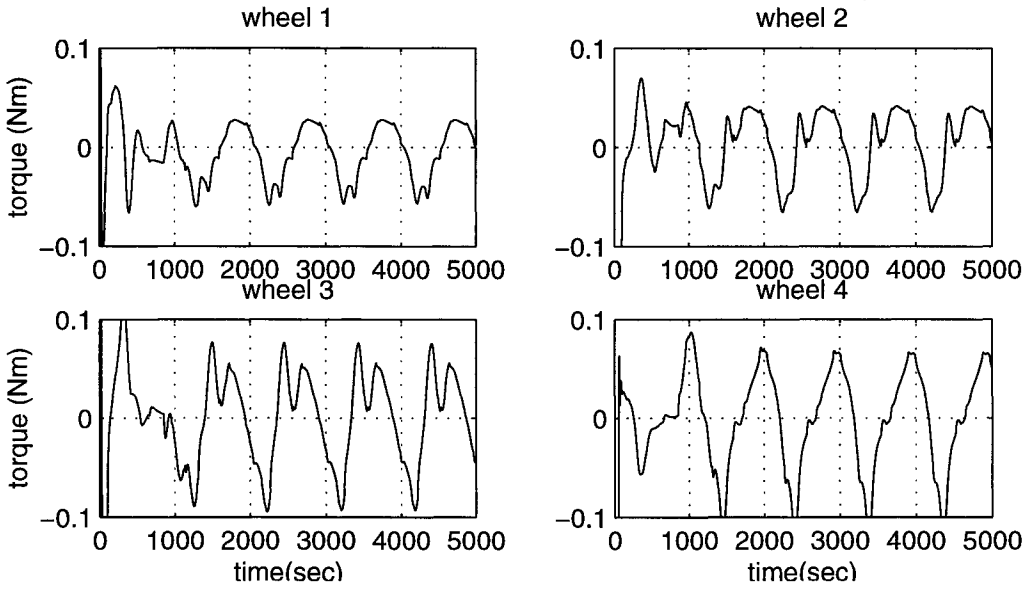


Figure 5: Wheel Torque Commands For Triana under 2-Axis Control in Non-ideal Conditions

Sun-Line Rate

Figure 4 shows that momenta of all four wheels in steady state exhibit slightly distorted sinusoidal fluctuations with the same frequency. These fluctuations imply that the net wheel momentum vector **h**, which is the sum of the angular momentum vectors of individual wheels relative to the spacecraft body, should also show similar distorted sinusoidal fluctuations. In fact, h_y and h_z, the Y and Z elements of **h**, respectively, also exhibit fluctuations with the same frequencies as can be seen in Figure 6. Notice that the desired sun line for Triana is the –X axis and thus the Y and Z axes are normal to the desired sun line. Since these two near sinusoids display roughly constant phase relations and similar magnitude, the curve obtained when one is plotted against another should be a distorted circle. Shown in Figure 7 are two traces: h_z vs. h_y and H_z vs. H_y, where H_y and H_z are, respectively, the Y and Z elements of

H, the 3 by 1 system momentum vector. Also shown on the two traces are markers corresponding to equally separated time points for the two traces. It can easily be observed from the markers that the component in YZ plane of **h** is tracking closely that of **H** in steady state. Notice that these traces repeat as time goes on in steady state; traces of any two dynamic variables when plotted one against the other will also repeat. This steady-state behavior is referred to as a periodic orbit in state space.[2]

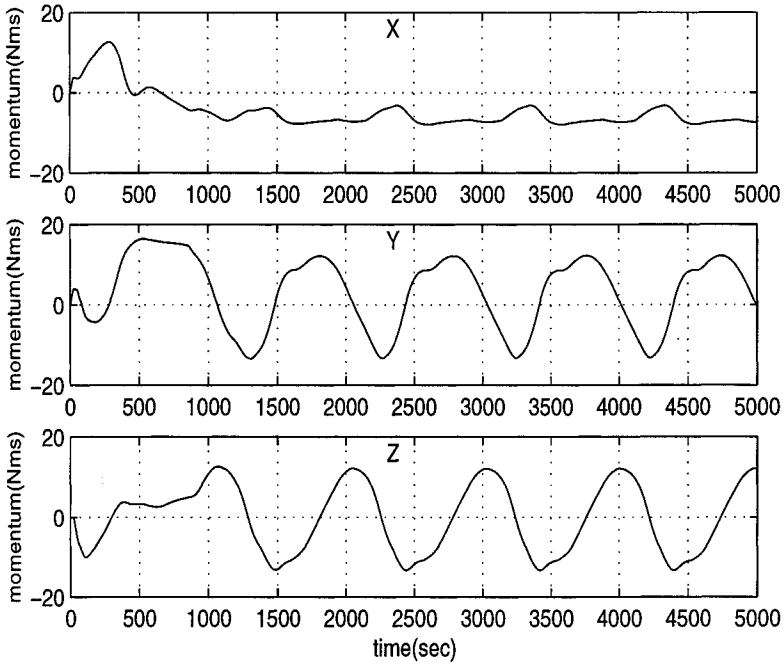


Figure 6: Net Wheel Momentum Vector Measurements under 2-Axis Control in Non-ideal Conditions

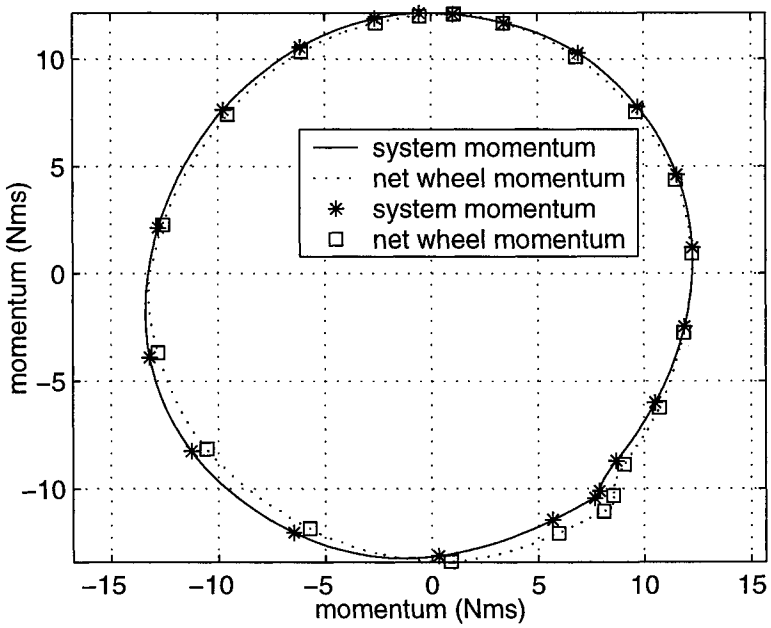


Figure 7: Z vs. Y Momentum under 2-Axis Control in Non-ideal Conditions

This tracking of momentum vector components in steady state can be explained with the use of the system momentum equation, $\mathbf{I}\omega + \mathbf{h} = \mathbf{H}$. Since external disturbance torque for Triana is very small, the "system momentum vector" can be considered as fixed in an inertial frame for the duration of interest. Because of this, the short term variations of **H** are entirely due to ω. If the spacecraft experiences a pure X rotation, then the graph of Z vs. Y elements of **H** will become a pure circle. For such pure X rotation to continue indefinitely, the Y and Z elements of the left hand side must be exactly equal to those of **H**. If ω is also constant then h_y and h_z must track H_y and H_z with perhaps a constant offset that is contributed by $\mathbf{I}\omega$. The steady-state motion of the spacecraft under 2-axis control is not a pure constant speed X rotation, but is not too far from it either as can be seen in Figure 8.

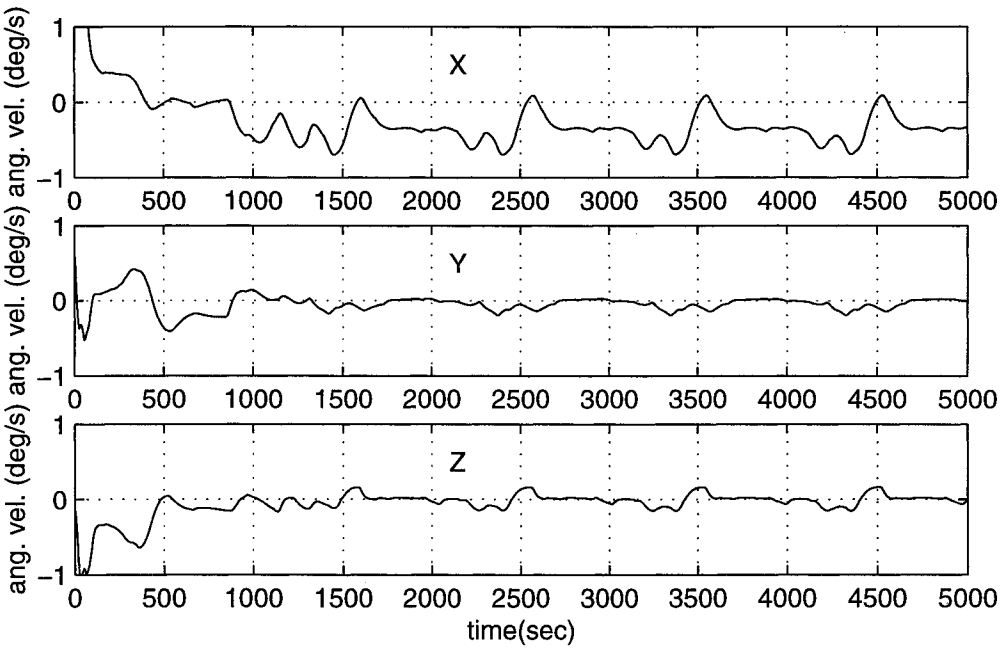


Figure 8: Spacecraft Angular Velocity Measurements under 2-Axis Control in Non-ideal Conditions

If, on the other hand, h_y and h_z do not track H_y and H_z sufficiently closely, their differences have to be absorbed by $\mathbf{I}\omega$ and thus cause change in ω, especially ω_y and ω_z, the Y and Z elements of ω, respectively. Non-zero ω_y and ω_z implies that the spacecraft –X axis moves relative to the sun vector or the sun angle fluctuates; the greater ω_y and ω_z are, the greater the sun angle fluctuates. It can thus be argued that the 2-axis control law keeps the component of **h** in YZ plane tracking that of **H** closely in steady state so that the sun angle remains relatively small.

Since the steady-state attitude motion of the spacecraft under the 2-axis control is roughly an X axis rotation, the component of **H** in the spacecraft YZ plane rotates with a rate approximately equal to the angular rate of the spacecraft about its X axis. Furthermore, the component of **h** tracks reasonably closely that of **H** in the YZ plane under the 2-axis control. Therefore, it is possible to use the measurable **h** for the estimation of the X angular rate, or the rate about the desired sun-line axis.

Because the rate of change of the angular momentum of a reaction wheel is approximately equal to the net axial torque, which is the applied motor torque minus drag torque, commanded wheel torque* can also be applied in sun-line rate estimation. The use of $\mathbf{T}_c$ and **h** together allows the estimation to be done without the need to differentiate any measured data. An example is a quantity **u**, which is roughly equal to the sun-line rate scaled by a non-negative multiplier, defined as

$$\mathbf{u} = (\mathbf{T}_c \times \mathbf{h}) \cdot \mathbf{s}_d$$

* Alan Reth of NASA GSFC suggested to the authors the use of commanded wheel torque for sun-line rate estimation.

The multiplier is roughly equal to $h_y^2 + h_z^2$. A similar idea of using wheel torque and momentum for estimation of the motion about the sun line had been proposed for the International Ultraviolet Explorer if all of its gyros had failed.[3] Figure 9 shows the comparison of **u** and the true sun-line rate with **u** scaled by the inverse of a multiplier. It is clear that **u** in steady state is approximately proportional to the sun-line rate and may be used to augment the 2-axis control.

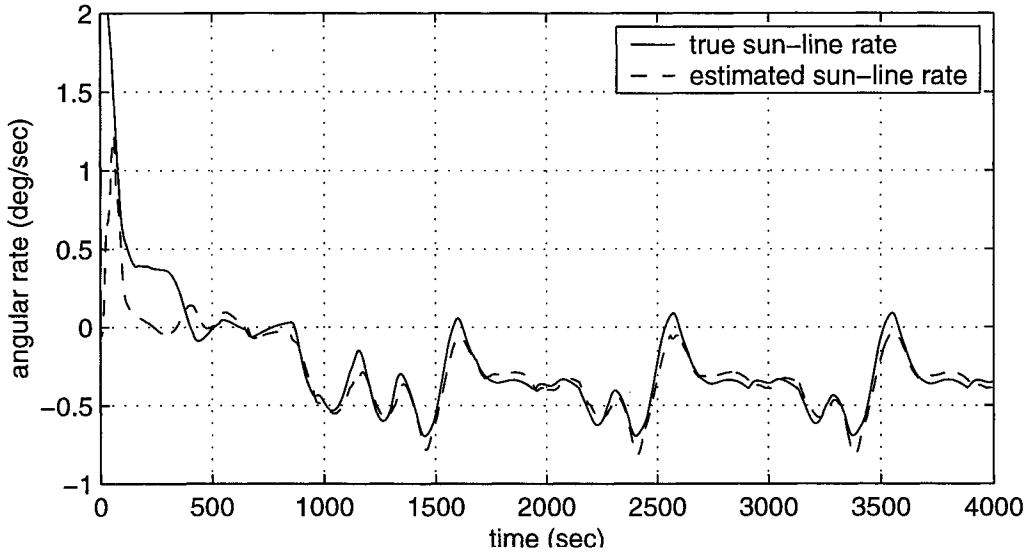


Figure 9: Estimated Sun-Line Rate Quantity vs. True Sun-Line Rate under 2-Axis Control in Non-ideal Conditions

Triana Safehold Control

The control law for Triana Safehold incorporates the rough knowledge of the sun-line rate in the 2-axis controller described above. With this control law the torque command becomes

$$\mathbf{T}_{c1} = \mathbf{I}\,[\mathbf{k}_p\ \mathbf{Limit}\ (\theta_e) + \mathbf{k}_v(\frac{\mathbf{d}}{\mathbf{dt}}\mathbf{s}) \times \mathbf{s}]$$

$$\mathbf{T}_{c2} = \mathbf{I}\,\mathbf{k}_w\,\mathbf{u}\,\mathbf{s}_d$$

$$\mathbf{u} = (\mathbf{T}_{c1} \times \mathbf{h}) \cdot \mathbf{s}_d$$

$$\mathbf{T}_c = \mathbf{T}_{c1} + \mathbf{T}_{c2}$$

$$\theta_e = \mathbf{s} \times \mathbf{s}_d$$

$$\mathbf{T}_{wc} = \mathbf{W}^+\,\mathbf{T}_c + \mathbf{T}_{mr} + \mathbf{T}_{dc}$$

where $\mathbf{k}_w$ is a scalar gain applied on **u**. Since the sun-line rate quantity **u** includes in it a non-negative variable scale factor, the value for gain $\mathbf{k}_w$ should be chosen with the scale factor taken into consideration. Furthermore, the fact that **u** is only a good indicator of sun-line rate in steady state means that $\mathbf{k}_w$ should be sufficiently small so that the transient behavior of the 2-axis control is not adversely affected. Simulating exactly the same case as the one with sun angle shown in Figure 3 with the Triana Safehold control law results in sun angle behavior shown in Figure 10. It is quite easily observed that the use of **u** in the control law can improve the steady-state sun angle performance without degrading the transient behavior. More simulation results generated with a high fidelity model of the Triana's control system are presented below.

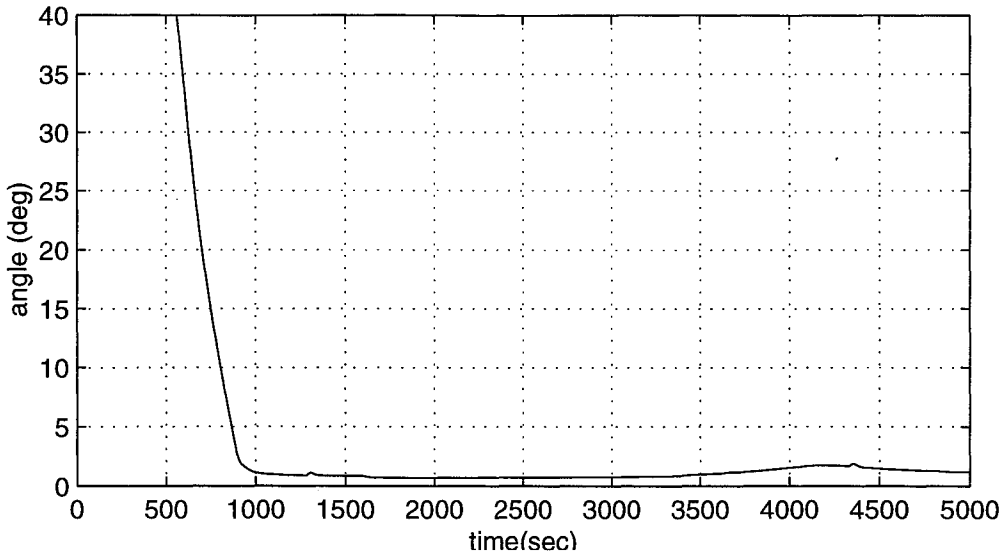


Figure 10: Sun Angle under Triana Safehold Control with Non-Ideal Condition

SYSTEM ANALYSIS

Using the commutation of wheel momentum to estimate the sun-line rate provides Triana with an effective Safehold using the minimum complement of hardware. However, this design also produces a coupled, non-linear system, which cannot be analyzed using the traditional linear stability analysis techniques. As more complex algorithms are implemented, the problem of analyzing a non-linear system is becoming more prevalent. To meet Triana's stability proof requirement, a non-conventional approach has been tried. It is briefly described here to provide the readers some understanding as to what has been done. The analyses address a coupled dynamic system, including both rigid and flexible dynamics.

Stability of Periodic Orbits

Similar to what is shown in Figure 7, Triana Safehold control also exhibits a periodic steady-state motion, but with much longer period. Given a system which exhibits a periodic orbit as its steady state, the dynamical behavior will be repeatable from cycle to cycle. If the system is stable, any perturbation away from this periodic steady state, will return to steady-state orbit. Such perturbation is demonstrated in Figure 11. By linearizing the perturbed equations of motion about this steady state, we can examine the system's stability. This linearized perturbed equations have periodic coefficients and can only be obtained from numerical simulation of the non-linear model. The process involves repeated simulation for one period with the initial values of the system states varied by small amounts from their steady state values. The changes in initial values are chosen to be orthogonal from run to run and to span the space of the states. The objective is to evaluate the state transition matrix at the end of one period. The largest magnitude of the eigenvalues of this state transition matrix, except the one corresponding to an eigenvector that is tangent to the orbit, is indicative of system stability for the given perturbation. This numerical solution must be repeated for a great number of perturbations, which would entail hundreds of successive numerical solutions, making this stability proof fairly tedious to complete, especially for a long orbit period. This leads the analysis team to pursue a more streamlined stability proof.

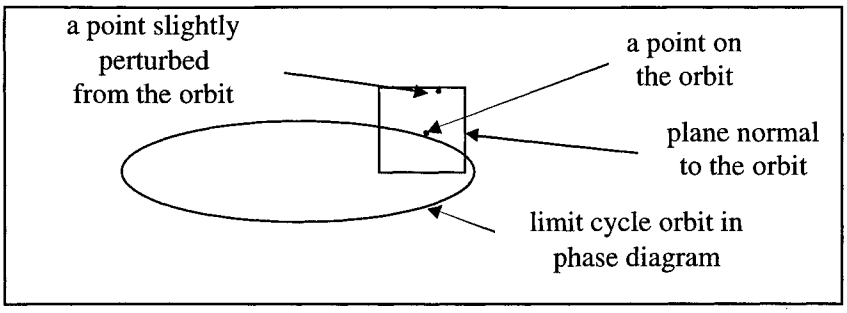


Figure 11: Perturbation from a Periodic Orbit

Stability of Equilibrium

In ideal conditions, a spacecraft under the Triana Safehold control behaves just like the 2-axis controller; it reaches an equilibrium with the desired sun line pointing at the Sun without any body rates. Equations of motion governing the perturbation from this equilibrium can be explicitly formulated and linearized into constant coefficient differential equations. Laplace transformation can thus be applied to the linearized perturbation equations. The resulting frequency domain equations are still a coupled, three-axis system, not amenable to conventional frequency domain stability criteria, such as Bode or Nyquist. However, the real parts of the closed-loop system poles determine the stability of the equilibrium. Examining the poles is much more computationally efficient than solving for the state transition matrix as required for determining stability for periodic orbit described above.

Using the equilibrium for Triana stability study requires some basic assumptions, so we will next consider the validity of these assumptions before proceeding to stability results. The effect of periodic orbit is most significant when the orbit period is very short. Simulation shows our Safehold will exhibit a 'fast' cyclic behavior under several conditions. First, if the sun-line rate is nearly unobservable, which occurs when the system momentum is nearly aligned with the sun vector, a non-nominal condition for Triana. Second, if Safehold is subject to very high system momentum, which is a condition outside the Safehold operating requirements. Finally, if Safehold has very high uncompensated wheel drag, which should not occur since the compensation is based on precise hardware data. For these cases, it is probably not a good idea to assume that equilibrium is the steady state motion. However, all of these cases are highly unlikely for the Triana Safehold, so the equilibrium assumption is reasonable.

To further support the equilibrium assumption, the values of control gains that cause instability have been shown to be about the same for both the non-ideal periodic orbit and ideal equilibrium for a few randomly chosen cases. This implies that analysis with both steady-state motions will result in similar stability margins.

Stability Margins

The steps taken to establish stability margins are as follows: For a given magnitude of system momentum, let its direction vary to span spherical space. Then, using the linearized perturbed equations around an equilibrium, find minimum damping poles of closed-loop system among all these directions. Vary gains and computation delays individually until the minimum damping poles for a given system momentum magnitude have negative damping. The gains and numbers of delay cycles at the point the closed loop poles change from positive to negative damping (from stable to unstable) are indicative of Safehold's gain and phase margins for that system momentum.

Vary the magnitude of system momentum and repeat these steps over a range of system momentum from zero to 27 Nms, which is twice the required operating range. By doing this it has been found that the Triana Safehold has margins of at least 6 dB in gain and 30 degrees in phase over this entire range, well within our stability requirements. The margin results were validated using the Triana simulation, by increasing the gains by the analytically predicted margins and watching the system become unstable as predicted.

SIMULATION AND RESULTS

This section delves into the closed-loop modeling for the Safehold control mode and then gives results from the simulation to demonstrate performance. Previous sections have already defined the Safehold control law and a brief description of the sensor and hardware complement. What remains is to define the fidelity by which the sensors,

actuators and their respective data are modeled. The Triana high fidelity (HiFi) simulator represents a closed loop modeling of the onboard systems, the spacecraft's environment and the physical laws that govern them. The HiFi simulator was implemented on a UNIX platform using a commercially available modeling and simulation package called Xmath/Systembuild.

The Safehold control law uses a minimal and simplistic sensor configuration, the Coarse Sun Sensor (CSS), to provide both rate and attitude error signals for control, as described above. There are six CSSs grouped in triads, thus two sets of three CSSs. The triad is arranged so that each individual CSS in the set is orthogonal to the other two and each CSS boresight makes the same angle from the X spacecraft axis. A set is mounted on the +X side (earthward) and the –X side (sunward) of the spacecraft. With this configuration the unit sun vector can be observed in any orientation except when the sun is shining in the blind rings of opposing pairs of CSSs. Because the half-cone angle of a CSS is 85 degrees, the blind ring of an opposing pair of CSSs has a width of 10 degrees. It is thus possible for only one or two of the CSSs to see the sun, and this then produces "jumps" in the CSS data when the sun leaves or enters these regions. Since differentiation of the unit sun vector is required in the control law, this discontinuity may cause undesirable effect on the control performance. If the triads were aligned with the spacecraft axes, one of these regions would be in the neighborhood of the desired sun line. This would have caused a problem in steady state when the desired sun line is nearly aligned with the sun vector.

In the HiFi simulator, each CSS is faithfully modeled as generating a positive voltage signal that represents the cosine of the angle between the CSS boresight and the sun unit vector. The CSS sensor model computes this by taking the dot product between each CSS boresight and the true sun vector. The CSS voltage measurements are then limited to values between zero and a threshold corresponding to the half-cone angle. To this voltage output is added Gaussian noise, with a standard deviation of 0.001. The voltage signals are then sent to the Utility Hub (Uhub) which combines the voltages from each triad, applies the appropriate sign (whether it is from the +X or –X triad) and outputs a representation in counts of the sun unit vector in body coordinates. The sensor processing routine then takes this vector and converts it to a normalized unit vector, which is the sun unit vector in body coordinates.

The control torques calculated with the Safehold control law are sent to the IRWA. The simulator computes realistic torque capacity based on current speed, and maximum voltage and power available, and limits the actual torque delivered to each wheel. The tach measurements are computed from the current wheel momentum and corrupted with a Gaussian noise process depending on the wheel speeds.

The parameter values used in the Safehold performance simulation define a fully deployed spacecraft configuration and a point in the Lissajous orbit. The spacecraft inertia matrix for this configuration is

$$\mathbf{I} = \begin{bmatrix} 251.116 & 3.322 & -36.779 \\ 3.322 & 271.04 & 0.707 \\ -36.779 & 0.707 & 217.5 \end{bmatrix} \mathbf{kg\,m^2}$$

Important initialization differences for the various cases are illustrated in Table 1.

Table 1: Defining Initialization Parameters for Safehold Performance Test Cases

	True Body Rates (rad/sec)	True System Mom. (Nms)	Initial Angle* (deg)	Cmd Sun Vector
Case 1	X = 0.0 Y = 0.0 Z = 0.0	X = 0.1484 Y = 3.6318 Z = 3.4332	96.7	Cmd X = -1.0 Cmd Y = 0.0 Cmd Z = 0.0
Case 2	X = -0.04872 Y = -0.00271 Z = -0.03942	X = -10.790 Y = -0.9228 Z = -6.784	96.7	Cmd X = -1.0 Cmd Y = 0.0 Cmd Z = 0.0
Case 3	X = 0.0 Y = 0.0 Z = 0.0	X = 2.5981 Y = 0.0 Z = 0.0	71.1	Cmd X = -0.9848077 Cmd Y = 0.1736482 Cmd Z = 0.0
Case 4	X = 0.0 Y = 0.0 Z = 0.0	X = 0.1484 Y = 3.6318 Z = 3.4332	179.85	Cmd X = -1.0 Cmd Y = 0.0 Cmd Z = 0.0

* Initial Angle is defined as the angle between initial sun vector and desired commanded sun vector.

Triana Safehold control is required to keep the sun pointing to within 15 degrees of the commanded sun vector and acquire the commanded pointing from any point in space within 15 minutes. Once acquired, the Safehold control laws should maintain this pointing indefinitely with minimal ground intervention. Sun angle performance for the four cases are shown in Figure 12.

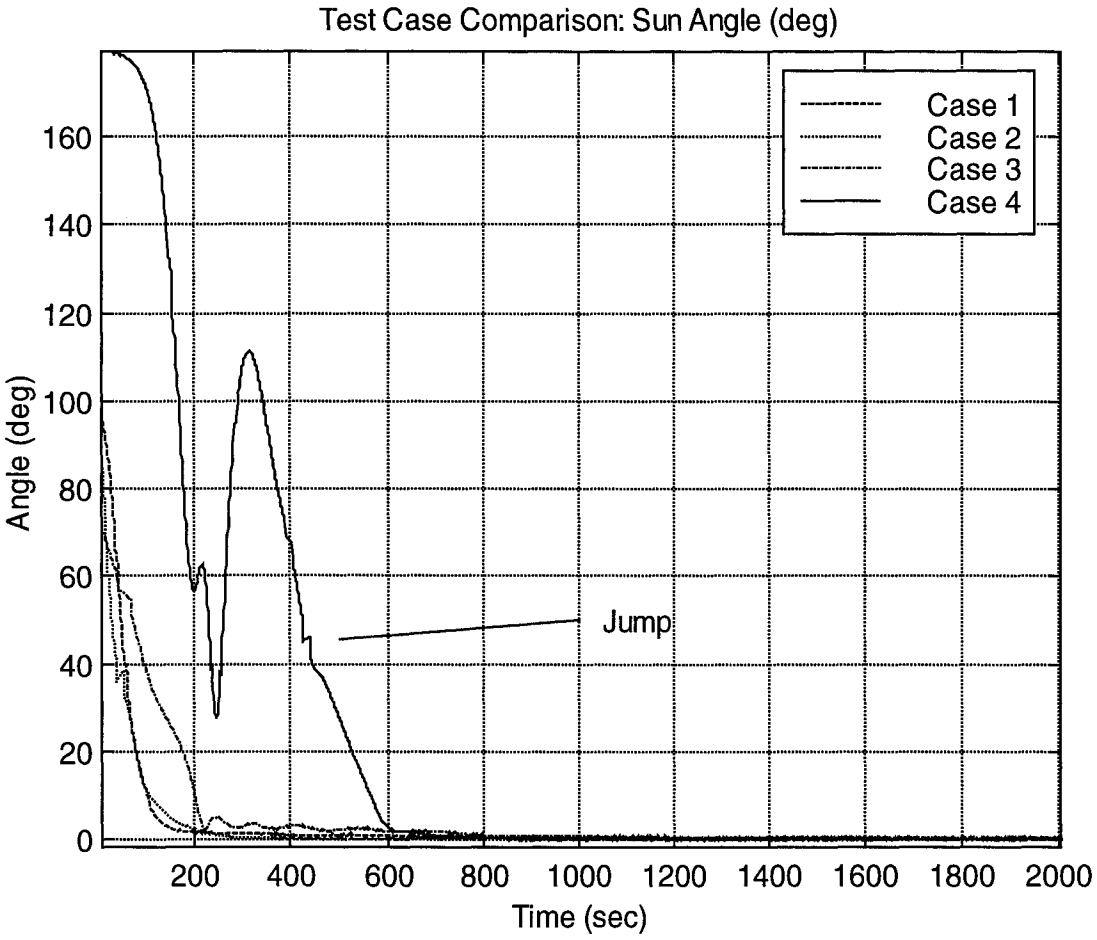


Figure 12: Sun Angle For Four Cases From HiFi Simulation

The line marked 'jump' on Case 4 points to an occurrence in which the sun vector is in one of the discontinuous regions where only one or two of the CSS in the triad are active. The sun angle appears to stay constant for a short period of time while the sun vector remains in the region. In fact, the true sun angle continues to change during the period. Figure 13 shows enlargement of the data near zero sun angle for a better assessment of steady-state performance. The small steady-state sun angle is due to the use of sun-line rate estimation in the control law.

DISCUSSION

Incorporation of a simple, albeit rough, sun-line rate estimation transforms a 2-axis control law into a pseudo 2-and-a-half-axis control law with satisfactory performance. The authors have recognized the potential of further performance improvement if an integral of the estimated sun-line rate quantity is also included in the controller. Cursory simulation supports the idea, but more studies are needed to ensure its validity.

It is also worth mentioning that flexibility in the choice of $\mathbf{s}_d$ exists in the Triana Safehold controller such that an axis other than the nominal sun-pointing axis can be directed toward the sun as demonstrated in case 3 of the simulation results. Although this flexibility is only meant to be used for small pointing offset in Triana operations, it may enable other uses of this control law.

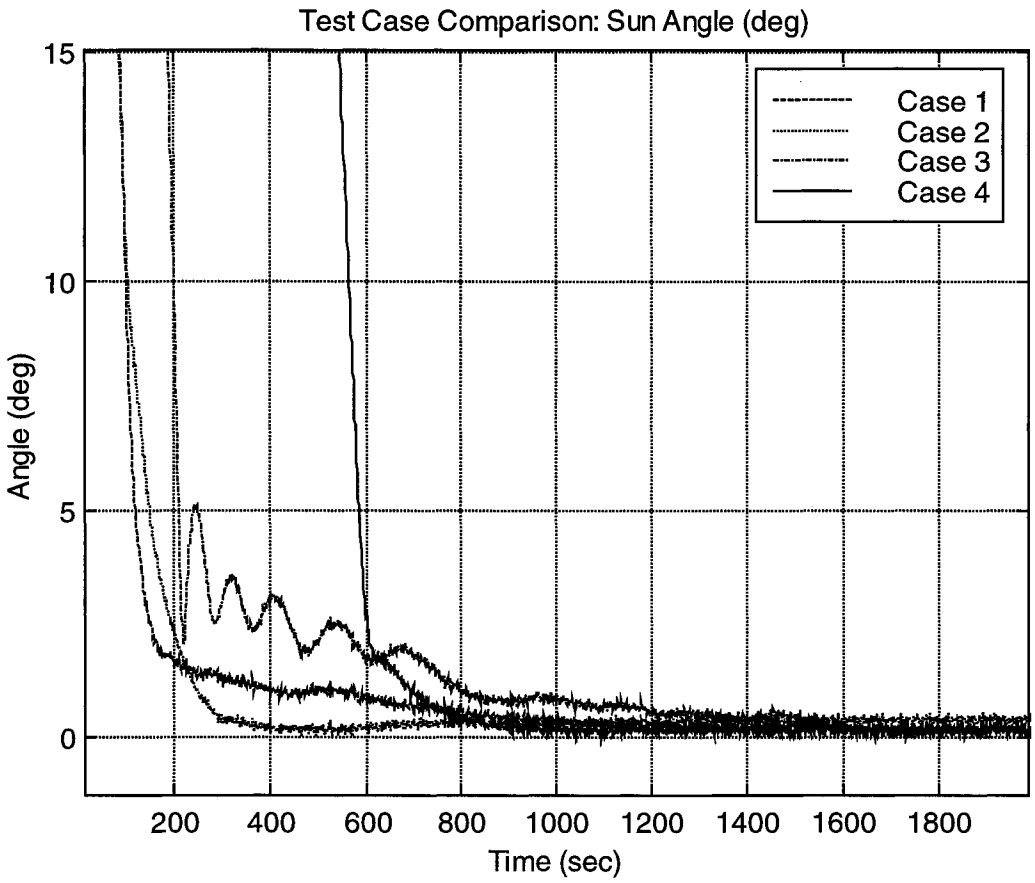


Figure 13: Steady-State Sun Angle For Four Cases From HiFi Simulation

SUMMARY

Because of Triana's operational plan, its attitude control system has to function within requirements even when system momentum is near the capacity of its wheel momentum storage. High system momentum, in addition to the worst-case uncompensated wheel drags, puts the gyroless Safehold controller in a particularly difficult situation. Incorporation of a simple, albeit rough, sun-line rate estimation transforms a 2-axis control law into a pseudo 2-and-a-half-axis control law with satisfactory performance. The concept of this estimation originates from observations of the simulation results of a 2-axis control law adapted from previous work. Conservation of system momentum and an expression for system momentum are then used to support the validity of this estimation concept. In addition, a simple algorithm for sun-line rate estimation is introduced and demonstrated to be reasonable in the steady state. Next, the steps taken to establish stability margins for the Triana Safehold control, with sun-line rate incorporated, are described. Results of simulations that includes a high fidelity model of the sun sensors and the wheels are then presented to show performance that satisfy requirements. Finally, suggestions for future improvements and the flexibility of this control law are briefly discussed.

REFERENCES

1. Andrews, S. F., Campbell, C. E., Erricsson-Jackson, A. J., Markley, F. L. and O'Donnell, J. R., Jr., "MAP Attitude Control System Design and Analysis," Proceedings of the Flight Mechanics Symposium, Goddard Space Flight Center, MD, May 19-21, 1997, pp. 445-454.

2. Hale, J. K. and Kocak, H., Dynamics and Bifurcations, Springer-Verlag, New York, 1991.

3. O'Donnell, J. R., Jr. and Hoffman, H. C., "Zero-gyro Control of the International Ultraviolet Explorer," AIAA Guidance, Navigation and Control conference, Monterey, CA, Aug. 9-11, 1993, Technical Papers Part 1, pp. 530-538.

Global Lyapunov Control of Spin Stabilized Spacecraft

Reid Reynolds
Spectrum Astro, Inc.

Glenn Creamer
Naval Research Laboratory

Abstract - Spin stabilization is the most basic technique available for the control of spacecraft. However, analyses of control laws for spin rate and spin axis precession often rely on linearization theory or heuristic arguments. In this paper, we introduce a new Lyapunov based control strategy for simultaneous spin, nutation, and precession control, which guarantees a continual downward trend to a unique, global minimum in the specified control cost. A variety of actuators may be used to implement the control scheme, including thrusters and magnetic torquers, and we examine practical implementation issues for each of these. Outstanding performance is demonstrated via simulations.

I. INTRODUCTION

Spin stabilization is the most basic technique available for the control of spacecraft. The stiffness provided by inertial torques keeps the spin axis fixed in inertial space, with decreasing sensitivity to disturbances as the angular momentum increases. With effective nutation damping, which can be provided by passive devices[1], this is overall an essentially passive and exceedingly robust control approach.

There are essentially three active control functions which can be implemented for spin stabilized spacecraft: spin control, precession control, and nutation damping. (Active nutation damping may be employed to augment or replace passive devices if an increased rate of energy dissipation is desired.) These different control objectives are often dealt with piecemeal in the literature, and usually employ some measure of the momentum error as the control quantity to be reduced[2,3]. In this paper, we develop a control strategy which accomplishes all three objectives simultaneously with very little effort. In addition, it is found that the rate error is the appropriate quantity with which to be concerned to produce this extremely simple and efficient control approach.

II. LYAPUNOV FUNCTION

The rigid body dynamics for the momentum are given by

$$\dot{\underline{H}} = -\underline{\omega} \times \underline{H} + \underline{\tau} \tag{1}$$

where $\underline{H} = \mathbf{J}\underline{\omega}$ is the momentum vector, $\mathbf{J}$ is the 3×3 inertia matrix, $\underline{\omega}$ is the angular rate vector, and $\underline{\tau}$ is the external torque vector. If we take $\hat{s}$ to be a unit vector in body space which represents the desired major axis pointing direction in inertial space, then

$$\dot{\hat{s}} = -\underline{\omega} \times \hat{s} \tag{2}$$

We construct the Lyapunov function[4,5]

$$V = \tfrac{1}{2}(\underline{H} - H_0\hat{s})^T(\underline{H} - H_0\hat{s}) + \tfrac{1}{2}\underline{H}^T\left(\bar{\lambda}\mathbf{J}^{-1} - \mathbf{I}\right)\underline{H} \tag{3}$$

where $\bar{\lambda}$ is the maximum eigenvalue of **J**, and note that $V=0$ implies a major axis spin with $\underline{H} = H_0\hat{s}$, where H_0 is a constant representing desired angular momentum magnitude. Then

$$\begin{aligned}\dot{V} &= (\underline{H} - H_0\hat{s})^T(-\underline{\omega}\times(\underline{H} - H_0\hat{s}) + \underline{\tau}) + \underline{H}^T\left(\bar{\lambda}\mathbf{J}^{-1} - \mathbf{I}\right)(-\underline{\omega}\times\underline{H} + \underline{\tau}) \\ &= \left(\bar{\lambda}\underline{\omega} - H_0\hat{s}\right)^T\underline{\tau} \quad = \quad \bar{\lambda}(\underline{\omega} - \omega_0\hat{s})^T\underline{\tau} \quad , \quad \omega_0 = H_0/\bar{\lambda}\end{aligned} \tag{4}$$

It is easily seen that if the torque vector is aligned such that the component in the direction of the rate error is always negative, $\dot{V}$ will be negative definite, and the final state will be $V=0$ with all that entails. Note particularly that alignment of the torque with the negative *rate error* is essential for the convergence to a major axis spin in the desired direction. Alignment with the negative *momentum error* will only guarantee convergence to the desired momentum, but it does not guarantee a steady state, major axis spin.

III. THRUSTER BASED CONTROL

Thruster based implementation of a globally stable control for a spin stabilized spacecraft is easy using equation (4). Informally, we shall refer to the major axis of inertia as the "spin axis" and the plane of the intermediate and minor axes of inertia as the "spin plane". In theory, so long as the nutation rate is not equal to the spin rate, as few as three thrusters could be used to implement the control, two for +/- spin axis torques, and one for spin plane torque to provide precession and nutation control. However, a minimal set of thrusters is not necessarily the most efficient approach in either time or propellant.

Propellant efficiency is, of course, the primary consideration. Thus, for spin plane thrusters, it would make sense to only fire when the angle between the torque axis and the rate error vector is as close to 180° as possible, as this ensures the maximum rate of decrease in V for a given thrust. In general terms, the efficiency also decreases as the rate error becomes smaller, so there is a practical lower bound on the achievable error, which should be enforced with a deadzone. If the deadzone is δ and the efficiency angle is ε, and we assume that the torque axes are either along the spin axis or located in the spin plane, the control algorithm becomes

Thruster Based Spin Stabilized Control Algorithm

$\hat{a}_i = i^{th}$ *thruster torque axis*
$\underline{\omega}$ = *rate vector*
ω_0 = *target rate*
$\hat{s}$ = *inertial target vector in body coordinates*
if $\hat{a}_i \bullet (\omega_0\hat{s} - \underline{\omega}) > \delta$ *and* $\hat{a}_i \bullet (\omega_0\hat{s} - \underline{\omega}) > |\omega_o\hat{s} - \underline{\omega}|\cos\varepsilon$, *fire thruster*

IV. MAGNETIC TORQUER CONTROL

Assuring convergence of the magnetic control implementation is a little more difficult than the thruster based implementation since the available torques are always orthogonal to the local magnetic flux density vector. Generally, we can apply a dipole moment as

$$\underline{D} = \sum_{i=1}^{N} m_i \hat{a}_i \tag{5}$$

where the number of magnetic torquers N is usually (but not necessarily) 3, m_i is the scaling of the ith dipole moment, and $\hat{a}_i$ is a unit vector in the ith dipole direction. The torque is $\underline{D} \times \underline{B}$, where $\underline{B}$ is the local external magnetic flux density vector so, from equation (4), the rate of change of the Lyapunov function is

$$\dot{V} = -\bar{\lambda} \sum_i m_i (\hat{a}_i \cdot \underline{\chi}) \tag{6}$$

where $\underline{\chi} = (\underline{\omega} - \omega_0 \hat{s}) \times \underline{B}$. Clearly a control policy which makes

$$sign(m_i) = sign(\hat{a}_i \cdot \underline{\chi}) \tag{7}$$

ensures a downward progression in the Lyapunov function.

Upon alignment of the rate error vector $\underline{\omega} - \omega_0 \hat{s}$ with the magnetic flux density vector $\underline{B}$, the Lyapunov rate goes to zero, the sign of the dot product in (7) becomes indeterminate, and the spacecraft will engage in torque free motion[6,7,8]. If it was possible for the magnetic flux density vector in an orbit to mimic the very regular dynamics of the rate error vector under these conditions, there might be a danger of convergence to an undesired equilibrium.

Such an equilibrium would be unstable in the sense that any perturbation of the magnetic flux density vector would result in a further decrease of the Lyapunov function. Given the significant irregularity of the Earth's magnetic field, nonconvergence of this kind is really not possible (though performance may not be spectacular if the desired spin rate is close to twice the orbit rate or the spacecraft is in an equatorial orbit). Simulations, such as those that follow, demonstrate that the control policy given above works very well in the magnetic field environment of low earth orbits at reasonable inclinations and at spin rates which are much greater than twice the orbit rate.

V. IMPLEMENTATION AND SIMULATION OF THRUSTER BASED CONTROL

To demonstrate performance of the global Lyapunov controller using thrusters we investigate the problem of simultaneous major-axis precession and nutation control of a spin-stabilized spacecraft. The assumed spacecraft parameters are provided in Table 1, along with thruster control deadzone limits and efficiency angle. It is also assumed that one thruster is utilized along each transverse axis to provide positive-valued torque for both precession and nutation control, and a three-axis rate gyro is utilized for body rate measurement and attitude propagation.

Given that the spacecraft is spinning at 30 rpm about its major axis with a 10-degree momentum precession error and 2 degrees of nutational motion, the goal is to simultaneously remove the precession error and nutation angle to within the limits defined by the control logic. Provided in Figure 1 are results from the maneuver, depicting the momentum error, spin-axis pointing error, and nutation angle.

It is observed that within 10 minutes the momentum error is entirely removed and a residual nutational motion of 0.2 degrees exists. Of course, any energy loss through fuel slosh or structural damping will remove residual nutation over time. For more rapid energy dissipation, a specially designed passive damping device would be recommended[1].

To assess sensitivity to efficiency angle, Figure 2 compares the time histories of the total transverse thruster on-times for this case ($\varepsilon = 20$ deg) and a reduced efficiency case in which the efficiency angle is tripled ($\varepsilon = 60$ deg). The angular momentum precessed per revolution is proportional to the sine of the efficiency angle, hence the time required to complete the precession maneuver is reduced by a factor of $\sin 20° / \sin 60° = 0.4$ or 60%.

Although the maneuver time is significantly decreased, a price is paid through an increase in propellant usage. The total propellant usage is proportional to the total time for the slew multiplied by the duty cycle. Hence, the propellent usage increases by a factor of $(\sin 20°/20°)/(\sin 60°/60°) = 1.19$ or 19%. Decisions on choice of efficiency angle are clearly mission-dependent.

VI. IMPLEMENTATION AND SIMULATION OF MAGNETIC CONTROL

To demonstrate performance of the global Lyapunov controller using magnetic torquers we investigate the problem of simultaneous major-axis spin stabilization and Sun acquisition of a spacecraft upon release from its launch vehicle. The High Energy Solar Spectroscopic Imager (HESSI) mission[9], set for launch in the Spring of 2001, represents a perfect example of such a system. The assumed spacecraft and orbit parameters are provided in Table 2. It is also assumed that a three-axis magnetometer is utilized to measure the local body-fixed field vector (simulated by an 8th-order geomagnetic field model), a two-axis Sun sensor is utilized to measure the desired body-fixed Sun direction, and body rates are either directly measured with a three-axis gyro or estimated from the Sun sensor and magnetometer.

Given initial body rates of 0.1 rad/sec in z-axis spin rate and 0.025 rad/sec in transverse rate, the goal is to simultaneously raise the spin rate to 0.3 rad/sec, remove nutational motion, and align the spin axis with the Sun. For this specific problem the control logic of equation (7) is satisfied by the three-axis feedback dipole

$$\underline{m} = k * (1 - \cos\theta_s) * sign(\underline{\chi}) \qquad (8)$$

where θ_s is the instantaneous Sun angle and k is a positive gain selected to saturate the torquers for a Sun angle of 2 degrees or greater. For a precession error of less than 2 degrees the torquers are commanded in the manner depicted above by adjusting the applied current to the torquer coils. Provided in Figures 3 and 4 are results from simultaneous spin rate control, nutation control, and precession control. It is observed that the nutation angle and Sun pointing error initially increase while the large spin rate error is being reduced, and then begin to decrease as the spin rate error gets smaller. All three errors are essentially removed after 6 to 7 hours. A plot of

the spacecraft dipole magnitude, shown in Figure 5, demonstrates the desaturation phase of the torquers upon reaching a precession error of 2 degrees.

VII. CONCLUSIONS

We have developed globally stable, Lyapunov-based control laws for simultaneous spin rate, precession, and nutation control of spin-stabilized spacecraft. The Lyapunov function was constructed using a combination of the momentum vector error and the eigenvector relation associated with the maximum eigenvalue of the spacecraft inertia matrix to guarantee convergence of the spacecraft major axis to the desired momentum vector. Simulations using both thrusters and magnetic torquers demonstrate the performance and utility of the control law for a variety of practical spacecraft missions.

VIII. REFERENCES

[1]Reynolds, R.G., "Dynamic Modeling of Ring Nutation Dampers," AIAA/AAS Astrodynamics Specialist Conference, Denver, Co., August 2000.

[2]Hubert, C.H., "Spacecraft Stabilization System and Method", U.S. Patent # US04758957.

[3]Holden, T.E., and Lawrence, D. A., "A Lyapunov Design Approach to Magnetic Nutation Damping," AIAA Guidance, Navigation, and Control Conference, Portland, OR, August 9-11, 1999, pp. 146-154.

[4]Khalil, H.K., *Nonlinear Systems,* Prentice-Hall, Upper Saddle River, N.J., 1996, pp 98-113.

[5]Slotine, J.J.E., and Li, W., *Applied Nonlinear Control,* Prentice-Hall, Upper Saddle River, N.J., 1991, pp 62-68.

[6]Greenwood, D.T., *Principles of Dynamics,* Prentice-Hall, Upper Saddle River, N.J., 1988, pp 408-420.

[7]Wertz, J.R., *Spacecraft Attitude Determination and Control,* Kluwer, Boston, 1978, pp. 636-654.

[8]Kaplan, M.H., *Modern Spacecraft Dynamics and Control,* Wiley, N.Y., 1976, pp. 57-61.

[9]Creamer, N.G., "The HESSI Magnetic Attitude Control System," AIAA Guidance, Navigation, and Control Conference, Portland, OR, August 9-11, 1999.

Table 1: Spacecraft and Control Parameters for the Thruster Controller

Spin Inertia (kg-m^2)	400
Transverse Inertia (kg-m^2)	350
Thruster Level (N)	4.45
Thruster Moment Arm (m)	0.5
Thruster Deadzone (rad/sec)	0.001
Efficiency Angle (deg)	20

Table 2: Spacecraft and Orbit Parameters for the Magnetic Controller

Spin Inertia (kg-m^2)	200
Transverse Inertia (kg-m^2)	150
Magnetic Torquer Strength (amp-m^2) (one aligned with each body axis)	100
Orbit Inclination (deg)	30
Orbit Altitude (km)	600

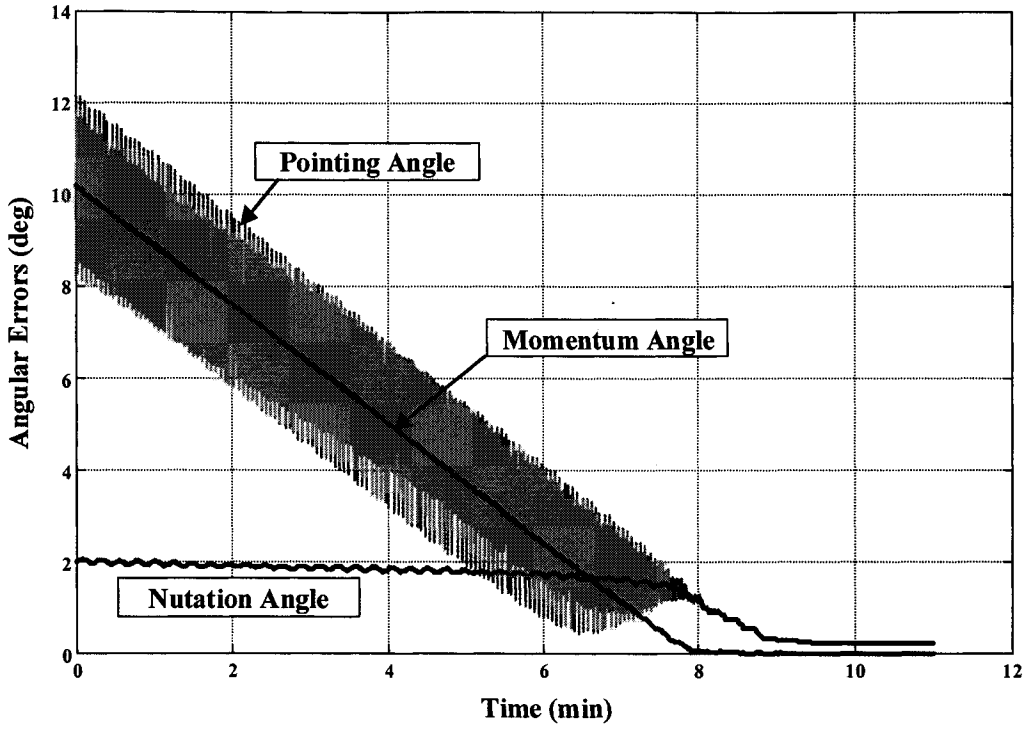


Figure 1: Spacecraft Attitude Motion During Thruster Control

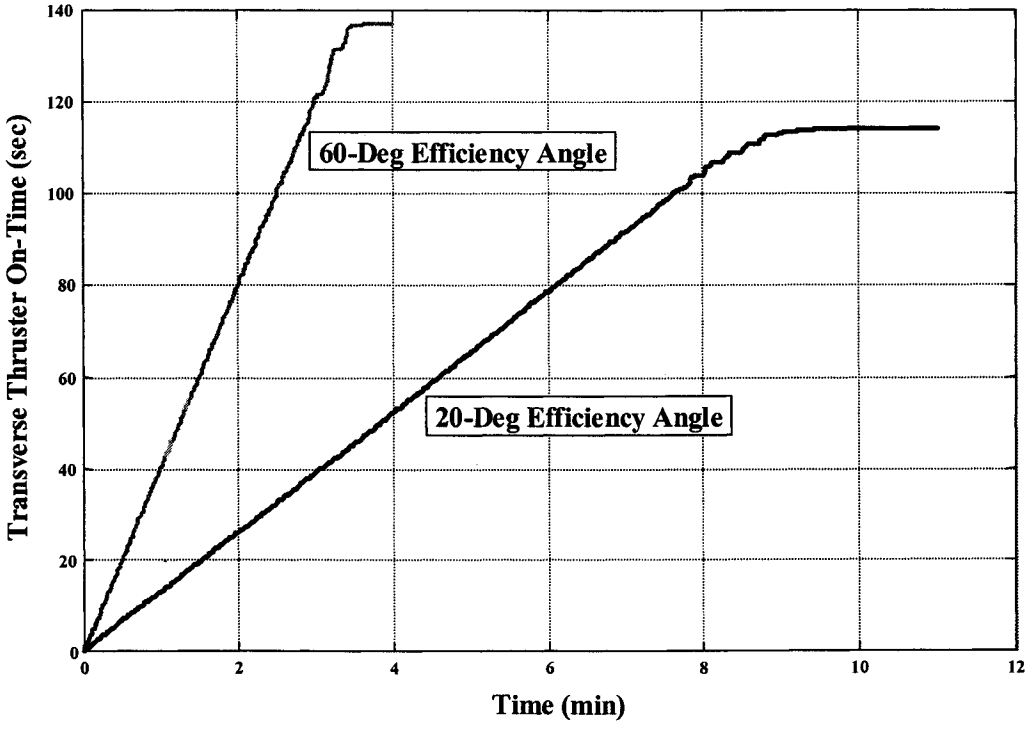


Figure 2: Effect of Efficiency Angle On Transverse Thruster On-Times

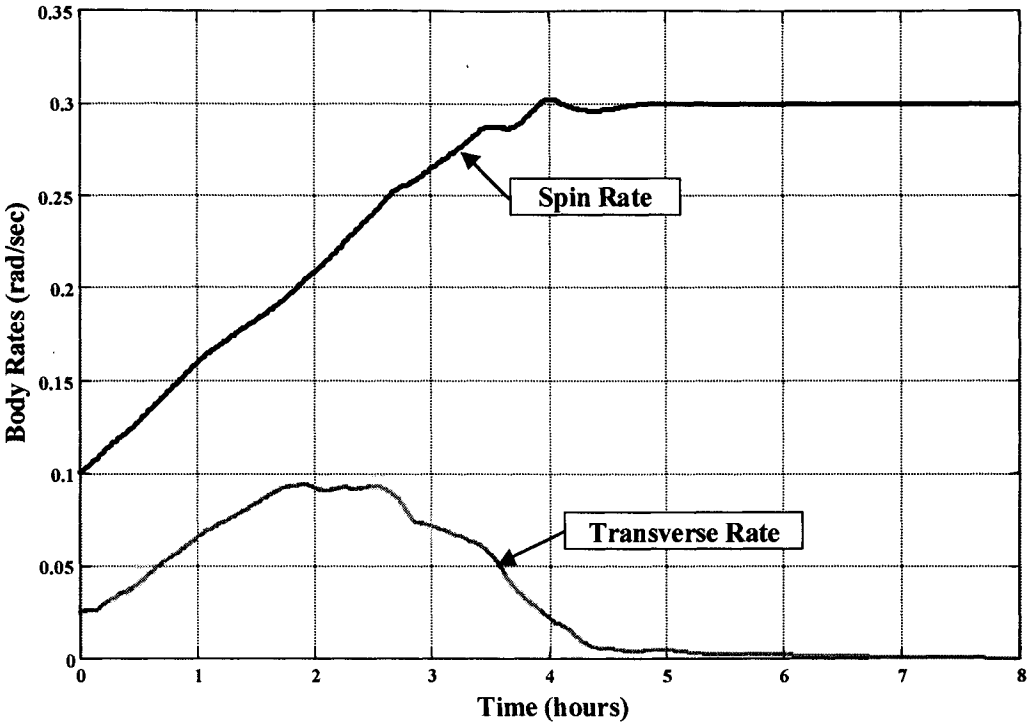


Figure 3: Spacecraft Rates During Magnetic Control

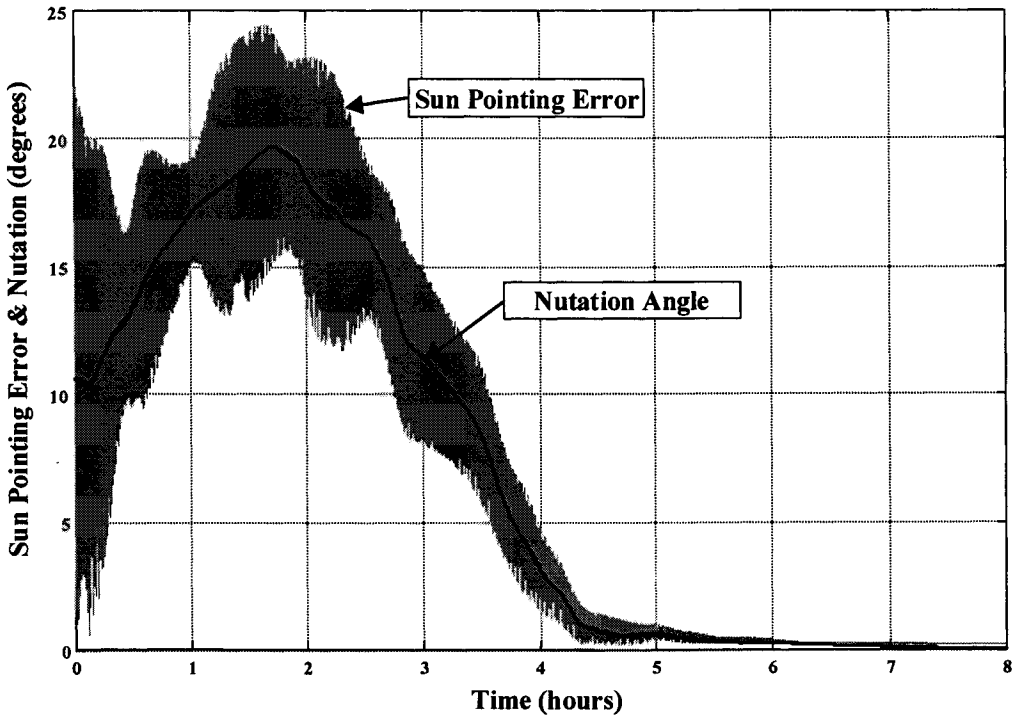


Figure 4: Spacecraft Precession and Nutation Errors During Magnetic Control

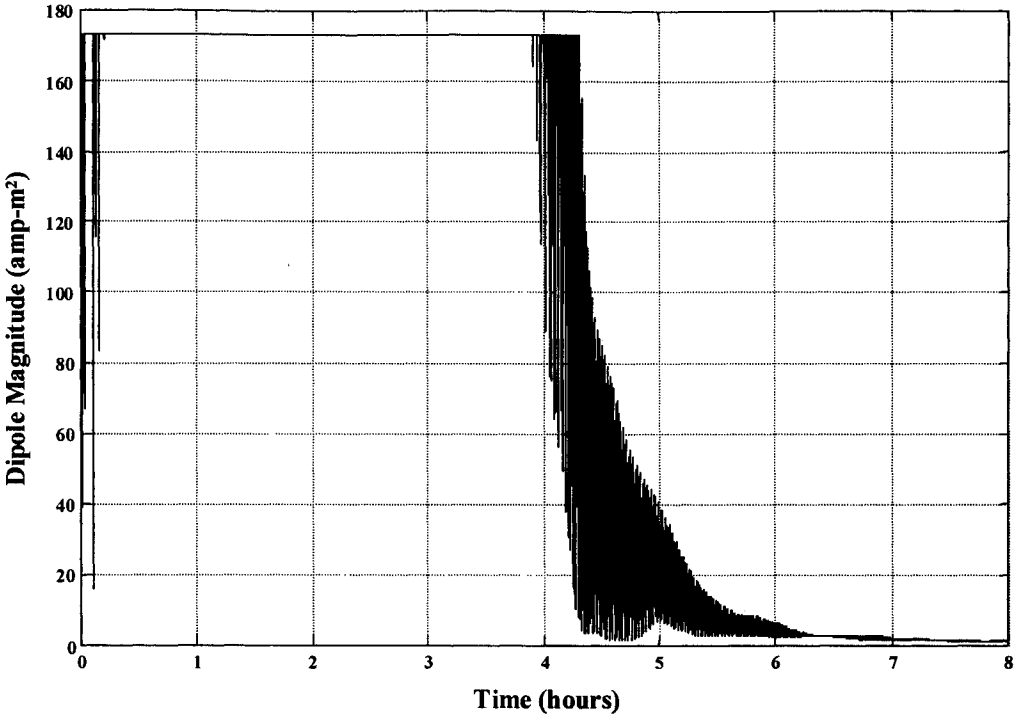


Figure 5: Spacecraft Dipole Magnitude

USING AUTOMATION TO IMPROVE THE FLIGHT SOFTWARE TESTING PROCESS

James R. O'Donnell, Jr., Ph.D.[1], Wendy M. Morgenstern[1], Maureen O. Bartholomew[2]
NASA Goddard Space Flight Center
Greenbelt, MD 20771 USA

One of the critical phases in the development of a spacecraft attitude control system (ACS) is the testing of its flight software. The testing (and test verification) of ACS flight software requires a mix of skills involving software, knowledge of attitude control and attitude control hardware, data manipulation, and analysis. The process of analyzing and verifying flight software test results often creates a bottleneck which dictates the speed at which flight software verification can be conducted. In the development of the Microwave Anisotropy Probe (MAP) spacecraft ACS subsystem, an integrated design environment was used that included a MAP high fidelity (HiFi) simulation, a central database of spacecraft parameters, a script language for numeric and string processing, and plotting capability. In this integrated environment, it was possible to automate many of the steps involved in flight software testing, making the entire process more efficient and thorough than on previous missions. In this paper, we will compare the testing process used on MAP to that used on other missions. The software tools that were developed to automate testing and test verification will be discussed, including the ability to import and process test data, synchronize test data and automatically generate HiFi script files used for test verification, and an automated capability for generating comparison plots. A summary of the benefits of applying these test methods on MAP will be given. Finally, the paper will conclude with a discussion of re-use of the tools and techniques presented, and the ongoing effort to apply them to flight software testing of the Triana spacecraft ACS subsystem.

INTRODUCTION

The development of the attitude control system (ACS) for the Microwave Anisotropy Probe (MAP) spacecraft (see Figure 1) included a number of firsts for the Goddard Space Flight Center. It was the first time that the MatrixX integrated simulation, analysis, and design toolkit was used at Goddard. Of greater note, it was the first time that automatically generated flight software, generated from the MAP high fidelity (HiFi) simulation using the AutoCode feature of the MatrixX toolkit, was used here. A number of papers (ref. 1, 2, 3) have been written about the use of the MatrixX toolkit on the MAP project, as well as on the special considerations that needed to be taken into account when testing automatically generated flight software.

During the development and testing of the MAP ACS subsystem, a number of tools were developed to automate portions of the ACS flight software test process. This set of MAP test tools, combined with some of the other integrated features of the MAP project, greatly increased the efficiency and thoroughness of the flight software testing. In the remainder of this paper, these tools will be discussed, along with a discussion of their application to the Triana spacecraft.

---

[1] Flight Dynamics Analysis Branch, Code 572
[2] Flight Software Branch, Code 582

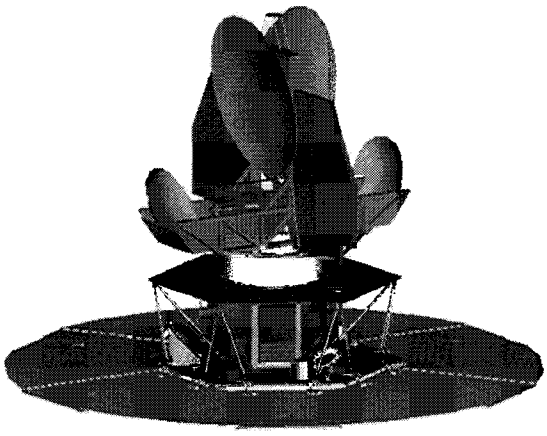


Figure 1: The MAP Spacecraft

FLIGHT SOFTWARE TEST PROCESS

On the MAP project, as with many past and present missions, the ACS flight software test process was divided up into a number of steps. At the lowest level, as the flight software developers wrote each piece of code, tests were performed at the unit test level. Once the flight software for the spacecraft was integrated onto the spacecraft or test facility, software testing was done first at the build test and then the acceptance test levels. During build testing, testers verified the correct operation of each specific flight software function. By and large, each of these tests was driven by a specified requirement of the software, and testing was used to verify that the requirement was met. At the acceptance test level, more realistic tests were run that verified that the spacecraft would operate in both nominal conditions and in the face of anomalies. These acceptance tests sought to cover the full lifetime of the spacecraft, from launch and early orbit operations to orbit maneuvers to the spacecraft science mode. Finally, in addition to the build and acceptance test levels, the MAP test team performed a number of sets of regression testing; this was an abbreviated set of build and acceptance tests performed after late additions or changes to the flight software.

For MAP, the test facility mentioned above was known as FlatSat. It consisted of engineering test units of the spacecraft main processor, attitude control electronics, and several other electronics boxes. The spacecraft's sensors and actuators, as well as the dynamics and other environmental modeling, was done using a hybrid dynamic simulator (HDS). The MAP HDS had very high fidelity, and it was possible to do very realistic testing with it on FlatSat.

Process Steps

The basic flight software test flow began with developing test scenarios. These scenarios defined the initial conditions of the test as well as the flow of the test, including mode transitions, telemetry verification and failure conditions. From these scenarios, test procedures were developed, then executed on the test environment. The test environment included a ground system that was used to send commands and receive telemetry from the flight system. It also included the flight system comprised of hardware and flight software. Ground support equipment such as the HDS was another critical component of the test environment. Once the tests were executed, the results needed to be analyzed in order to verify that the flight software met the test objectives. In general, the test results needed to be plotted for the analysis.

In parallel, the test scenario was duplicated in the high fidelity (HiFi) simulation, producing results which were used to compare with the flight software test output results from the FlatSat test environment. The HiFi results were the "truth" used to determine if the performance of the flight system was adequate.

There were several aspects of flight software testing critical to its timely success. First, the ability to replicate the initial conditions and test flow in the HiFi is extremely important since the results of the flight software test were compared with the HiFi results. Second, the flight software, the HDS, and the HiFi need to be consistent. There were many variables such as scale factors, biases, and alignments that existed in two or more of these systems and must be identical in each. Third, the ability to plot and analyze the results with ease was crucial.

"Old vs New" Comparison

In order to appreciate the benefits of the automation that was done on MAP, it is important to understand the process and bottlenecks on previous projects. In order to illustrate this, the testing process used for past missions, such as the Rossi X-Ray Timing Explorer (XTE), will be contrasted with that of MAP. It should be noted first, though, that XTE has been a very successful mission; the fact that bottlenecks or possible areas for improvement in their flight software testing process are identified below in no way diminishes from that fact.

XTE Flight Software Test Process

Figure 2 shows the flight software test process used for XTE. As denoted by the grey shaded arrows, there were many steps in this process that were manual and, therefore, prone to errors.

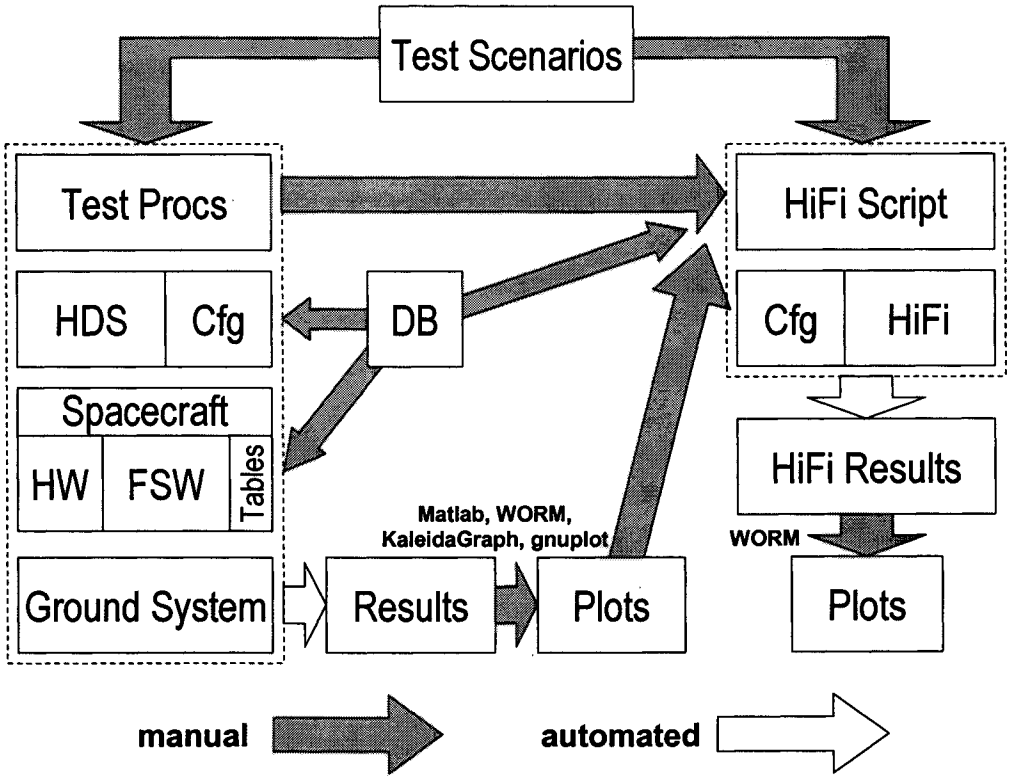


Figure 2: "Old" Flight Software Test Process Includes Many Manual Steps

On XTE, flight software test procedures were developed on the ground system. The initial conditions for these tests were then given to an ACS analyst to replicate the initial conditions in HiFi, which were used to define the expected test results. The flight software test procedures were executed in the flight software lab. The results were plotted using the test author's favorite plotting package, which was usually different from any of the other tester's favorite plotting package. The resulting plots from both the HiFi and the flight software test execution, plotted separately, were then held side to side and visually compared. Another critical piece to the flow was that there existed a plethora of variables that needed to be consistent between the HiFi, HDS, and flight software. On XTE, these variables were maintained in a spreadsheet. The variables were then manually entered into the HiFi, HDS, and flight software.

The pitfalls that the XTE process suffered were in each of those manual processes listed above, namely replicating the initial conditions, plotting of the results, and maintaining consistency between all of the variables. Especially in the infancy of the XTE flight software test program, it was not unusual to go through several iterations of a particular test, trying to match the initial conditions and variables. Inconsistencies would be discovered in the plot results review which would result in a discovery that, for example, the Kalman filter was not enabled in the flight software test but was enabled in the HiFi version of the test, creating different results. Sometimes the inconsistencies were more subtle, such as a number swap of one of the digits of a variable (e.g., flight software gyro bias = 3.0234, HiFi gyro bias = 3.0243). Time consuming iterations were also made in the plotting process. When comparing the flight software test plots with the HiFi plots, consistency in scaling and units are extremely important. Many times, the flight software plots would have to be recreated in order to scale the plots with enough detail to verify the results with HiFi or to change the units to facilitate comparison.

Finally, the separate nature of the flight software testing and HiFi verification paths above created a potential slowdown in the process. The HiFi verification runs could not be created except by a member of the ACS analysis team familiar with the HiFi, and could also not be done with any confidence until the flight software test results were viewed. The limited number of analysts, as well as the fact that the analyst would need to interact with the tester on multiple occasions for each test, frequently resulted in a bottleneck waiting for tests to be verified.

MAP Flight Software Test Process

On MAP, many of the manual processes were automated, streamlining the testing process. Figure 3 illustrates the flight software test flow that MAP followed, highlighting those areas that were automated. Unlike XTE, MAP's flight software testing process had tools that dissected a flight software test procedure, automatically producing a HiFi script that replicated the flight software test procedure flow. In addition, MAP had a centralized database that defined and linked all of the variables used in HiFi, HDS, and the flight software. With the press of a button, the database generated the source files for each of these systems which guaranteed consistency. Finally, the plotting process was streamlined since the MAP tools defined a standard set of plots that were used to plot the flight software, HiFi, and HDS test results on one plot.

In the next two sections of this paper, the specific tools that were developed to perform each of the tasks shown in Figure 3 will be discussed. Additionally, the other assumptions necessary for the correct operation of these tools will be shown; each of these assumptions involved a

specific standard way of writing procedures or formatting output that was established to improve the MAP testing flow.

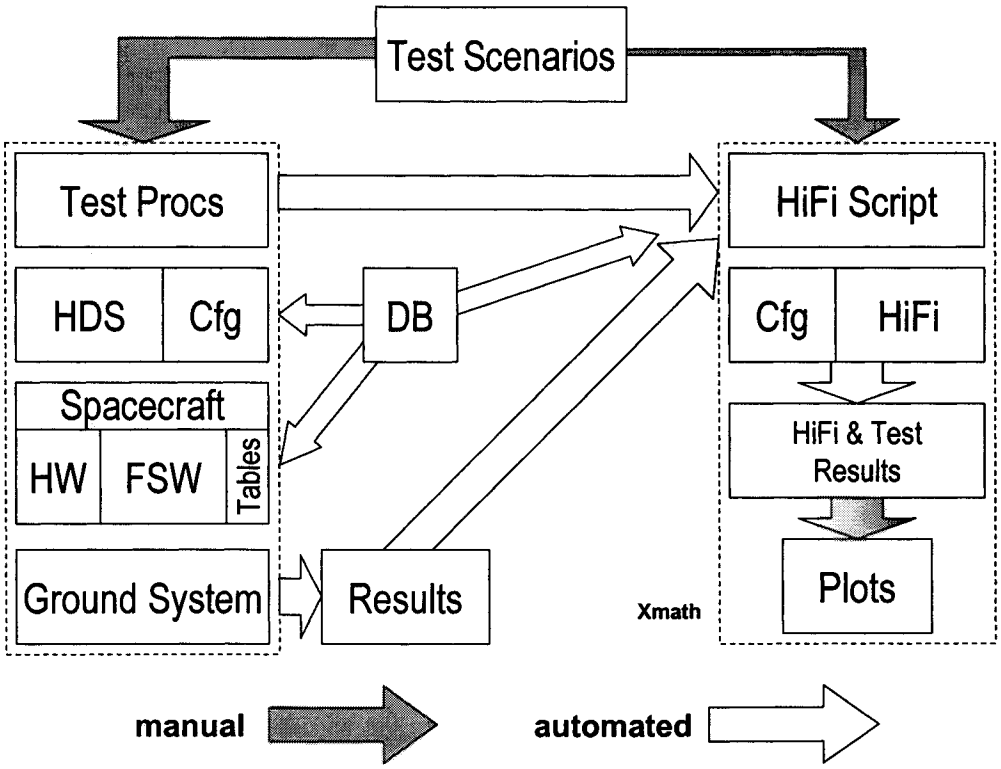


Figure 3: Automating Testing Steps Makes Test Process More Efficient

The different pieces of the test process flow shown in Figure 3 will be discussed in the next two sections, along with the software tools that implement them.

INTEGRATED DESIGN ENVIRONMENT

In order to be able to apply the automated test tools developed for the MAP flight software test effort to maximum effect, it was necessary to combine the elements of the subsystem and the testing tools into an integrated design and testing environment.

Parameter Database

One of the key elements to the success of the MAP flight software testing effort was the parameter database used by the ACS subsystem. This database was used to configuration manage virtually all of the variables, control gains, failure detection and correction (FDC) parameters, and other parameters used by the MAP ACS. As shown in Figure 3, the database fed into all of the ACS elements. As each parameter was placed in the MAP database, it was assigned to an appropriate subsystem engineer, in order for the database to be populated with the correct information and verified.

Upon a release each new version of flight software for either the spacecraft main processor or attitude control electronics, a corresponding release from the MAP database was created. Output templates from the database were created as header files for the flight software; additionally, script files for initializing the HDS and HiFi were generated at the same time. In this way, when flight software tests and HiFi verification simulations were performed, a consistent set of parameters across each test system component was assured.

Scripting Language with String Processing

The MathScript scripting language of MatrixX was used extensively with the MAP HiFi in order to set up simulations and to perform many data analysis functions. MathScript allows flight software and HiFi simulation to be analyzed, it interfaces with MatrixX's SystemBuild simulation environment to allow simulations to be created and run, and provides the mechanism for creating comparison plots between flight software and HiFi simulation verification data.

MathScript is a very complete scripting language for data processing, particularly matrix processing, and provides all of the functions necessary for interfacing numerical data and the HiFi simulation. One capability it lacks, however, is for doing extensive string processing. Because of the need to process the sequentially printed data output files from the MAP ground system, which are mixed numeric and text, and to process STOL procedures and RDLs (which are used to run test procedures and describe MAP data packets, respectively) as well, string processing is necessary.

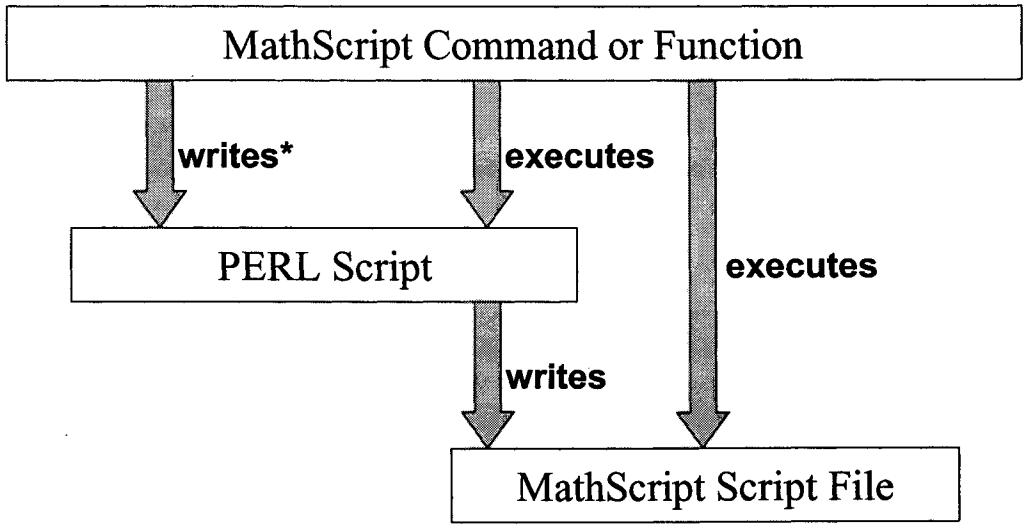


Figure 4: MathScript Architecture Allows Additional Capabilities to be Added

Figure 4 shows the solution implemented to add the string processing abilities necessary for the MAP test tools. The solution makes use of the ability that MathScript has to call external commands native to the underlying operating system. In order to add string processing to MathScript functions, the following steps were followed:

1. A MathScript function, using its `fprintf` statements, writes a PERL script out to the file system. (In some cases, a pre-existing PERL script is used instead, in which case this step is not necessary.)

2. Using a native operating system command, the PERL script from step 1 is executed.

3. The PERL script executes and processes the test procedure sequential print data output files, STOL procedures, or RDLs, as appropriate. As output, the PERL script writes out a new MathScript script file.

4. Finally, the original MathScript function calls the newly created script file, which might, for example, read the test data into MatrixX.

When the MAP test tools were originally developed, a combination of the Unix commands and utilities of bash, sed, and grep were used to implement these functions. Eventually, though, PERL was used instead because it could perform all of the functions that were divided between the other utilities, plus it is widely available on many different platforms.

Configured HiFi Simulation

Figure 5 shows the top-level block diagram of the MAP HiFi simulation, shown here to give a sense of what the simulation looks like as implemented in the MatrixX SystemBuild simulation environment. SystemBuild is a hierarchical block diagram editor made up of SuperBlocks, such as the ones shown in Figure 5, and lower level blocks that can be used to implement the different mathematical functions and dynamic elements needed in a simulation.

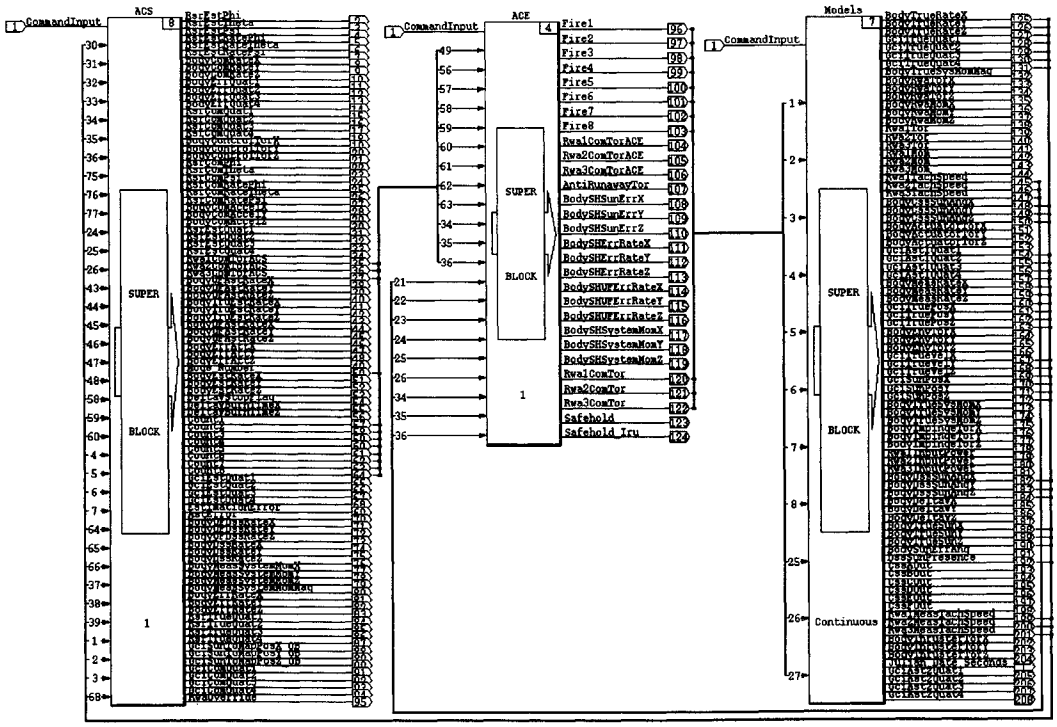


Figure 5: MatrixX Simulation Environment Lends Itself to Automation

The MatrixX integrated toolkit, its Xmath command environment (in which the MathScript scripting language is based), and its SystemBuild simulation component lend themselves especially well to the automation process devised in the MAP test tools. As described in the previous section, the scripting language acts as the glue that holds the simulation and testing process together. Through the string processing extensions, test data can be read in and processed. Using SystemBuild Access, an extension to Xmath that allows it to access and control SystemBuild simulations, MathScript functions can complete the process of data analysis and test verification by creating and running HiFi simulations.

It should be noted that while the MatrixX integrated toolkit lends itself particularly well to the design of automated test tools, such tools could be designed in other settings. The key ingredient that is necessary is a scripting language around which the rest of the system can be built. The Matlab/Simulink environment, using Matlab's m-file scripting, could also be used. Even a dedicated simulation such as one written in a language such as FORTRAN or C can be used with the automated test techniques described herein by using an environment such as Matlab or MatrixX as a front end interface.

Plotting Capability

The final capability needed to support the automated test tools is a flexible plotting capability. By being able to plot test data and HiFi verification simulation data on a common plot, the verification process is made much simpler. Along with the test tools, which will be described in the next section, that input test data and automatically set up HiFi simulations, the plotting of an arbitrary number of verification plots is supported. This becomes the final step in the test verification process; after comparison plots are created, they can be analyzed to verify a flight software test.

TOOL DESCRIPTIONS

In this section, the specific automated test tools that were developed for use on the ACS flight software testing for MAP will be discussed. These are the MathScript and PERL functions that implement the capabilities described in the previous section.

Test Data Processing

The main tool for processing the sequential print data output from flight software testing is called `pktproc`. This procedure, using the PERL-enhanced string processing methods described in the previous section, reads in STOL sequential print data and description files. It is able to automatically combine split packets (large packets must sometimes be split to accommodate ground system limitations). Each packet of data is formatted into an Xmath PDM (a data structure used by Xmath that is able to store a matrix along with auxiliary data, such as a time domain vector, and names applied to each column of data, in this case the packet telemetry names). The `pktproc` routine then writes out each set of packet data PDM for later processing by other Xmath test tools.

There are other test data processing functions used by the MAP test tools. These are used for post-processing flight software test data, calculating parameters not downlinked directly.

Synchronization and Data Analysis

Once the flight software test data has been processed and imported into the Xmath environment, the two most important functions of the automated test tools are implemented. The first is implemented with the `msync` function. This function finds the beginning of a test run and synchronizes the test data from each packet. This is necessary because when sequential prints are started at the beginning of a flight software test, each packet begins to be output at a different time. A standard test format was established for the MAP flight software testers in which all of the sequential prints were begun, and then the "restart" process was begun in the FlatSat flight software test facility which set the initial conditions for the test. The `msync` function analyzes the appropriate data packet for signs in the output data of this "restart". It then determines the offset into each packet that corresponds to this time. The offsets are saved for use by the other test tools. In this way, though the packets each begin at a different time, the "zero point" for each packet in the test is determined.

In addition to the `msync` function, the other critical test data analysis function for purposes of implementing the MAP automated test tools is the `transitions` function. The `transitions` function examines many of the different test data output packets for different aspects of the test run. This information is then used to *automatically* create a MathScript script file for a HiFi simulation that will match the test run. Some of the flight software test conditions that `transitions` looks for are the following:

- initial conditions: time, spacecraft position and velocity, initial spacecraft attitude quaternion, system momentum, and body rates
- Safehold and Safehold rate sensor transitions
- ACS control mode transitions
- Observing Mode spin down transitions
- thruster one-shots
- command quaternion transitions
- reaction wheel override commands
- Delta V Mode commanded burn times
- commands to enable or disable the Delta V impulse controller
- HDS commands to enable or disable sensor and actuator noise

In many ways, the `transitions` function is the heart and soul of the MAP automated test tools. By allowing software testers to automatically create matching HiFi simulations, it is possible for anyone, not simply ACS analysts familiar with the HiFi, to create HiFi verification runs for flight software tests. Analysts are still necessary for verifying the flight software tests,

but by making the creation of the HiFi verification run automatic, one potential bottleneck is removed from the test and test verification process.

Plotting and Test Verification

Once the flight software test data is imported into Xmath and analyzed, and a HiFi verification simulation run, the final step is to plot the test data along with the HiFi data for comparison. Because of the data synchronization and automatic HiFi simulation setup, it is a simple matter to create these comparison plots.

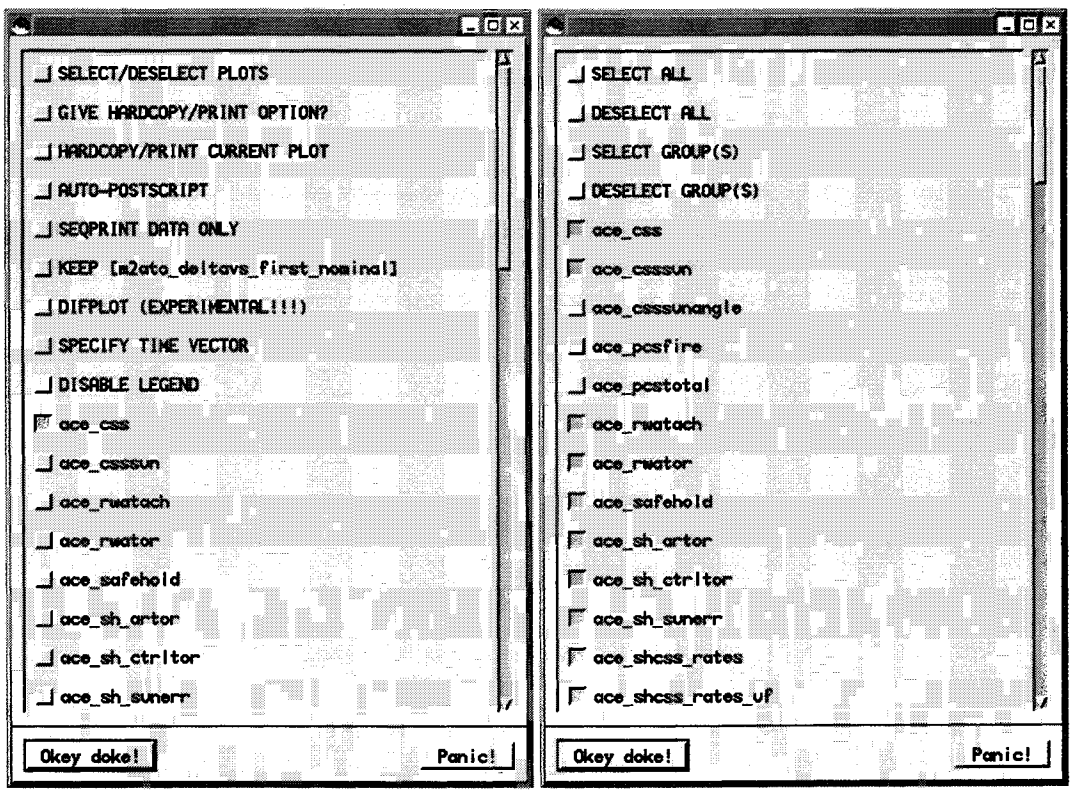


Figure 6: Plot Tools Offer Flexibility and Batch Processing

Figure 6 shows the graphical interface to the test tool `mapplot` function. This function gives the user the ability to create one or more of many predefined comparison plots. If a plot is selected, then test data and HiFi data are plotted together, such as in the plot shown in Figure 7. As shown on the left of Figure 6, there are a number of options that the user can select in addition to picking one of the plots. If the `SELECT/DESELECT PLOTS` option is selected, the window shown at right is given, which allows the user to select which plots are appropriate for a given run. Once back at the main screen, any of these plots can be created and, if desired, copied to a printer or a PostScript file. In addition, the `AUTO-POSTSCRIPT` option allows for all of the selected plots to be automatically created and copied to PostScript files.

The `mapplot` function included a number of globally defined plots that users could employ. By defining a standard set of plots, it became that much easier for each tester to produce a complete review package for each test run. In addition, mapplot supported a very simple format

for each user to define their own plots, either globally across all of their flight software tests, or specific to a given test.

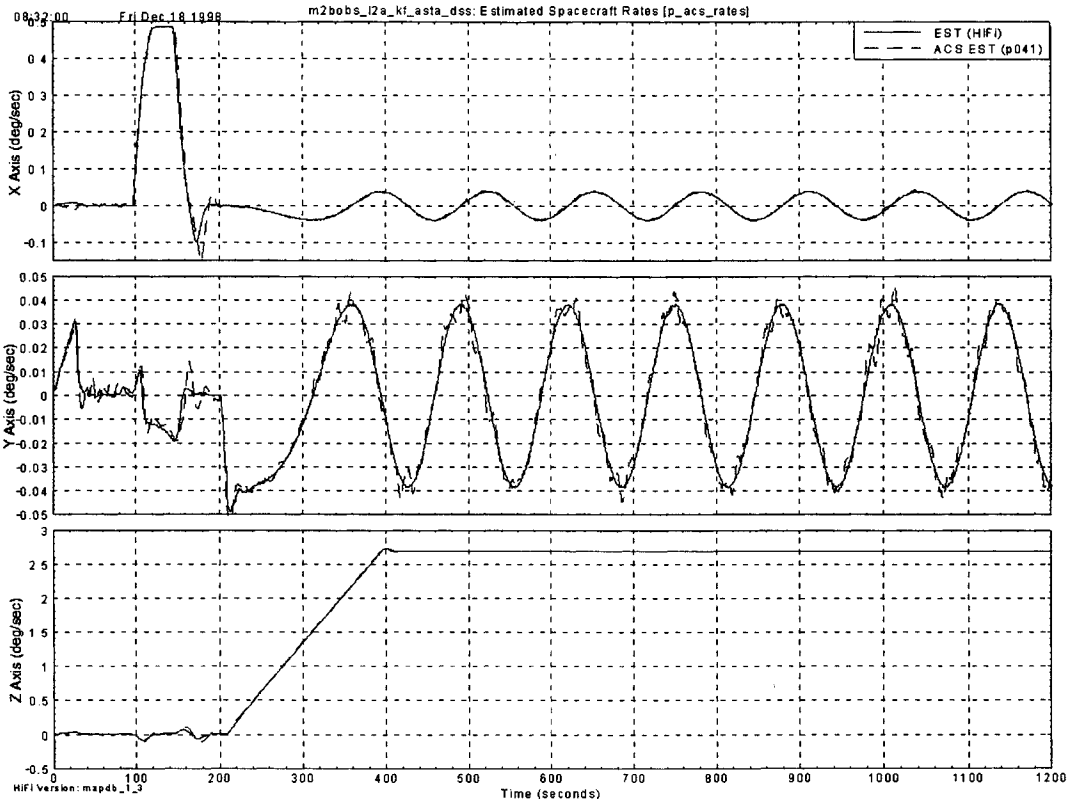


Figure 7: Synchronized Data on Common Plots Aids Test Verification

In addition to the `mapplot` function, a lower-level plotting function was provided in the test tools, called `pktplot`. Unlike `mapplot`, which provided a choice of a number of predefined plots, `pktplot` gave access to any telemetry point in any test data output packet, or any HiFi output variable. This ability was primarily meant for troubleshooting problems or anomalies in a given flight software test, or for producing a very specific plot.

One final piece of the plotting capabilities of the MAP automated test tools should be mentioned. In defining the plots used by the `mapplot` function, the tools only make it necessary to specify the telemetry point desired. How then do the tools know what data packet a given telemetry point is in? This information is kept in the "packet key". A `setpktkey` function was written which operated in the background, calling a PERL script to process through the standard set of sequential print STOL procedures that all flight software tests employed. By parsing out the telemetry points from each STOL procedure, a packet key matching telemetry points and packets was established. This packet key was then used by the other test tools to figure out where to find each needed telemetry point.

Summary of Flight Software Test Verification Steps

Referring back to Figure 3, which detailed the automated test process used on the MAP project, the following short list of steps reflect everything needed to be done to perform a flight software test and its verification:

1. From a flight software test scenario, write a STOL procedure to test that scenario on FlatSat.
2. Copy the sequential print output files to the MatrixX computer.
3. Within MatrixX, run the `pktproc` function to process the data; `pktproc` automatically calls `msync` and `transitions` to synchronize the data and set up the HiFi simulation.
4. After a review of the HiFi simulation script created in the previous step, run the simulation.
5. Use the `mapplot` function to create an appropriate set of comparison plots.
6. Review the comparison plots to determine whether or not the flight software test passed.

As can be seen in the above process, the only major manual step is in step 1, where the original flight software test procedure is created. Other than possible manual tweaks that might be required for the HiFi verification simulation in step 4, everything else is automated by the MAP test tools until the final step where the flight software test data is compared with the data from the HiFi simulation and a PASS or FAIL is given to the test.

REUSABILITY

The flight software test tools used for MAP ACS flight software testing and described above were not the result of a formal development process. Rather, the tools began as an assortment of testing shortcuts from one of the MAP testers; as the collection grew, these shortcuts were assembled into a toolset that eventually allowed much of the test verification process to be automated. Because the tools were written in an assortment of MatrixX's MathScript scripting language and the Unix utilities and languages of bash, grep, sed, and PERL, and not thoroughly documented, they were exclusively geared towards use on the MAP project.

When development of the Triana spacecraft began at Goddard, the Triana ACS subsystem team chose to base their high fidelity simulation on MAP's. Because this meant that the Triana HiFi was closely related to the MAP HiFi, it was logical to assume that Triana could also benefit from the test tools developed for MAP, after they were converted for Triana's use.

Tool Conversion

The decision was made to reuse the MAP test tools after the Triana lead ACS analyst advocated their use. The MAP test tools were demonstrated for the flight software test lead, the analytical team and the ACS flight software lead. Everyone was enthused, and agreed that their use would be beneficial for Triana. Input was solicited on what people wanted to keep and what they didn't want, and based on this input, Dave McComas, a member of the MAP flight software development team, worked on reshaping the tools to Triana's wants and needs. While not the

original developer of the tools, he was familiar enough with MAP, MatrixX, and the Unix utilities to be a good candidate for adapting them for Triana.

The Triana team identified the following features of the MAP test tools as the ones they would be most interested in:

1. the ability to time synchronize test data
2. the ability to plot data, and to batch produce a set of plots
3. the ability to automatically make a matching HiFi verification simulation

Further, due to time and resource constraints, Triana chose to adopt only the first and second features shown. In terms of the specific MAP test tools, Triana versions of `pktproc`, `mapplot`, `msync`, and `pktplot` were developed. The routine used to automatically produce a HiFi script corresponding to the test, `transitions`, was not adapted.

Because of the nature of the MAP test tools, the first step necessary before adapting them to use on Triana was to "reverse-engineer" and document them, learning how they work. This was done with the assistance of the original tool developer. The tools and techniques themselves were not difficult to translate, once the underlying assumptions about standards were understood. As mentioned, a big reason that the MAP test tools worked as well as they did was because of the widespread use of standard test procedure and data output formats established for MAP testing. Once the assumptions about these standards were understood, it became possible to adapt them to Triana by establishing corresponding standards there.

The greatest difficulties in translating the MAP test tools for Triana was caused by differences in the ground systems used on the two projects. In particular, the Triana system worked with different time formats; in order to be able to synchronize data it is vital to be able to work with compatible times, so establishing a compatible standard was important. Also, because the Triana ground system had to deal with data rates much higher than with MAP (10 Hz vs 1 Hz data) and would sometimes drop data packets more frequently, the Triana test tools needed to be able to deal with that.

After discussions with the Triana test team and a number of iterations, a workable set of tools was produced for use on that project. Additionally, unlike with the "home grown" set of MAP test tools, a complete user's guide was produced.

Experience with the Triana Test Tools

The experience of using the automated test tools on Triana ended up being significantly different than MAP. In spite of the early enthusiasm on the part of the ACS flight software and flight software test leads, as well as the ACS analysis team, the entire group did not "buy in" to the use of the tools as much as would have been desired.

On the positive side, there were some features of the automated test tools that everyone liked, and the tools saw their most use for these things. Especially useful was the tools' ability to synchronize test data from different packets to allow it to be easily "lined up" within a data plot.

The ability to batch output an arbitrarily long series of standardized plots for a test data run was also found to be useful. Finally, the point and click user interface to plotting (as shown above in Figure 6) was liked.

For Triana, there ended up being more negative feelings toward the test tools than positives for many members of the test team. Because of the lack of integration of the flight software and ACS teams, there was a greater division of labor than was present on MAP. Combined with the fact that the `transitions` procedure used to "close the loop", automatically producing a HiFi script to correspond to a given test data set, was not adapted for Triana, this meant that there was much less perceived benefit from using the tools. Further, because many members of the team were not familiar with MatrixX or Unix, and were not provided with the tools until after the project was already underway, they were less likely to want to change the way they were doing things in order to use them.

In summary, the data synchronization and batch plotting abilities of the test tools were found to be invaluable to the Triana project. However, because of the less-integrated flight software test and ACS teams on Triana, the "imposed" nature of the tool use resulting in less buy in from the team, and the fact that many members of the team had established their own procedures before the tools became available, meant that the automated test tools did not make as big an impact on the Triana project as on MAP.

CONCLUSION

What began as an *ad hoc* development of software "shortcuts" for ACS flight software testing on the MAP spacecraft evolved over time into a very useful set of tools for greatly increasing the efficiency of the testing process. By standardizing interfaces and test procedures and developing new data analysis tools, the MAP project was able to leverage its existing ACS subsystem-wide parameter database, an integrated test team consisting of flight software developers, testers, and ACS analysts, and the MatrixX integrated simulation, analysis, and plotting toolkit, to remove many of the testing bottlenecks that slowed down projects in the past. As a result, MAP was able to go through build and acceptance testing, as well as a number of rounds of regression testing to accommodate late software changes and additions, in a very timely, efficient, and thorough manner.

Because the HiFi simulation used for the Triana project also used MatrixX and was directly based on the MAP HiFi, it would have seemed that Triana was an ideal candidate for also making use of the MAP test tools. However, the experience of trying to apply them for use on Triana showed some of the potential pitfalls and limitations when attempting this sort of reuse. Because the MAP test tools were not the result of a formal development process, they were not well-documented at first. It took some time for the conversion and documentation process to be performed, which meant that the tools were not available to Triana from the beginning. Additionally, on many levels the Triana project was not as integrated as MAP, with no central parameter database for all components of the testing environment, separate ACS and flight software test teams, and more difficulty establishing the standard test procedure and data output formats upon which the MAP tools depend. Nevertheless, while not used as widely or universally on Triana as on MAP, the tools did provide some benefit to the Triana's ACS flight

software testing effort. It is clear that the techniques inherent in the MAP test tools, if not the actual tools themselves, can be applied to great benefit in future projects.

REFERENCES

[1] McComas, David C., James R. O'Donnell, Jr., Ph.D., and Stephen F. Andrews, "Using Automatic Code Generation in the Attitude Control Flight Software Engineering Process", *23rd Software Engineering Workshop*, NASA Goddard Space Flight Center, Greenbelt, MD, 1998.

[2] Ward, David K., Stephen F. Andrews, David C. McComas, and James R. O'Donnell, Jr., Ph.D., "Use of the MatrixX Integrated Toolkit on the Microwave Anisotropy Probe Attitude Control System," *21st AAS Guidance and Control Conference*, Breckenridge, CO, 1999.

[3] O'Donnell, James R., Jr., Ph.D., Stephen F. Andrews, David C. McComas, and David K. Ward, "Development and Testing of Automatically-Generated Flight Software for the MAP Spacecraft," *14th International Symposium on Space Flight Dynamics*, Iguassu Falls, Brazil, 1999.

A TWO-WHEEL OBSERVING MODE FOR THE MAP SPACECRAFT

Scott R. Starin and James R. O'Donnell, Jr., Ph.D.
NASA/GSFC, Flight Dynamics Analysis Branch, Greenbelt, MD 20771 USA

Abstract

The Microwave Anisotropy Probe (MAP) is a follow-on to the Differential Microwave Radiometer (DMR) instrument on the Cosmic Background Explorer (COBE). Due to the MAP project's limited mass, power, and budget, a traditional reliability concept including fully redundant components was not feasible. The MAP design employs selective hardware redundancy, along with backup software modes and algorithms, to improve the odds of mission success. This paper describes the effort to develop a backup control mode, known as ObservingII, that will allow the MAP science mission to continue in the event of a failure of one of its three reaction wheel assemblies. This backup science mode requires a change from MAP's nominal zero-momentum control system to a momentum-bias system. In this system, existing thruster-based control modes are used to establish a momentum bias about the sun line sufficient to spin the spacecraft up to the desired scan rate. Natural spacecraft dynamics exhibits spin and nutation similar to the nominal MAP science mode with different relative rotation rates, so the two reaction wheels are used to establish and maintain the desired nutation angle from the sun line. Detailed descriptions of the ObservingII control algorithm and simulation results will be presented, along with the operational considerations of performing the rest of MAP's necessary functions with only two wheels.

Introduction

The MAP Observatory will nominally use three reaction wheels for most of its attitude control requirements. See references 1 and 2 for further information on MAP. In the event of a failure of one of MAP's three reaction wheel assemblies (RWAs), it is not possible to achieve three-axis control using the remaining two wheels. Because of this, two of the attitude control algorithms implemented on the MAP spacecraft will no longer be usable. The two are Inertial Mode, used for slewing to and holding inertial attitudes, and Observing Mode, which implements the nominal dual-spin science mode. As a result of a Red Team Review (Sept. 12-13, 2000), the pre-launch development of a strategy for completing the mission in the case of a wheel failure became an imperative. The bulk of this paper will discuss the design for a two-wheel science mode for the MAP spacecraft. However, in addition to this backup science mode, there are a number of other changes that need to be made to the MAP onboard flight software in order for it to be able to fulfill even a degraded science mission. In this section, the philosophy used in designing these changes is shown, followed by a discussion of the changes and additions themselves.

Wheel Failure Design Philosophy

In order to be able to deliver and implement a backup two-wheel control design in a timely fashion, a design philosophy was first adopted.[2] The elements of this philosophy are as follows:

- Wherever possible, existing control algorithms already implemented and tested would be used as is, or with as few changes as possible. ⇒ Reduces development and testing time.
- Where completely new algorithms are needed, such as in the new science mode, they would be implemented in a manner consistent with the current flight software design. This design makes extensive use of tables of the parameters needed for proper configuration. ⇒ Allows for flexibility on-orbit for configuration and tuning of the control algorithms.
- New and changed algorithms would be prioritized by when they are necessary, and the development and testing schedule set up to reflect this. ⇒ Most effectively uses available resources to maximize the chances of mission success in the event of a wheel failure.

Required Mission Functions and Implementation Plan

The following functions are required in order to be able to carry out the MAP mission. After the description of each function, the way in which this function is implemented is shown. The selected implementation was based on the philosophy discussed above.

- Safehold at Low System Momentum: The existing MAP Safehold/CSS controller, named for its use of coarse sun sensor signals for attitude and derived rate information, works as is with only two wheels.
- Two-Wheel Science Mode (ObservingII): This mode will be fully discussed below. It is important to note here that the two-wheel science mode works by first establishing a 20–25 newton·meter·second (Nms) momentum bias about the Sun line. The controller then increases the nutation angle to the desired value to approximate the dual spin of the nominal science mode controller.
- Establishing and Removing Momentum Bias: The existing MAP Delta H Mode, a thruster-based mode nominally designed for dumping system momentum, can be used to establish or remove the science mode momentum bias. The only necessary changes are to parameters in an existing flight software table.
- Thruster Operations for Orbit Maneuvers and Maintenance: In order to fulfill its mission, MAP must get to the L_2 libration point, which means that it must be able to perform a number of orbit maneuvers during phasing loops about the Earth (on the final phasing loop, MAP performs a lunar swingby which provides the final "push" to L_2). Once at L_2, orbit maintenance maneuvers must also be performed, nominally four times a year. The existing MAP Delta V Mode performs this function in both the nominal and two-wheel case.

- Thruster-Based Inertial Mode: With the loss of Inertial Mode, a wheel-based mode that allows MAP to be slewed to any inertial attitude, an alternative way is needed to get the spacecraft in the right attitude for thruster operations. With two relatively small changes to the existing Delta V controller, it will be possible to use it for this function, as a Thruster-Based Inertial Mode.
- Safehold with a Momentum Bias: The two-wheel Safehold/CSS also works with a momentum bias as long as the bias is close to its nominal orientation about the Sun line. **Figure 1** shows an example of the performance of this mode.
- Momentum Bias Adjustment: A big difference between the nominal and two-wheel science modes is that the nominal mode performs its dual spin about the Sun line, while the two-wheel mode performs its dual spin about a momentum bias, nominally applied about the Sun line. This momentum bias is fixed in inertial space, however, and will move relative to the Sun line approximately 1°/day. A new thruster control mode is needed to perform the small daily adjustments to the momentum bias that will be necessary in flight.

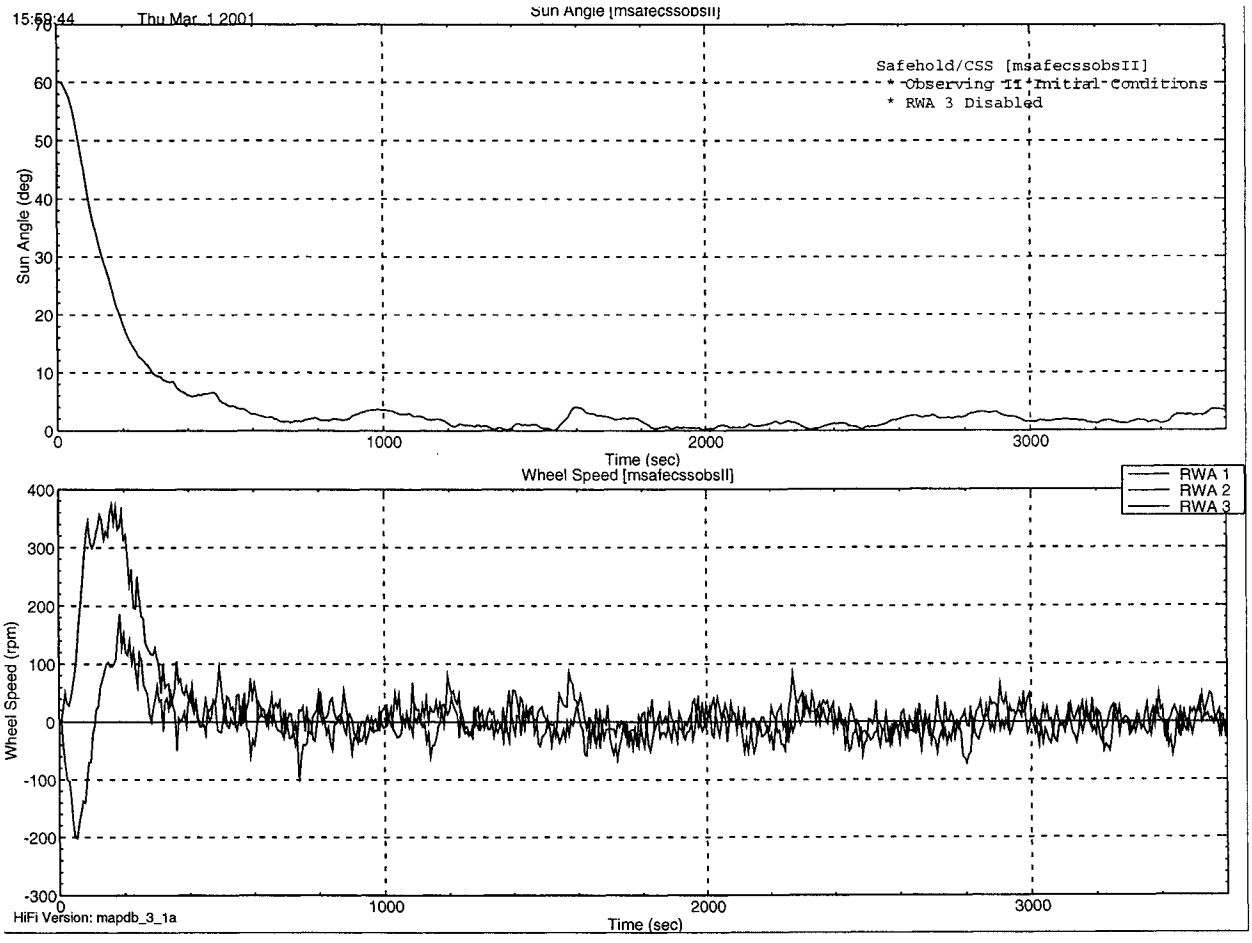


Figure 1: Safehold/CSS with Two Wheels and 20 Nms Sun line Momentum Bias

Upon reviewing the necessary functions needed for MAP to fulfill its mission, one thing becomes clear that both drives and simplifies the development schedule. Of the functions that are not supported by existing algorithms on the spacecraft, only the Thruster-Based Inertial Mode function is necessary in order for MAP to be able to reach the second Lagrange point (L_2). The two-wheel science mode and momentum bias adjustment functions are not needed until MAP reaches L_2, at least 90 days after launch. In order for the spacecraft to be able to get there, it is necessary to have a means to get the spacecraft into the correct attitude for its critical phasing loop thruster operations so that it can achieve the correct orbit for its lunar swingby. Because of this, the development of the backup algorithms and software is being done in two phases. Phase 1 includes what is needed to get to L_2 and will be completed before launch. It will be tested and available to be uploaded as a patch to the onboard software in the event of a wheel failure.

In the remainder of this paper, the three new or changed algorithms that are being designed for the MAP two-wheel contingency will be discussed. Fairly brief descriptions of the thruster-based inertial mode and momentum bias adjustment algorithms and implementations will be given, followed by extensive discussion of the two-wheel science mode.

Thruster-Based Inertial Mode: "Zero ΔV" Delta V

One of the design goals for the two-wheel contingency algorithms was to use existing control algorithms already implemented and tested on the spacecraft as much as possible, with as few changes as possible. Because of the nominal ACS flight software design, it was found that it was easier to implement a thruster-based inertial control mode through small changes to the existing Delta V Mode, rather than the existing Inertial Mode. The existing Delta V Mode was already set up to hold an inertial attitude during a burn; only two changes were necessary to allow for its use in the two-wheel contingency case for slewing to and holding inertial attitudes:

- The current Delta V Mode is meant for thruster operations, not simply holding an attitude, and automatically exits when a commanded burn is completed. In order for the mode to be used as a thruster-based inertial mode independent of an orbit maneuver, an option to allow for a "Zero ΔV" Delta V was necessary. If commanded in this fashion, the mode would only exit when it timed out (controlled by a flight software table value) or upon command into another control mode (or into a "conventional" Delta V orbit maneuver). A preliminary implementation of the "Zero ΔV" Delta V Mode was tested and showed acceptable performance.
- The existing Delta V Mode was not meant to execute large inertial slews; attempting to do so would result in gyro rate saturation because the algorithm does not include any rate-limiting action. In the two-wheel contingency case, an attitude limiter is added to the Delta V Mode controller to effectively limit its slew rate and allow it to be used for inertial hold and slews of arbitrary length.

Momentum Bias Adjustment

As mentioned above, the two-wheel science mode performs a dual spin about a momentum bias nominally applied about the Sun line. Because this momentum bias remains fixed in inertial space, the Sun line moves relative to the bias approximately 1°/day. Therefore, in order to keep the science mode spin about the Sun line, the momentum bias must be adjusted daily. Because MAP nominally only has one 37-minute ground contact per day, this momentum bias thruster operation must not be overly complicated.

In order to implement this, an algorithm has been designed that uses the instantaneous estimate of spacecraft attitude to determine when the body z-axis lies in the plane of the ecliptic, the plane in which the momentum bias should be applied or adjusted. See **Figure 2** for axis definitions. With a momentum bias of 20 Nms, this occurs once every 90 seconds. In order to move the direction of the momentum bias 1° inertially, when the body z-axis is in the ecliptic plane the momentum can be adjusted in either the body z-axis or the body xy-plane. Because there is more momentum in the body z-axis, it takes more Δ momentum in that axis to move the direction 1° than in xy. So, xy-adjustments can be made for five or six days, each of which slightly increase the magnitude of the momentum bias while adjusting its direction, followed by one day with a z-axis adjustment to return the momentum bias magnitude to its nominal value. The size of the momentum bias adjustment burns, which will be determined by comparing an average value of body momentum and the Sun line direction, are less than one second of burn per thruster for one or two thrusters. Fuel usage is on the order of a total of 30 seconds of thruster firing over 90 days, which is a fairly insignificant portion of the fuel budget.

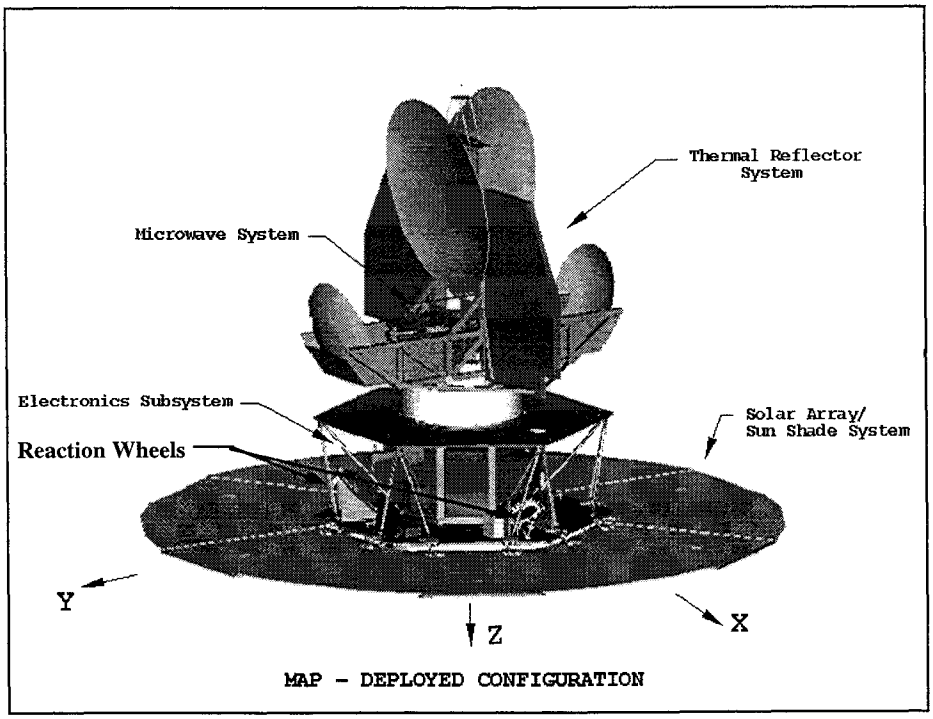


Figure 2: Configuration of MAP spacecraft with solar arrays deployed.

MAP Two-Wheel Science Mode

The MAP Observatory is equipped with three reaction wheels, which are arranged symmetrically about the spacecraft's z-axis. Wheel #1 lies in the +x/–z quarter-plane of the spacecraft body reference frame, and its unit torque vector is $n_1 = [-0.866\ 0\ -0.5]^T$. Wheels #2 and #3 have the same 30° cant with respect to the z-axis, with unit vectors $n_2 = [0.433\ -0.75\ -0.5]^T$ and $n_3 = [0.433\ 0.75\ -0.5]^T$, respectively. The nominal-mission Observing Mode uses the three reaction wheels to establish commanded Euler angle rates.[1] A 3-1-3 Euler angle rotation is used, with the angles named ϕ, θ and ψ, respectively. The rates to be commanded are:

$$\dot{\phi} = 1\ rph = 0.1^\circ/_s \tag{1}$$

$$\dot{\theta} = 0^\circ/_s; \quad with\ \theta = 22.5^\circ \tag{2}$$

$$\dot{\psi} = 0.464\ rpm = 2.8^\circ/_s \tag{3}$$

The angle θ, *i.e.* the angle between the Sun vector and the unit vector along the geometrical z-axis, is maintained at a constant value of 22.5°. The rate in ψ represents a spin about the z-axis, and the rate in ϕ is therefore a precession of the z-axis.

Since the precession rate is so much slower than the body spin rate, the body momentum is nearly parallel to the z-axis, and this momentum is precessed in the inertial space by the wheels. In other words, the spacecraft momentum associated with commanded rates is stored by the wheels, and nominal science operations are kept at very low system momentum. These rates are established by continually generating a small delta-quaternion, Δq, and then commanding the wheels to enact that Δq. Because the system momentum is small, tiny corrections in the calculated Δq values are sufficient to adjust for the rotation of the Sun vector in inertial space. This allows the precession of the z-axis to remain symmetric about the Sun line. Thus, the combination of three rates – spin, precession, and Sun vector rotation – sweeps the instrument boresights over the entire sky in a period as short as six months.

Because of the large difference between various rates, this nominal science mode may be said to operate under the condition of being able to store momentum of a specified magnitude, but of an essentially arbitrary direction with respect to the body frame. With the loss of the use of a reaction wheel, however, the ability to store momentum in any arbitrary direction is lost. The two remaining wheels are limited to storing momentum in directions parallel to the plane spanned by their unit torque vectors:

$$\begin{aligned} \vec{H}_{RWA} &= \frac{d}{dt}\vec{N} = \frac{d}{dt}\left(\underline{n}_\alpha J_\alpha \omega_\alpha + \underline{n}_\beta J_\beta \omega_\beta\right) \\ &= \underline{n}_\alpha J_\alpha \left(\frac{d}{dt}\omega_\alpha\right) + \underline{n}_\beta J_\beta \left(\frac{d}{dt}\omega_\beta\right) \end{aligned} \tag{4}$$

Here and throughout the paper, H denotes momentum; N, torque; $\underline{n}$, any characteristic unit vector (here denoting the wheel spin axes); J, the moment of inertia of a wheel about its spin axis; ω, angular velocity. The subscripts α and β refer to the two operable wheels after the failure of the third. If wheel #1 had failed, α would be 2 and β would be 3.

Since the wheels are arranged in a pyramid about and canted 30° to the z-axis, the nominal method of having the momentum nearly parallel to the z-axis becomes unfeasible. In the zero-momentum scheme of attitude control, two wheels may be shown capable of enacting an algorithm to reach any given attitude[3,4]. Unfortunately such an algorithm requires large slews for even very small attitude corrections and is therefore unacceptable for the purposes of smooth rate establishment. As for smooth feedback algorithms, Byrnes and Isidori[5] showed that "a rigid satellite with (one or) two independent actuators cannot be locally asymptotically stabilized using continuously differentiable static or dynamic state feedback." And though Morin *et al* [6] have shown that a time-varying or discontinuous feedback controller may stabilize the attitude using only two inputs, their design requires the use of external torques in the form of jets or other propulsion-based attitude control. Since every indication is that control of the full-state attitude of a rigid spacecraft is not possible using two internal torque inputs, a design approach that is completely different from the nominal design is required.

Dynamical Considerations

The basic concept for an alternate control approach using two wheels was presented by O'Donnell *et al.*[2] In the event of a wheel failure, the Safehold control mode is capable of directing the spacecraft z-axis, and therefore the solar arrays, toward the Sun and holding an instrument-safe, power-positive condition until new software can be uploaded.

The newly installed software would spin the spacecraft up to a momentum of approximately 20–25 Nms (2.3–2.9°/sec) about the spacecraft z-axis using thrusters. This value provides rates of about the same magnitude as the nominal Observing mode. Once this momentum vector is established, the backup algorithm would increase the Sun angle, θ, to the desired value of 22.5° and maintain that angle as well as possible. Since thrusters may be used to periodically realign the momentum vector with the Sun vector, such maintenance should require little effort since the Sun angle would then coincide with the nutation angle of the z-axis about the momentum vector.

The natural dynamics of the rigid body allows this approach because of the mass properties of the MAP observatory. The current estimates for end-of-life (EOL) moments of inertia about the x- and y-axes are nearly equal (I_{xx} = 572 kg·m^2 and I_{yy} = 580 kg·m^2), and the moment of inertia about the z-axis is smaller (I_{zz} = 496 kg·m^2), so that the body is prolate. EOL values are used because most of the MAP fuel budget is used when travelling out to its target orbit at L_2; the beginning-of-life (BOL) ratio is $\sigma = 0.84$.

This near symmetry greatly reduces any effects a non-zero value of the xy-product of inertia might have, so long as the z-axis is in close alignment with a principal axis. As established by launch vehicle requirements, the geometrical z-axis is within 0.25° of the nearest principal axis. Misalignments of this magnitude have proved negligible in the development of this backup mode. Since they have little bearing on the remainder of the paper, the differences between the geometrical axes and the principal axes will no longer be mentioned. The axis frame "x-y-z" will refer to both axis frames simultaneously.

If we define a transverse moment of inertia, $I_T \equiv \sqrt{I_{xx} \cdot I_{yy}} = 576$ kg·m^2, then the ratio of the z- and transverse moments of inertia is $\sigma = 0.86$. From rigid-body dynamics, the ratio between the inertial nutation rate(ϕ-rate; analogous to precession of nominal mode) and body nutation rate (ψ-rate; spin with respect to momentum vector) is:

$$\frac{\dot{\phi}}{\dot{\psi}} = \frac{1-\sigma}{\sigma}\cos\theta. \tag{5}$$

For a nominal Sun angle of $\theta = 22.5°$, this ratio is approximately 0.15 at EOL, or 0.18 at BOL. For a system momentum of 20 Nms, the inertial nutation rate would be about 0.33 revolutions per minute, and the body nutation rate would be about 3 revolutions per hour. See **Figure 3** below for a comparison of nominal and back-up scan patterns.

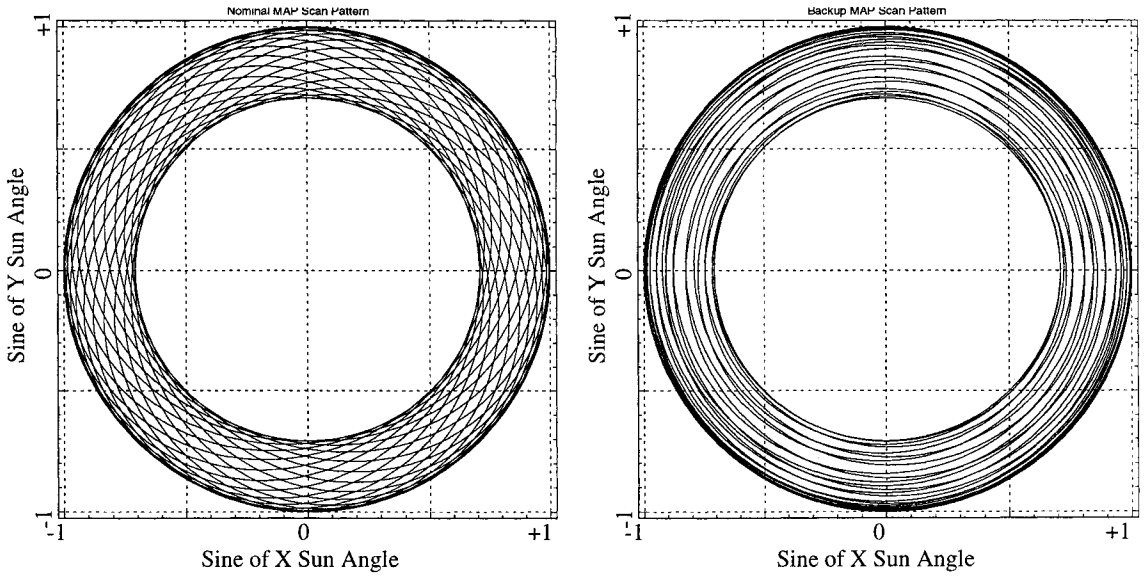


Figure 3: Comparison of scan patterns in nominal and two-wheel science modes.

Were there no internal sink for energy dissipation on the Observatory, whatever Sun angle was established by the control law would remain constant except for the gradual rotation of the Sun vector in inertial space. However, the fuel tank on MAP, with its elastomeric diaphragm, provides just such an energy sink in the form of fuel slosh. Therefore, the control law must be able to respond to and to counteract any increase in θ.

The inequality of the x- and y-moments of inertia, though small, is not negligible. This asymmetry results in a sinusoidal variation of the body and inertial nutation rates and the Sun angle. Specifically, the Sun angle varies in a peak-to-peak range of 1.5°. Intuitively, the variation is a result of the constant system momentum nutating through varying moments of inertia in the rotating spacecraft frame. To remove this wobble requires controlling the magnitude or direction of the system momentum in the spacecraft frame independent of that frame. Again, two obliquely mounted wheels are incapable of actuating control torques in arbitrary directions.

Because the original requirement of 22.5°±0.25° cannot be met using only two wheels, the goals for this backup control law have been changed to match the priorities

dictated by the science goals. Specifically, the heating of the instrument by sunlight is to be avoided above all other concerns, including the completeness of the current scan pattern. To avoid heating of the instrument by diffracted sunlight, the Sun angle, θ, will not be allowed to exceed 22.75° during normal establishment of, and operations during, ObservingII. If 22.75° is exceeded, the algorithm should reduce θ as quickly as possible. To avoid sunlight falling directly on the instrument, θ will never be allowed to exceed 25° under any circumstances. In the event that an anomaly places the Sun angle outside of 25°, the current Safehold mode (which operates using two wheels, as noted above) will take over attitude control.

Beyond the protection of the instrument, the largest, most constant Sun angle feasible is to be maintained by the control algorithm. The attitude determination should not suffer, since a complex scan pattern may still be processed as long as the attitude data is good. This requirement is currently the limiting factor for the rate selection; the star trackers provide less accurate attitude determination as the transverse rates increase. The star tracker accuracy would have to be balanced with the increased sky coverage granted by higher rates. Lastly, there is some indication that, due to possible limitations on contact times, the algorithm should be able to establish backup operations as quickly as can be done without risking the other objectives.

Algorithm Design

The control algorithm design process was more exploratory, rather than strictly analytical. This approach was driven by factors such as nonlinear dynamics, underactuated control, a reduced set of state variables to be controlled (only θ needs to be controlled), and sparse literature from which to start. Due to limited software development resources, the algorithm was to fit into the same space in the software, with the same inputs and outputs, as the nominal science mode. The main effect of this restriction on the ObservingII algorithm was that, since Sun angle was not already input into the nominal Observing mode, the backup ObservingII mode had to rely on the measurement of body and wheel rates to establish the angle between the z-axis and the momentum vector.

The first conceptual step in the design process was to consider controllability. The wheels, which are fixed in the body frame (x-y-z), can have no effect on system momentum. They can only provide momentum to the body by temporarily storing the opposite momentum, thus maintaining an inertially constant system momentum vector. This storage must be temporary because wheel drag will eventually bring each wheel to rest. It soon became clear that the movement of the system momentum vector through the body frame was vital to controllability; if the momentum were stationary in the body frame, any momentum stored by the wheels would necessarily be released along exactly the same direction and thus undo any actuation that had been commanded.

The general dynamics are governed by Euler's equation,

$$\dot{\bar{H}}_{sys} = -\bar{\omega}_b \times \bar{H}_{sys} \quad (body\ frame) \tag{6}$$

which may be expanded in the following manner:

$$\dot{\vec{H}}_b = -\dot{\vec{H}}_{RWA} - \vec{\omega}_b \times \vec{H}_b - \vec{\omega}_b \times \vec{H}_{RWA}, \tag{7}$$

$$where\ \vec{T} = \vec{T}_{command} + \vec{T}_{drag} = -\dot{\vec{H}}_{RWA}. \tag{8}$$

If the simplification is made that $I_{xx} = I_{yy}$, then the control-free motion may be parameterized using the angle between the system momentum and the spacecraft z-axis, θ", and the azimuthal angle for the body momentum, ψ":

$$H_x = H \sin\theta" \cos\psi" \tag{9a}$$

$$H_y = H \sin\theta" \sin\psi" \tag{9b}$$

$$H_z = H \cos\theta" \tag{9c}$$

$$\psi" = \dot{\psi}" t + \psi"_{t=0}. \tag{9d}$$

Here, H is the magnitude of the system momentum and remains constant as the system momentum vector sweeps a cone-like surface in the rotating body frame; if θ is constant, then the surface is truly a cone. The angles θ" and ψ" are analogous to the Euler angles with respect to the Sun—specifically θ, the Sun angle; and ψ, the angle between the x-axis and the xy-projection of the Sun vector. The two sets of angles are identical when the system momentum is exactly aligned with the Sun vector.

The effect of $I_{xx} \neq I_{yy}$ is that the cone-like surface, which is circular as parameterized above, is in fact elliptical. Since no further benefit is gained by using the less familiar elliptical parametric equations, the circular equations were used in the design of the algorithm.

Note that, since the motion of the body momentum vector is nearly symmetric about the z-axis, and the wheels are also symmetric about the z-axis, a geometrical simplification may be used to make one general control law work in nearly the same way no matter which wheel fails. The failed wheel must be identified and its identity entered into the software. Then, the algorithm may redefine the body frame such that the failed wheel is located on the new positive x-axis, which may be called the x'-axis. The operable wheels are then situated symmetrically about the x'-axis. The z-axis is identical with the z'-axis, and the y-axis is rotated into the y'-axis accordingly. The –y-wheel is referred to by the subscript α, and the +y-wheel is referred to by β. While there were differences in the configuration of the algorithms for the failure of each wheel, there were no problems large enough to warrant designing an entirely different algorithm for each wheel.

The first attempt at a control law design was to establish an open-loop rule for torque input to be provided by a single wheel. The body momentum vector could be calculated by measuring the body rates and multiplying by the measured spacecraft inertia matrix. When the body momentum vector was most advantageously oriented with respect to the torque direction, the wheel would torque according to the error in the Sun angle, θ. Then, when the wheel torque would have relatively little effect, the wheel speed was restored to zero, often passively by drag torque. This method appeared to perform

appropriately, but the open-loop nature of the control algorithm was undesirable. So, other solutions were sought.

Most of the design effort was directed toward the use of the Lyapunov theorem to reduce a set of error variables to zero. The effort to find a Lyapunov function of all appropriate variables was, in the end, unsuccessful. One attempt was to use the following error variables to find a Lyapunov function:

$$\begin{aligned} &\textit{Transverse error}: e_T \equiv H_b \sin\theta - H_{b,desired} \sin\theta_{desired}, \\ &Z-\textit{axis error}: \quad e_Z \equiv H_b \cos\theta - H_{b,desired} \cos\theta_{desired}. \end{aligned} \tag{10}$$

Functions of these variables were chosen that were positive definite, and the wheel speeds were included. The time derivatives of the function tried, however, contained terms that were independent of the wheel torques, and so could not be influenced by the control algorithm. Finally, the Lyapunov approach was abandoned.

Though the search for a true Lyapunov function was unsuccessful, the various attempts resulted in several simple one- and two-wheel algorithms based on the error functions defined above. A few of the algorithms acted to stabilize the closed-loop system about a desired Sun angle for initial angles of 25° or less. Some algorithms could operate outside of this range, but such capability would be useless, since the Safehold mode would take over control. The control algorithms were composed of three parts: an initial kick to escape limit cycles about the Sun vector, an error-reducing function of the measured body momentum, and an anti-runaway restriction on the speed of the wheels. Since the error-reducing functions had certain similarities, a simple expression was found which included the most promising candidates. This expression, along with the other two components of the algorithm, are listed here in order of precedence (*e.g.* first listed is always commanded):

$$\textit{If } \omega_i \geq 50\,rpm,\ i \in \{\alpha,\beta\},\ \textit{then } T_i = -K\omega_i. \qquad \text{Anti-Runaway} \tag{11}$$

$$\textit{If } \theta < 10°,\ \textit{then } T_\alpha = T_\beta = T_{escape}. \qquad \text{Sun Vector Escape} \tag{12}$$

$$T_\alpha = k_{Z\alpha} e_Z + k_{T\alpha} e_T \cos(\psi + \varphi_\alpha) \qquad \text{Error-Reduction - } \alpha \tag{13a}$$

$$T_\beta = k_{Z\beta} e_Z + k_{T\beta} e_T \cos(\psi + \varphi_\beta). \qquad \text{Error-Reduction - } \beta \tag{13b}$$

According to simulations, the best-performing of these functions has been the one defined by:

$$\{k_{Z\alpha} = 0.007 \text{ Nm/Nms},\ k_{T\beta} = 0.010 \text{ Nm/Nms},\ \varphi_\beta = 30°,\ k_{Z\beta} = k_{T\alpha} = 0,\ \varphi_\alpha = 0°\} \tag{14}$$

This control law, which may be called ObservingII-A, has been tested using high fidelity simulations in order to gain confidence in the stability of the controller. At the time this paper is being written, Monte Carlo simulations are being developed to further establish the stability of the controller. Other means of proving stability should be forthcoming, but none are complete as yet.

Two-Wheel Performance

The figures presented at the end of this section show simulated operation of ObservingII-A. **Figure 4** and **Figure 5** show the case in which Wheel #1 has failed, the spacecraft has been recovered, and the thrusters have established a momentum bias of 20 Nms parallel to the Sun vector. The algorithm shows good stability and performance for initial conditions over the full range of the cone defined by $\theta \leq 25°$; in this case the Sun angle is started at 10°. When the natural motion reduces the Sun angle below the 10° Sun Escape limit, the Sun Escape torques quickly bring the wheels up to their speed limit. At the speed limit, the anti-runaway portion of the algorithm strongly counteracts any increase in wheel speed, and a back-and-forth torquing pattern begins which drives the Sun angle up again.

Starting from Sun angles less than the desired value, the algorithm establishes the desired motion in approximately 6–10 minutes, as is shown in Figure 3. Since the nominal ground contact during science operations is supposed to last 37 minutes, this time is well within the goal of operating within a contact so that results may be immediately observed, and corrections may be applied before returning to autonomous operations for the day.

The simulated performance of the controller varies slightly for the different wheel failure cases. **Figure 6** shows establishment of science operations in each failure case from the same initial conditions. To prevent the maximum Sun angle from exceeding the 22.75° limit, the target Sun angle, $\theta_{desired}$ (see Eq. 10), must be set at a slightly different, empirically determined value for each case. This fine-tuning adjusts for small body asymmetries that were neglected by generalizing the algorithm using rotational symmetry about the z-axis. There are some differences in performance that depend on the geometry of the initial conditions. For example, the case in which wheel #3 fails shows the controller taking somewhat longer to establish the science mode. This delay occurs at some initial conditions in each case, but is never more than the half-period of the inertial nutation cycle.

Though analytical modeling of fuel slosh indicates that the energy dissipation caused by slosh during this natural motion mode would be very small, it is possible that those models are in error. Such energy dissipation would cause the Sun angle to increase. **Figure 7** shows a case where the Sun angle has exceeded the allowable range for science and is close to the dangerous range ($\theta > 25°$) where direct sunlight would heat the instrument. Ideally, the Sun angle would never reach this high a value; nevertheless, the control algorithm must be and is able to stabilize the Sun angle quickly about the desired value. During operations at high Sun angle, the error reduction portion of the algorithm is generally the only active portion, as the wheel speeds do not exceed the anti-runaway limits.

Other non-ideal circumstances such as pinwheel torques or inaccurate momentum adjustments may arise. Also, changes to the drag torque model in the simulations indicate that the algorithm may be sensitive to changes in wheel drag that are not dependent on wheel speed. Because the performance of the control mode is very dependent on effects that cannot be well-understood until MAP has operated in space for some time, considerable adaptability has been built into the algorithm. By changing

certain flight software table values, several types of alteration to the algorithm may be implemented.

The controller may be set on a timer so that it does not activate until a certain amount of time after a microprocessor restart. A counter may be set to start when science operations have been established, as determined by the value of θ entering and remaining within a certain range. When the counter reaches a target value, the controller will zero its commands until θ leaves the acceptable range. Another optional addition uses a quasi-Lyapunov function of the errors and torque to determine if the commanded torques are likely to be useful; non-useful torques, as determined by the sign of the function, are zeroed out. Finally, the errors may be increased over certain ranges in order to speed up the error reduction portion. These addenda to the algorithm are not necessary according to the performance of ObservingII-A in simulation. However, since the simulations are incomplete compared to the unpredictability of deep space operations, it was felt that every practical means of improving the adaptability of the completed control software should be included.

Conclusion

Due to cost and mass budget considerations, the MAP Observatory has no redundant reaction wheel to back up the three nominally operational wheels. In the event of a wheel failure, it is possible to perform all of the necessary attitude control using thrusters and the two remaining wheels. This paper has briefly described the use of thrusters in slewing for critical burns; the necessary software changes will be available at launch for immediate upload should a wheel fail before the spacecraft reaches L_2.

A two-wheel science mode has been detailed in which a thruster-generated momentum bias is used to sweep the instrument boresights across the sky in a manner similar to the nominal science mode. However, this natural-motion science mode results in a decrease in data density over the sky. The momentum bias must point at the Sun as much as is feasible, so it must be adjusted by the thrusters often for the science mode to operate safely and effectively.

Because of the intractability of the attitude control problem, the science mode algorithm has been designed to allow for the adjustment of gains multiplying functions of the angle error. In addition to these gains, other parameters may be used to adjust various failsafe modes built into the algorithm. All of these parameters may be adjusted in a flight software table which is part of the existing architecture. This design philosophy allows a measure of security against unpredictable dynamical effects.

Acknowledgments

The authors would like to thank the engineers of the Flight Dynamics and Analysis Branch and David K. Ward of the GN&C Systems Engineering Branch for their contributions.

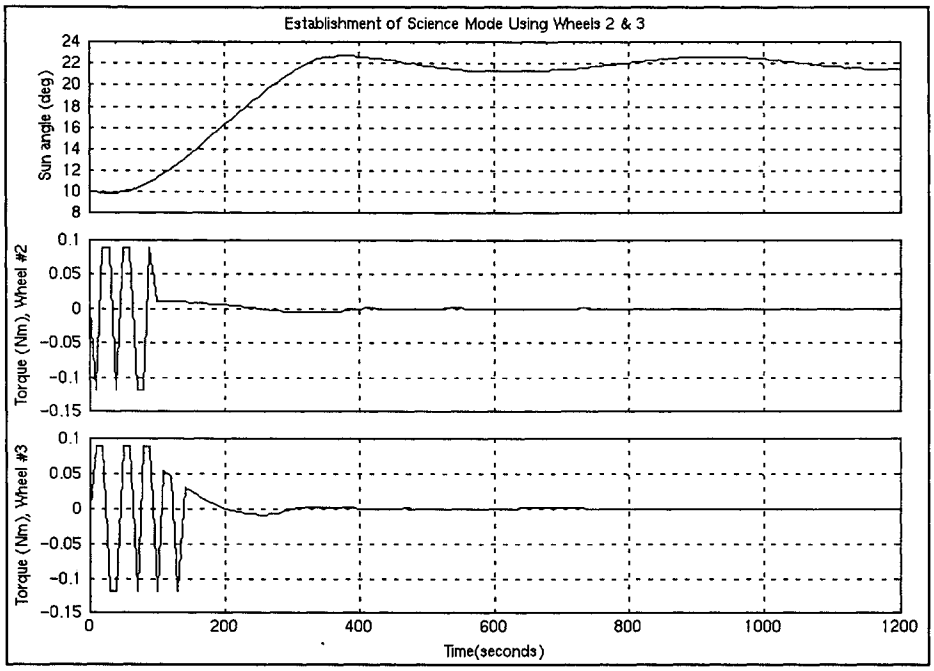


Figure 4: Case of wheel #1 failure. The jagged torque profile shows typical commands during Sun Escape, a portion of the ObservingII algorithm which is used to avoid limit cycles at low Sun angles. After error reduction, the mode is established, and torque commands become unnecessary unless the Sun angle leaves the operating range.

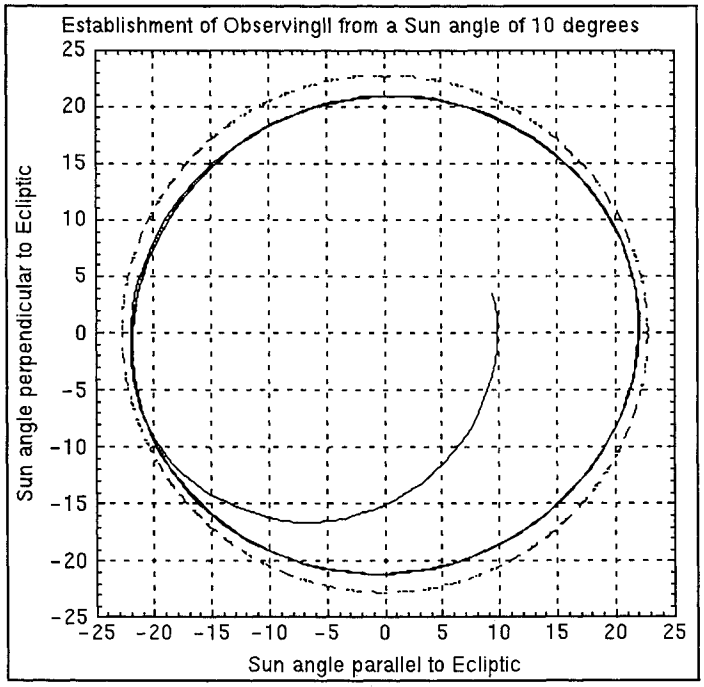


Figure 5: Case of wheel #1 failure. Establishment of mode showing the trajectory of the z-axis direction. The Sun is located at the origin; the dashed line is the 22.75° performance limit.

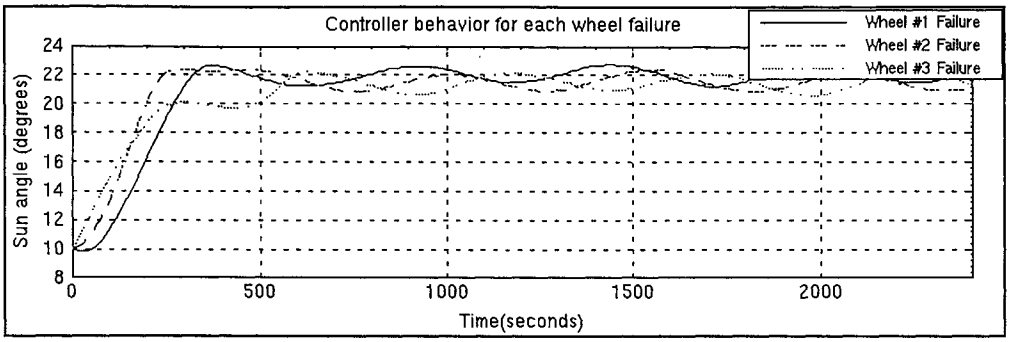


Figure 6: Establishment of ObservingII science mode in all three cases of wheel failure. Initial conditions for each case were identical.

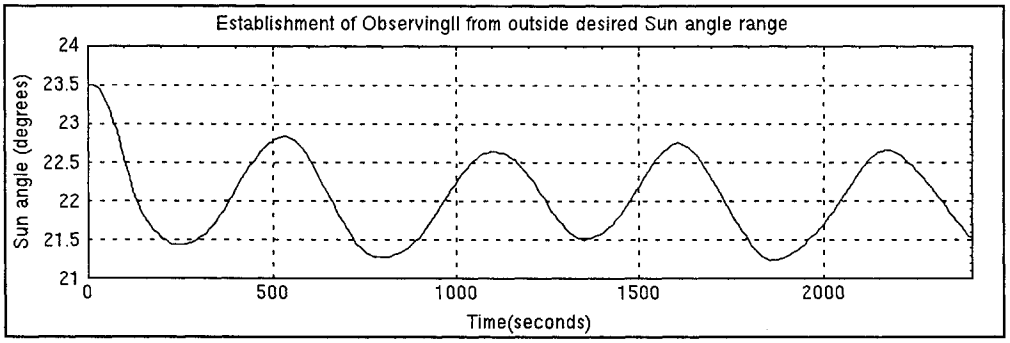


Figure 7: Because the spacecraft symmetry axis has the minimum moment of inertia, the control algorithm must be able to counteract Sun angle increases resulting from energy dissipation. From an initial Sun angle of 23.5°, science operations are restored within 10 minutes.

References

1. S. F. Andrews, C. E. Campbell, A. J. Ericsson-Jackson, F. L. Markley, J. R. O'Donnell, "MAP attitude control system design and analysis", *Flight Mechanics and Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, MD, May 1997.
2. J. R. O'Donnell, S. F. Andrews, A. J. Ericsson-Jackson, T. W. Flatley, D. K. Ward, P. M. Bay, "Backup attitude control algorithms for the MAP spacecraft", *Flight Mechanics Symposium*, Greenbelt, MD,1999.
3. H. Krishnan, N. H. McClamroch, M. Reyhanoglu, "Attitude stabilization of a rigid spacecraft using two momentum wheel actuators", *J. Guidance, Control, & Dynamics*, Vol. 18, No. 2, Mar-Apr 1995, pp. 255–263.
4. V. Coverstone-Carroll, "Detumbling and reorienting underactuated rigid spacecraft", *J. Guidance*, Vol. 19, No. 3, 1996, pp. 708–710.
5. C. I. Byrnes & A. Isidori, "On the attitude stabilization of rigid spacecraft", *Automatica*, Vol. 27, No. 1, 1991, pp. 87–95.
6. P. Morin, C. Samson, J.-B. Pomet, Z.-P. Jiang, "Time-varying feedback stabilization of the attitude of a rigid spacecraft with two controls", *Systems and Control Letters*, Vol. 25, 1995, pp. 375–385.

MAXIMUM TORQUE AND MOMENTUM ENVELOPES FOR REACTION WHEEL ARRAYS

R. G. Reynolds
SpectrumAstro, Inc., Herndon, VA

F. Landis Markley
NASA's Goddard Space Flight Center, Greenbelt, MD

ABSTRACT

Spacecraft reaction wheel maneuvers are limited by the maximum torque and/or angular momentum which the wheels can provide. For an n-wheel configuration, the torque or momentum envelope can be obtained by projecting the n-dimensional hypercube, representing the domain boundary of individual wheel torques or momenta, into three dimensional space via the $3\times n$ matrix of wheel axes. In this paper, the properties of the projected hypercube are discussed, and algorithms are proposed for determining this maximal torque or momentum envelope for general wheel configurations. Practical implementation strategies for specific wheel configurations are also considered.

INTRODUCTION

Many spacecraft employ more than three wheels for attitude control, both for redundancy and for the additional maneuvering capability of the extra wheels. The Swift Gamma Ray Burst Explorer, for example, will employ six reaction wheels for rapid slewing to enable observation of the initial stages of ephemeral gamma ray burst events[1]. In order to employ these wheels effectively, it is necessary to characterize the maximum torque available in any direction from the reaction wheels. Since long slews may be momentum-limited rather than torque-limited, it is equally necessary to characterize the maximum angular momentum storage capability in any direction. This knowledge is also important in designing momentum unloading strategies. Fortunately, the mathematical analyses required to identify the maximum torque and maximum angular momentum storage are identical. This analysis is the subject of this paper. Previous work on the optimal use of more than three reaction wheels has largely been limited to the four-wheel case[2-4].

In the case of equal wheel capabilities, the available torques fill the interior of an n-dimensional hypercube centered on the origin in reaction wheel torque space. If the torque capabilities vary from wheel to wheel, the torques fill the interior of an n-dimensional rectangular parallelopiped in this space, where each side has length equal to twice the magnitude of the maximum torque that a given wheel can provide. We will refer to this figure as a hypercube for brevity, at the expense of precision, since the case of equal wheel capabilities is the most common in practice.

The available torques in three dimensional space fill a polyhedron given by the mapping of this hypercube by the $3\times n$ reaction wheel torque distribution matrix, whose columns are the unit vectors parallel to the reaction wheel axes in three dimensional space. We will only consider "non-defective" reaction wheel configurations, by which we mean that no three of the reaction wheel spin axis directions are coplanar. As shown below, this polyhedron is the convex hull of the mapping of the vertices of the n-dimensional hypercube, which are the points at which all n wheels are saturated in the positive or negative direction. The exterior vertices of the three-dimensional envelope are connected by edges that are the boundaries of the two dimensional facets that form the envelope.

We have developed algorithms for defining the envelope and for finding the maximum torque in a given direction. The latter algorithm examines all the facets of the envelope to determine which facet limits the maximum torque, and hence which n–2 wheels have saturated torque commands. It is then straightforward to find the remaining two wheel torque commands necessary to produce the desired net torque. While the exact solution algorithm is not very efficient for onboard processing, it works for arbitrary wheel configurations. When the wheel configuration is known, it is possible to develop much more efficient algorithms.

GEOMETRY OF THE ENVELOPE

General Analysis

For brevity of exposition, we will only refer to torques in the following; the analysis for angular momentum is identical. The torques of an n-wheel system can be represented by an n-dimensional vector

$$\mathbf{T}_{\text{wheels}} = [T_1 \; T_2 \cdots T_n] \tag{1}$$

These torques fill the interior of an n-dimensional hypercube, where the length of the side along the ith axis is two times the maximum torque magnitude $T_{\max i}$ of the wheel along that axis. This can be the absolute maximum that the wheel can provide, or it can be some desired maximum for a given attitude maneuver. The vertices of the hypercube are points where all wheels are saturated, i.e., they supply either their positive or negative maximum torques. Each edge has all but one wheel saturated and is parallel to the axis of the unsaturated wheel. The facets of the hypercube have all but two wheels saturated.

The torque in three dimensional space is

$$\mathbf{T} = W\mathbf{T}_{\text{wheels}} \tag{2}$$

where

$$W \equiv [\hat{\mathbf{w}}_1 \; \hat{\mathbf{w}}_2 \cdots \hat{\mathbf{w}}_n] \tag{3}$$

is the torque distribution matrix. The available torques in three dimensional space fill a polyhedron specified by the mapping of the n-dimensional hypercube by Eq. (2). The vertices, edges, and facets of the polyhedron in three dimensional space are mappings of the vertices, edges, and facets of the hypercube. However, this mapping is not 1:1, since some of the vertices, edges, and facets of the hypercube will map to the interior of the polyhedron in three-dimensional space. Thus the vertices of the available polyhedron in three dimensional space are points where all wheels are saturated. The edges of the polyhedron have all but one wheel saturated, and an edge with wheel i unsaturated is parallel to the wheel's spin axis $\hat{\mathbf{w}}_i$. The facets have all but two wheels saturated, and the facet with wheels i and j unsaturated is a parallelogram with sides parallel to $\hat{\mathbf{w}}_i$ and $\hat{\mathbf{w}}_j$ and with normal vector in the direction of the cross product $\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j$.

We can identify the bounding planes of the polyhedron, and thus determine the envelope, by considering all pairs i, j of wheels in turn. There are $n(n-1)/2$ such pairs for an n-wheel configuration. The i, j pair defines a set of 2^{n-2} facets of the hypercube in the n dimensional reaction wheel command space. Each facet corresponds to a saturated command for the $n-2$ wheels other than wheels i and j, and with the torque commands for wheels i and j varying over their full range over the facet. Under the transformation into three dimensional space specified by Eq. (2), these facets map onto 2^{n-2} parallel planes with normal vector

$$\hat{\mathbf{n}}_{ij} \equiv (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j) \Big/ \left|\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right|. \tag{4}$$

On each of these planes, the wheel torque in the direction of the normal to the plane is

$$\mathbf{T} \cdot \hat{\mathbf{n}}_{ij} = \sum_{k=1}^{n} T_k [\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)] \Big/ \left|\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right| = \sum_{k \neq i,j} \sigma_k T_{\max k} [\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)] \Big/ \left|\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right|, \tag{5}$$

where $\sigma_k = \pm 1$ specifies the direction of saturation of the torque on wheel k. Wheels i and j can clearly be omitted from the sum, since they don't contribute to the torque perpendicular to the plane. This is indicated algebraically by the vanishing of the triple product $\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)$ for $k = i$ or j. This triple product does not vanish for any other k, since we have assumed a "non-defective" reaction wheel configuration, for which no three of the reaction wheel spin axis directions are coplanar.

The planes with the maximum and minimum torque in the direction of the normal are the planes for which

$$\sigma_k = \text{sign}[\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)] \tag{6a}$$

or

$$\sigma_k = -\text{sign}[\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)] \tag{6b}$$

for all k other than i and j. It is clear that these two planes and only these of the 2^{n-2} parallel planes for the i, j pair form part of the torque envelope. On each of these bounding facets T_i and T_j vary over the range between $\pm T_{maxi}$ and $\pm T_{maxj}$, respectively. Saturation of T_i and T_j give the four corners of a parallelogram, so we can identify eight of the vertices of the envelope for each i, j pair, corresponding to the two signs each for saturated T_i and T_j, and the overall sign of the other saturated wheel torques given by Eq. (6a) or (6b).

Carrying out the above procedure over all pairs i, j will produce a total of $4n(n-1)$ vertices. These aren't all distinct, however, for each vertex will be identified once for each bounding plane of the envelope that intersects there. This is useful information about the "structure" of the envelope, where "structure" denotes the number of vertices of the envelope, and the number of planes and edges that intersect at each vertex, as well as the specific information about which wheels are saturated in which direction at each vertex. It is interesting that this structure depends only on the unit vectors through their triple products $\hat{\mathbf{w}}_k \cdot (\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j)$, and not on the maximum torque capabilities of the wheels. If this torque capability is changed, the shape of the envelope will change, but not the structure as defined here. It is obvious that if the envelope has an inversion symmetry; for each vertex of torque $\mathbf{T}$ there is a corresponding vertex of torque $-\mathbf{T}$ obtained by reversing all the signs σ_k.

Specific Cases

First consider the easy case of three wheels. The envelope has the structure of a cube, with three planes meeting at each vertex. Thus each vertex will be identified three times by the procedure described above, and the number of distinct vertices is $4 \cdot 3 \cdot (3-1)/3 = 8$, the correct number of vertices of a cube. Figure 1 shows the torque envelope for three orthogonal wheels with equal torque capabilities. For this case, the wheel torque distribution matrix W is the 3×3 identity matrix and the envelope is actually a cube. The unit of length in this and subsequent figures is the maximum torque capability of a single wheel. The momentum envelope, scaled to the maximum momentum storage capability of a single wheel, is identical. Since the figures represent both envelopes, they are labeled Envelope, rather than Torque Envelope or Momentum Envelope.

We next consider the often-employed case of four reaction wheels are oriented in a tetrahedral configuration around the x axis. The torque distribution matrix is

$$W = \begin{bmatrix} s & s & s & s \\ 0 & c & 0 & -c \\ c & 0 & -c & 0 \end{bmatrix}, \tag{7}$$

where $s \equiv \sin\eta$ and $c \equiv \cos\eta$, with η being an angle chosen based on convenience of fabrication or on spacecraft inertia properties. The torque or momentum envelope for this configuration, again assuming identical wheels, is shown in Figure 2.

The next case we consider has the wheels oriented with hexagonal symmetry around the x axis. This is a good approximation to the reaction wheel configuration on the Swift spacecraft, whose moment of inertia is approximately axially symmetric about the x axis. The torque distribution matrix is

$$W = \begin{bmatrix} s & s & s & s & s & s \\ 0 & \frac{\sqrt{3}}{2}c & \frac{\sqrt{3}}{2}c & 0 & -\frac{\sqrt{3}}{2}c & -\frac{\sqrt{3}}{2}c \\ c & \frac{1}{2}c & -\frac{1}{2}c & -c & -\frac{1}{2}c & \frac{1}{2}c \end{bmatrix}. \tag{8}$$

This six-wheel configuration has an envelope with 32 vertices as shown in Table 1. Two of the vertices are defined by the intersection of six facets each, 18 by four facets each, and 12 by three facets each. The index pairs labeling the facets are the indices of the pair of wheels that are not saturated on the facet. Figure 3 illustrates the torque or momentum envelope for six wheels with identical capabilities in this configuration. This envelope has a sixfold rotational symmetry about the z axis.

Finally, Figure 4 shows the torque or momentum envelope for the case where the torque distribution matrix is obtained by deleting one column from the matrix of Eq. (8). This illustrates the case of a six-wheel configuration with one failed wheel. This case will be discussed further below.

Table 1. Vertices of envelope for the six-wheel configuration

Facets at this vertex	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
12, 23, 34, 45, 56, 16	+	+	+	+	+	+
15, 16, 56	+	+	+	+	+	–
45, 46, 56	+	+	+	+	–	+
34, 35, 45	+	+	+	–	+	+
23, 24, 34	+	+	–	+	+	+
12, 13, 23	+	–	+	+	+	+
12, 16, 26	–	+	+	+	+	
14, 15, 46, 56	+	+	+	+	–	–
35, 36, 45, 46	+	+	+	–	–	+
24, 25, 34, 35	+	+	–	–	+	+
13, 14, 23, 24	+	–	–	+	+	+
12, 13, 26, 36	–	–	+	+	+	+
15, 16, 25, 26	–	+	+	+	+	–
13, 14, 36, 46	+	+	+	–	–	–
25, 26, 35, 36	+	+	–	–	–	+
14, 15, 24, 25	+	–	–	–	+	+
13, 14, 36, 46	–	–	–	+	+	+
25, 26, 35, 36	–	–	+	+	+	–
14, 15, 24, 25	–	+	+	+	–	–
14, 15, 46, 56	–	–	–	–	+	+
35, 36, 45, 46	–	–	–	+	+	–
24, 25, 34, 35	–	–	+	+	–	–
13, 14, 23, 24	–	+	+	–	–	–
12, 13, 26, 36	+	+	–	–	–	–
15, 16, 25, 26	+	–	–	–	–	+
15, 16, 56	–	–	–	–	–	+
45, 46, 56	–	–	–	–	+	–
34, 35, 45	–	–	–	+	–	–
23, 24, 34	–	–	+	–	–	–
12, 13, 23	–	+	–	–	–	–
12, 16, 26	+	–	–	–	–	–
12, 23, 34, 45, 56, 16	–	–	–	–	–	–

OPTIMIZING THE TORQUE COMMANDS

Optimality Criterion

Optimal use of the torque capability of a wheel configuration requires that equal demands be made on all wheels. In the case of wheels with identical capacities, this is accomplished by minimizing the maximum value, or L_∞ norm, of the set of individual wheel torques which will produce the desired net torque. This minimax value T_{max}, which must be determined, determines the torque envelope, as in the previous section, and the required torque commands; it clearly cannot exceed the wheel capability of any wheel. The generalization to wheel assemblies with varying capabilities is straightforward, but we will not consider this uncommon case.

Numerical Algorithm For Minimizing The L_∞ Norm of a Torque Command

Finding the L_∞ norm requires identifying which facet of the wheel envelope contains a point along the desired torque axis. We can search for this facet by considering all pairs i, j of wheels in turn until we find the correct facet. This procedure is quite general but is generally too slow for application onboard a spacecraft. The Swift spacecraft employs an algorithm tailored to the six-wheel configuration using the azimuth and elevation of the desired torque vector to determine the optimal facet via a rapidly executed series of inequalities. The following discussion applies to both strategies.

Define a scaled torque n-vector $\boldsymbol{\psi}$ as

$$\boldsymbol{\psi} \equiv T_{\max}^{-1} \mathbf{T}_{\text{wheels}}, \tag{9}$$

where $T_{\max}$ denotes the L_∞ norm of the wheel torque vector, which is to be determined. It is clear from this scaling that the magnitude of the components of $\boldsymbol{\psi}$ is bounded by unity. Let the desired torque axis be represented by the unit vector $\hat{\mathbf{v}}$. Suppose, by previous knowledge or as a hypothesis, that the facet in the direction of $\hat{\mathbf{v}}$ has all wheels except those with indices i and j saturated. Then Eq. (6a) or (6b) immediately determines the magnitude and polarity of the other $n-2$ wheel torques. We tentatively assign a 1 or a -1 to each of these entries of $\boldsymbol{\psi}$ according to Eq. (6a). If this is the correct assignment, there is some positive constant γ such that

$$\gamma\, \hat{\mathbf{v}} = T_{\max}^{-1} \mathbf{T} = T_{\max}^{-1}\, W\, \mathbf{T}_{\text{wheels}} = W\, \boldsymbol{\psi} = \sum_{k=1}^{n} \psi_k \hat{\mathbf{w}}_k \tag{10}$$

which may be rearranged to

$$\begin{bmatrix} -\hat{\mathbf{w}}_i & -\hat{\mathbf{w}}_j & \hat{\mathbf{v}} \end{bmatrix} \begin{bmatrix} \psi_i \\ \psi_j \\ \gamma \end{bmatrix} = \sum_{k \neq i,j} \psi_k \hat{\mathbf{w}}_k \equiv \mathbf{z} \tag{11}$$

If the matrix on the left side has full rank, the solution of this equation is

$$\begin{bmatrix} \psi_i \\ \psi_j \\ \gamma \end{bmatrix} = \frac{1}{\left(\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right) \cdot \hat{\mathbf{v}}} \begin{bmatrix} \left(\hat{\mathbf{w}}_j \times \hat{\mathbf{v}}\right) \cdot \mathbf{z} \\ \left(\hat{\mathbf{v}} \times \hat{\mathbf{w}}_i\right) \cdot \mathbf{z} \\ \left(\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right) \cdot \mathbf{z} \end{bmatrix} \tag{12}$$

A negative value of γ means that the signs of Eq. (6b) should have been used rather than those of Eq. (6a), which is easily accomplished without repeating the computations by simply negating both γ and $\boldsymbol{\psi}$. The L_∞ norm value $T_{\max}$ for a given desired torque is given by Eq. (10) as

$$T_{\max} = T_{\text{desired}} / \gamma, \tag{13}$$

and the individual wheel torques are given by Eq. (9). Clearly, the maximum torque that can be applied in the direction $\hat{\mathbf{v}}$ is γ times the maximum torque per wheel.

If we had identified the appropriate facet for the desired torque direction *a priori*, as in the Swift onboard algorithm, the computation is complete. If we are searching over wheel pairs for the correct facet, it is possible either ψ_i or ψ_j has magnitude greater than unity or that the determinant $\left(\hat{\mathbf{w}}_i \times \hat{\mathbf{w}}_j\right) \cdot \hat{\mathbf{v}}$ is zero. This means that the facet corresponding to the pair i, j is not the facet that determines the L_∞ norm, so we must proceed to consider the next pair. We continue this search until we find a pair that gives an acceptable solution to Eq. (12).

A MATLAB algorithm implementing this procedure for determining $\boldsymbol{\psi}$ is contained in the box below. This algorithm may be utilized in a loop to determine the wheel envelopes shown in Figures 1 – 4 for configurations of 3, 4, 6, and 5 wheels. As illustrated by Figure 4, a single wheel failure in the six wheel case destroys the symmetry of the figure and significantly reduces the envelope. The maximum torque available in the direction of minimum capability drops by factor of 1.5 from the full six-wheel configuration.

```
function [psi,gamma] = maxwhltrq(v,w)

% Function for determining L-infinity optimal wheel distribution
% v = desired torque direction
% w = 3 X N wheel matrix
% psi = torque distribution vector
% gamma = maximum torque parameter

small_number = 1e-8;
N = size(w,2);

%%%%% Cycle through until requirements satisfied
for i = 1:(N-1)
   for j = (i+1):N
      X = cross(w(:,i),w(:,j));
      psi = sign(w'*X);
      psi([i j]) = [0 0]';
      z = w*psi;
      d = v'*X;
      if d ~= 0
         psi([i j]) = (cross(v,z)'*[-w(:,j) w(:,i)])/d;
         gamma = (z'*X)/d;
         if gamma < 0
            gamma = -gamma;
            psi = -psi;
         end
         if all(abs(psi([i j])) < 1+small_number)
            return
         end
      end
   end
end
```

CONCLUSIONS

We have presented algorithms for determining the envelope of available torques and angular momenta in three dimensional space for general configurations of n reaction wheels. These fill a polyhedron given by projection into three dimensional space of an n-dimensional hypercube in reaction wheel space. This polyhedron is the convex hull of the projection of the vertices of the hypercube, the points at which all the wheels are saturated in the positive or negative direction. We have also presented an algorithm for finding the optimal torque commands to produce any desired torque. The optimality criterion is that these commands minimize the maximum value, or L_∞ norm, of the vector of individual wheel torques. The algorithm also determines the maximum torque available in any given direction. While the exact solution algorithm is not very efficient for onboard processing, it is applicable to arbitrary wheel configurations.

REFERENCES

1. http://swift.sonoma.edu/
2. Kawaguchi, J., Maeda, K., Matsuo, H., and Ninomiya, K., "Closed Loop Momentum Transfer Maneuvers Using Multiwheels," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 4, July-Aug. 1995, pp. 867-874.
3. Steyn, W. H., "Near-Minimum-Time Eigenaxis Rotation Maneuvers Using Reaction Wheels," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 5, Sept-Oct. 1995, pp. 1184-1189.
4. Alonso, Roberto, and Crassidis, John, "Reaction Wheel Configuration Analysis," submitted to the *Journal of Guidance, Control, and Dynamics*

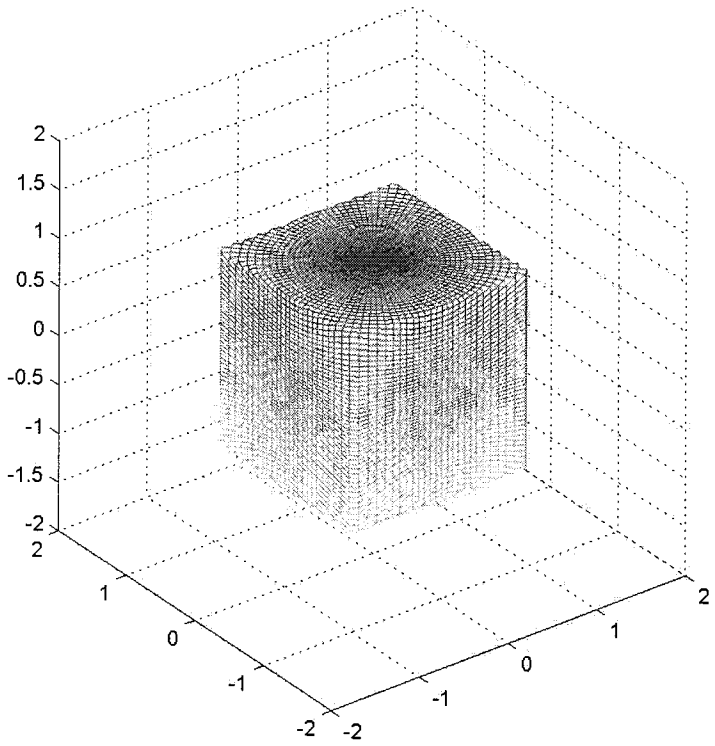


Figure 1. Three Orthogonal Wheel Configuration Envelope

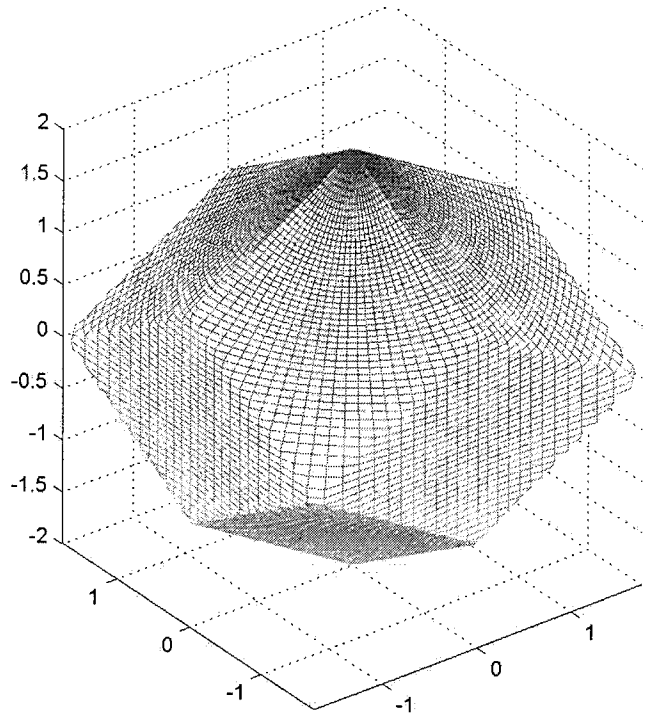


Figure 2. Four Symmetric Wheel Configuration Envelope

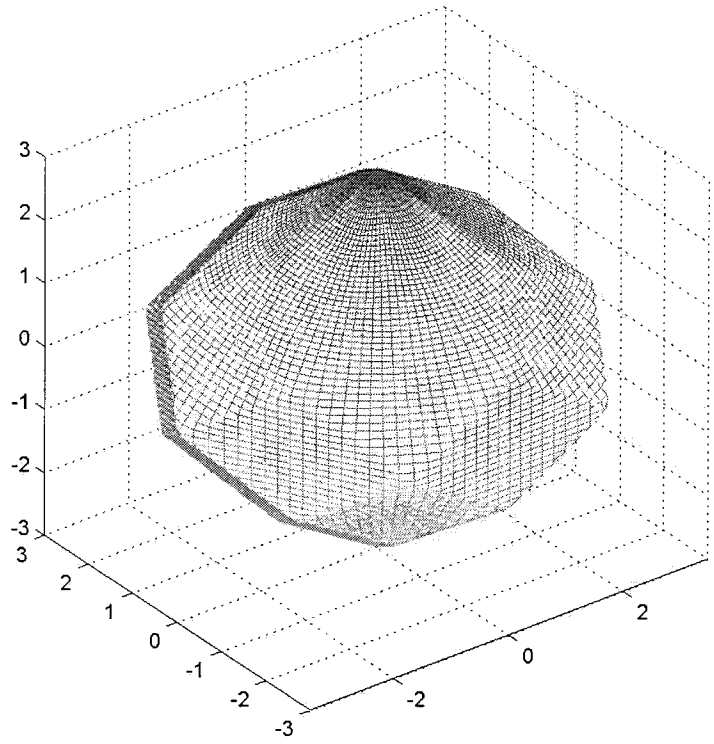


Figure 3. Six Symmetric Wheel Configuration Envelope

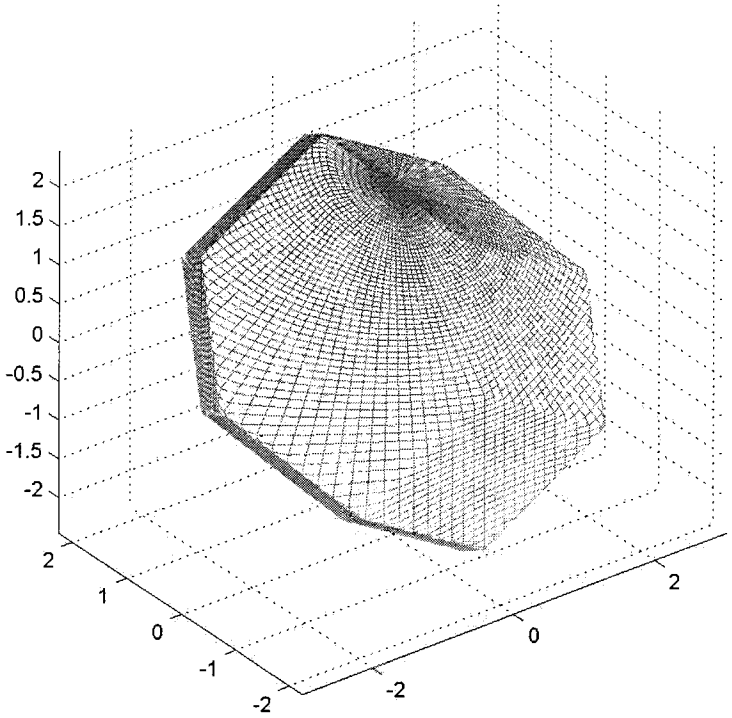


Figure 4. Envelope of Six Wheel Configuration with One Wheel Failed

A FULLY COUPLED MULTI-RIGID-BODY FUEL SLOSH DYNAMICS MODEL APPLIED TO THE TRIANA STACK [1]

K.W. London
Swales Aerospace, 5050 Powder Mill Rd., Beltsville, MD 20705

ABSTRACT

A somewhat general multibody model is presented that accounts for energy dissipation associated with fuel slosh and which unifies some of the existing more specialized representations. This model is used to predict the nutation growth time constant for the Triana Spacecraft, or Stack, consisting of the Triana Observatory mated with the Gyroscopic Upper Stage or GUS (includes the solid rocket motor, SRM, booster) . At the nominal spin rate of 60 rpm and with 145 kg of hydrazine propellant on board, a time constant of 116 s is predicted for worst case sloshing of a spherical slug model compared to 1,681 s (nominal), 1,043 s (worst case) for sloshing of a three degree of freedom pendulum model.

INTRODUCTION

It is a common practice to resort to simple pendulum-like models as a means of capturing the effect of internal fluid motion on the attitude behaviour of spacecraft. Although approximate, these models do provide physical insight and do permit time history simulation of interactions that include energy dissipation. A common approach is to represent the fluid, often the fuel or the propellant, as a spherical, rigid, dissipative slug centered at the vehicle center of mass [1,2] . In Reference 3, the analysis is extended to accommodate filled ellipsoidal tanks.

A number of the slosh models are variations of the simple pendulum. Reference 4, for example, allows a single rigid body spacecraft rotation and a single pendulum rotation, that is two degrees of freedom (DOF). A parallel development is followed in Reference 5 but, in this instance, only a part of the propellant mass is assigned to the pendulum with the balance remaining fixed in the vehicle. In addition, a torsional stiffness and viscous rate-dependent torque are added at the pivot to allow for a tank with an elastomeric diaphram propellant management device (PMD). Another model like that of Reference 4 is that given in Reference 6, except now translational motions of the vehicle are included. Reference 7 has effectively 3 DOF, one a rigid body spacecraft rotation and two lateral pendulum rotations. While no equations are given, Reference 8 appears to account for coupled vehicle translation, rotation and 2 DOF for each of four pendulums. There is no fixed component of propellant mass and no stiffness at the pivot, but a damping ratio of 0.01 is imposed. A symbolic dynamics model builder is employed in Reference 9 to capture 3-axis attitude behaviour coupled with what appears to be a 1 DOF pendulum for each of four tanks. A portion of the propellant is fixed and stiffness and viscous effects are present at the pivot. Reference 10 alludes to use of a different symbolic code generator to provide the coupled vehicle, pendulum dynamics. What is different for this case is that pendulum parameter estimates are derived using an independent Navier-Stokes fluid model (in this case for an ellipsoidal tank).

---

[1] Presented @ *2001 Flight Mechanics Symposium*, NASA Goddard Space Flight Center, Greenbelt, MD June 19-21, 2001. (This work was carried out under the NASA GSFC Triana contract 21[illegible]3-012.)

Presented in this paper is a general momentum-based multibodied dynamics model applicable to a wide range of slosh model configurations including all of those discussed above. The model can be applied directly, for simulation purposes, or it can be used to derive more simplified and,or linearized models. This is demonstrated by presenting equations, in velocity format, that result from applying the general equations to the case of a single spherical slug slosh model. Both the vehicle and the slosh body orientation kinematics are solved for using Euler parameters.

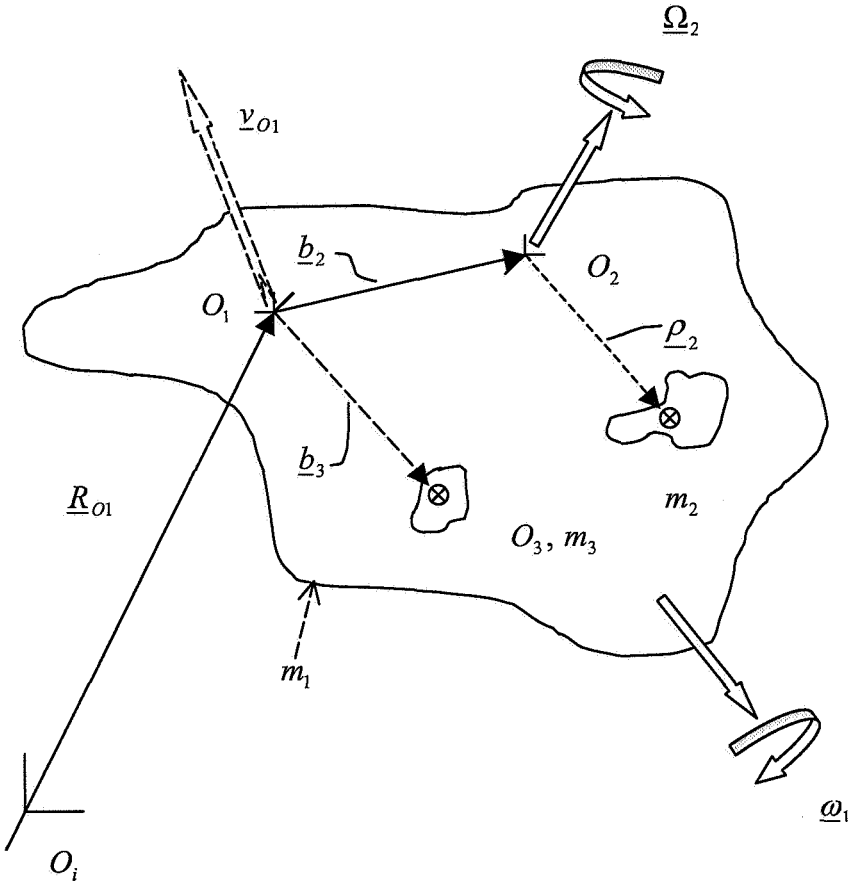


Figure 1 Position, velocity vectors of generic 3-body configuration. Reference point O_1, fixed in core body of mass m_1 and located by position vector $\underline{R}_{O1}$ relative to inertial reference at O_i, translates with absolute velocity $\underline{v}_{O1}$. Inertial angular velocity of core body is $\underline{\omega}_1$. Slosh mass m_2 has angular velocity $\underline{\Omega}_2$ relative to the core. The reference fixed to m_2 at O_2 is positioned with respect to O_1 by $\underline{b}_2$. Offset of the center of mass of body 2 from O_2 is $\underline{\rho}_2$ which, for a pendulum, becomes length L_2. Body 3 of mass m_3 represents any fixed, nonsloshing, portion of propellant mass. Offset vector of the center of mass of m_3 (O_3) from O_1 is $\underline{b}_3$.

These models are used to predict nutation time constant for the Triana spacecraft, an important design parameter for the Nutation Control System in use during orbit reboost. Effectiveness of the model is enhanced by adopting pendulum parameters extrapolated from test data already in existence for spherical tanks with a diaphram type PMD. Time constants produced by pendulum models are compared with those found for a maximum energy dissipating spherical slug slosh model. Sensitivity to parameters such as spin rate, stiffness and viscosity at the pivot are also examined.

SYSTEM DYNAMICS, KINEMATICS

Dynamic equations for both the general model and for the spherical slug model follow.

Momentum Rate Equations for the General 3-Body Model

The Newton, Euler formulation,as developed in Reference 11 for articulated interconnected rigid bodies, is used here to generate governing motion equations for a rigid 3-body model, one body representing the vehicle or spacecraft and two bodies for the propellant (one an articulated slosh mass and one fixed to the vehicle). Only a minor extension is required to the two body problem, contained in reference 11, to accommodate a third nonarticulating mass. Consequently, details of the derivation are not presented. Figure 1 depicts the generic configuration used here in order to have a model applicable to a wide range of configurations. The fully coupled motions consist of 3 DOF translation and 3 DOF rotation of rigid core body 1 with respect to inertial together with the 3 DOF rotation of body 2 relative to the core body. The momentum rate equations for the system and for body 2 about its attach point take on the following relatively straightforward appearance:

$$\dot{\underline{p}} = \underline{f}^{ext}; \qquad \text{(1a) system translation}$$

$$\dot{\underline{h}} + (\underline{v}_{O1} \otimes \underline{p}) = \underline{\Gamma}^{ext}; \qquad \text{(1b) system rotation}$$

$$\dot{\underline{h}}_2 + [\underline{v}_{O1} + (\underline{\omega}_1 \otimes \underline{b}_2)] \otimes \underline{p}_2 = \underline{\Gamma}_2^{ext} + \underline{\Gamma}_{21}; \qquad \text{(1c) slosh mass rotation}$$

where,

$$\underline{p} = m\underline{v}_{O1} - (\underline{c} \otimes \underline{\omega}_1) - (\underline{c}_2 \otimes \underline{\Omega}_2); \qquad \text{(1d) linear momentum of system}$$

$$\underline{h} = (\underline{c} \otimes \underline{v}_{O1}) + \underline{\underline{J}} \circ \underline{\omega}_1 + \underline{\underline{J}}_{12} \circ \underline{\Omega}_2; \qquad \text{(1e) angular momentum of system}$$

$$\underline{h}_2 = (\underline{c}_2 \otimes \underline{v}_{O1}) + \underline{\underline{J}}_{21} \circ \underline{\omega}_1 + \underline{\underline{J}}_2 \circ \underline{\Omega}_2; \qquad \text{(1f) angular momentum of slosh mass}$$

with,

$\underline{f}^{ext}$ external force on all bodies

$\underline{\Gamma}^{ext}$ external torques about body 1 reference point, $\underline{\Gamma}_2$ external torques about body 2 attach point

$\underline{\Gamma}_{21}$ torque exerted on body 2 by body 1 about O_2 (e.g. $-k_{pend}\,\theta_{pend}{}^{i} - c_{pend}\,\dot{\theta}_{pend}{}^{i}$)

c_{pend}, k_{pend} viscous coefficient and and torsional stiffness at body 2 attach point

$\theta_{pend}{}^{i}$ angular displacement about the ith pendulum axis; i = 1, 2, or 3

and, mass properties,

$m = m_1 + m_2 + m_3$; system mass

$\underline{c}_j = \int_{mj} \underline{r}_j \, dm$; first mass moment of body j about its reference point

$\underline{r}_j$; location of elemental mass in body j from its reference point

$\underline{c} = \underline{c}_1 + \underline{c}_2 + \underline{c}_3$; first mass moment of system

$\underline{\underline{J}}$; system moment of inertia about body 1 reference point

$\underline{\underline{J}}_2$; moment of inertia of slosh mass about its local reference point

$\underline{\underline{J}}_{12} = \underline{\underline{J}}_2 + \underline{c}_2 \circ \underline{b}_2 \underline{\underline{1}} - \underline{c}_2 \underline{b}_2$; coupling moment of inertia matrix (symmetric in matrix form).

Also,

$(\underline{\ })$ refers to a vector; $(\underline{\underline{\ }})$ refers to a dyadic; $\underline{\underline{1}}$ is the unit dyadic ; $\otimes$ stands for vector cross product;

$\circ$ is a dot product; and dot( ) represents time derivative with respect to inertial.

With these equations as a base, and by adopting more restrictive assumptions, one can generate simpler models, as may be desirable in controls design. For example, if translation is not of interest then $\underline{p} = m\underline{v}_{O1} = \underline{0}$. Reference points are made to coincide with the core body reference if $\underline{b}_j = \underline{0}$ and setting $\underline{c}_j = \underline{0}$ places body reference point at its center of mass. Note, while the differing DOF are ultimately coupled, the first order momentum rates above are uncoupled and amenable to numerical integration as is. Of course, this is done while simultaneously solving the algebraic velocity, momentum relations together with the kinematics discussed later.

Velocity Rate Equations for 2-Body Model with Spherical Slug

A common slosh model appearing in the literature is that in which the entire mass of fluid is replaced by a single rigid spherical slug with mass uniformly distributed throughout the tank. The slug exerts a viscous torque on the rigid body vehicle proportional to its angular velocity relative to the vehicle. The general model equations are applied to this case using a core body reference attached to system center of mass. Substituting momentum relations from equations (1d), (1e), (1f) into equations (1a), (1b), (1c) gives, after a certain amount of vector algebra, the equations in velocity form:

$$m\dot{\underline{v}}_{CM} = \underline{f}^{ext}; \qquad \text{(2a)} \quad \text{system translation}$$

$$(\underline{\underline{I}} - \underline{\underline{I}}_{SS}) \circ \dot{\underline{\omega}}_1 = \underline{\Gamma}^{ext} - \underline{\omega}_1 \otimes \underline{\underline{I}} \circ \underline{\omega}_1 + c_{SS}\underline{\Omega}_{SS}; \qquad \text{(2b)} \quad \text{system rotation}$$

$$\underline{\underline{I}}_{SS} \circ \dot{\underline{\Omega}}_{SS} = -\underline{\underline{I}}_{SS} \circ \dot{\underline{\omega}}_1 - \underline{\omega}_1 \otimes \underline{\underline{I}}_{SS} \circ \underline{\Omega}_{SS} - c_{SS}\underline{\Omega}_{SS}; \qquad \text{(2c)} \quad \text{slug rotation}$$

with,

$\underline{v}_{CM}$ velocity at mass center;

$\underline{\underline{I}}$ system moment of inertia about system center of mass;

$\underline{\underline{I}}_{SS}$ moment of inertia of spherical slug about its center of mass (same about all axes, all axes principal)

m_{SS} mass of spherical slug;

$\underline{b}_{SS}$ position vector to mass center of spherical slug from system center of mass;

c_{SS} constant viscous coupling coefficient giving rise to interbody torques, $\pm c_{SS}\underline{\Omega}_{SS}$, in equations (2b), (2c);

$\underline{\Omega}_{SS}$ angular velocity of spherical slug with respect to core body;

$\dot{\underline{\Omega}}_{SS}$ time rate of change in angular velocity of spherical slug with respect to core body;

and, $\underline{\underline{I}} = \underline{\underline{J}}_1 + \underline{\underline{I}}_{SS} + m_{SS}(\underline{b}_{SS}{}^2 - \underline{b}_{SS}\underline{b}_{SS})$; (3)

with,

$\underline{\underline{J}}_1$ moment of inertia of core about system center of mass.

As written here, the translation is relative to inertial, whereas the system rotation is with respect to axes fixed in the core body. Since the system center of mass was chosen as main reference and the reference for the slug is at its center of mass, the system rotational motion is uncoupled from translation. The rotation equations, however, remain coupled and are in agreement with those presented elsewhere [1,2]. They are valid for non-principal axes, as well, in which case off-diagonal elements appear in the system moment of inertia matrix.

By way of observation, note that the system rotation equation (2b) can be put in an even simpler form if only principal axes are used:

$$\underline{\underline{I}}_{EQ} \circ \dot{\underline{\omega}}_1 = \underline{\Gamma}^{ext} - \underline{\omega}_1 \otimes \underline{\underline{I}}_{EQ} \circ \underline{\omega}_1 + c_{SS}\,\underline{\Omega}_{SS}; \quad (4a)$$

where,

$\underline{\underline{I}}_{EQ} = \underline{\underline{I}} - \underline{\underline{I}}_{SS}$. (an 'equivalent' system moment of inertia) (4b)

Rotation Kinematics

The dynamics equations presented depend implicitly on body attitude. Here Euler parameters are employed to track attiude for both vehicle and slosh body. The Euler parameter rate equations, which depend on angular velocity, are integrated in step with the dynamics. This allows one to continuously update the transformations between differing coordinate frames. More information is needed in the case of the pendulum model which feeds back position and rate dependent torques. In this case, Euler angles and Euler angle rates are also calculated for the pendulum mass. For such an application one might also consider using a space-fixed rotation sequence, particularly if large amplitude oscillations are anticipated that might affect the joint torque parameters.

Software Implementation

Independent computer software solutions, are developed in a MATLAB environment, for both the general model and for the spherical slug model. Numerical integration of the equations described above provides time histories of response. Of particular interest is angle of nutation and its rate of growth as characterized by some time constant. Since, for cases considered here, nutation is expected to closely approximate exponential growth, the time period chosen is time required for amplitude to increase by a factor e = 2.7183. A linear curve fit to the logarithm of the nutation response is used to extract this constant.

APPLICATION TO THE TRIANA SPACECRAFT

Background

The Triana spacecraft shown schematically in Figure 2 is intended to be launched from the Shuttle Orbiter payload bay by the Italian Research Interim Stage (IRIS) cradle-spin-table-launcher assembly. The Spacecraft Axis System (SAS) of coordinates, fixed in the spacecraft, has an origin at the intersection of the IRIS/spacecraft separation plane (SEP) and the centerline of the STAR-48 SRM. The Observatory Axis System (OAS) has its origin at the intersection of the GUS/ Observatory separation plane and the centerline of the STAR-48 SRM. Triana is to be ejected at a nominal spin of 60 rpm about the 'z' axis as indicated in Figure 2. Nominally, on ejection, mass totals 2989 kg with principal moments of inertia about the center of mass of 1843.5195 $kg.m^2$ about the transverse ('x', 'y') axes and 575.4170 $kg.m^2$ about the spin direction.

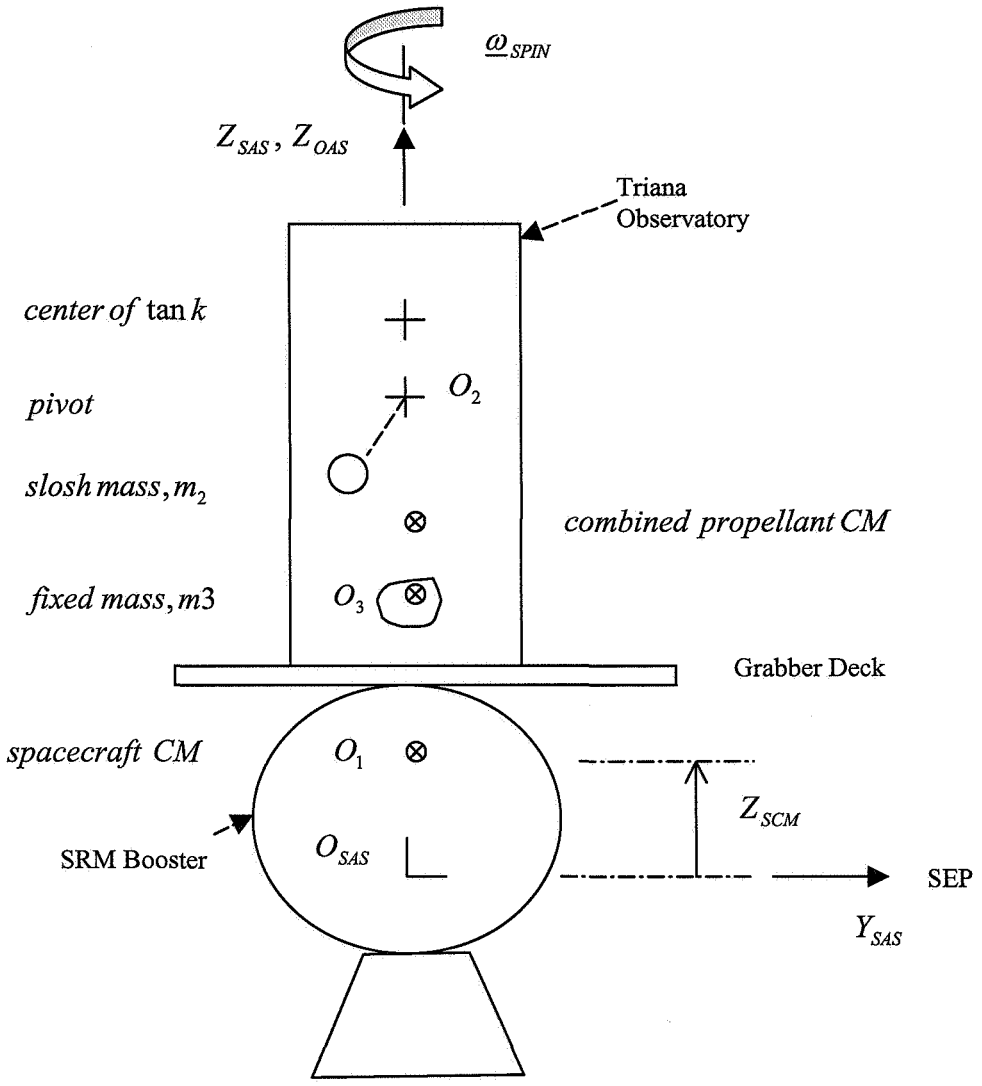


Figure 2 Schematic of Triana spacecraft showing key elements of the 2-body pendulum slosh model, coordinate system alignment and axis of spin.

Triana Observatory Propellant and Tank

Hydrazine thrusters provide active attitude control of the TRIANA Observatory. The liquid hydrazine is housed in a single 28 inch diameter pressurized spherical tank centered on the 'z' axis parallel to the direction of spin. What is referred to as an elastomeric diaphram (membrane, or bladder) Propellant Management Device (PMD), with its plane normal to 'z', is added to assist with expulsion of the propellant. This membrane is welded in place at the mid-section of the tank thus separating it into two distinct compartments; one containing the hydrazine and one containing a pressurized gas. Additional details are given in Reference 5.

A Pendulum Slosh Model for Triana

A complete transient dynamic analysis is complex and requires a rigid body rotational model for the tank coupled with the partial differential equations and associated boundary conditions of the fluid. Hence, it is common to resort to equivalent mechanical analogs, such as the pendulum, in an attempt to capture first order fluid force, torque interactions. For a spherical tank with diaphram, lateral slosh associated with surface waves is assumed to be the predominant influence. There is some measured data available for tanks of this type (Reference 12). Existing test data has been extrapolated to a number of different spacecraft as, for example, in References 9, 13 and 14 and, more specifically here, for the Triana Observatory in Reference 5. It is the work of the later reference that is made use of here. The same characteristics at the pivot are used since they depend on thickness of the diaphram (0.0625 inches) and not on fill level of the tank and they include a torsional stiffness of 54.74 Nm/Rad and a viscous rate coefficient of 4.569 Nm/(RAD/s). A number of other parameters depend on fill level, which is determined here, for known propellant mass of 145 kg contained in a spherical segment, to be 0.689 (as a fraction of diameter). Least squares fit relationships, based on earlier test data, are given in Reference 5 for pendulum mass, pendulum length (L2) and distance of the pivot from the center of the tank (h1), as well as a linear two-point fit for an equivalent rigid body inertia (Io) of the fixed mass component about its mass center. They are evaluated here for the above fill level giving:

L_2 = 0.1126 m; m_2 = 49.44 kg; m_3 = 95.56 kg; Io = 1.0281 $kg.m^2$

h_1 = 0.06854 m (below center of tank);

h_0 = 0.29847 m; (leaves original center of mass location for propellant unchanged at 0.25846 m from tank center).

Since the pendulum model here has 3 DOF, the rigid body Io about the 'z' axis is divided between pendulum and fixed mass in proportion to their mass (the fixed mass still has principal transverse moments of inertia of Io). Note, on implementation into the general model the core body reference was positioned at the system center of mass nominally 0.55613 m from SEP. Also, the center of the tank is 1.47617 m from SEP, thus completing the information needed to specify the pendulum model as used here.

Maximum Dissipation Spherical Slug Model

Analysis has been done to estimate maximum energy dissipation for a spherical slug slosh model with the slug at the system center of mass of a spinning axisymmetric rigid body[2]. It is possible to analytically solve for angular rates of the slug in terms of the steady rates of the nutating rigid body. Since applied torque depends on those rates and on viscous coefficient, the energy dissipation can be expressed as a function of the viscous torque coefficient. The value of this coefficient which renders the dissipation rate stationary is, as established by Flatley,

$$c_{SS}^{\ F} = (I_{SS}\, I_{SPIN}\, \omega_{SPIN}) / I_T ;$$

where,

ω_{SPIN} is spin rate;

I_{SPIN}, I_T refers to system moments of inertia about spin and transverse axes, respectively.

Also, equating this dissipation rate to that contained in the approximate Energy Sink model gives a minimum time constant:

$$\tau_{MIN}^{\ F} = 2I_T^{\ 2} / [I_{SS}\, \omega_{SPIN}\, (I_T - I_{SPIN})]$$

For the Triana tank $I_{SS} = (2/5) m_{SS}\, r^2 = 7.3342\ kg.m^2$.

---

[2] Flatley, T., Notes to Tobin re-'An Energy Dissipation Model', NASA GSFC 11/15/93 as attached to NASA GSFC (Code 712) document from: Houghton, Martin B. to: Ward, D. entitled: "FUSE Nutation Time Constant," December 15, 1993.

Table 1 shows viscous cofficient and minimum time constant, calculated using nominal Triana mass properties, for low, nominal and high end Triana spacecraft spin rates. These coefficients represent the maximum dissipation theoretically achievable for such a model.

Table 1
Viscous coefficient and related nutation time constant for maximum spherical slug dissipation

ω_{SPIN}, rpm	c_{SS}^{F}, NMs	τ_{MIN}^{F}, s
52	12.4658	134.21
60	14.3836	116.32
68	16.3014	102.63

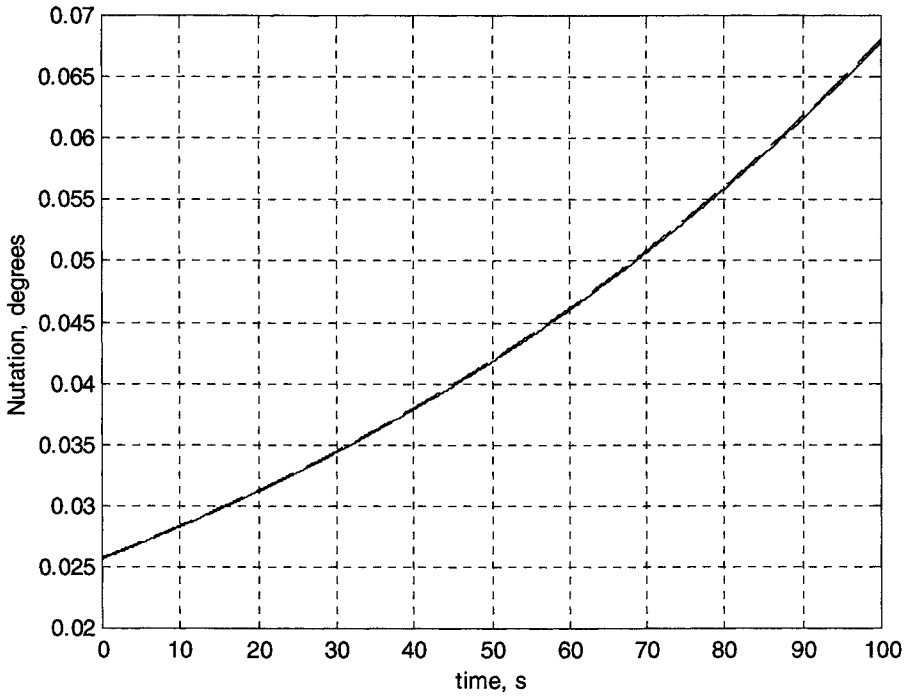


Figure 3 Spherical slug slosh: general model versus specialized 2-body model at 68 rpm.

Results for Spherical Slug Slosh Model

Results presented here are based on nominal parameter values unless otherwise noted.

Figure 3 compares nutation growth found using a 2-body spherical slug model, similar to that currently in use in the literature, with that predicted using the general model. Agreement is seen to be reasonable.

The general model is applied next to the FUSE spacecraft reported on by Houghton in the memo of footnote 2, where again, the agreement appears reasonable when using a spherical slug slosh model (Table 2).

Table 2
Comparison of Nutation Time Constants for the FUSE Spacecraft using a Spherical Slug Slosh Model (60 rpm)

	Spin Axis Moment of Inertia, kgm^2	Transverse Axis Moment of Inertia, kgm^2	Time Constant, s from Houghton	Time Constant, s from General Model
prior to booster firing	885	4602	2126	2164
post burnout	546	2096	1058	1065
post separation	500	1450	826	830

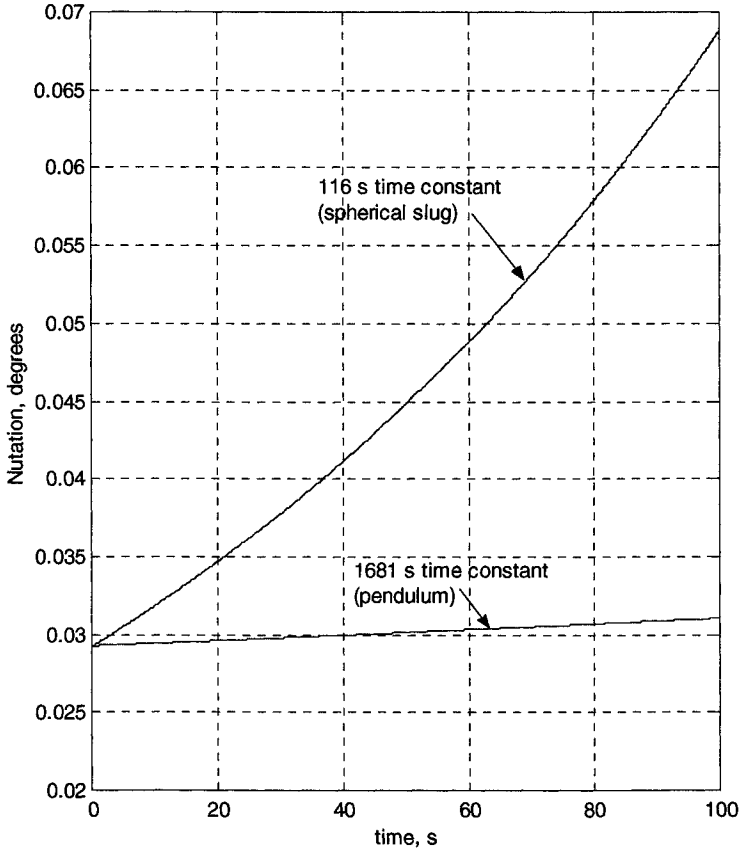


Figure 4 Nutation growth predicted using the general model with the tuned Triana 3 DOF pendulum versus the 2-body spherical slug with maximum dissipation at 60 rpm.

Results for Pendulum Slosh Model

Figure 4 compares nutation growth for the worst case spherical slug slosh model for the propellant with that predicted using a pendulum slosh model in the general program with parameters tuned to the Triana spacecraft. The growth rate is significantly higher for the spherical slug case as reflected in the time constant which is more than an order of magnitude below that for the pendulum model.

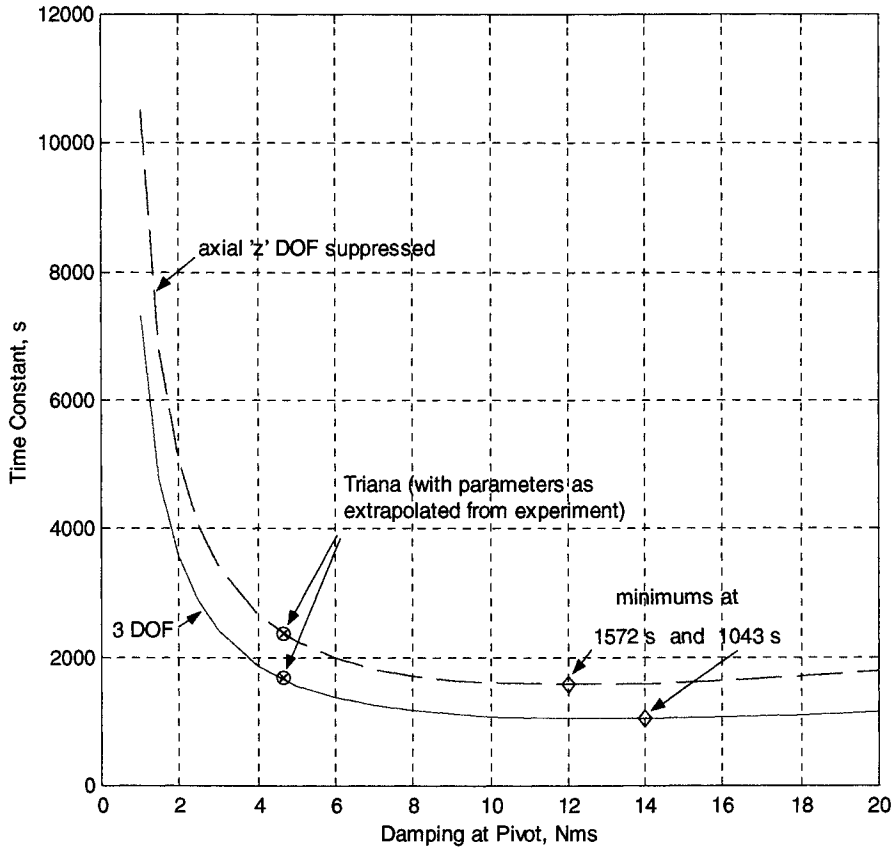


Figure 5 Sensitivity to viscous damping coefficient of pendulum slosh model for Triana spacecraft (60 rpm).

The effect of viscous damping at the pendulum pivot is illustrated in Figure 5. The nominal extrapolated value of 4.569 Nms yields a 1681s time constant. This curve allows one to gauge the effect changes in viscosity will have and it also gives a minimum, or worst case, of 1043 s. Suppressing the pendulum rotation (swirl) about its axis resulted in somewhat less dissipation. This effect was confirmed by cases run with close to zero moment of inertia about this axis for which the time constant appeared to converge to these same levels. Note, stiffness about this axis is nominally set to zero.

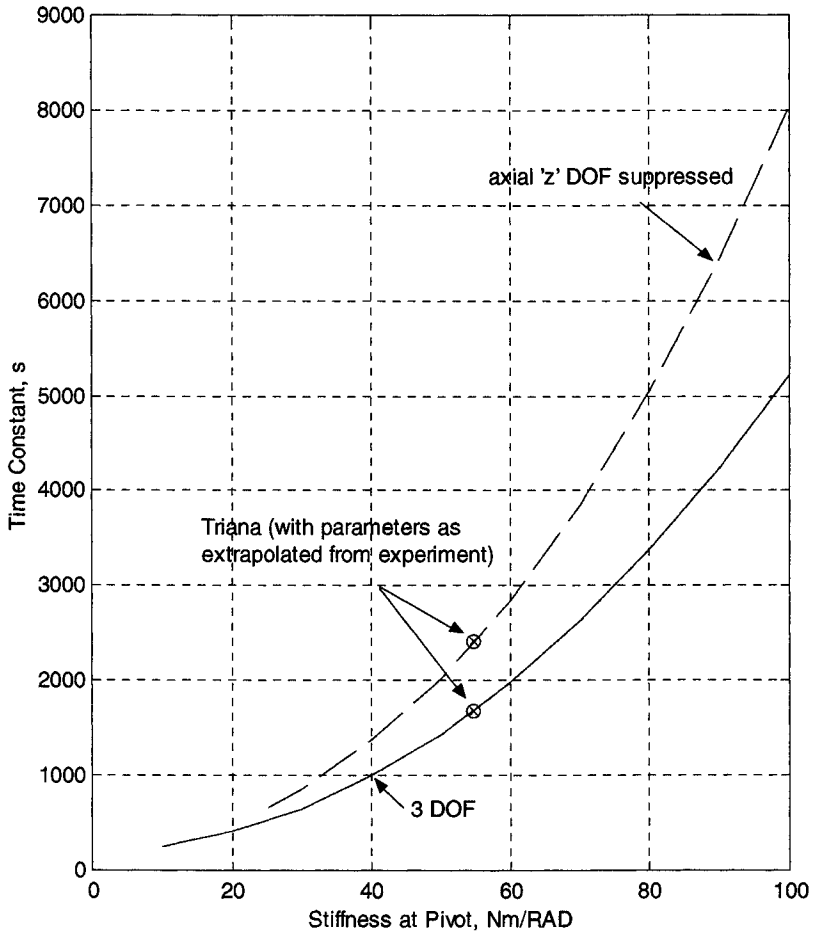


Figure 6 Sensitivity to stiffness at pendulum pivot for Triana spacecraft (60 rpm).

For the Triana pendulum the pivot and mass are nominally aligned along the axis of spin. Figure 6 points to increased dissipation with reduced stiffness. The 3 DOF case gives a more conservative time constant estimate compared to when the 'z' DOF is suppressed.

CONCLUDING COMMENT

A general 3-body slosh model is presented and validated for the case in which the propellant is represented solely as a spherical slug. The model can readily be extended to include any number of additional tanks as well as additional slosh modes if desired. It is applied to the Triana spacecraft using a pendulum slosh model as well. Sensitivity to some of the important parameters is demonstrated. The estimates should prove useful in gauging a worst case energy dissipation scenario. An independent 2-body maximum dissipation spherical slug model is used as well to provide a worst case nutation time constant prediction.

REFERENCES

1. Rahn, C.D. and Barba, P.M., "Reorientation Maneuver for Spinning Spacecraft," Journal of Guidance, Control, and Dynamics," Vol.14, No.4, 1991, pp. 724-728.

2. Livneh, Rafael and Wie, Bong, "Asymmetric Body Spinning Motion with Energy Dissipation and Constant Body-Fixed Torques," Journal of Guidance, Control, and Dynamics," Vol.22, No.2, 1999, pp. 322-328.

3. Kuang, Jinlu and Leung, A.Y.T., "H_∞ Feedback for Attitude Control of Liquid-Filled Spacecraft," Journal of Guidance, Control, and Dynamics," Vol.24, No.1, 2001, pp. 46-53.

4. Bryson, Jr., Arthur E., "Control of Spacecraft and Aircraft," Princeton University Press, Princeton, New Jersey, 1994.

5. Morgenstern, Wendy, "Control of the Triana Spacecraft in the Presence of Fuel Slosh Dynamics," Scholarly Paper, Department of Aerospace Engineering, University of Maryland, College Park, MD, December 1, 1999.

6. Cho, Sangbum and McClamroch, N. Harris, "Feedback Control of a Space Vehicle with Unactuated Fuel Slosh Dynamics," Paper No. A00-37036, AIAA Guidance, Navigation, and Control Conference and Exhibit, Denver, CO, 14-17 August 2000.

7. Or, A.C., "Rotor-Pendulum Model for the Perigee Assist Module Nutation Anomaly," Journal of Guidance, Control, and Dynamics," Vol.15, No.2, 1992, pp. 297-303.

8. Hill, Daniel E., Baumgarten, Joseph R. and Miller, John T., "Dynamic Simulation of Spin-Stabilized Spacecraft with Sloshing Fluid Stores," Journal of Guidance, Control, and Dynamics," Vol.11 No.6, 1988, pp. 597-599.

9. Sirlin, S.W., "Mars Pathfinder Launch Vehicle Nutation Analysis," IOM 3456-96-032, Jet Propulsion Lab, Pasadena, CA, July 19, 1996.

10. Saeed, Salma I., "Fuel Slosh and Active Nutation Control Interaction Analysis using Ellipsoidal Fuel Pendulums," Paper No. AAS 99-116, AAS/AIAA Space Flight Mechanics Meeting, Breckenridge, CO, Feb. 7-10, 1999.

11. Hughes, Peter C., "Spacecraft Attitude Dynamics," John Wiley & Sons, 1986.

12. Stofan A. and Sumner, I., "Experimental Investigation of the Slosh-Damping Effectiveness of Positive-Expulsion Bags and Diaphrams in Spherical Tanks," NASA Technical Note D-1712, June 1963.

13. Kana, Daniel D. and Dodge, Franklin T., "Study of Liquid Slosh in the Tracking and Data Relay Satellite Hydrazine Tanks," Final Report, SwRI Project 02-6539, September 25, 1981.

14. Dodge, Franklin T., "Propellant Dynamics and PMD Design for the Near Earth Asteroid Rendezvous (NEAR) Spacecraft," SwRI Project 04-6297, Southwest Research Institute, April 1994.

UNIFORM SAMPLING OF SO_3

Robert Bauer
Langhorne, PA 19047, www.BauerEngineering.com

Abstract – Uniformly distributed samples of SO_3 are needed to make an objective assessment of spacecraft attitude control performance over all attitudes. Random sampling algorithms are reviewed and stratified sampling algorithms are derived. Random sampling distributes samples in a probabilistic sense: the probability that a sample lies in any given region of SO_3 is directly proportional to the size of the region. Uniformly stratified sampling distributes samples in a geometric sense: SO_3 is partitioned into equal-size and nearly cubic cells and then a sample is placed at the center of each cell. Measure and probability theory on SO_3 is reviewed and used to define uniformity, verify the uniformity of the random sampling, and develop the stratified sampling algorithms. Statistical tests are used to validate the sampling algorithms as implemented.

INTRODUCTION

Suppose we want to demonstrate on a sample basis how well an attitude control system can achieve a desired attitude from any initial attitude. To make an objective assessment we need a *uniformly* distributed set of sample rotations to be used for the initial attitude. This paper considers two types of sampling: *random* and *stratified*. Uniformly distributed random (UDR) rotations are uniform in a probabilistic sense. Uniformly distributed stratified (UDS) rotations are uniform in a geometric sense. Figure 1 illustrates the difference for points on the unit circle (S_1). Random sampling suffers from clumping; i.e., there are regions with more or less samples than the average. Stratified sampling requires far fewer samples to guarantee the same minimum number of samples in any given region. Generating uniformly distributed samples over the rotation group, SO_3, is far more complicated than doing so on the unit circle. The paper gives the background in measure and probability theory needed to define uniformity on SO_3 and to develop the random and stratified sampling algorithms. The mathematics for stratified sampling is more involved but result in much more uniformly distributed samples.

a) 40 UDR Samples b) 40 UDS Samples

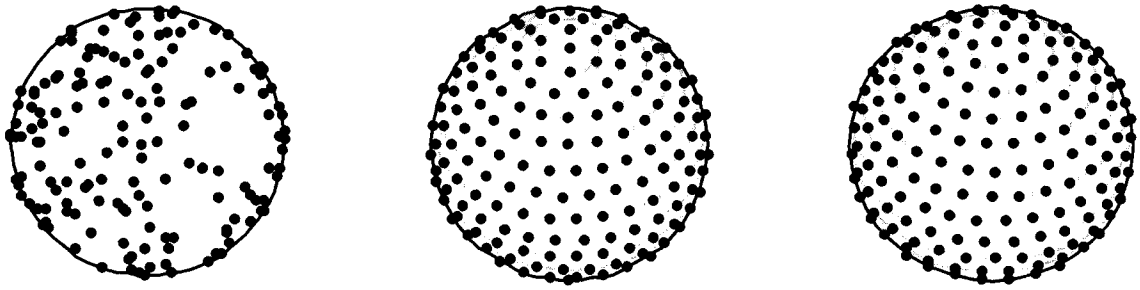
Figure 1 Uniformly distributed random (UDR) and uniformly distributed stratified (UDS) samples on S_1.

UDR orthogonal matrices have been used in a number of applications. Statisticians often inspect high-dimensional data by looking at low-dimensional projections. Asimov (Ref. 1) uses UDR orthogonal matrices to move a low-dimensional subspace around in an n-dimensional data space. Stewart (Ref. 2) uses UDR orthogonal matrices in an empirical study of two methods to estimate the condition number of a matrix. Computer graphics designers (Ref. 3) use UDR rotations to produce tumbling images in screen savers. We will review two salient algorithms to produce UDR members of O_n, the group of orthogonal matrices of order n, as they apply to the specific problem of generating a UDR member of SO_3, the group of special orthogonal matrices of order 3. We will also show that UDR unitary quaternions correspond to UDR rotations and summarize two algorithms to generate UDR unitary quaternions.

Our primary purpose is to develop two new algorithms that generate uniformly distributed stratified rotations. The general approach is to partition SO_3 into a large number of contiguous, equal-content and nearly cubic cells. Then one sample is taken from the middle of each cell. Figure 2 illustrates the algorithms as applied to S_2, the unit two-sphere imbedded in three-dimensional Euclidean space. The first algorithm, called *masonry sampling*, is more complicated, but is also more computationally efficient. First, SO_3 is cut into uniformly thick layers, such that the content of each layer is an integer multiple of the desired cell content. Then each layer is sliced into uniformly wide strips such that the content of each strip is again an integer multiple of the desired cell content. Finally the strips are diced into equal-content and nearly cubic cells. The second algorithm, called *igloo sampling*, borrows a technique used by the Inuit People to construct a dwelling from snow blocks (Ref. 4). Equal-content and nearly cubic cells are arranged in a spiral that completely and compactly fills SO_3.

To perform the partitioning we need parameterized measures of distance along a curve, angle between intersecting curves, and "volume" of regions in SO_3. Riemannian geometry (Ref. 5) provides these measures. A brief review of the requisite measure theory is given.

Finally, we use several well-known properties of uniformly distributed orthogonal matrices to show statistically that the sampling algorithms produce uniformly distributed samples.



a) Random sampling b) Masonry sampling c) Igloo sampling

Figure 2 Uniform sampling of S_2 (300 samples each)

RANDOM SAMPLING OF S_2

Before addressing uniform sampling of SO_3, it is instructive to look at uniform sampling of S_2, the two-dimensional sphere imbedded in three-dimensional Euclidean space. We define uniformity on S_2 using the *invariance principle*. Reasoning on physical grounds, if the direction of a random, unit length, physical vector is uniformly distributed, then the distribution of its coordinates should be invariant under a coordinate transformation. In other words, $\boldsymbol{v} \in S_2$ is uniformly distributed if and only if $\boldsymbol{v}$ and $\boldsymbol{v}' = \boldsymbol{B}\,\boldsymbol{v}$ have the same distribution for any fixed rotation $\boldsymbol{B}$. An equivalent definition is to say that a random point on S_2 is uniformly distributed if the probability that it lies in a region N_S is directly proportional to the area of N_S.

Projected Gaussian Ball Sampling on S_2

Projected Gaussian ball sampling (Ref. 6) is a very simple method to generate *uniformly distributed random* (UDR) samples of S_2. Choose u_1, u_2, and u_3 randomly and independently from a Gaussian distribution with zero mean and unity variance. With $\boldsymbol{u} = [\,u_1\ u_2\ u_3\,]^T$ form

$$\boldsymbol{v} = \boldsymbol{u} / |\boldsymbol{u}| \tag{1}$$

It is easy to show that $\boldsymbol{v}$ and $\boldsymbol{v}' = \boldsymbol{B}\,\boldsymbol{v}$ have the same distribution. Let $\boldsymbol{u}' = \boldsymbol{B}\,\boldsymbol{u}$. Then $\boldsymbol{v}' = \boldsymbol{u}'/|\boldsymbol{u}|$. Since $\boldsymbol{u}$ is comprised of independent, identically distributed Gaussian deviates and $\boldsymbol{B}$ is an orthogonal matrix, $\boldsymbol{u}'$ has the same distribution as $\boldsymbol{u}$. Thus, $\boldsymbol{v}'$ has the same distribution as $\boldsymbol{v}$.

Homolographic Random Sampling on S_2

Another simple method to generate UDR samples of S_2 uses the following *homolographic* mapping (Ref. 7).

$$\boldsymbol{v} = \begin{bmatrix} \cos u_1 \sqrt{1-u_2^2} \\ \sin u_1 \sqrt{1-u_2^2} \\ u_2 \end{bmatrix} \tag{2}$$

Choosing u_1 and u_2 randomly and independently from uniform distributions on the intervals $[0, 2\pi]$ and $[-1, 1]$ respectively results in a uniformly distributed $\boldsymbol{v}$, which we can show as follows. Let N_U be any region in the parameter space U and N_S be the corresponding region on S_2. The probability $P(N_U)$ that $\boldsymbol{u} = (u_1, u_2)$ lies in N_U equals the probability $P(N_S)$ that $\boldsymbol{v}(u_1, u_2)$ lies in N_S since the two events have a one-to-one correspondence. Next, and most importantly, the mapping specified by Eqn. (2) is *homolographic*, i.e. the area of N_S is directly proportional to the area of N_U. The general equation for a differential area element da_S on S_2 is (Ref. 5)

$$da_S = [\det(\boldsymbol{g})]^{1/2}\, du_1 du_2 \tag{3}$$

where $\boldsymbol{g}$ is the Riemannian metric tensor,

$$\boldsymbol{g} = (d\boldsymbol{v}/d\boldsymbol{u})^T\,(d\boldsymbol{v}/d\boldsymbol{u}) \tag{4}$$

and $\boldsymbol{u} = [u_1\ u_2]^T$. Applying Eqn. (3) to the mapping of Eqn. (2) we have

$$da_S = du_1 du_2 \tag{5}$$

Integrating, we see that the area of N_S equals the area of N_U. Finally, $P(N_U)$ is directly proportional to the area of N_U since $\boldsymbol{u}$ is uniformly distributed on U. Thus, $P(N_S)$ is directly proportional to the area of N_S and $\boldsymbol{v}$ is uniformly distributed on S_2 by definition.

STRATIFIED SAMPLING OF S_2

Lunar Sampling of S_2

Next we consider several algorithms to generate uniformly distributed *stratified* samples on S_2. *Lunar sampling* slices S_2 into lunes along longitudinal lines. The lunes are diced into equal volume cells and a sample is taken from the center of each cell. To achieve the slicing and dicing, the parameter space is partitioned into congruent and rectangular cells. Then, using a homolographic mapping, a sample corresponding to the center of each cell is generated. The sampling is uniform in the sense that one sample is taken per cell and the cells are equal-area. However, the spacing between neighboring samples is not at all uniform.

The general equation for a differential arc length on $\mathbf{S_2}$ is (Ref. 5)

$$(ds_S)^2 = \boldsymbol{du}^T \boldsymbol{g}\, \boldsymbol{du} \tag{6}$$

For the homolographic mapping of Eqn. (2), ds_S is given by

$$(ds_S)^2 = (du_1)^2 (1 - u_2^2) + (du_2)^2 / (1 - u_2^2) \tag{7}$$

As u_2 approaches ±1, the distance on $\mathbf{S_2}$ between samples becomes smaller in the u_1-direction and larger in the u_2-direction. This shortcoming is overcome with the following algorithm.

Masonry Sampling of $\mathbf{S_2}$

Masonry sampling generates uniformly distributed samples by arranging nearly square cells on $\mathbf{S_2}$, like bricks in a dome. First $\mathbf{S_2}$ is sliced along latitudinal lines into strips. Next, each strip is diced along longitudinal lines into cells. The width of each strip and number of cells in each strip are chosen so that the cells are equal-area and nearly square on $\mathbf{S_2}$. Finally samples corresponding to the middles of each cell are formed (Figure 2b). Since the cells are all nearly squares on $\mathbf{S_2}$ and since they are equal-area, the sampling is uniform and the spacing between neighboring samples is nearly the same in both the latitudinal and longitudinal directions.

The details of the algorithm are as follows. Let n be the number of samples to be generated. The desired area of each cell is

$$a_C = 4\pi/n \tag{8}$$

where 4π is the total area of $\mathbf{S_2}$. If the cells are to be nearly square on $\mathbf{S_2}$, then the length of the sides of the cells should be nearly equal to

$$\lambda^* = \sqrt{4\pi / n} \tag{9}$$

The following mapping from $\mathbf{U}$ to $\mathbf{S_2}$ is used:

$$\boldsymbol{v} = \begin{bmatrix} \cos u_2 \sin u_1 \\ \sin u_2 \sin u_1 \\ \cos u_1 \end{bmatrix} \tag{10}$$

The differential arc length and area are given by

$$(ds_S)^2 = (du_1)^2 + (du_2)^2 \sin^2 u_1 \tag{11}$$

$$da_S = \sin u_1\, du_1 du_2 \tag{12}$$

$\mathbf{S_2}$ is cut into m_1 strips in the u_1-direction such that the average width of the strips on $\mathbf{S_2}$ is nearly equal to λ^*. The total distance on $\mathbf{S_2}$ from $u_1 = 0$ to $u_1 = \pi$ is π. Therefore, m_1 must be set to approximately π/λ^*. The actual formula is

$$m_1 = 2\ \mathrm{floor}[\pi/(2\lambda^*)] + 1 \tag{13}$$

This formula always yields an odd number for m_1, which insures a convenience symmetry that is exploited later to reduce the number of computations required.

Next we determine the width of each strip. Let k_1 be the strip index, where $k_1 = 1$ indexes the topmost strip ($u_1 \approx 0$) and $k_1 = m_1$ the bottommost strip ($u_1 \approx \pi$). The width of the strips must satisfy two conditions:

a) The number of cells in the k_1-th strip must be an integer;
b) The width of the k_1-th strip should be very nearly equal the average width.

Let $N_1(k_1)$ be the total number of cells in strips 1 through k_1, and $A_1(k_1)$ be their summed area. According to Condition a) and Eqn. (8), A_1 is

$$A_1 = 4\pi N_1 / n \tag{14}$$

Now let $b_1(k_1)$ be the upper value of u_1 for the k_1-th strip, allowing us to write A_1 as

$$A_1 = 2\pi(1 - \cos b_1) \tag{15}$$

Solving for N_1 from the last two equations we have

$$N_1 = (n/2)(1 - \cos b_1) \tag{16}$$

To satisfy Condition b) we should have

$$b_1(k_1) \approx k_1\, \pi/m_1 \tag{17}$$

Substituting this last expression into Eqn. (16) and rounding to the nearest integer yields

$$N_1(k_1) = \mathrm{round}\{\, (n/2)[1 - \cos(k_1\, \pi/m_1)]\, \} \tag{18}$$

Next we find the u_1-coordinate for the samples in each strip. We set $u_1(k_1)$ to the value that dissects the strip into two equal-area parts:

$$u_1(k_1) = \arccos\{\, 1 - [N_1(k_1) + N_1(k_1{-}1)]/n\, \} \tag{19}$$

where $N_1(0) = 0$.

Next, each strip is diced into cells and the u_2-coordinate of the sample in each cell is found. The number of cells in the k_1-th strip is

$$m_2(k_1) = N_1(k_1) - N_1(k_1{-}1) \tag{20}$$

The total extent of u_2 is 2π. Let $u_2(k_1, k_2)$ be the u_2-coordinate of the sample in the k_2-th cell of the k_1-th strip. Then $u_2(k_1, k_2)$ is given by:

$$u_2(k_1, k_2) = [k_2 - \mu(k_1)][2\pi/m_2(k_1)] \tag{21}$$

where $\mu(k_1)$ is a phasing parameter used to improve the spacing of points between adjacent strips. μ is set to 0 if the strip index is even and 1/2 if odd.

Finally, $\boldsymbol{u}(k_1, k_2) = [u_1(k_1), u_2(k_1, k_2)]^T$ is mapped to a point $\boldsymbol{v}(k_1, k_2)$ on $\mathbf{S_2}$ via Eqn. (10).

There is a useful symmetry in the set of samples. It is easily shown that

$$v_{1,2}(m_1 - k_1 + 1, k_2) = v_{1,2}(k_1, k_2) \tag{22}$$

$$v_3(m_1 - k_1 + 1, k_2) = -v_3(k_1, k_2) \tag{23}$$

Thus we only need to find $v(k_1, k_2)$ for $k_1 = 1$ to ceiling$(m_1/2)$ and the respective values of k_2, and then apply Eqns. (22) and (23) to find v for the remaining samples.

The masonry sampling on $\mathbf{S_2}$ is summarized as follows. Find the number of strips, m_1, from Eqn. (13). For $k_1 = 1$ to $(m_1+1)/2$ do the following. Find the total number of cells in strips 1 through k_1 from Eqn. (18). Find the u_1-coordinates of the samples in the k_1-th strip from Eqn. (19). Find the number of cells $m_2(k_1)$ in the k_1-th strip from Eqn. (20). For $k_2 = 1$ to $m_2(k_1)$ find the u_2-coordinates of the samples from Eqn. (21). Find $v(k_1, k_2)$ from Eqn. (10). For $k_1 = (m_1+3)/2$ to m_1 and $k_2 = 1$ to $m_2(k_1)$ find $v(k_1, k_2)$ from Eqns. (22) and (23).

Igloo Sampling of $\mathbf{S_2}$

The Inuit People build dwellings from snow blocks by arranging the blocks in a spherical spiral to form a dome (Ref. 4). We borrow this architectural technique to arrange points uniformly on $\mathbf{S_2}$ (Figure 2c). An *igloo spiral* on $\mathbf{S_2}$ is described in terms of spherical coordinates by

$$\theta = L(\phi - \pi/2) + \theta_0 \tag{24}$$

where L is the spiral slope, θ_0 is a phasing parameter, and the range of ϕ is $0 \le \phi \le \pi$. The salient feature of this spiral is that turns are separated by a constant distance. When θ increases by 2π (one turn) ϕ increases by $\lambda = 2\pi/L$.

To form the igloo sampling of $\mathbf{S_2}$ we begin by cutting $\mathbf{S_2}$ along the spiral

$$\theta = L(\phi - \pi/2) - \pi \tag{25}$$

to form an *igloo peel*. Figure 3 shows the peel near $\phi = 0$. The light line is the cut forming the peel (Eqn. (25)) and the heavy line is the centerline of the peel, given by

$$\theta = L(\phi - \pi/2) \tag{26}$$

The width of the peel measured along a curve of constant θ is $\lambda = 2\pi/L$, except at the ends where the width tapers to zero.

The next step in the igloo sampling is to find the area, $a_I(\theta)$, along the igloo peel as a function of θ along the centerline. Conceptually, the peel is cut by a curve segment having a constant value of θ. Then the area from the tip of the peel at $\phi = 0$ to the cut is found by first integrating $d\phi$ across the width of the peel and then integrating $d\theta$ along the length of the peel. A differential area element on S_2 is given by

$$da = \sin\phi \, d\phi \, d\theta \tag{27}$$

We must be careful to account for the tapers at the ends of the peel. Table 1 gives the values of θ at the boundaries of the tapers. For $B_\theta \le \theta \le C_\theta$, the main body of the peel, $a_I(\theta)$ is by

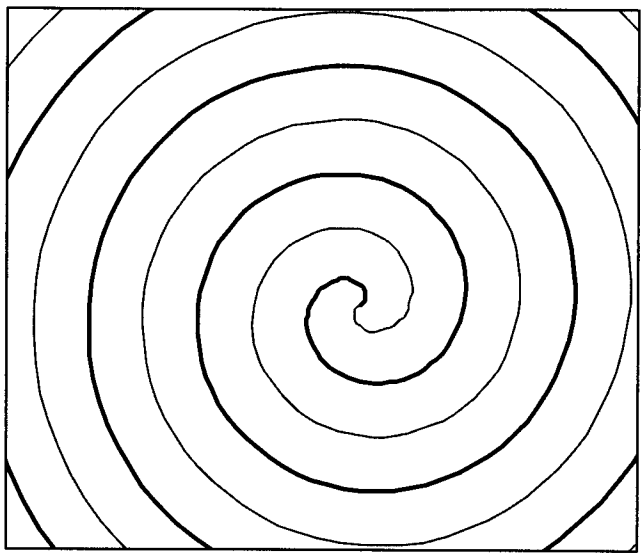


Figure 3 Portion of a peel on $\mathbf{S_2}$ formed by cutting along an igloo spiral.

Table 1 Taper Boundaries for Igloo Peel on $\mathbf{S_2}$

Description	Value
Tip of head taper at $\phi = 0$ end of peel	$A_\theta = -L\pi/2 - \pi$
Boundary between head taper and main body of peel	$B_\theta = -L\pi/2 + \pi$
Boundary between main body of peel and tail taper	$C_\theta = L\pi/2 - \pi$
Tip of tail taper at $\phi = \pi$ end of peel	$D_\theta = L\pi/2 + \pi$

$$a_I(\theta) = \int_{B_\theta}^{\theta} \int_{\theta'/L+\pi/2-\lambda/2}^{\theta'/L+\pi/2+\lambda/2} \sin\phi \, d\phi d\theta' + a_I(B_\theta) \tag{28}$$

$$= 2L \sin(\theta/L) \sin(\lambda/2) + L \sin(\lambda) + a_I(B_\theta)$$

The area of the head taper is

$$a_I(B_\theta) = \int_{\theta_A}^{\theta_B} d\theta \left\{ \int_0^{\theta/L+\pi/2+\lambda/2} \sin\phi \, d\phi \right\} \tag{29}$$

$$= 2\pi - L\sin(\lambda)$$

Substituting this expression into Eqn. (28) give us

$$a_I(\theta) = 2\pi[1 + \alpha \sin(\theta/L)] \tag{30}$$

where $$\alpha = \sin(\pi/L) / (\pi/L) \tag{31}$$

Next, the peel is diced into n equal-area cells (dashed lines in Figure 3) and then a sample is placed at the center of each cell. We index the cells by $k = 1$ to n, where $k = 1$ refers to the cell at the $\phi = 0$ end of the peel and $k = n$ refers to the cell at the $\phi = \pi$ end. The area of each cell is

$$a_C = 4\pi / n \tag{32}$$

where 4π is the total area of $\mathbf{S_2}$. We place a sample at the center of cell k by finding (θ_k, ϕ_k) according to

$$a_I(\theta_k) = (k-1/2)\, a_C = 4\pi\,(k-1/2)/n \tag{33}$$

$$\phi_k = \theta_k/L + \pi/2 \tag{34}$$

where ϕ_k is found by solving Eqn. (26) for ϕ. Substituting Eqn. (30) into Eqn. (33) yields

$$1 + \alpha \sin(\theta_k/L) = (2k-1)/n \tag{35}$$

Taking the cosine of Eqn. (34) yields

$$z_k = \cos(\phi_k) = -\sin(\theta_k/L) \tag{36}$$

Substituting this equation into Eqn. (35) yields an expression for z_k in terms of k:

$$z_k = (1/\alpha)\,[\,1-(2k-1)/n\,] \tag{37}$$

We solve Eqn. (36) for θ_k.

$$\theta_k = -L \sin^{-1} z_k \tag{38}$$

We need to determine the value of L that results in the cells being nearly square. Recall that the distance between turns of the spiral is $\lambda = 2\pi/L$. Now this distance is the cell height. If the cells are to be nearly square it should be set equal to the square root of the cell area:

$$\lambda = 2\pi/L = \sqrt{a_C} = \sqrt{(4\pi/n)} \tag{39}$$

Solving for L gives us the desired value:

$$L = \sqrt{(\pi n)} \tag{40}$$

The igloo sampling of $\mathbf{S_2}$ is summarized as follows. First, find L and α from Eqns. (40) and (31) respectively. Then for $k = 1$ to n find z_k and θ_k from Eqns. (37) and (38). Finally, substitute z_k and θ_k into to find the Cartesian coordinates for the k-th sample.

$$v_k = \begin{bmatrix} \cos\theta_k \sqrt{1-z_k^2} \\ \sin\theta_k \sqrt{1-z_k^2} \\ z_k \end{bmatrix} \tag{41}$$

MEASURE AND PROBABILITY ON $\mathbf{SO_3}$

We want to develop sampling algorithms on $\mathbf{SO_3}$ that are analogous to those developed for $\mathbf{S_2}$. To do so we need $\mathbf{SO_3}$ analogies to distance and area. In the present section we review relevant measure theory for continuous compact groups and specifically for $\mathbf{SO_3}$. We also review measure theory for the group of unitary quaternions, $\mathbf{H}$, since we will find the use of quaternions simplifies the arithmetic involved in generating uniformly distributed samples on $\mathbf{SO_3}$.

Haar Measure

Suppose $\mathbf{N}$ is a region (or more precisely, a *Baire set*) of a continuous group $\mathbf{G}$ (e.g. $\mathbf{SO_3}$ or $\mathbf{H}$) and $\mu(\mathbf{N})$ is a measure of the *content* or size of $\mathbf{N}$. *Content* is the general term in the sequence *length, area, volume, ...* In order for $\mu(\mathbf{N})$ to be a consistent measure in any reasonable sense, we require that it be positive when $\mathbf{N}$ is not empty and invariant when $\mathbf{N}$ is translated by an arbitrary group member $\boldsymbol{B}$. In other words we want

$$\mu(\mathbf{N}) > 0 \quad \text{for } \mathbf{N} \neq \varnothing \tag{42}$$

$$\mu(\boldsymbol{B}\mathbf{N}) = \mu(\mathbf{N}\boldsymbol{B}) = \mu(\mathbf{N}) \tag{43}$$

where $\boldsymbol{B}\mathbf{N}$ ($\mathbf{N}\boldsymbol{B}$) is the region generated by application of an arbitrary but fixed $\boldsymbol{B} \in \mathbf{G}$ to the left (right) of each member of $\mathbf{N}$. Such a measure is called a *Haar measure* (Refs. 8 and 9). Except for a scalar, the Haar measure is unique.

To perform integration of a function, $f(X)$, $X \in \mathbf{G}$, we need a differential form of the Haar measure, $d\mu(X)$. Hurwitz examined the construction of $d\mu$ at the end of the 19th century (Ref. 10). Assume that $\mathbf{G}$ is a three-parameter group, as are $\mathbf{SO_3}$ and $\mathbf{H}$. Let $X \in \mathbf{G}$ be parameterized by $\boldsymbol{u} = [u_1\ u_2\ u_3]^T$, i.e. $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{u})$. A differential content element $du_1du_2du_3$, in the parameter space $\mathbf{U}$, is related to $d\mu$ by

$$d\mu = \rho^G(\boldsymbol{u})\, du_1du_2du_3 \tag{44}$$

where ρ^G is the *group density*. Hurwitz derived two formulae for ρ^G: one based on a *Riemannian metric*, the other on the *group composition Jacobian*. The Riemannian metric formulation is the more powerful since it also gives us access to measures of distance and angle. (See Ref. 11 for a derivation of ρ^G using the group composition Jacobian.)

Riemannian Metrics for $\mathbf{SO_3}$ and H

A *Riemannian metric* (Ref. 5) allows one to measure distance and content in an m-dimensional curved subspace of an n-dimensional Euclidean space, where $m \le n$. In the case of $\mathbf{SO_3}$, a member $\boldsymbol{R}$ consists of nine elements having six constraints, namely $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I}^{3\times3}$. Hence, $\mathbf{SO_3}$ is a three-dimensional curved subspace of the nine-dimensional space of real 3×3 matrices*. Similarly, the unitary quaternion $\boldsymbol{h} \in \mathbf{H}$ consists of four elements having one constraint, $|\boldsymbol{h}| = 1$, making it a three-dimensional curved subspace of the four-dimensional space of quaternions. Let $\boldsymbol{R}(\boldsymbol{u})$ be a parameterization of $\mathbf{SO_3}$ where $\boldsymbol{u}$ is the three-element parameter vector. Assuming the elements of $\boldsymbol{R}(\boldsymbol{u})$ are differentiable, $d\boldsymbol{R}$ is defined as

$$\begin{aligned} d\boldsymbol{R} &= \boldsymbol{R}(\boldsymbol{u}+d\boldsymbol{u}) - \boldsymbol{R}(\boldsymbol{u}) \\ &= (\partial\boldsymbol{R}/\partial u_1)\, du_1 + (\partial\boldsymbol{R}/\partial u_2)\, du_2 + (\partial\boldsymbol{R}/\partial u_3)\, du_3 \end{aligned} \tag{45}$$

We want to assign a differential measure of length, ds, to $d\boldsymbol{R}$ to allow us to measure the length of a curve in $\mathbf{SO_3}$. We define ds by

* $\boldsymbol{R} \in \mathbf{SO_3}$ is also subject to the constraint that $\det(\boldsymbol{R}) = 1$. However, this constraint does not effect the dimensionality of $\mathbf{SO_3}$.

(46) $(ds)^2 = (1/2)\,\mathrm{trace}(d\boldsymbol{R}^{\mathsf{T}} d\boldsymbol{R}) = \sum_{\alpha,\beta=1,2,3} dR_{\alpha\beta}^2 /2$

It is easy verify that $(ds)^2 = |d\boldsymbol{\theta}|^2$, where $d\boldsymbol{\theta}$ is the differential rotation vector that carries $\boldsymbol{R}$ to $\boldsymbol{R} + d\boldsymbol{R}$. Thus the units for arc length are radians. Substituting Eqn. (45) into Eqn. (46) we have

(47) $$(ds)^2 = d\boldsymbol{u}^{\mathsf{T}} \boldsymbol{g}^R d\boldsymbol{u}$$

where $\boldsymbol{g}^R$ is the Riemannian metric tensor for $\mathbf{SO}_3$, whose elements are given by

(48) $$g^R_{\alpha\beta} = (1/2)\,\mathrm{trace}[(\partial \boldsymbol{R}/\partial u_\alpha)^{\mathsf{T}} (\partial \boldsymbol{R}/\partial u_\beta)]$$

The metric tensor can also be used to measure angle between two intersecting curves, $\boldsymbol{u}^a(t)$ and $\boldsymbol{u}^b(t)$ in $\mathbf{SO}_3$. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be the derivatives $d\boldsymbol{u}^a/dt$ and $d\boldsymbol{u}^b/dt$ respectively at the point of intersection. Then the cosine of the angle between the two curves in $\mathbf{SO}_3$ is (Ref. 5)

(49) $$\cos\psi = \frac{\boldsymbol{a}^{\mathsf{T}} \boldsymbol{g}^R \boldsymbol{b}}{\sqrt{(\boldsymbol{a}^{\mathsf{T}} \boldsymbol{g}^R \boldsymbol{a})(\boldsymbol{b}^{\mathsf{T}} \boldsymbol{g}^R \boldsymbol{b})}}$$

A parameterization is called *orthogonal* if at each point in $\mathbf{SO}_3$, the curves formed by varying each parameter one at a time are mutually orthogonal. Looking at Eqn. (49) we see that a parameterization is orthogonal if and only if the off-diagonal elements of $\boldsymbol{g}^R$ are zero.

The metric tensor can be used to measure the content of a region in $\mathbf{SO}_3$. A differential box in the parameter space, whose sides are du_1, du_2, and du_3, corresponds to a parallelepiped in $\mathbf{SO}_3$ having sides $d\boldsymbol{R}_1$, $d\boldsymbol{R}_2$, and $d\boldsymbol{R}_3$, where

(50) $$d\boldsymbol{R}_i = (\partial \boldsymbol{R}/\partial u_i)\, du_i,\ i = 1, 2, 3.$$

The differential content of this parallelepiped is (Ref. 5)

(51) $$d\mu^R = \sqrt{\gamma^R}\, du_1 du_2 du_3$$

(52) $$\gamma^R = \det(\boldsymbol{g}^R)$$

Comparing Eqn. (51) with Eqn. (44) we see that $\sqrt{\gamma^R}$ is the group density:

(53) $$\rho^R(\boldsymbol{u}) = \sqrt{\gamma^R(\boldsymbol{u})}$$

Using Eqn. (51), the measure of content of a region $\mathbf{N}_\mathbf{R}$ in $\mathbf{SO}_3$ is

(54) $$\mu^R(\mathbf{N}_\mathbf{R}) = \int_{\mathbf{N}_\mathbf{R}} d\mu^R = \int_{\mathbf{N}_\mathbf{U}} \sqrt{\gamma^R(\boldsymbol{u})}\, du_1 du_2 du_3$$

where $\mathbf{N}_\mathbf{U}$ is the region in the parameter space corresponding to $\mathbf{N}_\mathbf{R}$.

It is important to note that the metric tensor defined by Eqn. (48) is invariant under group translation, i.e. $\boldsymbol{g}^R(\boldsymbol{R}) = \boldsymbol{g}^R(\boldsymbol{BR}) = \boldsymbol{g}^R(\boldsymbol{RB})$, where $\boldsymbol{R}$ is a variable member of $\mathbf{SO}_3$ and $\boldsymbol{B}$ is a fixed but arbitrary member. This property leads immediately to the fact that the Riemannian form of the volume measure (Eqn. (51)) is a Haar measure.

Finding the metric tensor using Eqn. (48) is tedious. We must find three partial derivatives for each of the nine elements of $\boldsymbol{R}(\boldsymbol{u})$, and then form six different matrix dot products and simplify the resulting expression. Fortunately, there is a far easier way. We will show that the metric tensor for $\boldsymbol{R}(\boldsymbol{u})$ is directly proportional to the metric tensor for the corresponding unitary quaternion $\boldsymbol{h}(\boldsymbol{u})$. Finding the metric tensor for $\boldsymbol{h}(\boldsymbol{u})$ requires finding only twelve partials derivatives. Furthermore, the expressions for those partials are much simpler than in the case of the rotation matrix.

Let $\boldsymbol{h} = (w, x, y, z)$ be a member of the group of unitary quaternions, $\mathbf{H}$, with w being the real part and (x, y, z) the imaginary part of $\boldsymbol{h}$. Let $\boldsymbol{h}(\boldsymbol{u})$ be a parameterization of $\mathbf{H}$ where $\boldsymbol{u}$ is the three-element parameter vector. Assuming the elements of $\boldsymbol{h}(\boldsymbol{u})$ are differentiable the Riemannian metric tensor on $\mathbf{H}$ is

(55) $$\boldsymbol{g}^H = (\partial \boldsymbol{h}/\partial \boldsymbol{u})(\partial \boldsymbol{h}/\partial \boldsymbol{u})^{\mathsf{T}}$$

where, in a slight abuse of the notation, $\boldsymbol{h}$ is treated as the 4×1 matrix $[w\ x\ y\ z]^{\mathsf{T}}$. Using $\boldsymbol{g}^H$ instead of $\boldsymbol{g}^R$ in Eqns. (47), (49), and (51) we can define differential measures of length, angle, and content in $\mathbf{H}$. In particular, the differential Haar measure is

(56) $$d\mu^H = \sqrt{\gamma^H}\, du_1 du_2 du_3$$

(57) $$\gamma^H = \det(\boldsymbol{g}^H)$$

Comparing Eqns. (44) and (56) we see that the group density on $\mathbf{H}$ is

(58) $$\rho^H(\boldsymbol{u}) = \sqrt{\gamma^H(\boldsymbol{u})}$$

We now show that the metric tensor on $\mathbf{SO}_3$ is directly proportional to the metric tensor on $\mathbf{H}$. We start by noting that $\mathbf{H}$ is homomorphic to $\mathbf{SO}_3$ via

(59) $$\boldsymbol{R} = \begin{bmatrix} 2(x^2+w^2)-1 & 2(xy+wz) & 2(zx-wy) \\ 2(xy-wz) & 2(y^2+w^2)-1 & 2(yz+wx) \\ 2(zx+wy) & 2(yz-wx) & 2(z^2+w^2)-1 \end{bmatrix}$$

We rearrange the elements of $\boldsymbol{R}$ into a column vector $\boldsymbol{r}$ allowing us to write the elements of $\boldsymbol{g}^R$ as

(60) $$g^R_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial \boldsymbol{r}}{\partial u_\alpha}\right)^{\mathsf{T}}\left(\frac{\partial \boldsymbol{r}}{\partial u_\beta}\right)$$

Applying the chain rule to $\boldsymbol{r}(\boldsymbol{u}) = \boldsymbol{r}[\boldsymbol{h}(\boldsymbol{u})]$ yields

(61) $$g^R_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial \boldsymbol{h}}{\partial u_\alpha}\right)^{\mathsf{T}}\left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{h}}\right)^{\mathsf{T}}\left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{h}}\right)\left(\frac{\partial \boldsymbol{h}}{\partial u_\beta}\right)$$

The product $(\partial \boldsymbol{r}/\partial \boldsymbol{h})^{\mathsf{T}}(\partial \boldsymbol{r}/\partial \boldsymbol{h})$ can be written as

$$(62)\quad \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{h}}\right)^{\mathsf{T}}\left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{h}}\right) = 8\{\boldsymbol{h}\boldsymbol{h}^{\mathsf{T}} + 2\,[\mathrm{re}(\boldsymbol{h})\,\boldsymbol{h}^{\mathsf{T}} + \boldsymbol{h}\,\mathrm{re}(\boldsymbol{h}^{\mathsf{T}})] + \boldsymbol{I}_{4\times 4}\}$$

where $\mathrm{re}(\boldsymbol{h}) = [w, 0, 0, 0]^{\mathsf{T}}$ is the "real" part of $\boldsymbol{h}$. We know from differentiating $\boldsymbol{h}^{\mathsf{T}}\boldsymbol{h} = 1$ with respect to u_α that

$$(63)\quad \boldsymbol{h}^{\mathsf{T}}\left(\frac{\partial \boldsymbol{h}}{\partial u_\alpha}\right) = \left(\frac{\partial \boldsymbol{h}}{\partial u_\alpha}\right)^{\mathsf{T}}\boldsymbol{h} = 0$$

Substituting Eqn. (62) into Eqn. (61) and applying Eqn. (63) yields the result we seek

$$(64)\quad g^R = 4g^H$$

Probability on $\mathbf{SO_3}$ and H

Next we define the probability density function (pdf) for a random member of $\mathbf{SO_3}$ and **H**. Let $P_G(\mathbf{N})$ be the probability that a random $X_r \in \mathbf{G}$ lies in the region **N**, where **G** is either $\mathbf{SO_3}$ or **H**. Furthermore, let **N** be bounded by a sphere of radius ε centered on $X \in \mathbf{N}$. More precisely, for any $Y \in \mathbf{N}$, $d(Y, X) \le \varepsilon$, where $d(Y, X)$ is a distance function on **G**. We define the pdf, $p_G(X)$, by the following limit:

$$(65)\quad p_G(X) = \lim_{\varepsilon\to 0}[P_G(\mathbf{N})/\mu_G(\mathbf{N})] = dP_G(X)/d\mu^G(X)$$

The integral of $p_G(X)$ over **N** is $P_G(\mathbf{N})$:

$$(66)\quad P_G(\mathbf{N}) = \int_{X\in\mathbf{N}} p_G(X)\,d\mu^G(X)$$

We can easily derive an expression for the pdf of a random $X_r \in \mathbf{G}$ in terms of the joint pdf of is parameters. Let $\boldsymbol{u}_r$ be the random parameter vector corresponding to X_r and let $\mathbf{N_U}$ be the region in parameter space corresponding to $\mathbf{N} \subset \mathbf{G}$. The probability that $\boldsymbol{u}_r$ lies in $\mathbf{N_U}$ is

$$(67)\quad P_U(\mathbf{N_U}) = \int_{\boldsymbol{u}\in\mathbf{N_U}} p_u(\boldsymbol{u})\,du_1 du_2 du_3$$

where p_U is the joint pdf of $\boldsymbol{u}_r$. Since the two events $X_r \in \mathbf{N}$ and $\boldsymbol{u}_r \in \mathbf{N_U}$ have a one-to-one correspondence, their probabilities must be equal. Equating Eqns. (66) and (67), and applying Eqn. (44), yields

$$(68)\quad \int_{\boldsymbol{u}\in\mathbf{N}_u} p_G[X(\boldsymbol{u})]\rho^G(\boldsymbol{u})\,du_1 du_2 du_3 = \int_{\boldsymbol{u}\in\mathbf{N}_u} p_u(\boldsymbol{u})\,du_1 du_2 du_3$$

Equating integrands gives the final result:

$$(69)\quad p_G(X) = p_U(\boldsymbol{u})/\rho^G(\boldsymbol{u})$$

This equation represents the general expression for the pdf of a random rotation or unitary quaternion in terms of the joint pdf of its parameters.

In the next section we will show several algorithms to generate uniformly distributed random (UDR) members of $\mathbf{SO_3}$ and **H**. The four following definitions of uniformity on **G** are equivalent:

a) $p_G(X)$ is constant.
b) $p_U(\boldsymbol{u})$ is directly proportional to $\rho^G(\boldsymbol{u})$.
c) $P(\mathbf{N})$ is directly proportional to the Haar measure of **N**.
d) X_r and BX_r have the same distribution for any fixed $\boldsymbol{B}$.

The equivalence of these definitions is shown as follows. From Eqn. (69) $p_G(X)$ is constant if and only if $p_U(\boldsymbol{u})$ is directly proportional to $\rho^G(\boldsymbol{u})$. Thus a) $\Leftrightarrow$ b). If $p_G(X)$ is constant then from the integral of Eqn. (66) $P(\mathbf{N})$ is directly proportional to $\mu(\mathbf{N})$. Thus a) $\Rightarrow$ c). If $P(\mathbf{N})$ is directly proportional to $\mu(\mathbf{N})$ then $P(\mathbf{N})$ is in fact a Haar measure and $P(\mathbf{N}) = P(\boldsymbol{B}\mathbf{N}) = P(\mathbf{N}\boldsymbol{B})$, where $\boldsymbol{B}$ is any fixed rotation. In other words the distributions of X_r, $\boldsymbol{B}X_r$, and $X_r\boldsymbol{B}$ are the same. Thus c) $\Rightarrow$ d). If $P(\mathbf{N}) = P(\boldsymbol{B}\mathbf{N}) = P(\mathbf{N}\boldsymbol{B})$ then from Eqn. (65) $p_G(X) = p_G(\boldsymbol{B}X) = p_G(X\boldsymbol{B})$, which can only be true if $p_G(X)$ is constant. Thus d) $\Rightarrow$ a). We have just shown and any of the four definitions can be derived from any of the other three. Thus the definitions are equivalent.

A useful property of a UDR unitary quaternion, $\boldsymbol{h}_r$, is that it correspond to UDR rotation, $\boldsymbol{R}_r$. By definition if h_r is uniformly distributed then $\boldsymbol{h}_r$ and $\boldsymbol{b}\boldsymbol{h}_r$ have the same distribution, where $\boldsymbol{b}$ is any fixed unitary quaternion. Therefore, the rotations $\boldsymbol{R}(\boldsymbol{h}_r)$ and $\boldsymbol{B}\boldsymbol{R}(\boldsymbol{h}_r)$ have the same distribution, where $\boldsymbol{B} = \boldsymbol{R}(\boldsymbol{b})$. Since $\boldsymbol{B} = \boldsymbol{R}(\boldsymbol{b})$ covers $\mathbf{SO_3}$ (twice!), $\boldsymbol{B}$ represents an arbitrary fixed rotation and therefore $\boldsymbol{R}(\boldsymbol{h}_r)$ is uniformly distributed by definition.

RANDOM SAMPLING OF $\mathbf{SO_3}$

We now look at several algorithms to produce uniformly distributed random (UDR) members of the orthogonal group, $\mathbf{O_n}$, as they apply to the specific problem of generating a UDR member of $\mathbf{SO_3}$. We start with a simple but computationally expensive algorithm (Ref. 2). Generate a random $n\times n$ matrix $\boldsymbol{X}$ whose elements are independent, identically distributed (i.i.d.) Gaussian deviates. Find the factorization $\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q}$ is an orthogonal matrix and $\boldsymbol{R}$ is an upper triangular with positive diagonal terms. This factorization is used in the well-known Gram-Schmidt orthogonalization procedure to invert $\boldsymbol{X}$. We can easily prove that $\boldsymbol{Q}$ is uniformly distributed. Since the elements of $\boldsymbol{X}$ are i.i.d. Gaussian, the distributions of $\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{R}$ and $\boldsymbol{B}\boldsymbol{X} = \boldsymbol{B}\boldsymbol{Q}\boldsymbol{R}$ are the same, where $\boldsymbol{B}$ is any fixed orthogonal matrix. Thus, the distributions of $\boldsymbol{Q}$ and $\boldsymbol{B}\boldsymbol{Q}$ are the same, and $\boldsymbol{Q}$ is uniformly distributed on $\mathbf{O_n}$.

Anderson, Olkin, and Underhill (Ref. 12) have devised a very efficient algorithm to generate UDR orthogonal matrices using random Givens rotations. Applied to $\mathbf{SO_3}$ the algorithm is as follows. Choose independent random deviates u_1, u_2, and u_3 which are uniformly distributed on $[0, 2\pi]$, $[-1, 1]$, and $[0, 2\pi]$ respectively. Then form the rotation matrix $\boldsymbol{R}_r$ according to

$$(70)\quad \boldsymbol{R}_r(\boldsymbol{u}) = \boldsymbol{E}_x(u_1)\,\boldsymbol{E}_y(\arccos u_2)\,\boldsymbol{E}_x(u_3)$$

where $\boldsymbol{E}_v$ represents an Euler rotation about the v-axis. This algorithm, when properly implemented, requires only five transcendental operations per sample. We can show that it produces UDR rotations as follows. First, the unitary quaternion $\boldsymbol{h}_r(\boldsymbol{u})$ corresponding to $\boldsymbol{R}_r(\boldsymbol{u})$ is

$$\boldsymbol{h}_r(\boldsymbol{u}) = \begin{bmatrix} \cos[(u_1-u_3)/2]\sqrt{(1-u_2)/2} \\ \sin[(u_1-u_3)/2]\sqrt{(1-u_2)/2} \\ \cos[(u_1+u_3)/2]\sqrt{(u_2-1)/2} \\ \sin[(u_1+u_3)/2]\sqrt{(u_2-1)/2} \end{bmatrix} \tag{71}$$

From Eqns. (55), (64), (53) we find that $\rho^R = 1$. The joint pdf of $\boldsymbol{u}$ is $p_U = 1/(8\pi^2)$. Since p_U is directly proportional to ρ^R, $\boldsymbol{R}_r$ is uniformly distributed.

RANDOM SAMPLING OF H

Previously we proved that UDR unitary quaternions correspond to UDR rotations. We now give two algorithms to generate UDR unitary quaternions. The first is the *projected Gaussian ball sampling* (Refs. 3, 6). Simply generate a quaternion $\boldsymbol{q}$ whose elements are i.i.d. Gaussian deviates. Then set $\boldsymbol{h}_r = \boldsymbol{q}/|\boldsymbol{q}|$. Since the elements of $\boldsymbol{q}$ are i.i.d. Gaussian, $\boldsymbol{q}$ and $\boldsymbol{bq}$ have the same distribution for any fixed unitary quaternion $\boldsymbol{b}$. Therefore $\boldsymbol{h}_r$ and $\boldsymbol{bh}_r$ have the same distribution, and $\boldsymbol{h}_r$ is uniformly distributed.

Shoemake (Ref. 3) has devised a computationally faster algorithm based on the subgroup algorithm of Diaconis and Shahshahanna (Ref. 13). Choose independent random deviates u_1, u_2, and u_3 uniformly distributed on [0,1], [0, 2π], and [0, 2π] respectively. Then generate $\boldsymbol{h}_r$ according to

$$\boldsymbol{h}_r = \begin{bmatrix} \cos u_2\sqrt{1-u_1} \\ \sin u_2\sqrt{1-u_1} \\ \cos u_3\sqrt{u_1} \\ \sin u_3\sqrt{u_1} \end{bmatrix} \tag{72}$$

Note the similarity to Eqn. (71). Shoemake's algorithm requires six transcendental operations per sample, one more than needed to generate a UDR rotation matrix per Eqn. (70). However, if one wants to generate a UDR unitary quaternion rather than a UDR rotation matrix, Shoemake's algorithm is the most efficient. We can show that it produces UDR unitary quaternions as follows. From Eqns. (55) and (58) we find that $\rho^H = 1/2$. The joint pdf of $\boldsymbol{u}$ is $p_U = 1/(4\pi^2)$. Since p_U is directly proportional to ρ^H, $\boldsymbol{h}_r$ is uniformly distributed.

STRATIFIED SAMPLING OF $\mathbf{SO_3}$

Random sampling provides us with a sample set of rotations that are uniformly distributed in a probabilistic sense, but any given set suffers from the clumping discussed in the introduction. Next we develop two algorithms to generate uniformly distributed stratified samples of $\mathbf{SO_3}$, which do not suffer from clumping. Both algorithms partition $\mathbf{SO_3}$ into equal-content and nearly cubic cells and then place a sample in the middle of the cell.

Before delving into these algorithms, let us take a look at how many samples we need to achieve a desired sampling density. Let λ be the length of a side of a cell in the stratification. The cell content is approximately

$$\mu_C = \lambda^3 \tag{73}$$

The total content of $\mathbf{SO_3}$ is $8\pi^2$. Thus the number of samples needed so that the spacing between neighboring samples is approximately λ is

$$n = \text{round}[\, 8\pi^2/\lambda^3 \,] \tag{74}$$

For example, suppose we want to achieve a spacing of 10 degrees ($\lambda = 0.1745$ radians). Then we need 15000 samples.

Masonry Sampling of $\mathbf{SO_3}$

Masonry sampling partitions $\mathbf{SO_3}$ into nearly cubic, equal-content cells arranged in layers. First, we cut $\mathbf{SO_3}$ into uniformly thick u_1-layers, such that the boundaries between layers are surfaces of constant u_1. Next we slice each u_1-layer into uniformly wide u_2-strips. Finally, we dice each u_2-strip into equal-content and nearly cubic cells. In order for the cells to be nearly cubic the cuts, slices, and dices must be mutually orthogonal. To achieve this condition we use an *orthogonal* parameterization, i.e. one whose metric tensor is diagonal. We also want the cells to have equal-content. Thus we need a parameterization for which the thickness of the u_1-layers is uniform over u_2 and u_3, and the width of the u_2-strips is uniform over u_3. We can see from Eqn. (47) that to achieve these conditions g_{11} must be independent of u_2 and u_3, and g_{22} must be independent of u_3. Fortunately, there are a couple of mappings that satisfy these conditions.

We choose a parameterization for which the u_1-layers are homolographic to spherical shells. This choice allows us to use masonry sampling on $\mathbf{S_2}$ to distributed samples uniformly on the u_1-layers. Written in terms of a unitary quaternion, the parameterization is

$$\boldsymbol{h} = \begin{bmatrix} \cos u_1 \\ \sin u_1 \sin u_2 \cos u_3 \\ \sin u_1 \sin u_2 \sin u_3 \\ \sin u_1 \cos u_2 \end{bmatrix} \tag{75}$$

Note that $2u_1$ is the Euler angle and (u_3, u_2) are the spherical coordinates of the Euler axis of the corresponding rotation. We can see from Eqn. (75) that surfaces of constant u_1 are spheres of radius $\sin u_1$.

We only need to distribute points over the $w \geq 0$ hemisphere to cover $\mathbf{SO_3}$. The ranges for u_1, u_2, and u_3 are [0, π/2], [0, π], and [0, 2π] respectively. The metric tensor, differential arc length, and differential content on $\mathbf{H}$ are

$$\boldsymbol{g}^H = \text{diag}(1, \sin^2 u_1, \sin^2 u_1 \sin^2 u_2) \tag{76}$$

(77) $(ds^H)^2 = (du_1)^2 + \sin^2 u_1 (du_2)^2 + \sin^2 u_1 \sin^2 u_2 (du_3)^2$

(78) $d\mu^H = (1/2)[1 - \cos(2u_1)]\sin u_2 \, du_1 du_2 du_3$

The first step in the stratification is to find the desired cell content. Let n be the total number of samples to be generated. The content of the $w \geq 0$ hemisphere is π^2, as can be easily found by integrating Eqn. (78). Thus, if the cells associated with each sample are to be equal in content, then the content of each must be

(79) $$\mu_C = \pi^2/n$$

The ideal length of each side of any given cell is

(80) $$\lambda^* = (\mu_C)^{1/3} = (\pi^2/n)^{1/3}$$

Next we cut the $w \geq 0$ hemisphere into m_1 uniformly thick u_1-layers, where m_1 is

(81) $$m_1 = \text{round}[(\pi/2)/\lambda^*]$$

so that the average thickness of the shells is

(82) $$\lambda_1 = (\pi/2)/m_1 \approx \lambda^*$$

The thickness of each layer must be adjusted slightly from λ_1 so that the content of each layer is an exact integer multiple of μ_c. Let k_1 be the layer index, where $k_1 = 1$ refers to the smallest shell ($u_1 \approx 0$) and $k_1 = m_1$ refers to the largest ($u_1 \approx \pi/2$). Let $N_1(k_1)$ be the total number of cells in layers 1 through k_1. Then the content, $\mu_S(k_1)$, of the $N_1(k_1)$ cells is:

(83) $$\mu_S = N_1 \mu_C$$

Now let $b_1(k_1)$ be the upper value of u_1 for layer k_1. Integrating Eqn. (78) we find

(84) $$\mu_S = \pi[2b_1 - \sin(2b_1)]$$

Solving for N_1 from the last two equations we have

(85) $$N_1 = (\pi/\mu_C)[2b_1 - \sin(2b_1)]$$

We require the boundary between layers to be very close to an integer multiple of the average layer thickness. Therefore we should have

(86) $$b_1(k_1) \approx k_1 \lambda_1 = k_1\pi/(2m_1)$$

Substituting the last expression into Eqn. (85) and rounding to the nearest integer we have

(87) $N_1(k_1) = \text{round}\{(\pi/\mu_C)[k_1\pi/m_1 - \sin(k_1\pi/m_1)]\}$

Next we find the u_1-coordinate for the stratified samples in each layer. We want the samples to lie at the middle of each cell. Since the cells are constructed to be nearly cubic, the u_1-coordinate of the middle dissects the respective layer into two equal-content parts. Using the same approach used to derive Eqn. (85) we see that $u_1(k_1)$ must satisfy:

(88) $[N_1(k_1) + N_1(k_1-1)]/2 = (\pi/\mu_C)[2u_1 - \sin(2u_1)]$

This equation is solved iteratively for u_1. An initial guess, accurate to 1 part in 200, is

(89) $$u_1 \approx (0.75x + 0.154x^2)^{1/3}$$

(90) $$x = [N_1(k_1) + N_1(k_1-1)][\mu_C/(2\pi)]$$

A single Newton-Raphson iteration produces a solution that is accurate to 1 part in 10^7.

The number of cells in the layer k_1 is

(91) $$n_1(k_1) = N_1(k_1) - N_1(k_1-1)$$

where $N_1(0) = 0$. All that remains is to distributed $n_1(k_1)$ samples on layer k_1 using the masonry stratification of $\mathbf{S_2}$ described previously.

The masonry stratification algorithm is summarized as follows. Find λ^* according to Eqn. (80). Find the number of u_1-layers, m_1, from Eqn. (81). For each layer $k_1 = 1$ to m_1 do the following. Find the total number of cells in layers 1 through k_1 from Eqn. (87). Solve Eqn. (88) iteratively for the u_1-coordinate of the samples in layer k_1. Find the number of cells, $n_1(k_1)$ in the layer k_1 from Eqn. (91). Distributed $n_1(k_1)$ samples on layer k_1 using the masonry stratification of $\mathbf{S_2}$ described previously.

Igloo Sampling of $\mathbf{SO_3}$

In this section we develop the igloo sampling for $\mathbf{SO_3}$. The development is made easier by finding the igloo sampling of $\mathbf{H}$ and then mapping it to $\mathbf{SO_3}$. We use the following parameterization of $\mathbf{H}$.

(92) $$\boldsymbol{h} = \begin{bmatrix} \cos u_2 \cos u_1 \\ \sin u_2 \cos u_1 \\ \cos u_3 \sin u_1 \\ \sin u_3 \sin u_1 \end{bmatrix}$$

The metric tensor, differential arc length, and group density are given by

(93) $$\boldsymbol{g}^H = \text{diag}(1, \cos^2 u_1, \sin^2 u_1)$$

(94) $$(ds)^2 = (du_1)^2 + \cos^2 u_1 (du_2)^2 + \sin^2 u_1 (du_3)^2$$

(95) $$\rho^H = \sin(u_1)\cos(u_1) = (1/2)\sin(2u_1)$$

An igloo spiral on $\mathbf{H}$ is described in terms of u_1, u_2, u_3 by

(96) $$u_2 = L_1 u_1$$

(97) $$u_3 = L_2 \sin u_1$$

where $0 \leq u_1 \leq \pi/2$ and $1 << L_1 << L_2$ so that along the spiral u_1 varies the slowest, u_2 the next fastest, and u_3 the fastest. To visualize the spiral we map $\mathbf{H}$ onto two balls in three-dimensional Euclidean space, as shown in Figure 4. Note the left-handed coordinate system used for the left ball. The surfaces of the balls represent the $w = 0$ sphere.

Eqn. (96) on its own represents a *spiroid* – a surface of revolution formed by revolving a spherical spiral lying in the *wxy* hyper-plane about the *wx* plane. The spiroid can be viewed as a cone wrapped within itself, or a snake eating its own tail. The igloo spiral lies on the spiroid. From Eqns. (94) and (96) we see that wraps are separated by

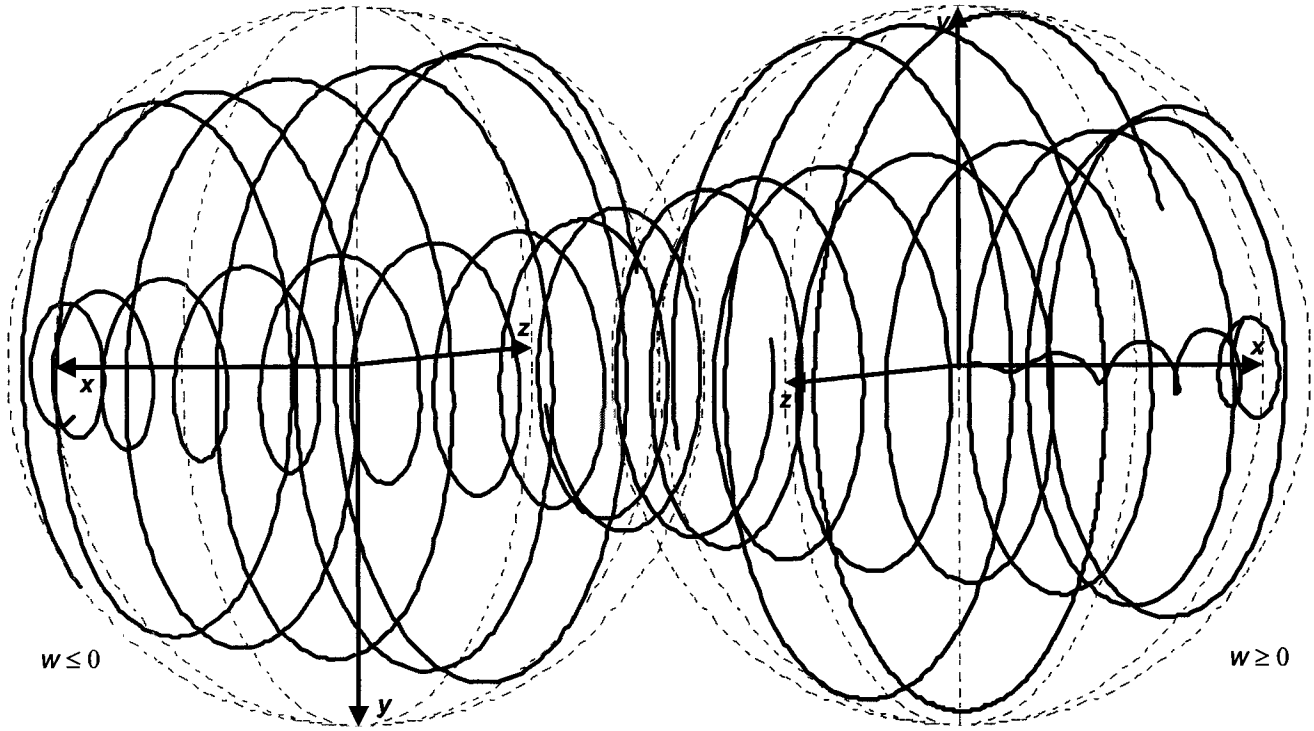


Figure 4 Igloo spiral on H visualized by projecting the spiral onto two balls in three-dimensional Euclidean space.

(98) $$\lambda_1 = 2\pi/L_1$$

The separation between turns of the spiral can be found by integrating along a curve of constant u_3 on the spiroid. The separation has a nearly constant value of

(99) $$\lambda_2 = 2\pi\sqrt{L_1^2 + 1}\ /L_2$$

The constant separation of wraps and turns justifies the terminology "igloo spiral on **H**".

We form the igloo sampling of **H** in a manner analogous to that for $\mathbf{S_2}$. We begin by cutting **H** along the spiroid

(100) $$u_2 = L_1 u_1 - \pi = L_1 (u_1 - \lambda_1/2)$$

to form a *delibroid* – a three-dimensional generalization of the peel on $\mathbf{S_2}$ used previously. The spiroid of Eqn. (96) is the *center-surface* of the delibroid. Next we cut the delibroid along the surface

(101) $$u_3 = L_2 \sin(u_2/L_1) - \pi,$$

$$(u_2 - \pi)/L_1 \le u_1 \le (u_2 + \pi)/L_1$$

to form an *igloo coil*, which is a tightly wound coil having a rectangular cross-section and completely filling **H**. The igloo spiral of Eqns. (97) and (96) is the centerline of the coil. The height and width of the coil measured along the u_1 and u_2 directions are λ_1 and λ_2 respectively. As with the igloo peel on $\mathbf{S_2}$, the cross section of the coil tapers to zero at the ends. Table 2 gives the boundaries of the tapers.

Table 2 Taper Boundaries for the Igloo Coil on H

Description	Value
Tip of head taper at $u_1 = 0$ end of coil	$A_3 = -L_2 \sin(\lambda_1/2)$
Boundary between head taper and main body of coil	$B_3 = +L_2 \sin(\lambda_1/2)$
Boundary between main body of coil and tail taper	$C_3 = L_2 \cos(\lambda_1/2)$
Tip of tail taper at $u_1 = \pi/2$ end of coil	$D_3 = C_2 + 2\pi$

The next step is to find the content $\mu_I(u_3)$ along the coil as a function of u_3 along the centerline. For $B_3 \le u_3 \le C_3$, the main body of the coil, $\mu_I(u_3)$ is given by

(102) $$\mu_I(u_3) = \int_{B_3}^{u_3}\int_{E_2}^{F_2}\int_{E_1}^{F_1} \rho^H du_1 du_2 du_3' + \mu_I(B_3)$$

(103) $$\rho^H = (1/2)\sin(2u_1)$$

(104) $$E_1 = (u_2 - \pi)/L_1$$

(105) $$F_1 = (u_2 + \pi)/L_1$$

(106) $$E_2 = L_1 \arcsin[(u_3' - \pi)/L_2]$$

(107) $$F_2 = L_1 \arcsin[(u_3' + \pi)/L_2]$$

Performing the integration yields

(108) $$\mu_I(u_3) = \alpha_B (u_3^2 - B_3^2) + \mu_I(B_3),$$

$$\text{for } A_3 \le u_3 \le B_3$$

(109) where $\alpha_B = \pi L_1 \sin\lambda_1 / L_2^2$

Next we find $\mu_I(u_3)$ for the tapers at ends of the coil. The exact expressions are unwieldy. Therefore we use approximations. For $A_3 \le u_3 \le B_3$, the head taper, we use

(110) $$\mu_I(u_3) \approx \mu_P[u_2(u_3)]$$

where from Eqns. (96) and (97) we have

(111) $$u_2 = L_1 \arcsin(u_3/L_2) \approx (L_1/L_2)u_3.$$

and $\mu_P(u_2)$ is the content of that part of the delibroid defined by Eqn. (100) from the tip of the delibroid to the surface of constant u_2:

(112) $$\mu_P(u_2) = \int_0^{2\pi}\int_{-\pi}^{u_2}\int_0^{F_1} \rho^H du_1 du_2' du_3$$

$$= (L_1\,\pi/4)\{[2(u_2+\pi)/L_1] - \sin[2(u_2+\pi)/L_1]\}$$

Substituting Eqn. (111) into Eqn. (112) gives us

(113) $$\mu_I(u_3) \approx \alpha_A (u_3 - A_3)^3,$$

$$\text{for } A_3 \le u_3 \le B_3$$

(114) where $$\alpha_A = (\pi/3)(L_1/L_2^3)$$

The shape of the tail taper can be roughly approximated by a torus centered on $u_1 = \pi/2$. The radius of the cross section is chosen so that the content of the torus is the same as that of the taper, namely $[2\pi^2 - \mu_I(C_3)]$. We then approximate $\mu_I(u_3)$ by

(115) $$\mu_I(u_3) \approx \alpha_C (u_3 - C_3) + \mu_I(C_3),$$

$$\text{for } C_3 \le u_3 \le D_3$$

(116) where $$\alpha_C = [2\pi^2 - \mu_I(C_3)]/(2\pi)$$

Next we dice the coil into n equal-content cells and then place a sample at the center of each. Cell 1 is at the head of the coil and cell n is at the tail. The content of each cell is

(117) $$\mu_C = 2\pi^2 / n$$

where $2\pi^2$ is the total content of **H**. We place a sample at the center of cell k by finding u_{3k} according to

(118) $$\mu_I(u_{3k}) = (k - 1/2)\,\mu_C$$

and then u_{1k} and u_{2k} from the igloo spiral equations, Eqns. (96) and (97): Eqn. (118) represents the content of cells 1 through "$k - 1/2$". Its solution depends on which part of the coil the sample is place. For $A_3 \le u_3 \le B_3$ we use Eqn.(113):

(119) $$u_{3k} = [(k - 1/2)\,\mu_C/\alpha_A]^{1/3} + A_3,$$

$$\text{for } 1 \le k \le k_B$$

(120) where $$k_B = \text{floor}\{[\mu_I(B_3)/\mu_C] + 1/2\}$$

For $B_3 \le u_3 \le C_3$ we use Eqn. (108):

(121) $$u_{3k} = \{[(k - 1/2)\,\mu_C - \mu_I(B_3)]/\alpha_B + B_3^2\}^{1/2},$$

$$\text{for } k_B + 1 \le k \le k_C$$

(122) where $$k_C = \text{floor}\{[\mu_I(C_3)/\mu_C] + 1/2\}$$

For $C_3 \le u_3 \le D_3$ we use Eqn. (115):

(123) $$u_{3k} = [(k - 1/2)\,\mu_C - \mu_I(C_3)]/\alpha_C + C_3$$

$$\text{for } k_c + 1 \le k \le n$$

All that remains is to find the spiral parameters L_1 and L_2 that result in the proper spacing of the samples. We find L_1 and L_2 in terms of λ_1 and λ_2 from Eqns. (98) and (99):

(124) $$L_1 = 2\pi/\lambda_1$$

(125) $$L_2 = 2\pi\sqrt{L_1^2 + 1}\,/\lambda_2$$

Recall that the separation between wraps of the igloo spiral on **H** is λ_1 and separation between turns is approximately λ_2. Also, according Eqn. (59), the unitary quaternions $\boldsymbol{h}$ and $-\boldsymbol{h}$ map to the same rotation $\boldsymbol{R}$. Thus for purposes of generating uniformly distributed rotations the left and right balls shown in Figure 4 are equivalent. Conceptually we slide the two balls together along the x-axis until they coincide. Wraps of the igloo spiral become interwoven so that the separation between wraps is $\lambda_1/2$. If we want the separation between wraps on $\mathbf{SO_3}$ to be the same as the separation between turns then we require that

(126) $$\lambda_1/2 = \lambda_2$$

By construction the cells are nearly rectangular boxes on **H**. In terms of the height λ_1, width λ_2,, and depth λ_3 of the cells the content is

(127) $$\mu_C = \lambda_1\,\lambda_2\,\lambda_3$$

We require that

(128) $$\lambda_3 = \lambda_2$$

so that the separation between consecutive samples is the same as the separation between turns. We substitute Eqns. (126) and (128) into Eqn. (127), and the result into Eqn. (117). We then solve for λ_1 to find that

(129) $$\lambda_1 = (8\pi^2/n)^{1/3}$$

We summarize the igloo sampling as follows. First find μ_C, λ_1, and λ_2 from Eqns. (117), (129), and (126). Then find L_1 and L_2 from Eqns. (124) and (125). Find A_3, B_3, and C_3 according to the equations in Table 2. Find $\mu_I(B_3)$ and $\mu_I(C_3)$ from Eqns. (113) and (108). Find k_B and k_C from Eqns. (120) and (122). For $1 \le k \le k_B$ find u_{3k} from Eqn. (119). For $k_B + 1 \le k \le k_C$ find u_{3k} from Eqn. (121). For $k_C + 1 \le k \le n$ find u_{3k} from Eqn. (123). For $u_{3k} \le 0$ set $u_{1k} = 0$ and find u_{2k} from Eqn. (111). For $0 < u_{3k} < L_2$ find u_{1k} and u_{2k} from Eqns. (96) and (97). For $u_{3k} \ge L_2$ set $u_1 = \pi/2$ and $u_2 = 0$. Finally find $\boldsymbol{h}_k$ from Eqn. (92) and R_k from Eqn. (59).

STATISTICAL TESTS FOR UNIFORMITY

We have developed several algorithms to generate uniformly distributed rotations. The random algorithms have been shown mathematically to generate uniformly distributed rotations in a probabilistic sense. The stratified sampling algorithms by their very construction are sure to generate uniformly distributed stratified rotations without the clumping inherent in random sampling. However, a simply mistake during coding could render any of these algorithms ineffective. Therefore we need a way to validate the algorithms as implemented. Below we give three statistical tests to be used for algorithm validation.

First we make an interesting observation regarding the elements of a UDR rotation, $\boldsymbol{R}_r$. From Eqn. (70) we see that R_{31}, the (3,1) element of $\boldsymbol{R}_r$, is uniformly distributed on [–1, 1]. We can choose $\boldsymbol{B}_1$, $\boldsymbol{B}_2 \in \mathbf{SO_3}$ so that any given element of $\boldsymbol{B}_1\boldsymbol{R}_r\boldsymbol{B}_2$ equals R_{31}. Since $\boldsymbol{R}_r$ is uniformly distributed on $\mathbf{SO_3}$, $\boldsymbol{B}_1\boldsymbol{R}_r\boldsymbol{B}_2$ and $\boldsymbol{R}_r$ have the same distribution. Thus every element of $\boldsymbol{R}_r$ is uniformly distributed on [–1, 1]. (Although we will not make use of it, it can also be shown that every column and every row of R_r is uniformly distributed on $\mathbf{S_2}$.) We can validate our sampling algorithms by checking the distribution of the elements of the rotations. Figure 5 shows the cumulative distributions of all nine elements for sets of 15000 samples generated by four sampling algorithms given previously. We see that the distributions are very nearly ideal.

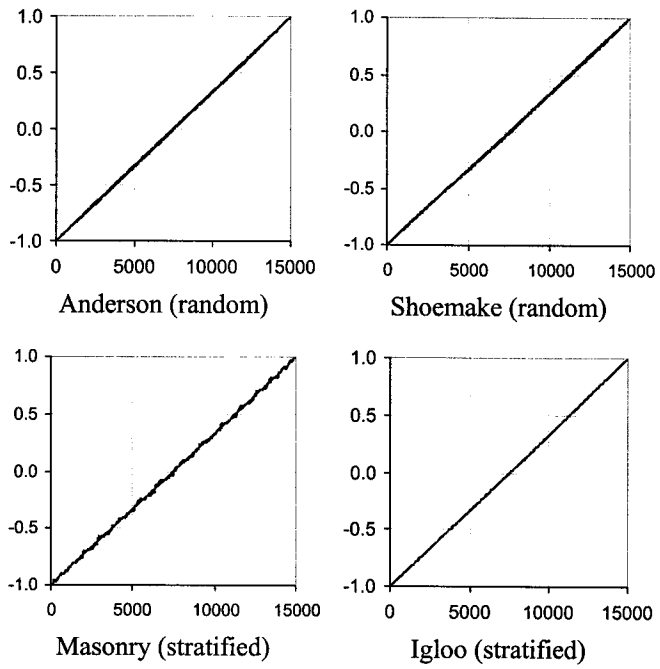


Figure 5 Cumulative distributions of R_{ij} generated by several sampling algorithms (n = 15000)

Next we validate the sampling algorithms using the *orthogonality relations* for irreducible unitary representations of compact groups (Ref. 14). For our purposes these relations tell us that the elements of a uniformly distributed rotation matrix are uncorrelated. More specifically, the expected value of the product of any two elements is

$$(130) E(R_{ij}R_{kl}) = \int_{\mathbf{SO_3}} R_{ij}R_{lm}/(8\pi^2)\, d\mu(\boldsymbol{R}) = (1/3)\,\delta_{il}\delta_{jm}$$

where $\delta_{ab} = 1$ if $a = b$ and 0 otherwise. To show that the elements of the rotation matrices generated by a given sampling algorithm are uncorrelated, we find the statistical correlation coefficient according to (Ref. 15)

$$(131)\ c_{ij,lm} = \sum_k R_{ij}(k)R_{lm}(k) \Big/ \sqrt{\sum_k R_{ij}^2(k) \sum_k R_{lm}^2(k)}$$

where the summations are taken over the n samples indexed by k. Assuming we have already shown that the elements are uniformly distributed on [–1, 1], the denominator becomes $n/3$ and we have

$$(132) \qquad c_{ij,lm} = (3/n) \sum_k R_{ij}(k)R_{lm}(k)$$

Table 3 shows the minimum, maximum, and root-mean-square (RMS) values of the correlation coefficients for sets of 15000 samples generated by four sampling algorithms developed previously. These sample sets are the same ones examined in Figure 5. We see that the correlation coefficients are very nearly zero, validating the algorithm implementations. Furthermore, the correlation coefficients for the stratified samples are an order of magnitude smaller than for the random samples, attesting to the better uniformity of stratified samples.

Table 3 Summary of rotation matrix correlation coefficient for several sampling algorithms (n = 15000)

	Random		Stratified	
	Anderson	**Shoemake**	**Masonry**	**Igloo**
Minimum	–0.0120	–0.0212	–0.0058	–0.0025
Maximum	0.0126	0.0304	0.0049	0.0011
RMS	0.0067	0.0126	0.0013	0.0009

Finally we validate the sampling algorithms using a polling technique. We determine the number of samples, p_m, that are within each of a set of M = 2500 *polling spheres*. Each sphere is the region of $\mathbf{SO_3}$ that is within a rotation angle δ of a *polling center*, $\boldsymbol{R}^C{}_m$.

We need polling centers that have been validated to be uniformly distributed in order to validate that our algorithms produce uniformly distributed samples. To get around this conundrum we generate a reference set using projected Gaussian ball sampling since it is the simplest and most intuitive of the sampling algorithms presented here, and hence the one least likely to suffer a mistake in

implementation. We then validate the uniformity of the polling centers using the previous two methods.

We choose the size of the polling spheres as follows. The average number of samples in the polling spheres is

$$p_A = (8\pi^2/\mu_P)n = n\pi/(\delta - \sin\delta) \quad (133)$$

where $\mu_P = 8\pi(\delta - \sin\delta)$ is the content of each polling sphere (Ref. 11). We set $\delta = 0.3986$ radians so that $p_A = 50$.

Figure 6 shows the normalized frequency of polling spheres versus the number of samples in the spheres. As one would expect the distributions for random sampling (diamonds and squares) follows a Poisson distribution (solid line). The distributions for stratified sampling are much narrower than for random sampling, validating that stratified sampling does not suffer from clumping.

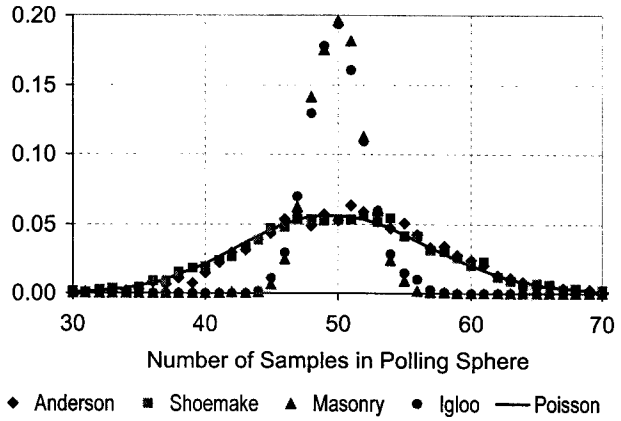


Figure 6 Normalized frequency of polling spheres versus number of samples in the spheres.

CONCLUSIONS

Algorithms to generate uniformly distributed random samples of $\mathbf{SO_3}$ are well known and very easy to implement. However, they suffer from clumping. There are regions with more or less samples than the average. The two stratified sampling algorithms develop here are more complicated to implement but they produce much more uniformly distributed samples. The igloo sampling algorithm, the simpler of the two to implement, produces the most uniformly distributed samples, as can be seen by examining the distribution and correlation coefficients of the rotation matrix elements and the distribution of the samples over $\mathbf{SO_3}$. It is recommended for general use.

REFERENCES

1. Asimov, D., "The grand tour," *SIAM Journal of Scientific and Statistical Computing*, 1983, Vol. 6, No. 1, pp. 128-143.

2. Stewart, G. W., "The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimation," SIAM Journal of Numerical Analysis, Vol. 17, No. 3, June 1980, pp. 403-409.

3. Shoemake, Ken, "Uniform Random Rotations," in *Graphics Gems III*, David Kirk, editor, © 1992 by Academic Press, Inc., pp. 124-132.

4. Steltzer, Ulli, *Building an Igloo*, 1995 Henry Holt and Company Inc., ISBN 0-8050-3753-5.

5. Kreyszig, Erwin, *Differential Geometry*, published 1991 by Dover Publications, Inc, Mineola, New York, ISBN 0-486-66721-9, Chapter III.

6. Muller, M. E., "A note on a method for generating points uniformly on N-dimensional spheres," *Communications of the Association for Computing Machinery*, 1959, Vol. 2, No. 1, pp. 19-20.

7. Feller, William, An Introduction to Probability Theory and Its Applications, Vol. II, 2nd Ed., © 1966 by John Wiley and Sons, pp. 29-33.

8. Halmos, Paul R., *Measure Theory*, © 1950 by D. Van Nostrand Company, Inc., pp. 250-265.

9. Haar, Alfred, "Der Massbegriff in der Theorie der kontinuierlichen Gruppen," *Annals of Mathematics*, Vol. 34 (1933), pp. 147-169.

10. Hurwitz, A., "Ueber die Erzeugung der Invarianten durch Integration", *Nachrichten von der Koeniglichen Gesellshaft der Wissenshaften zu Goettingen*, 1897, pp. 71-90.

11. Hamermesh, Morton, *Group Theory and its Application to Physical Problems*, © 1962 by Morton Hamermesh, published 1989 by Dover Publications, Inc., New York, ISBN 0-486-66181-4, pp. 313-317, 325-332.

12. Anderson, Olkin, and Underhill, "Generation of Random Orthogonal Matrices", Technical Report No. 216, Department of Statistics, Standford University, August 1985.

13. Diaconis, P., and Shahshahani, M., "The Subgroup Algorithm for Generating Uniform Random Variables", Technical Report No. 257, September 1986, Department of Statistics, Stanford University.

14. Sugiura, Mitsuo, *Unitary Representations and Harmonic Analysis*, 2nd ed., © 1990 by Kodansha Ltd, distributed by Elsevier Science Publishing Co., Inc., ISBN 0-444-88593-5, p. 18.

15. Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, 8th ed., © 1989 Iowa State University Press, Ames, Iowa, ISBN 0-8138-1561-4, p. 177.

SESSION 5: IN-FLIGHT ORBIT EXPERIENCE

GETTING THE MOST OUT OF FOUR THRUSTERS ON THE EARTH OBSERVING-1 SPACECRAFT

Kathie Blackman
the Hammers Company

Teresa Hunt, Paul Sanneman
Swales Aerospace

ABSTRACT

The NASA New Millennium Program (NMP) Earth Observing-1 (EO-1) spacecraft was launched in November 2000 on its primary mission to validate advanced remote sensing instruments. One of the critical mission requirements is formation flying with respect to the LandSat-7 mission for instrument image comparison. Due to the nature of the small spacecraft design, only four thrusters could be accommodated on the EO-1 spacecraft. This presented a challenge to the design of the Delta-V controller. This paper presents the design, development and on-orbit performance of this thruster based control mode. The control algorithm utilizes an a-priori open loop firing pattern combined with closed loop feedback control. The observed attitude performance has been well within the five degree requirement, and the delivered Delta-V has been within 1% of the goal. This success will allow the EO-1 Enhanced Formation Flying experiment to proceed with a higher degree of accuracy and precision than would have otherwise been possible.

INTRODUCTION

The goal of NASA's New Millennium Program (NMP) is to identify, develop, and perform flight validation of key breakthrough technologies. Future spacecraft can then take advantage of these technologies without assuming the risks inherent in their first use. The NMP technology development and validation process will also provide a significant return of valuable science data. Thus the NMP flights offer immediate benefits to the scientific community, in addition to a steady stream of breakthrough technologies for future science missions. The EO-1 mission is one of the first in the NMP series of smaller, faster, and cheaper Earth observing spacecraft. NASA's Jet Propulsion Laboratory manages the NMP, while Goddard Space Flight Center has responsibility for the EO-1 mission. The Goddard Space Flight Center awarded Swales Aerospace the prime contract for the EO-1 spacecraft.

The EO-1 spacecraft was launched from Vandenberg Air Force Base on November 21, 2000 into a circular, sun-synchronous polar orbit at an altitude of 705 kilometers. EO-1 was co-manifested on a Boeing Delta II launch vehicle with the SAC-C satellite developed by Argentina. The EO-1 orbital inclination (98.2 degrees) and descending nodal crossing time puts it in "formation flight" with Landsat-7 and EOS-AM (Terra). With all three satellites following the same ground track, EO-1 will be flying "behind" Landsat-7 but "ahead of" EOS-AM (Terra). Due to the formation flight aspect of the mission, the EO-1 spacecraft requires the capability for orbit adjustment maneuvers as well as precision orbit maintenance.

EO-1 SPACECRAFT EVOLUTION

Propulsion System Selection

The EO-1 spacecraft design and development was initiated in early 1996, with the original concept to be suitable for a Taurus launch vehicle; as such, early launch mass for EO-1 was projected near 250 kg. Since the primary EO-1 mission goal required co-flight with Landsat-7 to be maintained for an 18-month period, the Reaction Control System (RCS) was a critical requirement for the satellite system design. Although the initial RCS design proposed eight thrusters, system-level trades quickly reduced that number to six then to four thrusters. The Attitude Control

Subsystem (ACS) engineers accepted the challenge to make the most of a four thruster arrangement, with the lack of redundancy dictated by this single-string mission. During the spacecraft system development in 1996, it became apparent that the RCS was on the critical path of satellite integration. This required that the propulsion system components be selected and procurement initiated for the entire RCS relatively early in the program. As a consequence of this, the propulsion system components were selected based on existing qualification status and availability.

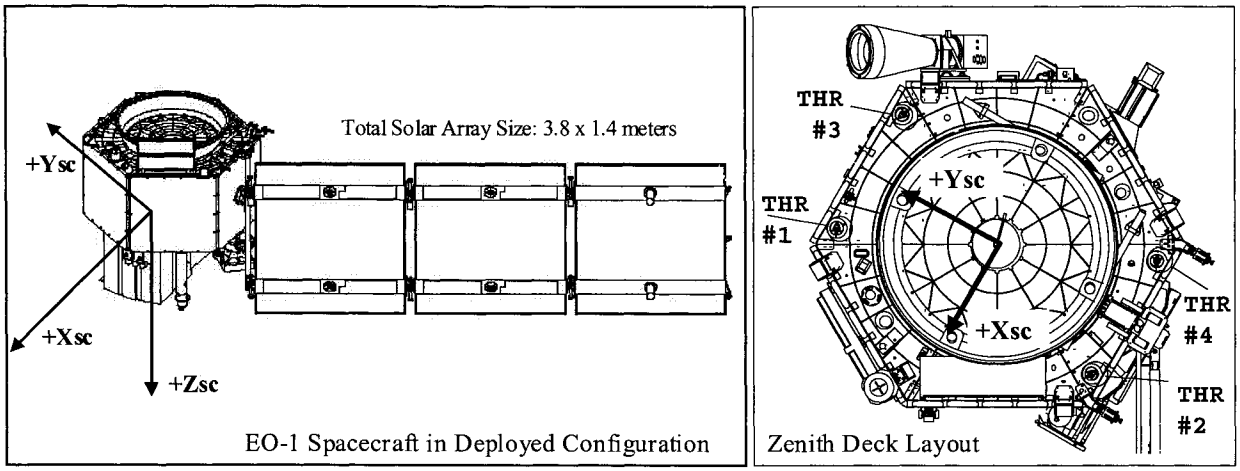


Figure 1 EO-1 Spacecraft and Zenith Deck Thruster Locations

The EO-1 RCS is a mono propellant hydrazine system operating in blow-down mode. The four thrusters are located on the zenith deck, as shown in Figure 1. Note that each of the four thrusters is canted 15 degrees from the –Z axis about a rotation axis drawn from the geometric center of the XY plane to the thruster location. The EO-1 RCS components are pictured in Figure 3. The propellant tank from Pressure Systems Inc. has a propellant capacity of 0.0225 m^3 (1375in^3) and a Maximum Expected Operating Pressure (MEOP) of 2.21E6 Pascals (320 PSI). An AF-E-332 elastomeric diaphragm inside the tank provides positive propellant expulsion using nitrogen as a pressurant. The thruster valves from Primex Aerospace Co. provide dual coil seat valves in series to mitigate risk of valve leakage. The latch valve from Moog is used to isolate the thrusters from the propellant tank.

There are three operational modes of the RCS: Nominal (10W), Standby (36W) and Delta-V (69W). In RCS Nominal mode, only the pressure transducer and thermal control heaters are powered. RCS Standby includes RCS Nominal and adds the load of catalyst bed heaters. RCS Delta-V mode consists of all RCS Standby loads in addition to the operation of all four thrusters. The EO-1 RCS component total mass is 19.6 kg. This mass includes the tank, four thrusters and all other filters, lines, brackets, and pressure transducer, but does not include propellant. The tank pressure, when filled with the 22.3 kg propellant at the Beginning of Life is near 2.055E6 Pascals (298 PSI) at a temperature of 22C. The thruster force for this BOL condition is 0.8 N. The propellant budget is shown in Table 1.

Mass Properties Evolution

As more New Millennium Program technology was added to the EO-1 mission, the Taurus launch vehicle was replaced with a dual-manifest Delta-II rocket. The launch vehicle change significantly increased the satellite mass allocation.

Figure 2 provides an illustration of the mass growth over time. As the spacecraft design matured, its mass properties, which account for mass, inertia and the Center-of-Mass (CM) to Center of Thrust (CT) offset, changed considerably. These mass property changes were of critical importance to the Delta-V mode control design since the thruster locations and orientations remained "frozen" in their original configuration.

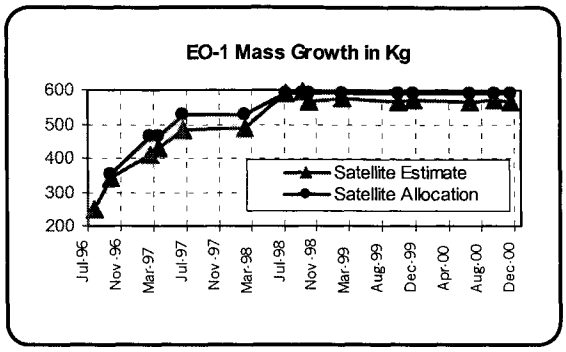


Figure 2 EO-1 Mass Growth

Launch Vehicle Injection Errors	5.0 kg
Orbit Phasing to Formation with Landsat-7	3.0 kg
Formation Flying with Landsat-7 (18 months)	3.0 kg
Deorbit Requirement	9.0 kg
Reserve Margin	2.3 kg
Total Propellant at Launch	**22.3 kg**

Table 1 Propellant Budget

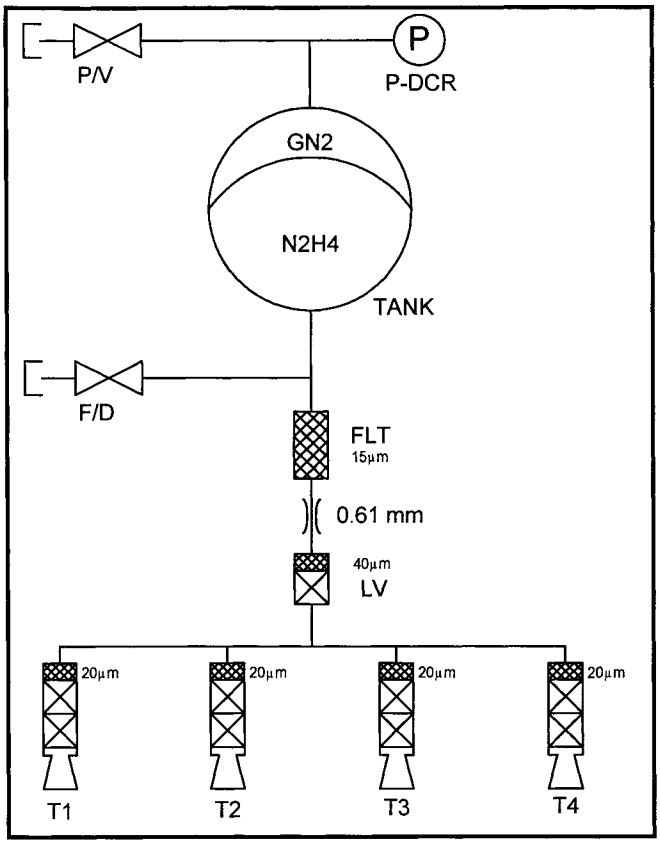


Figure 3 EO-1 RCS Schematic

ATTITUDE CONTROL ELECTRONICS (ACE) PROPULSION I/O BOARD

The basic purpose of the ACE is to provide the low-level electrical functions and interfaces required for the ACS software to the physical hardware of the attitude sensors and attitude control actuators. The Propulsion I/O board, shown in Figure 5, provides power functions, control functions and telemetry measurements for the RCS. Power functions include passing +28V power to the propellant Tank Pressure Sensor (TPS). The power required by the propellant tank pressure sensor flows through the Propulsion I/O board, but on/off control for this +28V power is provided by an ACE switched service. If this power feed from the ACE Low Voltage Power Converter (LVPC) is off, then tank pressure telemetry will be unavailable unless +28V power to the pressure sensor is provided through the T-0 umbilical connector. The analog tank pressure resolution is about 3447 Pascals (0.5 PSIA) per count. Control functions include +28V pulses for opening or closing the latch valve (between fuel tank and fuel lines), and +28V pulses for turning each of four thrusters on or off.

The power required to operate the four hydrazine thrusters flows through the Propulsion I/O board. On/off control for this +28V power is provided by a switched service in the power system. If this service from the power system is off, then the hydrazine thrusters cannot be turned on by ACE software. In order to protect against inadvertently turning a thruster on, the thruster firing circuits are designed with multiple logic and power stages that are electrically connected for ANDed operation. First, there is a separate Enable/Disable bit for each of the four thrusters. After the thrusters are enabled, they are activated with a two-step ARM and FIRE scheme. The two-step

scheme allows verification of the arm command before causing it to actually fire the thrusters. After the thruster ARM bits have been written and verified, the ACE software issues a thruster "execute" command. This command actually turns on the ARMed thrusters.

The ACE software operates on 25 Hz cycles. During thruster firing operations, the software will update the ARM bits and issue a new "execute" command to the Propulsion I/O board every 40 milliseconds. In order to protect against a thruster being unintentionally left on due to a software malfunction or a loss of communication between the ACE RSN and the Propulsion I/O during thruster operations, a re-triggerable analog time-out circuit is used. If this circuit times out, it will automatically turn off all the thrusters.

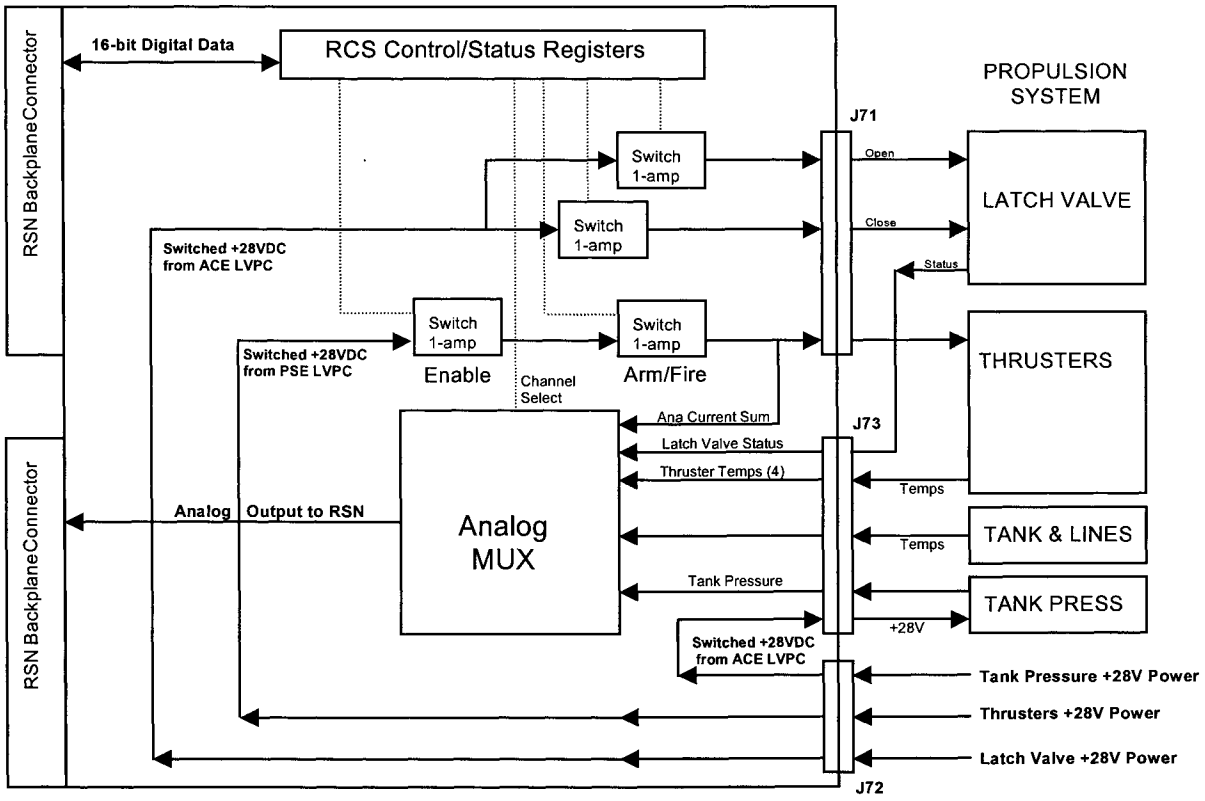


Figure 5 ACE Propulsion I/O Board Block Diagram

The power required for the spacecraft to open or close the latch valve flows through the Propulsion I/O board, but on/off control for this +28V power is provided by an ACE switched service. The latch valve relay is a latching relay, so after an open pulse it will stay open until a close pulse is issued. Likewise, if closed, it will remain closed until an open pulse is sent. Latch valve status (open or closed) is measured via an analog telemetry voltage that is digitized on the RSN board.

Additional propulsion system telemetry includes two propellant tank temperatures, two propellant line temperatures, four thruster valve temperatures, and a Propulsion I/O board temperature. Temperature telemetry is generated using thermistors. The Propulsion I/O board has a 1 mA current source that can be switched into the various thermistors to generate a voltage signal which is digitized at the RSN board. For calibration purposes, the Propulsion I/O can switch its 1 mA current source to a fixed resistor, and also pass a local ground reference back to the RSN for digitization.

DELTA-V CONTROL MODE DESIGN & ANALYSES

Thrust to Torque Mapping Development

Table 3 identifies thruster pair combinations for positive/negative control torque about each of the spacecraft axes. This assumes that the spacecraft center of mass resides along the vehicle Z-axis centerline, the thrusters are precisely located and oriented, and that each combination pair of thrusters is perfectly matched. Under these conditions, the controller dynamics would be decoupled in the three axes.

Thruster Number	Control Axis					
	+Roll	-Roll	+Pitch	-Pitch	+Yaw	-Yaw
1	X			X		X
2		X		X	X	
3	X		X		X	
4		X	X			X
Note: "X" Indicates Thruster Firing						

Table 3 Thruster Phasing with Center of Mass along Spacecraft Centerline

As seen in Figure 6, the spacecraft center of mass location has non-zero X and significant non-zero Y components. Furthermore, it can be seen that the X axis component is significantly affected by the solar array angle. This center of mass offset from the spacecraft centerline results in substantial coupling between control axes for the thruster pairings shown in Table 3. It is also known that thrust level variations and fabrication alignment tolerances in individual thruster orientations contribute to axis-to-axis coupling; however, these should be second-order perturbations compared to the influence of the center of mass offset.

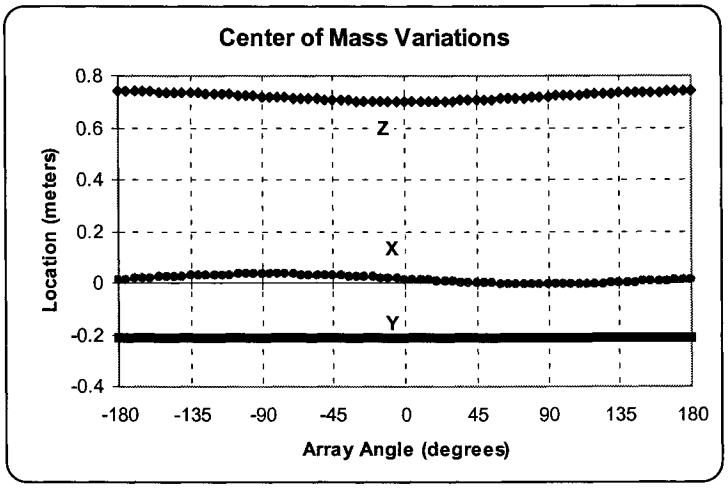


Figure 6 Center of Mass Variations with Solar Array Angle

A method was developed to determine thrust levels required by each of the four thrusters to achieve a decoupled system, similar to that in Table 3. Specifically, this method determines the thrust levels required for each thruster to achieve a unit magnitude torque about each of the spacecraft axes, in both positive and negative directions. This presupposes that thrust levels can be varied, which is not the case. However, thruster duty cycles can be varied, with the same net effect as varying thrust.

To determine the torque generated by each thruster, the thruster location with respect to the center of mass, Ri, must be crossed with the directional force vector:

$$\vec{\mathbf{T}}_i = \vec{\mathbf{R}}_i \times \vec{\mathbf{F}} = (\vec{\mathbf{R}}_{th_i} - \vec{\mathbf{R}}_{cm}) \times \vec{\mathbf{F}}_i \qquad i=1,2,3,4 \tag{1}$$

The axial components are summed to determine the total torque on the spacecraft generated by all four thrusters:

$$\vec{\mathbf{T}} = \sum_{i=1}^{4} \vec{\mathbf{T}}_i = \sum_{i=1}^{4} (\vec{\mathbf{R}}_{th_i} - \vec{\mathbf{R}}_{cm}) \times \vec{\mathbf{F}}_i \tag{2}$$

or

$$T = [\vec{R}_1 \times \vec{U}_1 \quad \vec{R}_2 \times \vec{U}_2 \quad \vec{R}_3 \times \vec{U}_3 \quad \vec{R}_4 \times \vec{U}_4] \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix} \tag{3}$$

where the unit vectors, Ui, provide the individual thrust vectors in the spacecraft coordinate frame:

$$\vec{F}_i = \vec{U}_i F_i \qquad i=1,2,3,4 \tag{4}$$

If the 3x4 matrix of cross products on the right side of the equation is represented by **A**, and the 4x1 thrust level vector $[F_1 \quad F_2 \quad F_3 \quad F_4]^T$ is represented as **F**, then:

$$T = AF \tag{5}$$

The EO-1 thruster values used in Equations 1-5 to compute the thruster torque vectors are shown in Table 4. Note the 15 degree thruster cant indicated by all four Z components of the unit vector orientation equal to the cosine of 15 degrees or 0.96593.

Basic matrix algebra would initially indicate that the thrust levels required for any desired torque could be determined by defining the torque vector and pre-multiplying by the pseudo-inverse of A:

$$F = [A]^{-1} T \tag{6}$$

Note, however, that this linear transformation maps rational, real torques into rational real thrust values, when what is needed is a transformation to rational, real *positive* thrust values. Therefore, the pseudo-inverse cannot be used to map torque to thrust because of the positive space-constraint on thrust value solutions.

Center of Mass Location (meters)

X	Y	Z
0.01728	-0.20000	0.70358

Thruster Locations on Spacecraft (meters)

	X	Y	Z
Thruster 1	0.26040	0.58280	0.20960
Thruster 2	0.26040	-0.58280	0.20960
Thruster 3	-0.26040	0.58280	0.20960
Thruster 4	-0.26040	-0.58280	0.20960

Thruster Moment Arms (Ri)

	X	Y	Z
Thruster 1	0.24312	0.78280	-0.49398
Thruster 2	0.24312	-0.38280	-0.49398
Thruster 3	-0.27768	0.78280	-0.49398
Thruster 4	-0.27768	-0.38280	-0.49398

Thrust Unit Vector Orientations (Fi)

	X	Y	Z
Thruster 1	0.23630	-0.10558	0.96593
Thruster 2	0.23630	0.10558	0.96593
Thruster 3	-0.23630	-0.10558	0.96593
Thruster 4	-0.23630	0.10558	0.96593

Thruster Torque Vectors (Ti)

	X	Y	Z
Thruster 1	0.70397	-0.35156	-0.21064
Thruster 2	-0.31760	-0.35156	0.11612
Thruster 3	0.70397	0.38495	0.21429
Thruster 4	-0.31760	0.38495	-0.11977

Table 4 Thruster Torque Table Calculation

In order to design the desired torque to thrust mapping, a Microsoft Excel program was created to solve for a set of positive thrust values that would produce a desired unit torque. The program is based on the principle that the four thruster torques divide the space around the spacecraft center of mass into four regions, each one bounded by three surfaces defined by three of the four thruster torque vectors (Reference 1). The desired control torque vector can always be attained using only three of the four thrusters. Such a solution is either the maximum or minimum for some cost function, and if one of the thrusters is zero, there is a unique solution for the remaining three thrusters. The Excel program constructs the four 3x3 combinations of three-thruster sets, calculates the matrix inverse and determines which transformation provides the desired axial torque with positive thrust values. In most cases, the resulting solution represents the minimum fuel consumption rate for the desired body torque. Table 5 provides an example of the process performed for the –Roll torque mapping, which is repeated using all four 3x3 combinations for the other five torque vectors. Instead of the 4x3 pseudo-inverse discussed above, this results in a 4x6 thruster map matrix described as:

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix} = [4\text{x}6\ \text{Thruster Map}] \begin{bmatrix} \begin{bmatrix} T_{+R} \\ T_{-R} \end{bmatrix} \\ \begin{bmatrix} T_{+P} \\ T_{-P} \end{bmatrix} \\ \begin{bmatrix} T_{+Y} \\ T_{-Y} \end{bmatrix} \end{bmatrix} \tag{7}$$

where $\mathbf{T}_{+R}$ is the magnitude of the positive roll torque command, and $\mathbf{T}_{-R}$ is the magnitude of the negative roll command [the roll doublet], and similar for the other two pairs. Note that only one element in each of the three doublets can be non-zero since the per axis torque command cannot be positive and negative at the same time.

Given the thruster locations, thrust vector orientations, and the spacecraft center of mass, it is possible to calculate the individual thrust values that would provide pure linear acceleration without torque disturbances. These four thrust levels could be applied in an open-loop, a-priori sense to perform the delta-v maneuver on the satellite. This a-priori estimate is called the Open Loop Firing Pattern (OLFP). The OLFP is essentially the sum of all six unit torque commands, with each of the four resulting elements divided by the largest element. The idea here is that if all

+/- unit torques were commanded simultaneously in each axis, there would be no net torque on the spacecraft, and at least one thruster would be commanded at a 100% duty cycle (full on). The name OLFP implies a bias thruster duty cycle profile in the absence of the closed loop attitude control commands. Table 6 shows the 4x6 thruster map matrix and OLFP computed for EO-1.

Step 1: Develop matrix of thruster torque vectors

	THR1	THR2	THR3	THR4
Tx	0.7040	-0.3176	0.7040	-0.3176
Ty	-0.3516	-0.3516	0.3849	0.3849
Tz	-0.2106	0.1161	0.2143	-0.1198

Step 2: Form 3x3 without thruster 3

A3	0.7040	-0.3176	-0.3176
	-0.3516	-0.3516	0.3849
	-0.2106	0.1161	-0.1198

Step 3: Calculate matrix inverse C3

C3	-0.0352	-1.0154	-3.1702
	-1.6696	-2.0494	-2.1595
	-1.5569	-0.2012	-4.8674

Step 4:

Evaluate [C3][-1.0 0.0 0.0]' to determine if –Roll torque can be provided by positive thrusts from THR1, THR2, and THR4
[0.0352 1.6696 1.5569]'

Step 5:

Since all components are positive, the resultant vector is the –Roll torque-to-thrust map with TH3 zero (A3 constraint)

	(-) X
THR1	0.0352
THR2	1.6696
THR3	0.0000
THR4	1.5569

Table 5 Sample Torque-to-Thrust Mapping Calculation

Thruster Map for Solar Array Angle = 0 Deg (Orbit Noon)					
	Axis	Thruster 1 (N)	Thruster 2 (N)	Thruster 3 (N)	Thruster 4 (N)
	Pos Roll (+X) (1 Nm)	0.7131	0.0284	0.6887	0.0000
	Neg Roll (-X) (1 Nm)	0.0284	1.7350	0.0000	1.6378
	Pos Pitch (+Y) (1 Nm)	0.0000	0.2279	0.9065	1.9278
	Neg Pitch (-Y) (1 Nm)	0.9760	2.0931	0.0000	0.2279
	Pos Yaw (+Z) (1 Nm)	0.0000	5.1119	2.9187	1.8291
	Neg Yaw (-Z) (1 Nm)	3.1425	2.3614	0.0000	5.1119
		duty cycle	duty cycle	duty cycle	duty cycle
	OLFP	0.4205	1.0000	0.3905	0.9288

Table 6 Final Total Torque-to-Thrust Mapping

Algorithm Design and Development

The thruster control algorithm is comprised of four components or steps described below and illustrated in Figure 7.

Component 1
Component 1 is comprised of a PD controller, acting on the position and rate error to generate a control torque command. The position error is formed between the orbit reference frame and spacecraft attitude. The controller bandwidth is 0.015 Hz with a damping ratio of 0.707. The control torque is also passed through a digital second-order low-pass filter with a 0.1 Hz break frequency.

Since the Delta-V mode attitude hangoff error can induce a small amount of undesired cross track velocity, the orbit reference frame is formed from a propagated ephemeris originating before the delta-V burn, rather than from the Global Positioning System (GPS) receiver navigation data during the burn. This is required to prevent a positive feedback effect of velocity vector updates from GPS and the formation of the orbit reference frame: i.e. avoid chasing the undesired cross track velocity and artificially resetting the cross track attitude error.

Position and rate control error is monitored by the ACS failure detection and correction (FDC) software which will abort the delta-V burn if predetermined limits are exceeded. No limiting is performed on the position and rate error within the PD control logic itself.

Component 2
Component 2 uses the thruster map, OLFP and assumed steady state thruster levels to map the commanded torque into a per-thruster force command. Two sets of commands are computed, the first being associated with the closed loop PD control. Any available thruster "headroom" left by this control command is then used to scale the OLFP thrust levels. The scaling maintains the relative proportions of the OLFP. The closed loop command and the scaled OLFP command are superimposed, resulting in at least one of the thrusters being commanded at 100% duration for the associated 1 Hz control cycle.

Component 3
The third component maps the commanded thrust levels into specific thruster duty cycle commands. Since the thrusters cannot be "throttled", the thrust level command from the second component cannot be implemented directly. The third component converts the momentum which would have been generated by the commanded thrust level over the 1 Hz control cycle to the thruster duty cycle which most closely matches this momentum. The result is an integer number of 40 ms control cycles, ranging from 0 to 25, for which each specific thruster is to be fired over the 1 Hz control cycle. The single parameter required by this component is the per-thruster rated thrust (4x1 vector) already used in Component 2. Alternately, a momentum scaling parameter vector can be defined as:

$$h\_scale_i = \frac{1}{rated\_thru\ st_i} * \frac{1}{0.04} \qquad i = 1,2,3,4 \tag{8}$$

When the thrust command from Component 2 is scaled by this parameter, it is mapped into an equivalent number of ACE control counts (0-25) over the 1 Hz control loop. The number is quantized to the closest integer number of counts.

Component 4
The final component resides in the ACE and is executed at 25 Hz. The integer command from Component 3 is compared against the value of the ACE 25 Hz counter. If the value of the command is greater than or equal to the counter, the command is sent to the appropriate Electronic Valve Driver to fire the thruster. The logic will result in no firing if the command from Component 3 is 0, and will result in 100% thruster duty cycle if the command is 25.

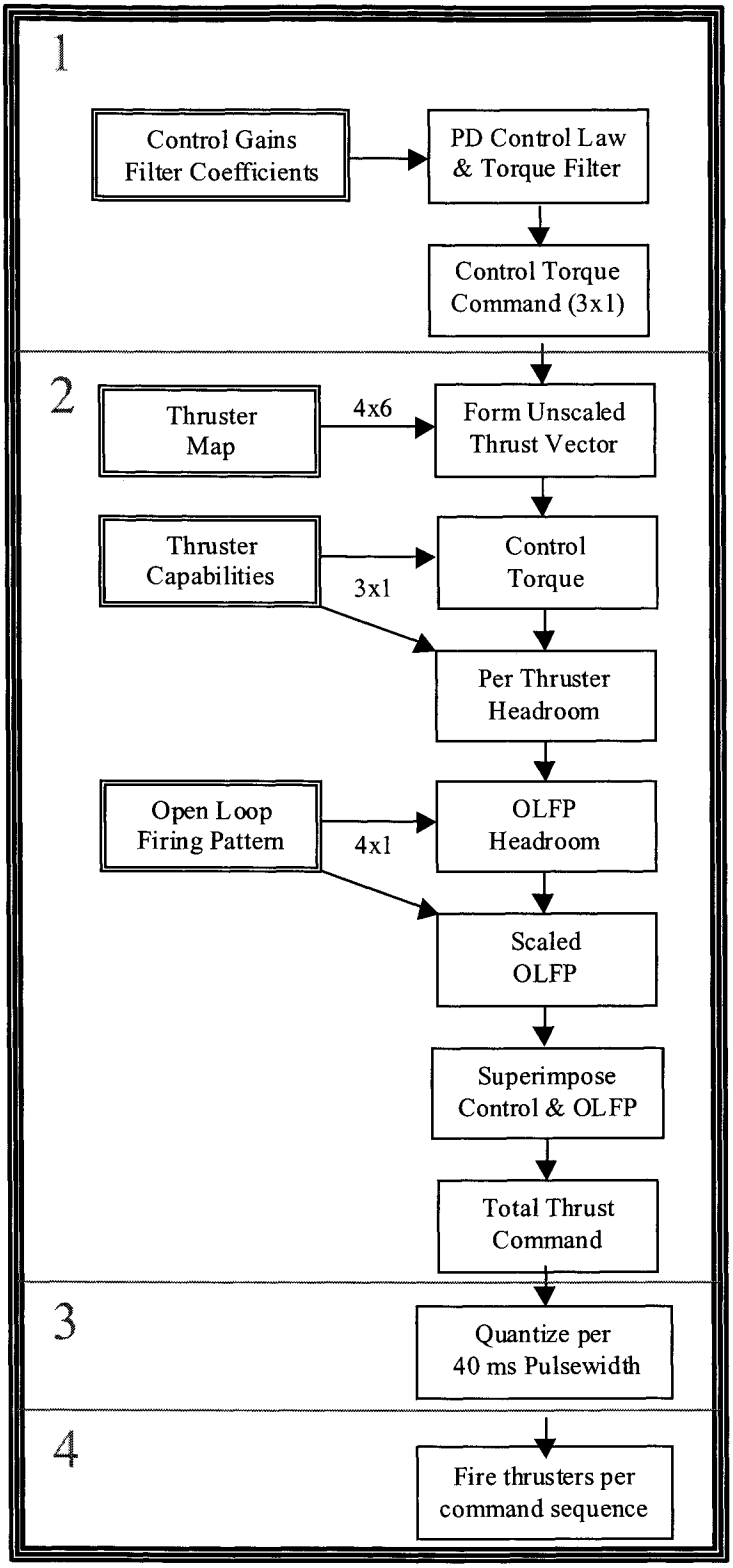


Figure 7 Thruster Control Algorithm

Robustness Validation via Simulation

A closed loop simulation was developed in the MATRIXx/SystemBuild environment to predict end-to-end performance of the Delta-V thruster control scheme. This simulation does not incorporate high-fidelity attitude determination sensor models, since the impact of sensor noise and quantization should be small compared to the thruster duty cycle quantization.

A circular orbit model is used to establish the control reference frame. This emulates the process that occurs in flight wherein the orbit is propagated based on pre-burn initial conditions (i.e. GPS navigation updates during the burn are not utilized). In addition to the circular orbit model used for the control frame, the simulation also implements a point source gravity model to represent true spacecraft position R and velocity V in the Earth Centered Inertial (ECI) frame. Additionally, the actual orbit reference frame was formed in the model, designated in the simulation as ORF. This frame is defined with the Z axis aligned with the –R vector, Y orthogonal to the orbit plane, and X in the RV plane, orthogonal to Y and Z. The velocity output from the orbit model is mapped into both the control reference frame and the orbit reference frame ORF.

In addition to the three degree of freedom (DOF) linear dynamics modeled in the orbit model, the simulation is configured with 3 DOF rotational dynamics. The spacecraft flexible body dynamics are included in the simulation in a state space model using thruster translational DOFs as input forcing functions, and IRU rotational and solar array tip translational deflections as output nodes. The IRU rotational modal responses are superimposed on the rigid body dynamics in order to investigate potential control/structure interactions. Since the spacecraft modal response varies with solar array orientation, the flexible body dynamics were developed for each of three solar array angles: 0, -90 and +90 degrees. The first nine flex modes for each axis and each array position were included as part of the robustness evaluation.

In addition to the software algorithm models operating at 1 Hz and 25 Hz, and the 6 DOF dynamics, the simulation includes a high fidelity thruster model developed to emulate actual propellant mass flow build-up/tail-off profiles generated in thruster tests. There is a small amount of excess momentum generated by a thruster pulse, when compared to an ideal rectangular pulse profile. The inclusion of the high fidelity thruster model in the simulation provides an accurate representation of actual delta-V performance, as opposed to simply multiplying the rated thrust by the pulse width to estimate momentum.

Although a describing function analysis was considered for evaluating the controller robustness to plant and actuator uncertainties, the high-fidelity simulation program described above was the primary tool for investigating sensitivity to parameter variations. Table 7 provides a summary of the simulation sequences utilized to demonstrate robustness.

Parameter Variation	Description of Significance
Propellant tank feed pressure: Beginning of Life versus End of Life	Fuel depletion causes a lowering of system pressure and thus results in lower thrust force at each valve
Initial conditions for attitude and rate errors	Demonstrate ability to capture and attenuate residual control errors from previous control mode
Variation of modal frequencies (±20%) and admittances (±200%) in the flexible body dynamics	Without a modal survey of full spacecraft, frequency resonances and mode shapes are more uncertain
Individual thrust level imbalances (hot/cold thrusters)	Thrust performances vary with each unit, with temperature, line pressure, and age/duty cycles
Incorrect Open Loop Firing Pattern	Demonstrate capability to recover from incorrectly calculated or uploaded OLFP ratios, including the case of starting with all thrusters at 100%

Table 7 Delta-V Simulation Summary

ON-ORBIT PERFORMANCE

The orbital maneuver campaign to position the EO-1 satellite in orbit with respect to the Landsat-7 satellite was initiated on Day 4 of the mission. A 60 second calibration burn was performed to evaluate thruster operation and confirm the expected attitude control trajectories with respect to pre-launch simulations. Since the orbital separation between EO-1 and Landsat-7 continued to drift on a daily basis, the maneuver campaign required performing the first four burns relatively close to one another. The delivered delta-velocity accuracy for the calibration burn was within 5% of the prediction and all subsequent burns have been performed within 0 to 2% error. A summary of the burns performed to date is provided in Table 8. The corresponding fuel mass usage over time is shown in Figure 12.

The derived requirements for the Delta-V mode control were 5.0 degrees attitude error, and 0.2 deg/sec rate error, 3σ. The on-orbit delivered accuracy has been better than 2.0 degrees attitude error and 0.05 deg/sec rate error. Figures 7 through 10 illustrate the attitude control performance parameters for the 17 minute delta-v executed on mission day 21 (00-346).

Due to significant changes in the satellite mass properties following the initial integration of the thrusters, as discussed earlier, all of the thrusters are offset with respect to the center of mass. Thrusters 2 and 4 have the shortest moment arms to the center of mass and are typically on at or near 100% of the time. Thrusters 1 and 3 have longer moment arms and are typically on 30 to 40% of the time, as illustrated in Figure 9.

Mission Day	UTC Day	Duration (sec)	Burn Type	Delta-V (m/sec)	Fuel Used (kg)
4	00-329	60	ALT+	0.21	0.058
6	00-331	514	INC	1.81	0.491
8	00-333	831	ALT+	2.71	0.735
10	00-335	901	ALT+	2.77	0.743
19	00-344	430	ALT-	1.32	0.342
21	00-346	1015	ALT-	2.83	0.759
23	00-348	75	INC	0.22	0.053
25	00-350	182	ALT+	0.53	0.129
25	00-350	114	ALT-	0.33	0.08
31	00-356	12	ALT+	0.04	0.009
		6	ALT-	0.02	0.004
45	01-004	49	ALT-	0.15	0.037
		25	ALT+	0.08	0.019
59	01-018	17	ALT+	0.05	0.012
80	01-039	31	ALT+	0.10	0.023
101	01-060	6	ALT+	0.02	0.0044
		20	ALT+	0.06	0.0146
122	01-081	12	ALT-	0.04	0.0086
		37	ALT+	0.11	0.0265
129	01-088	6	ALT+	0.02	0.0044
		18	ALT+	0.06	0.0131

Table 8 On-Orbit Delta-V Summary

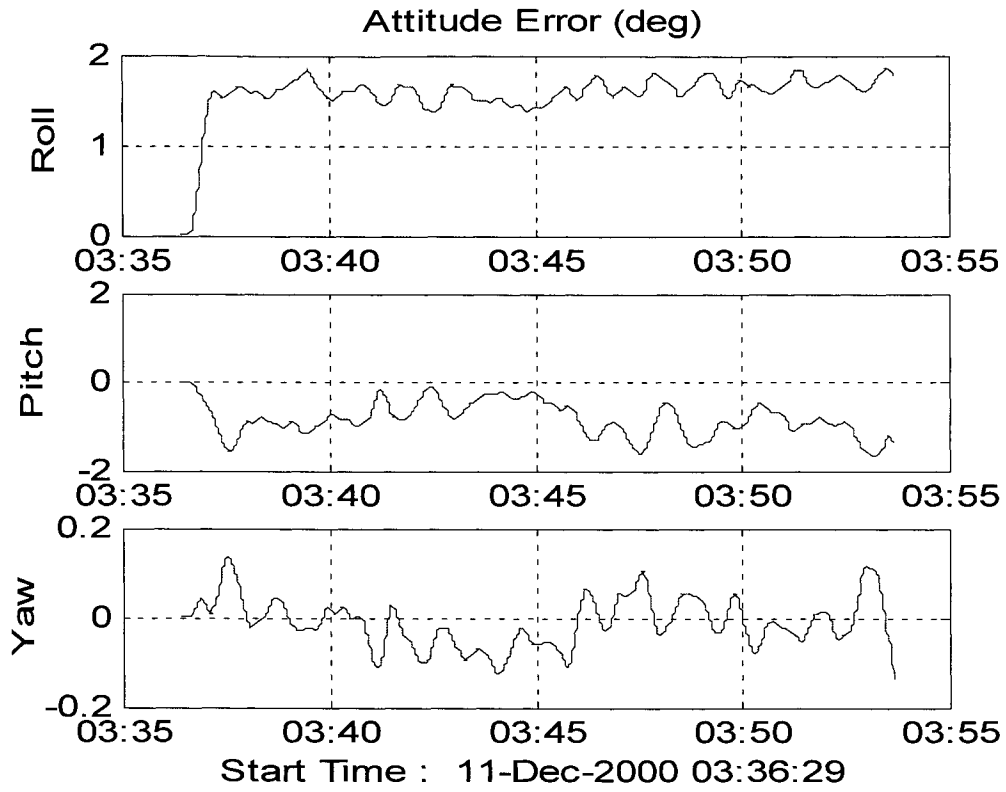


Figure 7 Delta-V Attitude Errors

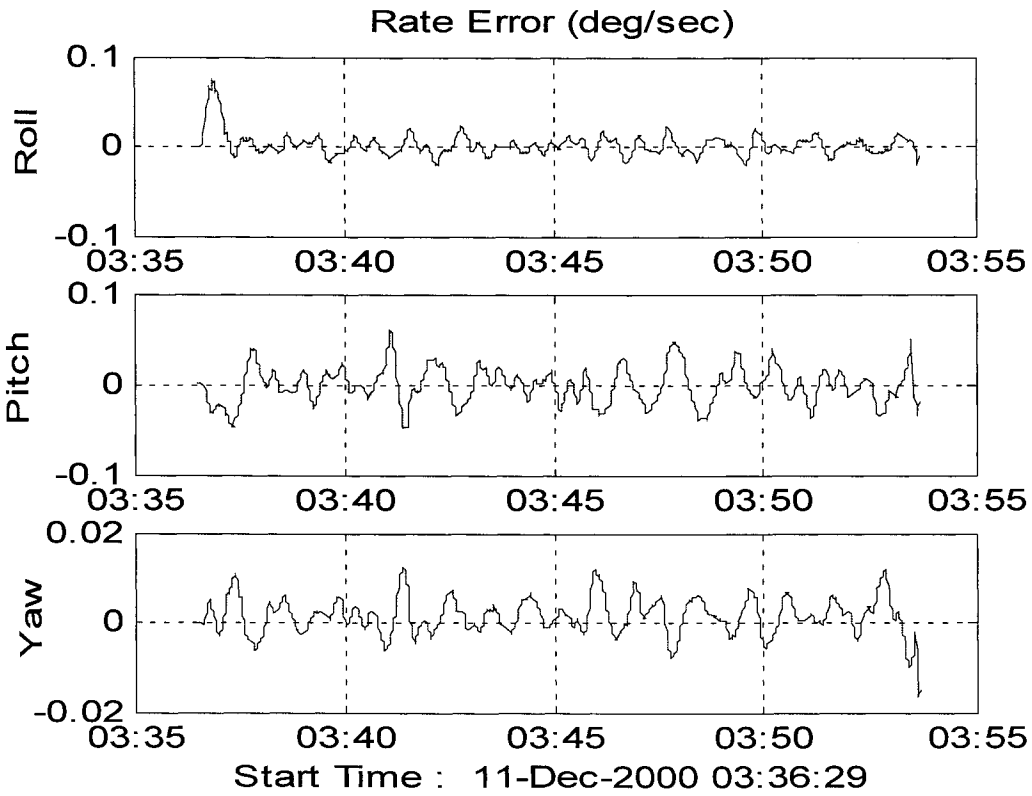


Figure 8 Delta-V Rate Errors

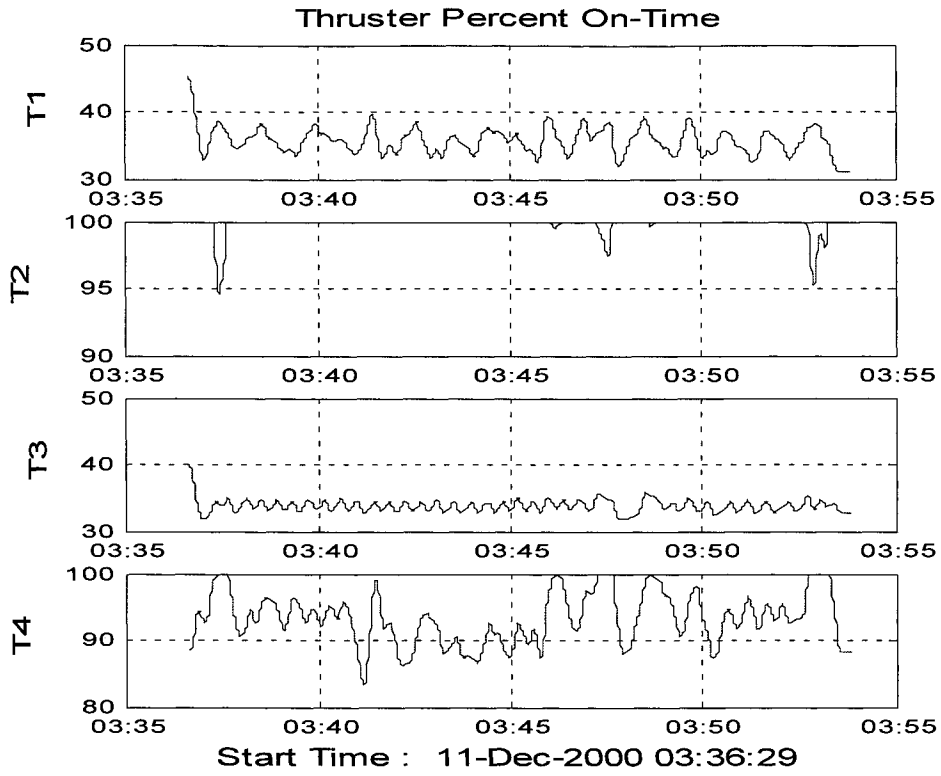


Figure 9 Delta-V Thruster On-Time

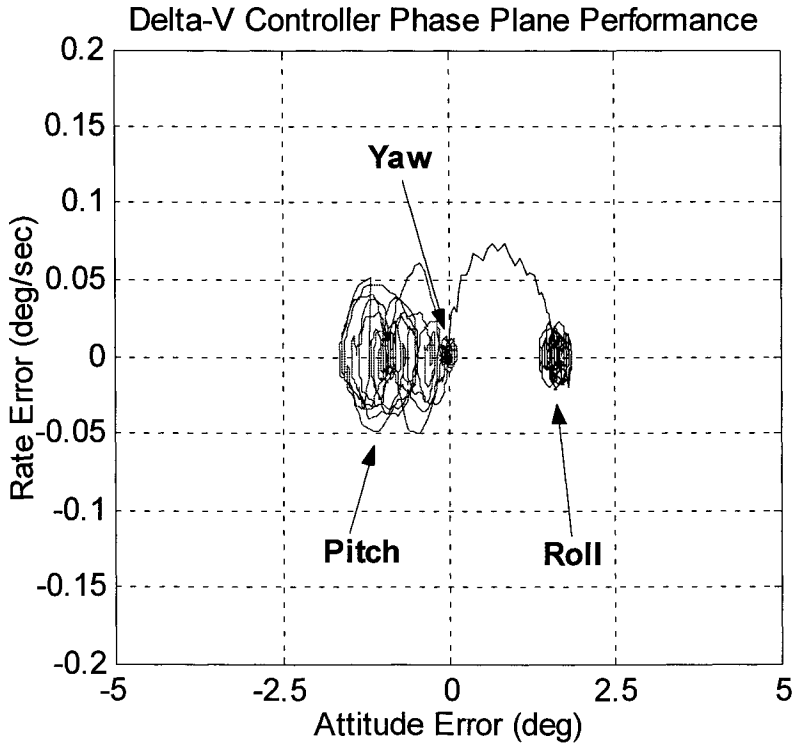


Figure 10 Delta-V Phase Plane

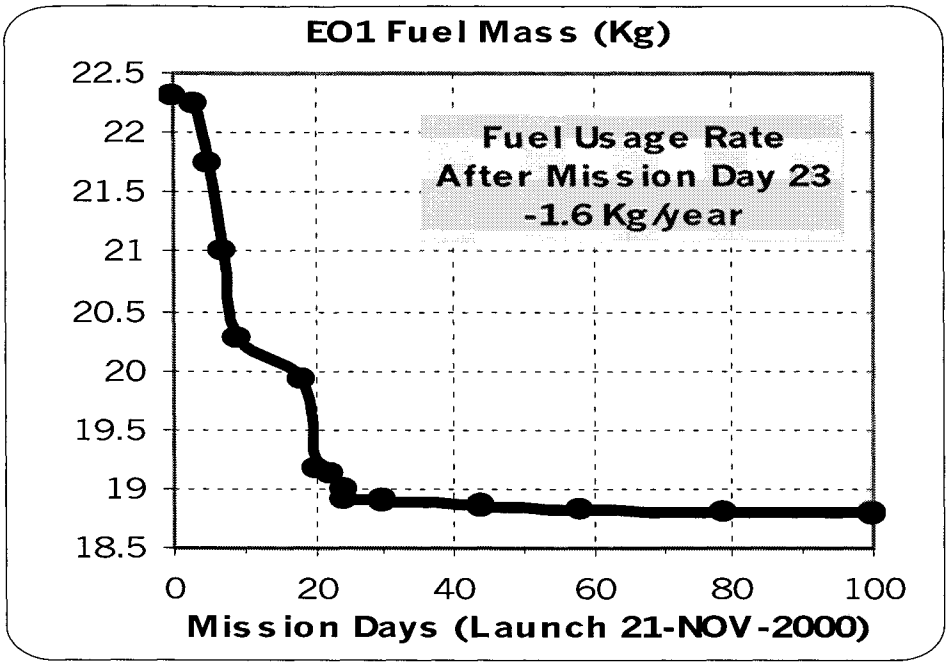


Figure 12 EO-1 Fuel Mass

CONCLUSION

The EO-1 Reaction Control Subsystem implementation and Delta-V controller design had numerous challenges as part of a 'faster, better, and cheaper' mission. The derived requirements for the Delta-V mode control were 5.0 degrees attitude error, and 0.2 deg/sec rate error, 3σ. The on-orbit delivered accuracy has been better than 2.0 degrees and 0.05 deg/sec despite the evolution of the spacecraft mass properties. Further analytical studies have demonstrated that the employed algorithm could be made even more fuel efficient by including the four 3x3 matrices in the calculation for each control cycle (similar to the matrix calculated in Step 3 of Table 4). The performance and predictability of the Delta-V mode has conserved propellant usage and set the stage for extended mission operations with the Enhanced Formation Flying experiment.

REFERENCES

1. M. Sidi, "Spacecraft Dynamics and Control", Cambridge University Press, 1997, pgs 287-289.

PRECISE ORBIT DETERMINATION FOR GEOSAT FOLLOW-ON USING SATELLITE LASER RANGING DATA AND INTERMISSION ALTIMETER CROSSOVERS

Frank G. Lemoine, David D. Rowlands, Scott B. Luthcke
Space Geodesy Branch, Code 926, NASA GSFC

Nikita P. Zelensky, Douglas S. Chinn, Despina E. Pavlis
Raytheon ITSS Corp., Lanham, MD

Gregory C. Marr
Flight Dynamics Analysis Branch, Code 572, NASA GSFC

ABSTRACT

The U.S. Navy's GEOSAT Follow-On Spacecraft was launched on February 10, 1998 with the primary objective of the mission to map the oceans using a radar altimeter. Following an extensive set of calibration campaigns in 1999 and 2000, the US Navy formally accepted delivery of the satellite on November 29, 2000. Satellite laser ranging (SLR) and Doppler (Tranet-style) beacons track the spacecraft. Although limited amounts of GPS data were obtained, the primary mode of tracking remains satellite laser ranging. The GFO altimeter measurements are highly precise, with orbit error the largest component in the error budget. We have tuned the non-conservative force model for GFO and the gravity model using SLR, Doppler and altimeter crossover data sampled over one year. Gravity covariance projections to 70x70 show the radial orbit error on GEOSAT was reduced from 2.6 cm in EGM96 to 1.3 cm with the addition of SLR, GFO/GFO and TOPEX/GFO crossover data. Evaluation of the gravity fields using SLR and crossover data support the covariance projections and also show a dramatic reduction in geographically-correlated error for the tuned fields. In this paper, we report on progress in orbit determination for GFO using GFO/GFO and TOPEX/GFO altimeter crossovers. We will discuss improvements in satellite force modeling and orbit determination strategy, which allows reduction in GFO radial orbit error from 10-15cm to better than 5 cm.

INTRODUCTION

The launch of the GEOSAT Follow-On (GFO) satellite February 10, 1998 marks the beginning of the Navy program to develop an operational series of low-cost altimeter satellites for maintaining continuous ocean observation via the GEOSAT exact repeat orbit (Table 1). GFO provides real-time measurements of the relative ocean heights for tactical applications and absolute heights post-processed for large-scale ocean modeling. Its inclination and ground-track repeat period serve to complement altimeter datasets collected by other missions such as TOPEX, ERS1 and ERS2.

GFO carries a single frequency (13.5 GHz) radar altimeter, a dual frequency water vapor radiometer, a dual frequency Doppler beacon for operational tracking, a laser retro reflector array (LRA) and four Global Positioning System (GPS) dual-frequency receivers for precision orbit determination (POD).

The measured quantity of interest, the ocean surface above the reference ellipsoid, is in fact a combination of two measurements: the ocean surface with respect to the satellite as observed by the altimeter, and the satellite height above the reference ellipsoid determined from the satellite tracking. GFO's capability to produce precise observations of the ocean surface thus depends critically on the accuracy of the orbits produced from the Doppler, SLR, or GPS tracking. GFO pre-launch analysis anticipates an accurate altimeter product (Table 2).

Since the GPS receivers delivered only limited data[1], SLR tracking has provided the only means for computing highly accurate orbits, and has been designated as the primary tracking system for GFO POD. The 5-cm radial orbit error estimate for SLR tracking shown in Table 2 was derived in a pre-launch simulation study[2] It is the Root Mean Square (RMS) error over one day.

The Space Geodesy Branch at Goddard Space Flight Center (GSFC) has been given the task of improving GFO POD. This work has included pre-flight orbit error analysis, tuning a "macro-model" of the approximate spacecraft geometry and surface properties in order to better model the nonconservative forces, tuning the gravity model, computing the SLR based Medium Precision Ephemeris (MOE) on a daily basis for use on the NAVY NGDR and NOAA IGDR altimeter products, and providing the SLR based precise ephemeris (POE) for the CalVal evaluation efforts.

This paper reviews the analysis of GFO tracking data (SLR, Doppler, altimeter crossover) and tuning of the various models. Previously, using untuned models, GFO POD radial orbit accuracy was at the 10cm level[3]. Now, thanks to improved force modeling and the use of altimeter crossover data, GFO radial accuracy is believed to be below 5 cm.

ORBIT MODELING AND ANTICIPATED ERRORS

Orbit determination can be stated as the adjustment of the orbit state, force, and measurement model parameters to minimize, in a least squares sense, the weighted difference between the actual tracking observations and their modeled values. The accuracy of the computed orbit depends on the accuracy and completeness of the force models, the measurement models, and the precision and coverage of the tracking data. GEODYN[4], a state-of-the-art least squares orbit determination and geodetic parameter recovery program, developed and maintained at GSFC, is used for GFO POD. Table 3 shows a summary of the POD models.

Several gravity fields were tested, EGM96[5], TEG3[6], JGM3[7], and PGS7609G, a GSFC combination model based on EGM96 but with additional TDRSS satellite tracking data from the EUVE, ERBS, XTE, GRO, and TRMM satellites. PGS7728 and PGS7727 are two preliminary fields, tuned using PGS7609G and GFO SLR, Doppler, and GFO/GFO altimeter crossover data. In addition to the data used for PGS7728, TOPEX/GFO crossover data was also used to tune PGS7727. Although covariance projections indicate that orbit error due to gravity will be only 1-3 cm (Table 4), the error structure will be complex, and include a geographically correlated component. By spherical harmonic order, the radial orbit error due to gravity is highest at order 1, and in the vicinity of the k=2 resonance (near order 29) (Figure 1). Tuning with GFO tracking data reduces this error

Nonconservative forces acting on GFO consist of radiative forces and atmospheric drag. Radiative forces include solar radiation pressure, the Earth's albedo (reflected light) and infrared radiation, and other secondary effects such as thermal imbalance in emission from spacecraft surfaces. Secondary effects are not modeled for GFO. The macro-model approximates GFO's surface geometry and material properties using eight plates (Figure 2). Each plate has been assigned a body-fixed orientation, area, and specular and diffuse reflectivity coefficients based on pre-launch engineering information. All plate interaction effects, such as self-shadowing and multiple reflections, are ignored. The total acceleration with respect to the center of mass (CoM) is computed by summing vectorially the force acting on each plate, taking into account each plate's area, angle of incidence and material properties. Throughout the orbit and over a Beta prime cycle, radiation will be incident to a changing orientation of the macro-model as computed using an analytical attitude model. Beta prime is the angle to the sun from the orbit plane (Figure 3), and for GFO shows a period of about 336 days.

As shown in Figure 4, the largest nonconservative force acting on GFO is by far due to solar radiation pressure. Since the solar radiation pressure is so large, even a small error will have a significant impact. The error for the macro-model should be 10 to 20 percent of the radiative force. For instance, the *a priori* macro-model for TOPEX was meticulously constructed using finite element modeling and could only account for 90% of the radiative forces[8]. However, after tuning, the TOPEX macro-model is believed to account for over 95% of the radiative forces[9]. The approach taken for the *apriori* GFO macro-model construction was much simpler and without application of finite element modeling. Even a 5% mismodeling of the solar radiation pressure would constitute a considerable source of error, requiring the adjustment of sufficient empirical acceleration parameters for reducing orbit error to an acceptable level for POD[10]. It has been shown in a previous study (Ref. 3) that given the 1999 level of SLR tracking, orbit error is due primarily to mismodeling of the radiative forces acting on the satellite. The same study (Ref. 3) also shows that following adjustment of the LRA offset to spacecraft Center of Mass (CoM), all error in the SLR measurement modeling remains very small including error in the analytical attitude model.

TRACKING DATA AND POD STRATEGY

GFO POD relies on SLR tracking provided by a global network of NASA and foreign stations (Figure 5). Operational tracking Doppler data from the three stations (Guam; Point Mugu, California; and Prospect Harbor, Maine) although noisy (2 cm/sec) is abundant, and serves to slightly strengthen the SLR solution. After 40% of the data is edited, typically nine Doppler passes per day remain. The Doppler station positions have been adjusted to the SLR frame using three months of Doppler data and SLR-determined orbits that were held fixed in the solution.

The recent increase in SLR tracking has been very welcome news for GFO POD, growing from an average of 7 passes/day for 1998 and 1999 to over 9 passes /day for 2000 and 2001. The 33% increase in tracking for 2000 is accompanied by a 22% increase in the number of tracking stations (see Table 5 and Figure 5). The long awaited site at Hartebeesthock, South Africa, became operational late last year, 2000. This new NASA site, an important addition to the SLR Network, brings better balance to the North – South tracking distribution.

Given the SLR tracking density, an arc length of five days was selected over shorter arcs to increase the dynamic strength of the solution[i]. Arc lengths of nine and ten days would also be suitable, however the frequency of satellite events over 1998 and 1999, such as computer resets or maneuvers which are not modeled for POD, have allowed only a few uninterrupted ten day spans.

Altimeter crossover data, computed by differencing altimeter ranges from two intersecting passes interpolated to a common geographic point, can be used to supplement the SLR and Doppler data for POD. Continuous altimeter tracking only began in mid-December '99. GFO crossovers provide dense spatial coverage and promise a high accuracy product [ii] . Crossovers used in POD are edited in regions which have high sea surface variability (greater than 20cm) and in shallow seas (1000m or less). It is possible to combine GFO with TOPEX altimeter data to form TOPEX/GFO altimeter crossover data (Table 6). From the well-defined and accurate TOPEX reference it may be possible to better calibrate the GFO altimeter corrections, to better tune the GFO macro-model and gravity field, and to better determine GFO orbits. The effective use of intermission altimeter crossover data to improve POD has been demonstrated for the ERS-1 and ERS-2 satellites [11] In the very least, the quantity of crossover data is increased 3-fold with the addition of TOPEX/GFO crossovers, further strengthening the orbit solution.

The solution strategy, with the objective to minimize orbit error, was developed considering the strength of the tracking data. Several parameterization schemes were tested and the one finally selected for SLR is summarized in Table 3. According to this strategy orbit error is minimized by adjusting, in addition to the orbit state, atmospheric drag scale coefficients and empirical one cycle per revolution (1cpr) accelerations for both the along-track and cross-track components. The empirical and drag terms can absorb much of the residual accelerations which remain from the mismodeling of the various forces, and greatly reduce orbit error[12, 13]. With perfect tracking data, adjustment of empirical accelerations can remove all orbit error with a sufficiently dense time series of the adjusted parameters. Imperfection in data and coverage limit the capability of empirical acceleration parameters to remove orbit error. Since the adjusted empirical acceleration terms capture information about the residual accelerations, they can also reveal characteristics of the mismodeled forces.

Sparse SLR tracking restricts the number of empirical accelerations which can be adjusted in the orbit solution, thereby limiting the capability for removing these residual accelerations. Including altimeter crossover data significantly strengthens the solution and can lead to far better orbits (Figure 6), however the GFO crossover data should be used with caution. It is believed about 40%-60% of the highly variable ionosphere refraction effect is not removed from the altimeter data with the IRI95 model. The ionosphere exhibits a high/low day/night effect. Since the GFO groundtrack repeats every 17 days and 15 minutes, the local time for a given pass will only advance by 15 minutes every 17 days. In our tests using July '99 data nearly all descending passes occurred during the day. Thus over extended periods the residual ionosphere day/night (descending pass / ascending pass) effect will directly contribute to a once per orbit revolution error (1/rev) when crossovers are used in the solution (Table 7). Nonetheless the orbit improvement gained appears to overshadow any orbit error induced with the use of crossover data (Figure 6). Thus including crossover data allows the adjustment of more empirical parameters to better remove non-conservative force model error. It is important to note that for solutions which include and which do not include crossover data, orbit accuracy remains correlated to the number of SLR points present (Figure 7). In the combination solution the presence of SLR data probably acts to constrain the effect of the ionosphere error contained in the GFO altimeter crossovers.

No single test can uniquely gauge orbit accuracy. This analysis uses SLR residuals, or the misclosure between the highly precise observed and computed ranges, altimeter crossover differences in which the non-orbit signal is believed to dominate (Table 8), and orbit differences between arcs sharing one day of overlapping data, to indicate the level of orbit error. Overlap orbit differences identify the least amount of orbit error shared by the two arcs across the overlap period, as common errors will cancel. Altimeter crossover differences represent the time-varying error component of both the altimeter measurement and orbit, as geographically correlated error is cancelled. Thus GFO/GFO crossover differences do not contain geographically correlated orbit error, but the TOPEX/GFO crossovers differences are expected to contain such orbit error for both satellites, as geographically correlated orbit error varies with satellite altitude. Since TOPEX orbit error is believed to be small, it is anticipated that TOPEX-GFO crossover data will be very useful for tuning the gravity field and gauging geographically correlated error for GFO. The GFO altimeter crossover non-orbit component estimate of 5.5 cm (Table 8) was actually derived from several tests using both GFO and TOPEX altimeter crossover data. The "non-orbit" is that component of the crossover signal, which could not be absorbed by the adjustment

---

i. In cases of sparse tracking one can usually rely on the fidelity of the dynamic force models to determine a better orbit over a longer span. Over a shorter arc, the solution may be ill determined, and the orbit error very large over periods with no data.

[ii] GFO IGDR altimeter data obtained from John Lillibridge, NOAA.

of any number of orbit parameters and must be due to ocean variability, residual ionosphere and altimeter noise. The TOPEX altimeter crossover non-orbit signal was estimated to be 4.9 cm.

MACRO-MODEL TUNING

The macro-model represents the GFO spacecraft as an eight surface composite (Figure 2). It approximates the spacecraft geometry and surface material properties to better model the surface force effects due to solar and terrestrial radiation pressure, and due to atmospheric drag. Each surface (or plate) had been assigned an orientation with respect to the satellite fixed frame, an area, and a specular and diffuse reflectivity coefficient based on pre-launch engineering specifications. The material properties of each plate are assumed to be homogenous, representing an average value. In tuning, these average values are adjusted to best fit the GFO tracking data using an orbit determination (OD) solution strategy to insure the mismodeled nonconservative forces are not absorbed in empirical parameter adjustments. Therefore the macro-model is tuned to the residual satellite acceleration history which is based on orbit errors sensed from the spacecraft tracking data.

OD parameterization suitable for macro-model tuning adjusts the orbit state, and one drag coefficient (C_D). The solar radiation pressure coefficient (C_R), which should remain constant, is fixed to a value of 1.0. Upon solution convergence, GEODYN writes out the normal equations for the orbit (state, C_D) and panel (area, specular, diffuse) parameters for each arc. These normal equations were combined from arcs sampled over the Beta prime cycle and the selected panel parameters estimated using Bayesian least squares. A preliminary sensitivity study was performed using the combined normal matrix from four well-spaced arcs to help identify panel parameters that were to be estimated. Assuming a specified allowed percent change in each respective panel parameter *a priori* value, and using only the left-hand side diagonal (variance) terms of the normal matrix, the resulting "uncorrelated weighted variance" is computed in order to compare parameter sensitivity, or change in residual variance, with respect to parameter adjustments. The *a priori* surface area assigned to each plate is believed to be relatively well determined with about a 10% error. There is much greater uncertainty for the *a priori* specular and diffuse reflectivity coefficients, computed as an aggregate average of these properties for each surface. The area is allowed to change by 10% and the reflectivity coefficients by 100% for the sensitivity analysis. As shown in Figure 8, specular coefficients for four parameters representing the solar array, the bottom plate (+z facing Earth), and the top and bottom sides of the altimeter antenna reflector, are likely candidates for the macro-model tuning adjustment.

The solar array specular reflectivity coefficient was adjusted using 31 SLR+Doppler arcs, 8 of which include Crossover data, spanning over 20 months (May 22, 1998 to February 6, 2000) or well over the 336 day Beta prime period. The preliminary tuned macro-model shows improvement in SLR fits, even for solutions adjusting empirical parameters (Table 9). Note that the *apriori* macro-model also shows improvement over the "cannonball" or spherical model, which would have been used in the absence of a macro-model (Table 9).

Even though the tuned macro-model shows improvement in SLR fits, the recovered empirical acceleration amplitudes and phases (Figures 9 and 10) are strongly correlated with Beta prime. This indicates that the solar radiation pressure still remains the largest mismodeled force, and that further tuning may be warranted.

As the absolute value of Beta prime increases from zero to 80+ degrees the solar radiation pressure (and the mismodeled effect) will change its projection from predominately along-track and radial directions to cross-track (Figure 4). The adjusted empirical accelerations should thus decrease in magnitude in the along-track component (Figures 9 and 10), and increase in the cross-track. The associated phase (with respect to orbit angle) will remain constant from arc to arc until the spacecraft enters the full sunlight regime. The observed phase coherence (Figure 10) indicates that the force error preserves the same orientation with respect to orbit plane from arc to arc, which in fact solar radiation pressure does. As the spacecraft reaches full sunlight (near |65°| Beta prime), the recovered along-track acceleration magnitude becomes very small and for which the phase is not well determined. The along-track acceleration changes phase between increasing/decreasing Beta prime (Figure 10). In another study tuning the TDRSS macro-model[14], a continuous phase was also observed in the recovered 1cpr along-track acceleration prior to tuning. After tuning, the recovered acceleration magnitudes were small and the phases showed no coherence.

GRAVITY MODEL TUNING

Two preliminary gravity models were determined using SLR, Doppler, and GFO/GFO and TOPEX/GFO altimeter crossover data. Twenty arcs spanning one year from June 1999 to June of 2000 were included, whose arc length ranged from two to ten days (the average arc length was 6.8 days). An average of 5967 TOPEX/GFO altimeter crossovers and 2728 GFO/GFO altimeter were included in each arc. The average RMS of fit over these 20 arcs prior to gravity tuning was 7.07 cm for the TOPEX/GFO crossover data, 7.71 cm for the GFO/GFO crossover data, and 5.68 cm for the SLR data. In the solution derivation, the TOPEX

orbit was first determined using SLR and DORIS data, and then held fixed in both the GFO orbit adjustment and gravity field tuning. Two solutions were estimated: PGS7728, which included the GFO SLR, Doppler and GFO/GFO altimeter crossover data, and PGS7727, which included these data, as well as the TOPEX/GFO crossover data. The entire set of GFO data was calibrated using the Lerch method of subset calibrations[15]. An overall calibration factor of 0.78 was obtained, indicating the data were conservatively weighted in the tuned solutions. The radial orbit error due to the geopotential projected from the two new solutions to 70x70 show a significant improvement in orbit accuracy, particularly at order 1, and near order 29 (see Figure 1). The total predicted radial orbit error due to the geopotential on the GFO orbit is 1.31 cm with PGS7727. Looking at several gravity fields, the SLR and crossover data RMS of fit for several test arcs, show a marked improvement for the tuned fields (Table 10)

Since the TOPEX geographically correlated orbit error is believed to be less than 1 cm (Ref. 10), it may be possible to differentiate the GFO geographically correlated error by geographically projecting the TOPEX-GFO altimeter crossover residuals. Indeed, after averaging crossover residuals from five 10-day arcs over 5° x 5° bins, geographic structure becomes apparent for PGS7609G and to a lesser degree for PGS7728 and PGS7727 (Figure 12). Note the large peak-to-peak amplitude of ± 8 cm for the C(2,2) type signal present for PGS7609G, is progressively reduced to a peak-to-peak amplitude of only ± 3 cm . The introduction of the TOPEX/GFO altimeter data appears to have largely removed this C(2,2) effect. Orbit difference projections with respect to PGS7727 orbits (Figure 11) correspond to the improvements shown with the TOPEX-GFO crossover differences (Figure 12). Notice the difference in scale for Figure 11a and Figure 11b. Figure 12 illustrates a significant reduction in geographically correlated orbit error. It is likely that the geographically correlated error for PGS7727 is indeed close to 1 cm (Table 4).

Orbit error can be estimated from altimeter crossover fits. For example, PGS7727 shows a GFO crossover fit of 7.6 cm (Table 10). Assuming 5.5 cm of this signal is not due to orbit error (Table 8), leaves 5.2 cm representing the time-varying orbit error combined from the ascending and descending pass differences. Assuming the time-varying error is evenly distributed between ascending and descending passes and assuming a geographically correlated error component (1.2 cm for PGS7727 from Table 4) will give us a total radial error estimate:

$$\text{radial error estimate using crossovers} = \sqrt{(crossover^2 - non\_orbit^2)/2 + geographically\_correlated^2} \quad (1)$$

where

crossover	is the total altimeter crossover difference RMS value
non_orbit	is the non-orbit component of the altimeter crossover difference value
geographically_correlated	is the geographically correlated radial orbit error

Thus from the GFO crossover fit one may estimate PGS7727 orbits have a radial error of 3.9 cm, and that the orbit error probably remains dominated by non-conservative force mismodeling (Table 4).

REDUCED DYNAMIC APPROACH

POD can benefit from the spatially dense altimeter crossover coverage by adjusting more empirical parameters. According to perturbation theory, most orbit error due to force mismodeling is of a resonant nature. The linearized equations of satellite motion, Hill's equations, further suggest that the orbit acts as a narrow bandwidth filter, smoothing the effects of complex acceleration perturbations, and that with the adjustment of nine parameters (A_R, B_R, C_R, A_L, B_L, C_L, A_C, B_C, C_C) together with the state, orbit error will be dramatically reduced (Ref 12, 13).

$$\text{Acceleration}_R = A_R \cos(\omega t) + B_R \sin(\omega t) + C_R \quad \text{(Radial)}$$
$$\text{Acceleration}_L = A_L \cos(\omega t) + B_L \sin(\omega t) + C_L \quad \text{(Along-track)}$$
$$\text{Acceleration}_C = A_C \cos(\omega t) + B_C \sin(\omega t) + C_C \quad \text{(Cross-track)}$$

or $\text{Acceleration} = \text{Amplitude} \cos(\omega t + \text{phase}) + \text{Constant}$

where

A's and B's	represent the adjusted Sine and Cosine amplitude terms (or 1cpr acceleration amplitude / phase terms)
C's	represent the adjusted constant acceleration terms
ω	is the orbital frequency
t	is the time.

Our dynamic solution strategy calls for the adjustment of four parameters - the along-track and cross-track 1cpr terms. GFO tracking data cannot support the adjustment of Hill's entire set of nine empirical parameters. Even the adjustment of the four 1cpr

parameters over an arc dramatically reduces orbit error to the decimeter level (MOE dynamic strategy). At the 10-cm level, orbit error will largely consist of 1/rev terms and include the non-resonant terms. With the inclusion of altimeter crossover data it becomes possible to adjust the 1cpr acceleration parameters more frequently - over each day of the arc. This will further remove 1cpr orbit error as well as some daily and modulated signals, reducing the radial error to about 5-cm (POE dynamic strategy).

Our "reduced dynamic" approach, can reduce all orbit error signal, by allowing a more frequent adjustment of the 1cpr acceleration amplitude and phase parameters, roughly approximating a time varying empirical acceleration of the form:

$$\text{Acceleration}(t) = \text{Amplitude}(t)\cos(\omega t + \text{phase}(t))$$

Deficiencies in the tracking data are accommodated by suitably constraining the parameter adjustments. The success of a reduced dynamic approach depends on precision and density of the tracking data and on the accuracy of the dynamic force models, especially the gravity field. The GEODYN "reduced dynamic" implementation has been shown to improve the Space Shuttle precision orbit [16]. A subset of the GEODYN implementation is through the least-squares adjustment of a time series of 1cpr empirical acceleration parameters, which have explicitly correlated constraining equations forcing greater continuity between the adjacent 1cpr amplitude and phase terms. For any one such parameter P_j, the closer in time it is to another parameter P_k, the tighter is the constraint forcing both P_j and P_k to adjust to the same value. The weight used in the constraint equation between two parameters at time T_j and at time T_k, is computed in GEODYN as follows:

$$\text{weight}(j,k) = (e/\text{sigma}^2)\, e^{-(|Tj - Tk|/\text{correlation time})}$$

where

T_j	is the the mid-point of the jth acceleration parameter interval
sigma	is the process noise input by user
correlation time	is the correlation time input by user
e	is the base of the natural logarithms (2.718 ...)

The acceleration parameter interval, correlation time, and sigma, selected from preliminary empirical tests (tuning), are 25 minutes, 12.5 minutes, and 1.e-9 m/s**2 respectfully. Such a reduced dynamic approach processing SLR, Doppler and GFO/GFO crossover data offers some improvement over the standard dynamic approach (Table11). The improvement is especially significant for the cross-track and along-track components as evidenced by the improved orbit overlap consistency (Table 11). Orbit error cannot be less than the orbit overlap differences, which indicate the radial error cannot be less than 1.7 cm. From the GFO crossover fits (Table 11), a radial error of 4.7 cm for the dynamic and 4.2 cm for the reduced-dynamic solutions, is estimated (Equation 1).

In these tests the TOPEX/GFO crossovers are used as independent data to measure the orbit accuracy. Radial orbit error may be estimated from TOPEX/GFO crossovers once the magnitude of the altimeter+TOPEX orbit component is identified in the crossover fits. Using a high and a low estimate for this component (Table 12), GFO radial orbit error is estimated to be between 4.6 cm – 5.6 cm. for the dynamic (7.9 cm crossover RMS from Table 11) and between 4.3 cm – 5.3 cm for the reduced dynamic solutions (Table 11). From these tests it appears that altimeter crossover data constrain the radial orbit component quite well for the dynamic solutions. However the dynamic orbit retains significant nonconservative force model error. This error may be further reduced with a reduced-dynamic strategy given adequate tracking. POD may be further improved with the inclusion of TOPEX/GFO crossover data, and with the application of other "tuning" combinations.

SUMMARY

The Space Geodesy Branch at Goddard Space Flight Center (GSFC) has been given the task of improving GFO precision orbit determination. Since GPS receivers delivered only very limited tracking data, SLR tracking offers the only means with which to compute precise orbits for this spacecraft. SLR data in combination with altimeter crossover data was successfully used to tune the preliminary macro-model and gravity field. Use of intermission altimeter crossover data (TOPEX/GFO) significantly contributed to reducing geographically correlated error for the tuned gravity field. The radial orbit accuracy is estimated to be 4-5 cm using the latest models and including altimeter crossover data in the orbit solution.

With the termination of the CalVal phase, better altimeter corrections have become available which will favorably impact the GFO orbit determination. For example, the estimate of the sea state (EM) bias has been refined, and the wet troposphere correction from the on-board radiometer has been validated. In addition, improved tidal solutions have been derived primarily from TOPEX altimeter data (eg GOT00.2) and will be applied in future work. Use of TOPEX/GFO crossover data is anticipated to further help GFO macromodel tuning and, in combination with the reduced-dynamic approach, to further help GFO precision orbit determination. Future work should include a review of the nonconservative force model, and whether more detailed modeling is warranted, as for example including shadowing and thermal radiation effects.

TABLES AND FIGURES

Table 1 GEOSAT Exact Repeat Orbit

Orbit parameter	value
Altitude	800 km
Eccentricity	0.0008
Inclination	108 deg
Repeat Period	244 revs in 17 days

Table 2 GFO Pre-Launch Altimeter Error Budget [iii]

Component	Source	Error (cm)
Altimeter		
instrument noise	Ball	1.9
biases	Ball	3.0
sea surface (EM & skewness)	TOPEX	2.3
Media		
troposphere	Ball	2.0
ionosphere	Ball	1.7
SLR POD (radial orbit)	GSFC	5.0
Total RSS		7.1

Table 3 GFO Precise Orbit Determination Modeling

Model Category	Description
Geophysical models	
Gravity	PGS7727 (PGS7609G + GFO SLR, Altimeter Crossover, and Doppler tracking data)
Ocean/Earth Tides	PGS7723C [i] resonant + Ray '99 background terms [17]
Atmospheric density	MSIS-86 [18]
Spacecraft geometry and surface forces	GFO Preliminary tuned macro-model
Station Coordinates	CSR95L02 SLR solution frame[ii]
Earth Orientation Parameters	CSR95L02 from LAGEOS tracking
Planetary Ephemeris	DE403
Measurement Model	
SLR	*A priori* CoM, estimated LRA offset, analytical attitude
Doppler	*A priori* CoM, *a priori* beacon offset, analytical attitude
Altimeter Crossover	*A priori* CoM, analytical attitude, GEODYN Dynamic Crossover model
Tracking Data Weights	
SLR	10 cm
Doppler	2 cm/sec
Altimeter Crossover	10 cm
Estimated Parameters	Orbit state, Atmospheric drag C_D per day (or more frequently data permitting) Along-track 1cpr empirical acceleration per arc Cross-track 1cpr empirical acceleration per arc Doppler measurement and troposphere bias per pass

[i] PGS7723C is a preliminary field determined from PGS7609G + GFO SLR/Doppler and GFO/GFO Altimeter Crossover data.

ii. CSR95L02 is the SLR station position and velocity frame used to compute the TOPEX/POSEIDON precise orbits, Richard Eanes, CSR, 1995.

iii. An official GFO altimeter system error budget has yet to be published. The values shown here have been compiled from an internal Ball document provided by Scott Mitchell, from the T/P Mission Plan, and error simulations performed at GSFC (Ref 2).

Table 4. Gravity Orbit Error Covariance (to 70x70) Projection

Gravity Field	GFO Orbit Error (cm)			
	geographically correlated Radial	Radial	Cross-Track	Along-Track
JGM3	4.53	4.97	23.80	42.61
TEG3	3.30	3.48	21.42	42.76
EGM96	2.35	2.61	8.94	17.72
PGS7609G	2.35	2.61	8.93	16.44
PGS7728	1.49	1.66	8.57	14.84
PGS7727	1.16	1.31	8.42	14.40

Table 5. SLR Tracking Summary

Year	Days	Stations	Passes
1998	254	33	1829
1999	365	32	2625
2000	366	39	3485
2001	59	27	543

Table 6. Altimeter Range Modeling for TOPEX/GFO Crossover Processing

Model	TOPEX[i]	GFO[ii]
Ocean tide + tidal loading	CSR 3.0 (GDR)	same (IGDR)
Earth Tide	Cartwright (GDR)	same (IGDR)
Dry troposphere	FMO (GDR)	NCEP (IGDR)
Wet troposphere	TMR (GDR)	NCEP[iii] (IGDR)
Ionosphere	dual frequency (GDR)	IRI95 (IGDR)
Inverse barometer	f(dry troposphere)	same f function
EM bias	Walsh (GDR)	5% SWH
Orbit	fixed [iv]	adjusted
Range bias	fixed [v]	adjusted
Timing bias	fixed (GDR time tag)	adjusted

Table 7 POD Sensitivity to Altimeter Ionosphere Error (July 28 – August 6 1999 data)

Satellite and (Beta prime angle) [a]	ionosphere altimeter correction	SLR + Doppler Orbit held fixed		SLR+Doppler+altimeter crossover Orbit adjusted	
		crossover rms (cm)	variance reduction (cm^2)	crossover rms (cm)	radial orbit difference(cm)
TOPEX	none	7.46		6.75	
(-30 °)	measured	5.97	20.00	5.77	1.50
GFO	none	9.04		6.81	
(-73 °)	IRI95	8.80	4.28	6.75	0.82

a) the largest ionosphere corrections are anticipated for Beta prime close to zero

[i] TOPEX GDR
[ii] GFO IGDR from NOAA
[iii] TMR value also available
[iv] POE has 2-3 cm radial accuracy
[v] routinely computed per cycle by PODPS based on the POE orbit

Table 8. GFO/GFO Altimeter Crossover Error Budget

Error Source	time varying error (cm)	
	range	crossover difference
Non-orbit		
altimeter	3.5	5.0
ionosphere	1.7	2.4
Orbit	3.7	5.2
RSS total		7.6

Table 9. GFO Macro model Tuning

spacecraft surface model	solar array (SA) reflectivity coefficient	SLR fits over 23 dependant arcs (cm)	SLR fits over 57 independent arcs (cm)	SLR fits over 80 arcs total [b] (cm)
cannonball	----	13.23	12.88	12.99
a-priori macro model	.160	13.11	12.89	12.95
tuned SA macro model [a]	.144	13.04	12.80	12.87

a) tuned using 23 SLR+Doppler and 8 SLR+Doppler +Crossover arcs spanning 980522 – 000206
b) 80 consecutive arcs spanning 980422 – 990603

Table 10. Gravity Field Tests

gravity field	radial orbit error projected from 70x70 gravity covariance (cm)	data RMS (cm) combined results over five 10-day arcs			
		TP crossover	TP/GFO crossover	GFO crossover	GFO SLR
JGM3	4.97	6.17	8.45	8.51	7.42
TEG3	3.48	6.20	7.31	7.95	6.89
EGM96	2.61	6.14	7.71	8.27	6.97
PGS7609G	2.61	6.16	7.74	8.26	6.75
PGS7728	1.66	6.14	7.17	7.68	5.64
PGS7727	1.31	6.13	7.02	7.59	5.53

Table 11. GFO Orbit Solution Strategy

Solution Strategy	TOPEX/GFO Crossover RMS (cm) (independent data)	GFO tracking data RMS *combined results over seven 5-day arcs (Jan 17 '00 – Feb 13 '00)*			orbit overlap difference RMS (cm)		
		Doppler (cm/sec)	SLR (cm)	GFO Crossover (cm)	radial	cross track	along track
moe (dynamic) 1 drag/24 hours, 1 cpr/5 days	11.2	2.1	31.1	12.8	6.7	23.0	83.0
poe (dynamic) 1 drag/08 hours, 1 cpr/1 days	7.9	2.0	5.2	8.5	2.3	24.5	23.2
reduced dynamic 1 drag/24 hours, 1 cpr/25 min	7.7	2.0	4.1	7.9	2.4	9.0	16.1

Table 12. TOPEX/GFO Altimeter Crossover Error Budget

Error Source	High GFO orbit estimate (cm)			Low GFO orbit estimate (cm)		
	TP range	GFO range	crossover	TP range	GFO range	crossover
altimeter time varying	3.5	3.5	4.9	3.5	3.5	4.9
ionosphere	0	1.7 [a]	1.7	0	2.8 [b]	2.8
orbit	2.0	5.3	5.7	3.0	4.3	5.2
RSS total			7.7			7.7
RSS altimeter + TP orbit			5.6			6.4

a. Ball estimate

b. RSS difference(measured ionosphere for TP – IRI95 modeled for GFO; see Table 7)

Figure 1. GFO Gravity Orbit Error Covariance (to 70x70) Projection

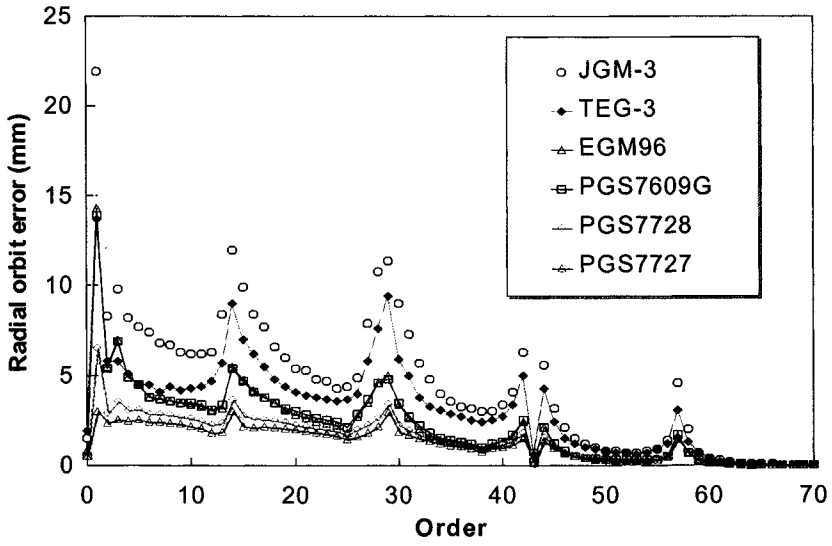
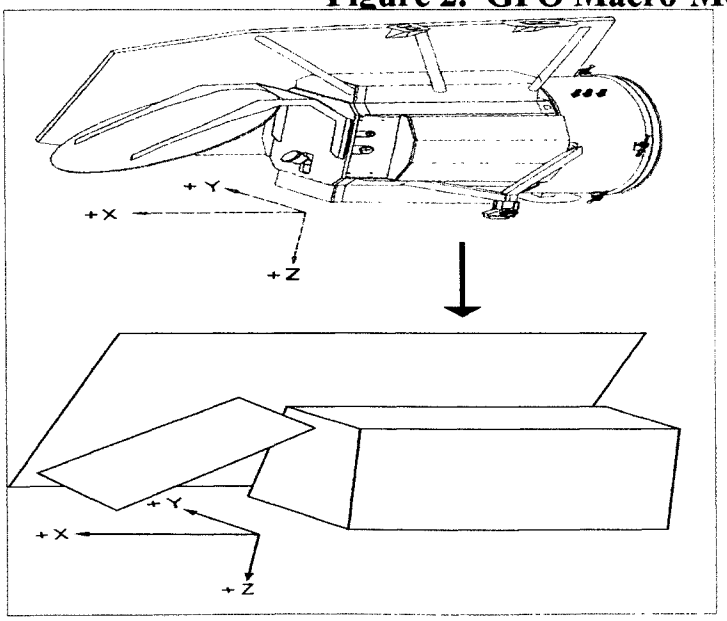


Figure 2. GFO Macro-Model Approximation



Acceleration due to radiation pressure on a flat plate:

$$\Gamma = -\frac{\Phi A \cos\theta}{Mc}\left[2\left(\delta/3+\rho\cos\theta\right)\mathbf{n}+(1-\rho)\,\mathbf{s}\right]$$

where

Γ	=	acceleration (m/s^2)	
Φ	=	radiation flux from source	
A	=	surface area of flat plate (m^2)	*
θ	=	incidence angle	
M	=	satellite mass (m)	
c	=	speed of light (m/s)	
δ	=	diffuse reflectivity	*
ρ	=	specular reflectivity	*
n	=	surface normal unit vector	
s	=	source incidence unit vector	

* are the adjustable macro model parameters

Figure 3. Orbit geometry

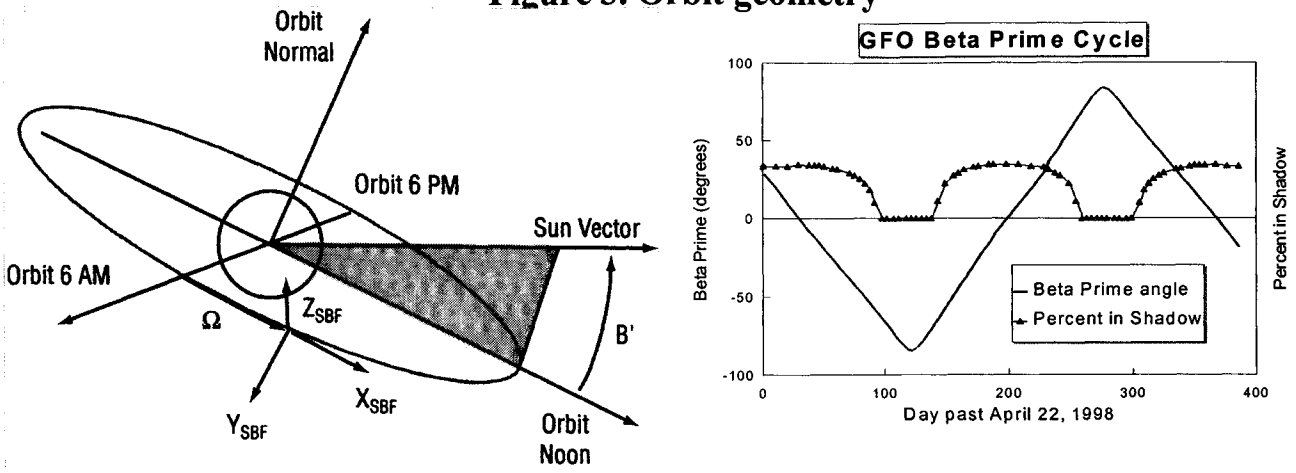


Figure 4. Non-conservative forces acting on GFO

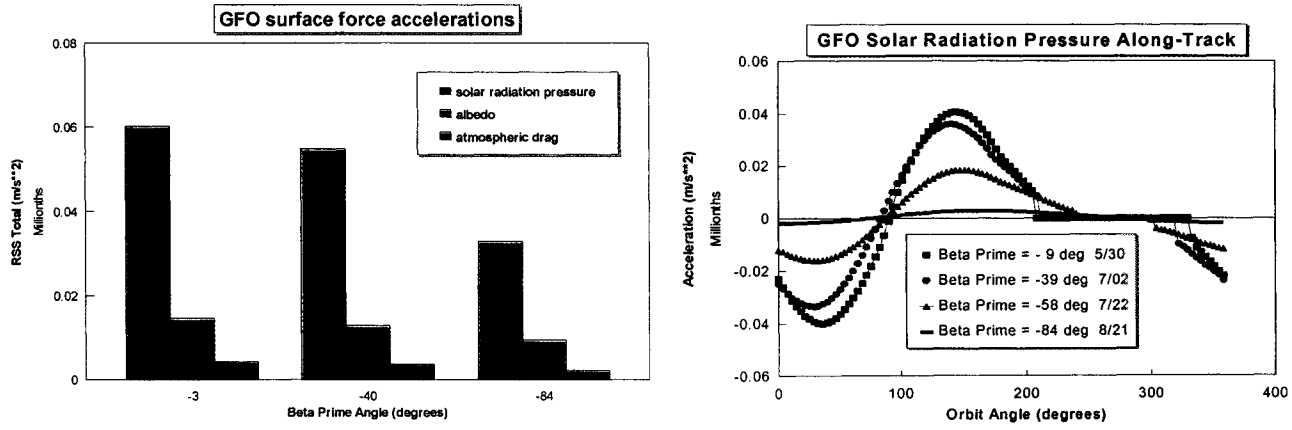


Figure 5. GFO Tracking Network

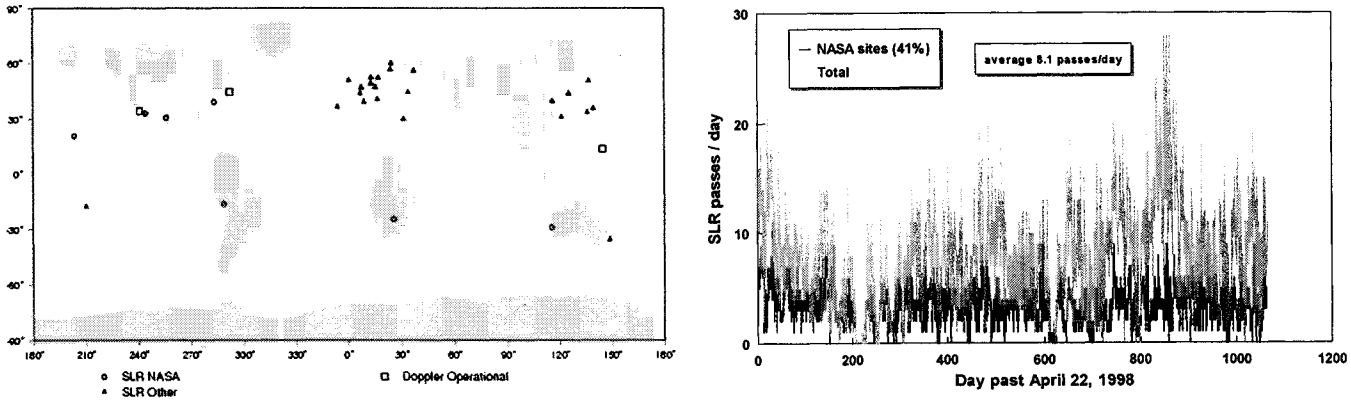


Figure 6. GFO Orbit Solution Strategies

nominal: 1drag/day, 1crp/5day; enhanced: 3drag/day, 1cpr/1day

Combined nine arcs spanning Jan 6 –Feb 13 2000

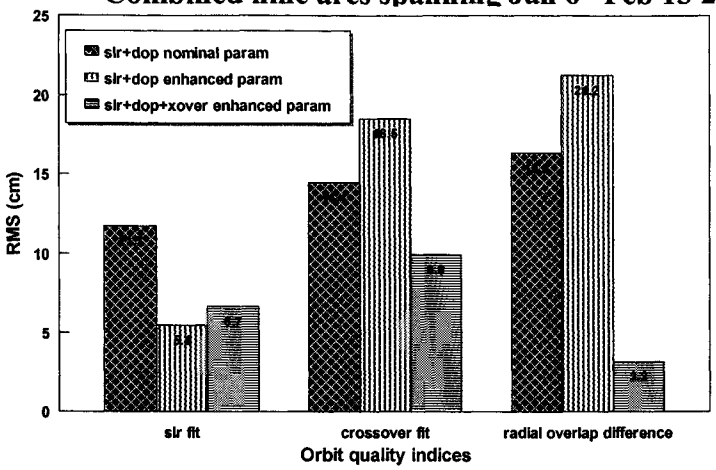


Figure 7. Orbit accuracy limited by sparse SLR tracking

GFO-GFO Altimeter Crossover Differences

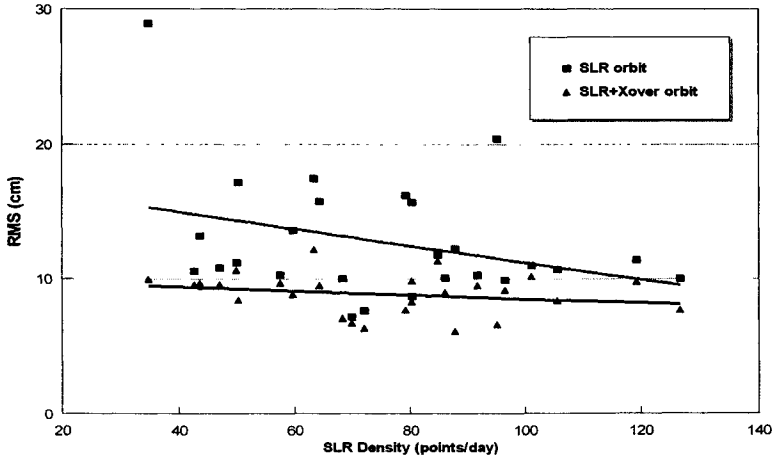


Figure 8. GFO macro-model parameter sensitivity

Uncorrelated Weighted Variance

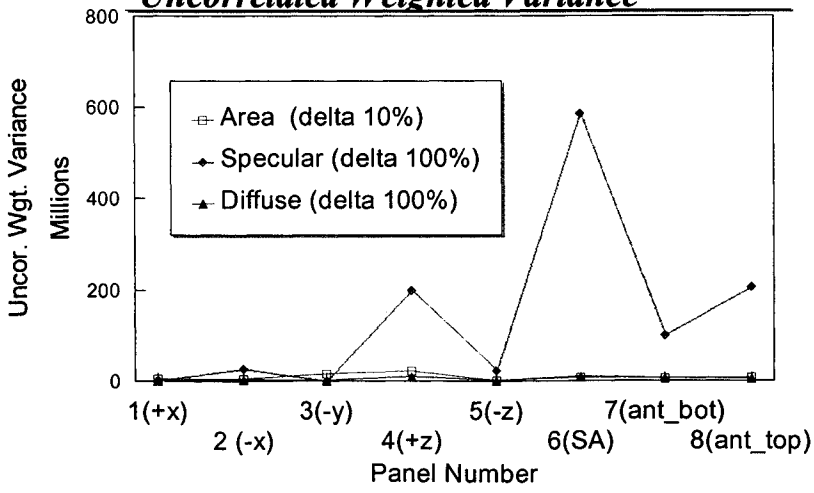


Figure 9. Recovered Empirical Accelerations vary with Beta'

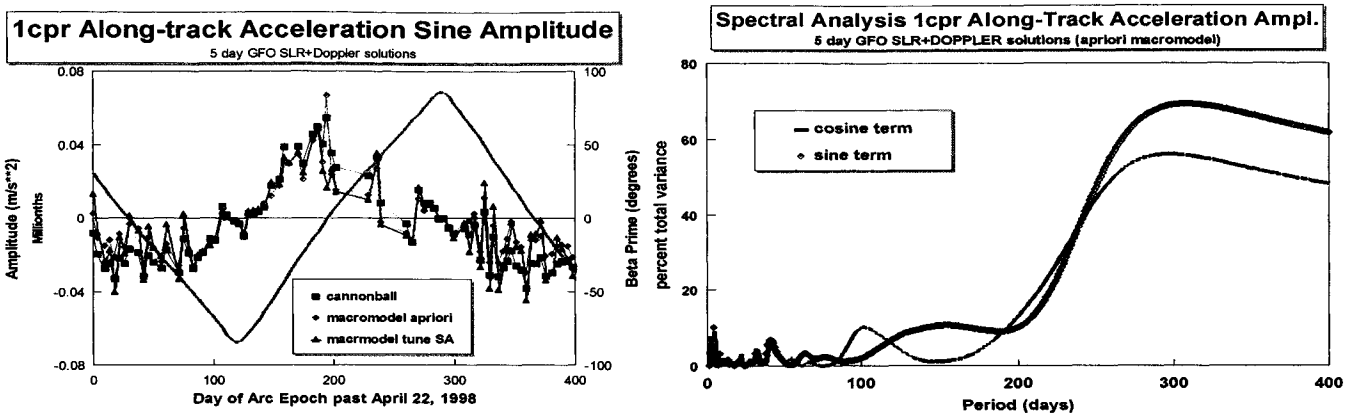


Figure 10. Empirical Acceleration Amplitude and Phase retain strong Beta' signal

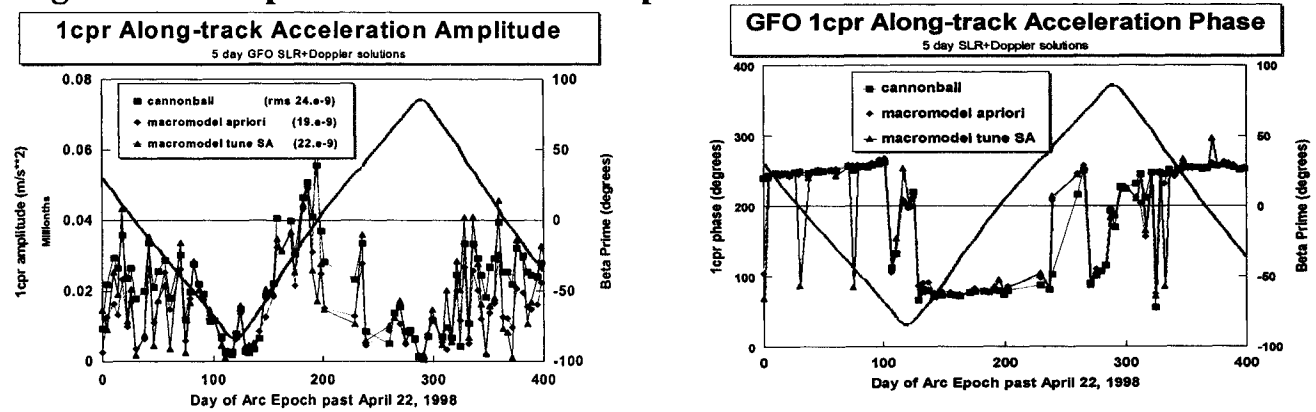


Figure 11. GFO geographically projected orbit difference

PGS7727 - PGS7609G radial orbit difference (cm)

PGS7727 - PGS7728 radial orbit difference (cm)

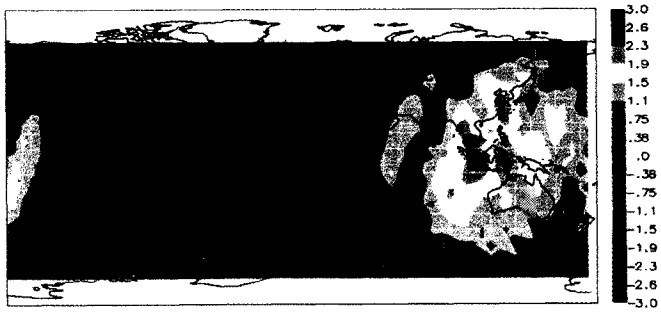
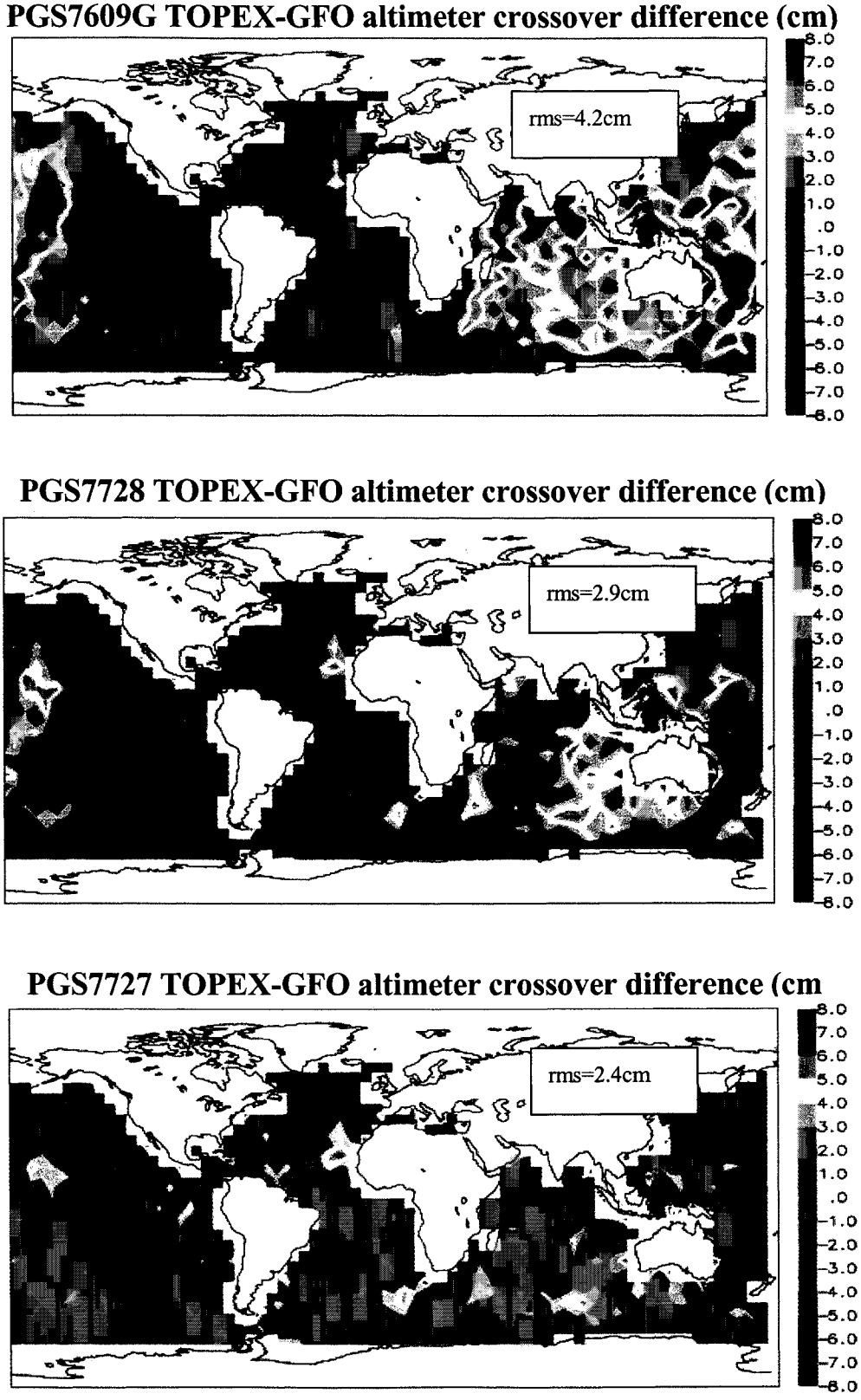


Figure 12. GFO geographically correlated orbit error



REFERENCES

1. Frazier W., Mitchell S., Weiss M., Wiener D., "Initialization and Early On-Orbit Performance of the Geosat Follow-On Satellite", *22nd AAS Guidance and Control Conference*, Feb '99, Breckenridge, Colorado, AAS 99-072.
2. Zelensky N.P., Luthcke S.B., Gehrmann L., Rowlands D.D., Marshall J.A., Lemoine F.G., "Error Analysis of the GEOSAT Follow-On Satellite Orbit Determined Using SLR and GPS Tracking (abstract)," *Annales Geophysicae EGS XXII General Assembly*, Vol 15, suppl. 1, pp. C187, April 1997.
3. F.G. Lemoine, N.P. Zelensky, D.D. Rowlands, G.C. Marr, S.B. Luthcke, C.M. Cox, "Precise Orbit Determination for the GEOSAT Follow-On Spacecraft", 1999 Flight Mechanics Symposium Proceedings, NASA GSFC NASA/CP-1999-209235, pp. 495-507, May 1999
4. Pavlis, D.E., S. G. Poulose, S. C. Rowton, J. J. McCarthy, and S. B. Luthcke, GEODYN operations manuals, Raytheon ITSS contractor report, March 15, 2000.
5. F.G. Lemoine, S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, and T.R. Olson, *The Development of the Joint NASA GSFC and NIMA Geopotential Model EGM96*, NASA/TP-1998-206861, July, 1998, NASA Goddard space Flight Center, Greenbelt, MD.
6. B.D. Tapley, C.K. Shum, J.C. Ries, S.R. Poole, P.A.M. Abusali, S.V. Bettadpur, R.J. Eanes, M.C. Kim, H.J. Rim, and B.E. Schutz, "The TEG-3 geopotential model," in *Gravity, Geoid, and Marine Geodesy* J. Segawa, H. Fujimoto, and S. Okubo (eds.), Vol 117, *International Association of Geodesy Symposia*, Springer-Verlag, Berlin, 453-460, 1997.
7. B.D. Tapley, M.M. Watkins, J.C. Ries, G.W. Davis, R.J. Eanes, R. Poole, H.J. Rim, B.E. Schutz, C.K. Shum, R.S. Nerem, F.J. Lerch, J.A. Marshall, S.M. Klosko, N.K. Pavlis, and R.G. Williamson, The Joint Gravity Model-3, *J. Geophys. Res.*, 101(B12), 28029-28049, 1996.
8. J.A. Marshall and S.B. Luthcke "Modeling Radiation Forces Acting on TOPEX/POSEIDON for Precision Orbit Determination," *Journal of Spacecraft and Rockets*, Vol 31, No. 1, 1994, pp. 99-105.
9. J.A. Marshall and S.B. Luthcke "Radiative Force Model Performance for TOPEX/POSEIDON for Precision Orbit Determination," *Journal of Astronomical Sciences*, Vol. 42, No. 2, 1994, pp. 229-246.
10. J.A. Marshall, N.P. Zelensky, S.M. Klosko, D.S. Chinn, S.B.Luthcke, K.E.Rachlin, R.G. Williamson, "The temporal and spatial characteristics of TOPEX/POSEIDON radial orbit eror," JGR, Vol. 100, No. C12, pp. 25331-25252, Dec. 1995.
11. R. Scharroo, P. Visser, "Precise orbit determination and gravity field improvement for the ERS satellites," *J Geodhys. Res.*, vol 103, no. C4, pp 8113-8127, April 15, 1998.
12. Colombo, O.L., "Ephemeris errors of GPS satellites," *Bull. Geod.*, 60, 64-84, 1986.
13. Cretaux, J, F Nouel, C Valorge, P Janniere, "Introduction of emperical parameters deduced from the Hill's equations for satellite orbit determination," Manuscripta Geod., 19, 135-156, 1994.
14. S.B. Luthcke, J.A. Marshall, S.C. Rowton, K.E. Rachlin, C.M. Cox, R.G. Williamson, "Enhanced Radiative Force Modeling of the Tracking and Data Relay satellites", *JAS* Vol. 45, No. 3, July-September 1997, pp. 349-370
15. F. J. Lerch, R. S. Nerem, D. .S. Chinn, J.C. Chann, G. B. Patel, and S. M. Klosko, "New error calibration tests for gravity models using subset solutions with independent data: Applied to GEM-T3," *Geophys. Res. Lett.*, 20(2), 249-252, 1993a.
16. D. D. Rowlands, S.B. Luthcke, J.A. Marshall, C.M. Cox, R.G. Williamson, S.C. Rowton "Space Shuttle Precision Orbit Determination in Support of SLA-1 Using TDRSS and GPS Tracking Data," *JAS* Vol. 45, No. 3, July-September 1997, pp. 349-370
17. Ray, R.D, "A Global Ocean Tide Model from TOPEX/POSEIDON Altimetry: GOT99.2," NASA/TM-1999-209478, NASA/GSFC Sep 1999.
18. Hedin, A.E., "The atmosphere model in the region 90 to 200 km," *Adv. Space Res.*, 8(5), 9-25, 1988

TRACKING AND DATA RELAY SATELLITE (TDRS-3) RANGE BIASES AND MOMENTUM UNLOAD MODELING FOR TERRA (EOS-AM1)*

Douglas T. Ward
Computer Sciences Corporation

ABSTRACT

The Flight Dynamics Facility (FDF) reports its performance in meeting Tracking and Data Relay Satellite (TDRS) predicted ephemeris accuracy requirements with TDRS-3. The Terra (Earth Observing System AM-1) satellite has 3σ TDRS requirements of 75 meters for total position accuracy predicted over 1 day onboard. The study sample includes selected cases over 21 months after Guam Remote Ground Terminal (GRGT) support started in June 1998. For daily solutions with a 1.5-day prediction span, predicted results of the study were below the Terra requirement by at least 12 meters.

Refined range bias estimation and modeled momentum unloads are needed to meet Terra's requirements for TDRS-3. Maintained at 275 degrees west longitude over the zone of exclusion, TDRS-3 is analyzed separately from other TDRSs because of its unique tracking data. Only the Bilateration Ranging Transponder (BRT) at Alice Springs (ALS), Australia, and the Telemetry, Tracking and Command (TT&C) system at Guam are used for routine operational tracking data for TDRS-3. Simultaneous batch orbit solutions with three TDRSs and either the Compton Gamma Ray Observatory (GRO) or Terra were done with the Goddard Trajectory Determination System (GTDS) to periodically refine the TT&C and BRT System (BRTS) range biases. As new biases were determined, significant changes were made in estimating the absolute position. FDF achieved similar results using a sequential filter with all operational TDRSs and four user satellites. Definitive accuracy (3σ) is expected to be below 50 meters.

The White Sands Complex (WSC) performs momentum unloads to maintain three-axis stabilized attitude of TDRSs. The relationship between velocity changes (delta-V) and reaction wheel speed changes was empirically determined for roll/yaw unloads. A theoretical relationship was verified and used for pitch unloads. Modeling both pitch and roll/yaw momentum unloads is necessary to meet the 75-meter requirement. Moving the orbit solution epoch an hour before a momentum unload can improve delta-V optimization and prediction accuracy over 1.5 days.

INTRODUCTION

In this paper, the Flight Dynamics Facility (FDF) at the Goddard Space Flight Center reports performance in meeting Tracking and Data Relay Satellite (TDRS) predicted ephemeris accuracy requirements with TDRS-3 for the Terra (Earth Observing System AM-1) satellite. The 3σ requirements are 75 meters total positional accuracy and 5.5 millimeters per second total velocity accuracy predicted over 1 day onboard (Reference 1).

TDRS-3 is maintained at 275 degrees west longitude over the TDRS System (TDRSS) zone of exclusion (excl.). Figure 1 is a map of TDRS longitudes and ground tracking and relay sites. TDRS-3 is analyzed separately from other TDRSs because of its unique tracking data. Since June 1998, only the Bilateration Ranging Transponder (BRT) at Alice Springs (ALS), Australia, and the Telemetry, Tracking and Command (TT&C) system at Guam are routinely used for operational tracking data for TDRS-3. While other operational TDRSs track BRTs at two sites, only one BRT and the Guam Remote Ground Terminal (GRGT, referred to here as Guam) are used for TDRS-3 tracking data. This paper describes extra analysis done to assess TDRS-3 range biases for both tracking systems. Although the desired consistencies were achieved for TDRS-3 with the initial range biases, significant changes in estimating the absolute position were made as new biases were determined.

* This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under Contract NAS 9-98100.

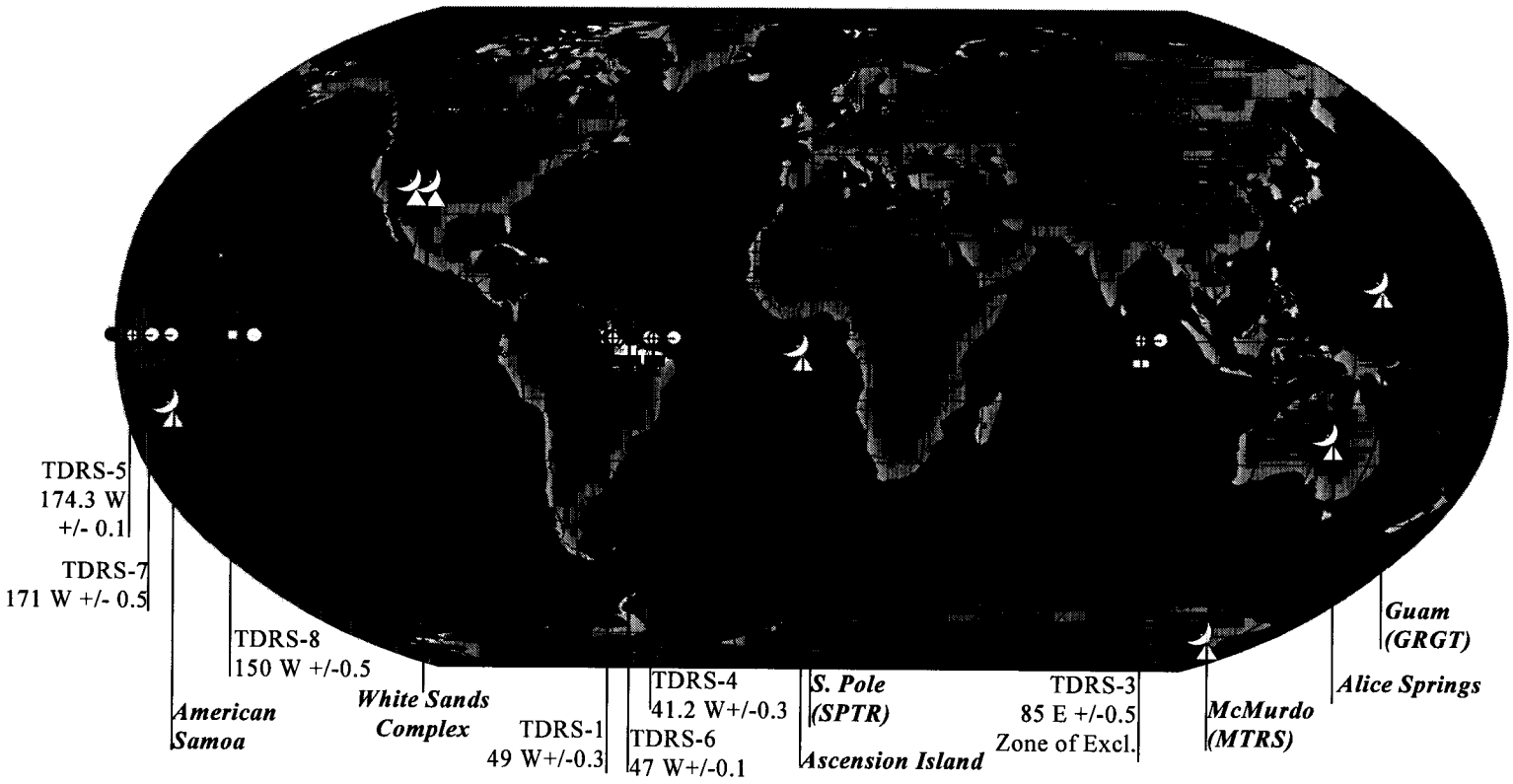


Figure 1. Map of TDRS Longitudes and Ground Tracking and Relay Sites in July 2000

In December 1999, Reference 2 reported progress toward meeting these requirements by modeling both pitch and roll/yaw momentum unloads for all other operational TDRSs. This TDRS-3 study includes results of modeling both pitch and roll/yaw momentum unloads.

The first section presents background information on modeling, requirements, accuracy, and momentum unloads. Then methods and results of orbit determination, including range bias estimation, are discussed, followed by methods and results of momentum unload modeling. Finally, a summary and recommendations are given.

1. BACKGROUND

TDRS-3 Characteristics and Modeling

Table 1 gives the current TDRS-3 location and function status, stationkeeping longitude and box size, and the inclination on April 29, 1998. The current inclination is listed at http://mmfd.gsfc.nasa.gov.

Table 1. TDRS-3 Information

TDRS Status	Position		Inclination
	Longitude (degrees)	Box (+/-degrees)	on April 29, 1998 (degrees)
Zone of Exclusion	275 West (85 East)	0.5	3.8

The standard modeling used for TDRS-3 during this study, which includes modeling both pitch and roll/yaw momentum unloads, is listed in Table 2. The **boldface** parts are different from what was used for other TDRSs in Reference 2, which describes the first known delta-V calibration of TDRS roll/yaw unloads. As recommended in Reference 2, operational TDRS solutions are being updated on a daily schedule.

Relationship between Position and Velocity Requirements

The first momentum unload study by FDF considered only pitch unloads and showed that, when the positional accuracy requirement (75 meters, 3σ) was met, the velocity accuracy requirement (5.5 millimeters per second, 3σ) was also met (Reference 4). Therefore, this study only addresses FDF's ability to meet the positional accuracy requirement.

Definitive Accuracy

Definitive solutions are used as a baseline from which to assess the predicted errors of ephemerides. Therefore, the definitive accuracy of a solution must be estimated before the measured predicted accuracy can be interpreted as an absolute accuracy.

For the current study, the average definitive consistency was computed from the maximum position differences between overlapping definitive ephemerides over 1/2 orbit (12 hours), at the typical end of overlapping tracking data. Results were adjusted to remove the effect of updates to the difference of Universal Time based on earth's rotation (UT1) and Coordinated Universal Time (UTC). Because of the uncertainty of the TDRS-3 range biases, the definitive consistency may be small (10 meters) while the definitive accuracy could have a large (100-meter) constant offset. The definitive accuracy is also addressed in the range bias estimation part of Section 2.

Table 2. TDRS-3 Standard Modeling Parameters

Parameter	Value
Data arc length	**1 day + 18 hours (42 hours)**
Geopotential model	75x75 JGM-2 truncated to 8x8, with constant J_2 term over time
Noncentral bodies	Sun and Moon
Coordinate integration reference system	Mean of J2000.0
Coordinate integration system	Keplerian
Integration type (step size)	Cowell fixed step (300 seconds)
Atmospheric density model	Not Applicable
Atmospheric drag coefficient	Not Applicable
Tropospheric refraction model	Saastamoinen/Niell/Radomski model for TDRSS refractive delays (Reference 3)
Solar reflectivity coefficient (C_R) (estimated)	between 1.35 and 1.47
Satellite geometry model	Sphere with cross-sectional area of 40 meters^2
Timing delays applied through GTDS	–54.7 nanoseconds for American Samoa BRTS
Estimated parameters	State vector, C_R, composite range bias
Tracking data types	BRTS range **and TT&C range**
Input range biases	**see section 2**
Polar Motion	On
Tides	Off
Antenna offsets	GTDS 99.01 defaults
Covariance constraints	10^{-12} degree^2 for both inclination and right ascension of ascending node
Shadow modeling	Conical umbra/penumbra

Studies before momentum unload modeling began indicated that definitive dual BRTS-based solutions had 3σ accuracies of approximately 100 meters (References 5 and 6). For TDRS-3, an error analysis estimated an achievable accuracy of less than 54 meters (3σ) with K-band TT&C and BRT System (BRTS) data from ALS (Reference 7). If, however, the TT&C bias was uncalibrated and had an uncertainty of 60 meters, position errors would vary from 78 to 444 meters. The ionosphere was indicated as the next major error source after range biases. The estimation of the range biases is discussed in section 2 below.

Predicted Accuracy

Predicted accuracy is a measure of how well a solution holds up with propagation over a length of time. It is assessed by comparing a predicted ephemeris to a definitive ephemeris. Predicted accuracies of TDRS solutions were assessed for 1.5-day spans and 2.5-day spans. These two spans are relevant for either the daily solution schedule, or three-sevenths (Tuesday, Thursday, and Saturday) of a Monday-Wednesday-Friday (M-W-F) solution schedule with no solutions done on Tuesday, Thursday, Saturday, and Sunday, when only ephemerides would be

updated (Reference 4). The daily solution schedule has been in effect since December 6, 1999, which was 12 days before the launch of Terra.

It is expected that the measured predicted accuracy, which is computed *relative* to the definitive ephemerides, will still be applicable, even though the *absolute* definitive accuracy may undergo a significant correction by changing the range biases.

Momentum Unloads

The White Sands Complex (WSC) performs momentum unloads so momentum wheels will maintain an appropriate three-axis stabilized attitude of a TDRS. In preparation for Terra support, WSC began sending the times and both predicted and definitive reaction wheel speed (RWSP) changes of momentum unloads to FDF. These unloads are either pitch or roll/yaw. In Figure 2, +P and –P indicate pitch thrusters; while +R and –R indicate roll thrusters. Both pitch and roll thrusters have a significant thrusting in the –Z direction, while the yaw thrusters should all be entirely in the XY-plane. Z1 and Z3 form a yaw thruster pair; Z2 and Z4 form another yaw thruster pair. Each triangle represents a pair of thrusters. +X points towards the velocity vector, +Y points near the south celestial pole, and +Z points towards the Earth.

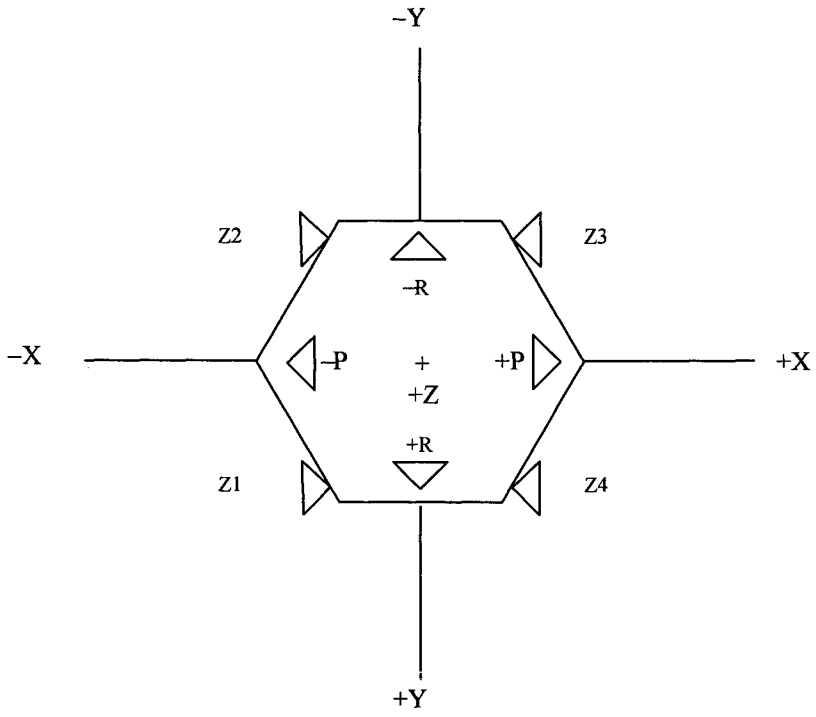


Figure 2: TDRS Momentum Unload Thrusters

From pitch thruster pointing, a linear relation of a pitch unload RWSP change and the delta-V in the along-track direction was derived[1]. Because of the good performance of this theoretical model, no adjustments are made operationally for other TDRSs (See Reference 4).

[1] Computer Sciences Corporation, internal report on momentum unload modeling, T. Lee, February 1998.

By being performed with paired thrusters (Z1/Z3 or Z2/Z4), roll/yaw unloads are intended to have no effect on TDRS orbits. However, misalignments or unequal thruster performance could induce orbital effects. Gas bubbles in the roll/yaw thruster fuel lines may affect thruster performance (Reference 8). These effects vary between TDRSs and between thruster pairs on the same TDRS (See Reference 4). Therefore, an empirical calibration for both roll/yaw thruster pairs is needed for each TDRS.

TDRS-3 has the least roll/yaw unloads with a 12-day average; TDRS-4 has the most with a 3-day average (See Table 3).

Table 3. Days Between TDRS Roll/Yaw Momentum Unloads
8/22/97 – 12/22/98

	TDRS-1 (days)[1]	TDRS-3 (days)[2]	TDRS-4 (days)	TDRS-5 (days)	TDRS-7 (days)
Average	9.2	**11.8**	2.9	4.7	8.3
Standard Deviation	9.6	**7.6**	0.8	2.9	8.1
Maximum	41	**28**	5	18	50
Minimum	0	**1**	0	0	1
Sample Size (counts)	48	**20**	168	103	50

1. Fourteen of the TDRS-1 roll/yaw unloads occurred on June 13 and 14, 1998, after an emergency time out (ETO).
2. The TDRS-3 sample above includes about 7 months: mainly 3 months after Guam support started in June 1998 and from December 8, 1999 until April 7, 2000.

2. ORBIT DETERMINATION METHODS AND RESULTS

First, the study period and selected samples are described, along with the software that was used. Then the estimation of range biases is explained and results are presented.

Study Period and Sample Selection

Since June 1998, TDRS-3 has been supported from Guam. The study period began after this time and extended until January 18, 2000. Several days were assessed around selected momentum unloads, including three cases that did not model roll/yaw unloads and failed to meet requirements in December 1999. This was when operational TDRS solutions reverted to a daily schedule just before the launch of Terra on December 18. Results of eighty-three cases are reported, including eight pitch unloads and twelve roll/yaw unloads. TDRS-3 averages a pitch unload every 20 days and a roll/yaw unload every 12 days. This would make up to five pitch unloads and eight roll/yaw unloads in a typical 83-day span. This sample has over fifty percent more unloads than the average. Because the worst accuracies occur when modeling momentum unloads, our results may therefore be worse than a more representative sample.

Software

The Goddard Trajectory Determination System (GTDS) 99.01 versions up to Delta-2 were used to perform orbit determination and ephemeris generation.

Range Bias Estimation

Both TT&C and BRTS biases were estimated (solved for) in orbit solutions. Two challenges were encountered in this process: the biases were always strongly correlated, and an occasional sudden change in one bias often degraded the other bias (See Reference 9). The latter challenge is discussed first, followed by correlations and the analysis done with other tracking data, including simultaneous solutions with satellites tracked by TDRS-3.

During the summer of 1998, BRTS range biases twice changed by 30 meters, once positive and once negative. The TDRS-3 MA and SA services were scheduled on every other day, and their biases changed a day apart of each other. This challenge was addressed by using an input standard deviation (SD) of 30 meters for the BRTS bias, and an input SD of 5 meters for the TT&C bias. These values helped to favor the TT&C data more than BRTS.

For 2 years, no other BRTS bias 30-meter changes occurred, but TT&C biases changed by integrals of 300 meters occasionally. The major range tone used for TT&C data for TDRSs is a sine wave at 500 kHz modulated on the uplink carrier. The maximum possible phase difference corresponds to a delay of 2 μsec, which is a 1-way range ambiguity of approximately 300 meters[2]. To reduce vulnerability to TT&C range ambiguity, the input TDRS-3 SDs were swapped to 5 meters for BRTS and 30 meters for TT&C on August 10, 1999. Sometimes the TDRS-3 TT&C bias change is immediately corrected by WSC, and sometimes it persists for a day.

On October 5, 2000, the BRTS SA range bias appeared to change by –10 meters. Three days later, the MA range bias followed, with the net result being a joint bias averaging 0 instead of 6 meters. Therefore, the input TDRS-3 SDs were switched back to again favor TT&C data over BRTS. Six days later on October 14, the MA bias seemed to jump back up. On October 25, 2000, a request went into effect to only schedule only one service type (either MA or SA) for TDRS-3 BRTS events. This should make bias modeling with GTDS simpler until an enhancement separates MA and SA biases.

An investigation was attempted to uncover the cause of the BRTS bias changes. WSC and Guam personnel were contacted, and the WSC Daily Operations Summaries were reviewed from October 3 through 8. Guam had power outages on October 5, 8 hours after the SA bias dropped, and on October 8, 4 hours before the MA bias drop was seen. No explanation for the bias changes was found.

The TT&C and BRTS range biases were always highly correlated. Both sites are east of TDRS-3 and are at similar latitudes in opposite hemispheres (See Figure 1).

Initial attempts to assess the TT&C range bias used the GRO Remote Terminal System (GRTS) range and Doppler data with BRTS and TT&C data. When GRTS began in December 1993, its range data was assessed relative to the collocated Deep Space Network site (DSS-46) in Canberra, Australia. Using the GRTS data as a reference point in June 1998, the initial coarse bias estimates at Guam were –25 meters for TT&C and +20 meters for BRTS. The input or "a priori" bias strongly influences the estimated or solve-for bias in TDRS-3 TT&C and BRTS-based solutions, because of the limited viewing geometry and the small dynamics of a geosynchronous satellite relative to a ground site.

Later bias determination used simultaneous solutions with GRO, TDRS-3, -4, and -5, or with Terra, TDRS-3, -7, and either -4 or -6. Other than the International Space Station, which has frequent maneuvers, GRO and Terra were the two users for which we have the most TDRS-3 tracking data. The simultaneous solutions used both range and Doppler data and had great variety in both viewing geometry and dynamics.

Table 4 shows consistent bias results from GRO simultaneous solutions between March and May 1999. The analytic calibration of biases (ACB) method (See Reference 10) provides a technique of range calibration for each of several components using six solutions to achieve the best TDRS-user solution. The timing delay in Table 2 for

[2] Stephen D. Hendry, "TDRS Direct Tone Ranging", Flight Dynamics Facility internal document, April 3, 2001

American Samoa BRTS filled the primary need of ACB solution 1 for other TDRSs. Because a composite BRTS-TDRS-ground terminal bias is used operationally, ACB solution 2 is deemed sufficient to obtain the range biases for the purposes of this study. The solutions' time spans are listed, along with the TDRS-3 TT&C and BRTS (ALS) range biases measured from Guam. The average biases and SDs are shown. The SD column is from the GTDS estimation of the range bias, while the Bias SD column is computed from the listed biases. The weighted average (wt. avg.) is obtained by weighting a bias by the inverse of its variance (square of SD), and dividing the sum of the weighted biases by the sum of the inverse variances (Reference 11).

Table 4. TDRS-3 Bias Summary from Spring 1999 GRO-Simultaneous Solutions

	Bias	SD	Bias	SD	Bias	SD	**Bias**	**Bias**	**Bias**
Span (m/dd.hh)	3/16.06-3/19.15		5/10.09-5/13.17		4/5.0 - 4/8.0		**Avg.**	**SD**	**Wt. Avg.**
TT&C	–43.5	0.7	–42.2	0.4	–42.5	0.4	**–42.7**	0.7	**–42.5**
ALS	–23.1	1.2	–18.6	0.7	–16.0	0.7	**–19.2**	3.6	**–18.1**

Input biases of –43 and either –18 or –19 meters were used on all data in this study before August 24, 1999. When the one-way, one-leg TT&C input bias was changed by 18 meters in May 1999, the TDRS-3 position changed by 135 meters in the along-track direction. This change in position is a factor of 7.5 greater than the range change.. The range measurement of a TDRS is primarily in the radial direction.

The positions of the other TDRSs in the simultaneous solutions usually agreed with the operational BRTS-based positions within 30 meters.

GRO simultaneous solutions in August and December 1999 had significantly different biases for TDRS-3. The TT&C and ALS biases determined in August were –13 and –3 meters, respectively. The TT&C and ALS biases determined in December were 3.3 and 24.5 meters, respectively.

The simultaneous solutions with GRO in January and February 2000 gave results that were even more varied. Simultaneous solutions with the recently launched Terra gave results more consistent with the August GRO results. The weighted average biases of the 2000 GRO solutions agree well with the Terra solutions. Table 5 lists the biases, SD, and weighted average bias using GRO in early 2000, which includes times of high atmospheric density fluctuations after increased solar-terrestrial activity. (GRO was deorbited in June 2000.) Table 6 lists the same parameters using Terra in early 2000. Because the weighted average biases agreed so well between the GRO and the Terra simultaneous solutions, new input biases of –7 and +6 meters for TT&C and ALS, respectively, were used starting on February 11, 2000. Most of the Terra data in the first 7 hours of the span from February 5 to 7 was rejected because of Terra thrusting. Thereafter the Terra data was very stable, and the TDRS-7 simultaneous solution was within 18 meters of the operational solution on February 9, 2000.

Table 5. TDRS-3 Bias Summary from Early 2000 GRO Simultaneous Solutions

	Bias	SD	Bias	SD	Bias	SD	Bias	SD	**Bias**	**Bias**	**Bias**
Span (m/dd.hhmm)	1/1.12-1/4.0430		1/6.16-1/8.0840		1/15.08-1/17.2350		2/5.0-2/8.0		**Avg.**	**SD**	**Wt. Avg.**
TT&C	–3.2	0.6	65.0	2.0	–20.0	0.4	7.5	0.5	**12.3**	36.9	**–7.5**
ALS	10.0	1.0	135.0	3.0	–19.5	0.8	31.2	0.9	**39.2**	67.2	**7.9**

Table 6. TDRS-3 Bias Summary from Early 2000 Terra Simultaneous Solutions

	Bias	SD	Bias	SD	**Bias**	**Bias**	**Bias**
Span (m/dd.hhmm)	1/3.06-1/5.0745		2/5.0-2/7.12		**Avg.**	**SD**	**Wt. Avg.**
TT&C	–6.8	1.4	–6.5	0.2	**–6.7**	0.2	**–6.5**
ALS	3.9	2.4	5.6	0.3	**4.8**	1.2	**5.6**

With the change of biases in February 2000, the GTDS estimate of the TDRS-3 position changed by 71 meters westward. Combining the March, April, and May 2000 Terra solutions with the prior Terra solutions precisely confirmed these values; these Terra solution TT&C biases were all within 3 meters of each other.

Figure 3 displays the input BRTS and TT&C biases used for TDRS-3 since Guam support began. The current input BRTS bias last changed from 6 to 0 meters on October 13, 2000, after MA and SA biases changed. The early evolution of the input biases over time may be more indicative of solution variations than of bias changes.

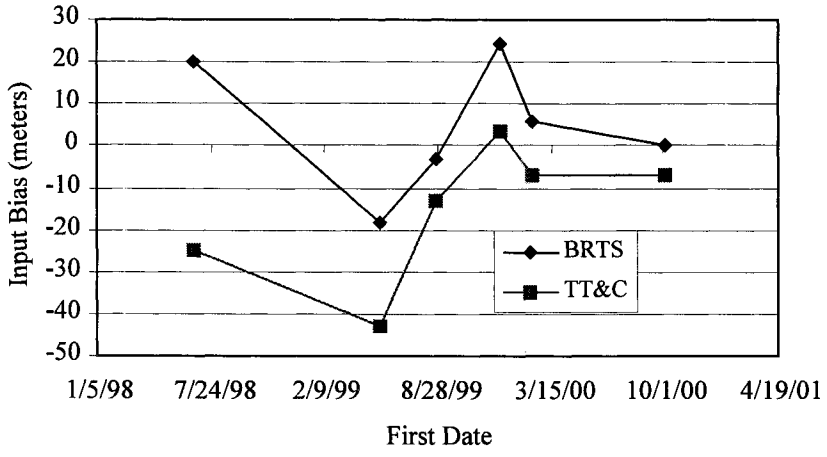


Figure 3. TDRS-3 Input Biases

Occasionally DSS-46 is used to track TDRS-3, providing an independent tracker with typically a small bias. On August 14 and 15, 2000, the mean DSS-46 range residual relative to an operational solution with TT&C and BRTS data was –8.2 meters. The default measurement SDs are 20 meters for DSS Universal Tracking Data Format (UTDF) range data and 10 meters for both BRTS and TT&C. Before April 2001, this and other DSS-46 range passes have confirmed the operational TT&C and BRTS biases well within the DSS 1σ of 20 meters. The TT&C and DSS range values used in GTDS are from ground to TDRS, whereas the BRTS range values are from ground to TDRS to BRTS.

In early 2000, FDF used a sequential filter to do orbit solutions for TDRS-3 simultaneously with all operational TDRSs and four user satellites: the Compton Gamma Ray Observatory (GRO), the Upper Atmospheric Research Satellite (UARS), the Hubble Space Telescope (HST), and the Extreme Ultraviolet Explorer (EUVE) (Reference 12). This sequential filter system has had evaluation and implementation by FDF for operational use and provides

another source for estimating the range biases. With the filter, the TDRS-3 BRTS average residuals were approximately 1 meter, and the TT&C average residuals were approximately −11 meters.

By comparing the GTDS simultaneous solutions of other TDRSs with operational results, and the TDRS-3 simultaneous solutions with the filter results, it is believed that the routine TDRS-3 absolute 3σ errors with GTDS are now less than 50 meters.

3. MOMENTUM UNLOAD MODELING METHODS AND RESULTS

Plane constraints are listed, followed by delta-V optimization for momentum unloads, choice of solution epoch, calibration of delta-V with RWSP, and the frequency of solutions. Then pitch unload modeling results are presented, followed by calibrating delta-V from roll/yaw optimization and all solution results.

Plane Constraints

The same covariance constraints were used for the TDRS-3 study sample as were used for the other TDRSs (See Table 2). Just after the study period, a secular cross-track error was observed for TDRS-5 and -6, so the operational constrained plane covariances were increased on February 10, 2000, by a factor of four to 4×10^{-12} degree^2. For 3 weeks, these looser constraints were used operationally, until a large cross-track error was seen again for TDRS-6, possibly from solar sailing or effects of frequent roll/yaw unloads. During the same time, no growth in the TDRS-5 cross-track error was seen. Results similar to TDRS-5's were expected for other TDRSs. Since March 14, 2000, automated switching is done using either the original 10^{-12}-degree^2 constraints or no constraints, if no delta-V is in the solution arc. This retains the improvements from constraining the plane when an unload is modeled and determines the plane afresh when there are no unloads.

Modeling of Momentum Unloads

GTDS does not directly estimate delta-V in TDRS orbit solutions. Consequently, momentum unloads are modeled by applying an impulsive along-track delta-V, which is the major change to the orbit. WSC does not provide delta-Vs; instead, it provides the unload times and RWSP changes. Therefore, the relationship between delta-V and RWSP change must be empirically determined before GTDS can model roll/yaw unloads.

Delta-V Optimization for Roll/Yaw Unloads

The empirical relationships were determined using the optimization process described in Reference 2. Starting with zero, delta-V was adjusted to reduce the growth of the along-track difference of the predicted and the definitive ephemerides. The process was repeated until the 2.5-day along-track error changed less than 1 meter from the prior iteration or no further improvement could be obtained. The delta-V that met this criterion is called an "optimal delta-V".

Choice of Solution Epoch

In the prior roll/yaw analysis for other TDRSs, obtaining an optimal delta-V was difficult when the unload was before the orbital solution epoch and the predicted compare span, from which the delta-V was computed, was after the epoch. In early 2000, it was found that moving the solution epoch an hour (several integration steps) before the unload can improve the delta-V optimization. Epochs within 6 hours (1/4 orbit) of an unload yield similar results. A moved solution epoch may also significantly change solution parameters, especially the solar reflectivity coefficient (C_R) or possibly the range bias, and improve prediction accuracy for 1.5 days. Because the same C_R was used for the full propagation span, its disturbance for a solution containing momentum unloads significantly alters the propagation over the full span. The magnitude of the change for TDRS-1 through -7 is near 5 meters per day for a change in C_R of 0.01. Placing the epoch before the unload can aid in optimizing delta-V and in improving prediction errors over 1.5 days, even when no other change is done. Because this change in epoch sometimes

improves and rarely degrades forward predictions for 1.5 days, this new procedure has been implemented when bias or C_R tolerance failures occur operationally. Bias tolerances are usually within 4 meters of the input range bias for TT&C and within 8 meters of the input range bias for BRTS. C_R tolerance limits are 1.35 and 1.47. Applying an average bias usually improves predictions when bias or C_R tolerance failures persist with an adjusted solution epoch.

Calibration of Delta-V with RWSP

After optimization, the RWSP change and the corresponding optimal delta-V value were plotted. When the actual RWSP change was not available, the predicted RWSP was used. Separate relationships were determined for positive RWSP changes and negative RWSP changes because different thruster pairs were used. Linear relationships were determined, as was done in the roll/yaw study for the other TDRSs (See Reference 2).

After the relationships between the RWSP change and delta-V were determined, these relationships were tested. First, the linear equation was used to compute a calibrated delta-V from a RWSP change. In the calibration phase, when a predicted RWSP change was not available, an actual RWSP was used. Next, this delta-V was used to model the momentum unload in the solution and ephemeris. Finally, the resulting ephemeris was compared to the definitive ephemeris. This comparison evaluated the predicted accuracy of the calibrated delta-V. Sometimes the predicted value was used instead of the actual, and sometimes only an actual value was available when a predicted value would have been used.

Daily versus Monday-Wednesday-Friday Solutions

Currently, orbit solutions for operational TDRSs are performed daily. This study assessed performances for both a daily and a Monday-Wednesday-Friday (M-W-F) solution update schedule. The current study contains predicted accuracies after 1.5 days and after 2.5 days. Accuracy of ephemerides generated on a daily schedule is indicated by the 1.5-day predicted accuracy measurements.

Accuracy of ephemerides generated on an M-W-F schedule is more complex to assess. On this schedule, there are three instances per week when predicted accuracy was assessed after 1.5 days of prediction and three instances when predicted accuracy was assessed after 2.5 days. The Sunday 3.5-day predicted span was omitted. In other words, each week had the same numbers of 1.5-day predictions as of 2.5-day predictions. Therefore, the accuracy of ephemerides generated on an M-W-F schedule was estimated by simply combining the accuracies of the two groups.

Pitch Unload Modeling Results

Reference 4 contains the standard modeling for TDRS pitch unloads, which is a linear relation of RWSP change and delta-V. Because of the good performance of this theoretical model, no adjustments are made operationally.

There were two cases of pitch unloads that were not close in time to any other momentum unloads during the prime period of this study before June 1999. Four other pitch unloads were used after November 1999. All six cases had comparisons below 35 meters over 1.5 days and below 38 meters over 2.5 days. The average comparisons without momentum unloads were 22 and 25 meters over the same spans. Based on this good performance, no adjustment to the theoretical delta-Vs is done for TDRS-3 pitch unload modeling.

Calibrating Delta-V from Roll/Yaw Optimization Results

Figure 4 gives results from the optimization portion of this study. Only the subset of study cases that were used for the optimization is plotted. For each TDRS, linear equations and correlation coefficients (R^2) are shown separately for negative and positive RWSP changes. Slopes are in units of millimeters per second (mm/s) per revolutions per minute (rpm).

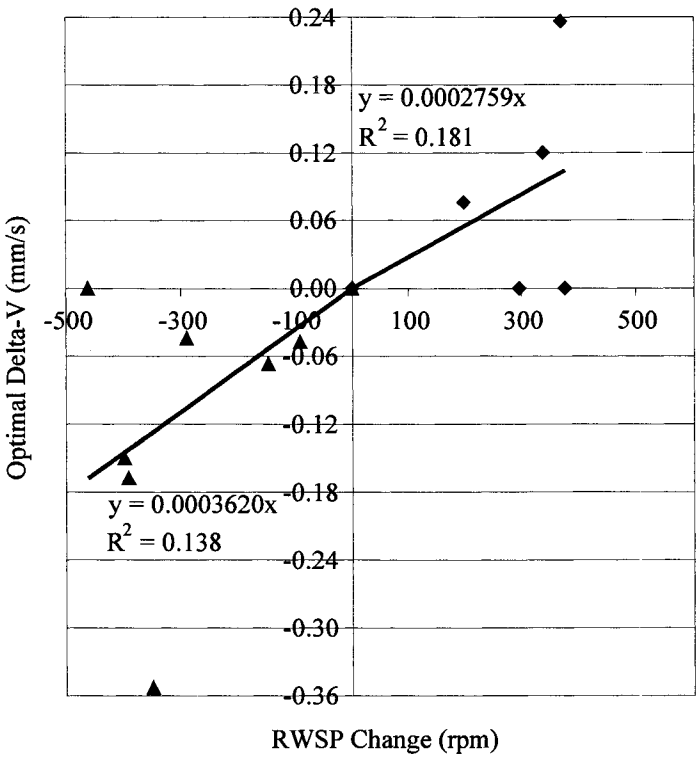


Figure 4. TDRS-3 Optimal Delta-V and Reaction Wheel Speed (RWSP) Change

For TDRS-3, the relationship between negative RWSP change and applied delta-V is modeled as a linear slope of 0.0003620 $mm \cdot s^{-1} \cdot rpm^{-1}$, with an R^2 correlation coefficient of 0.14. The relationship between positive RWSP change and applied delta-V is modeled as a linear slope of 0.0002759 $mm \cdot s^{-1} \cdot rpm^{-1}$, with an R^2 of 0.18. All points in the graph were used in calibrating the relationship. The mass used was 1762 kilograms. The ratio of operational mass to the former mass should be used to rescale the slopes, especially after a significant change in mass. This applies to roll/yaw unloads for all TDRSs. The pitch unload modeling inherently uses the current mass. The steeper TDRS-3 roll/yaw slope, 0.0003620, would induce a change of 47 meters over 1 day after the largest TDRS-3 roll/yaw unload (-489 rpm) in our sample.

In two cases in July 1998, solution noise was comparable to the effect of a roll/yaw unload. These two of the fourteen study cases were not used for optimization and are discussed below. All cases used in the optimization part of the study were also used in the calibration part.

For TDRS-3, most predicted RWSPs were not available in our sample. However, the three points with the optimal delta-Vs farthest from either line for both thruster pairs did have predicted RWSP. Using the predicted RWSP for these three points would change the slopes by less than 2 percent, but would improve the correlation coefficient R^2 by over 30 percent.

After further consideration, it is deemed better to do both optimization and calibration using only predicted RWSP changes. Then, if there is any bias in the predictions, as TDRS-4 roll/yaw unloads have, this calibration would automatically account for the bias. The TDRS-4 positive roll/yaw unload RWSP predicted changes are generally

lower than the actual RWSP changes. Therefore, the calibration slope will be correspondingly steeper. If, on the other hand, the predictions have no bias, or only random noise in them, using the actual RWSP changes for optimizing and then the predicted RWSP changes for calibrating will be essentially the same as using predicted RWSP changes for both. So, using only the predicted RWSP changes in both optimization and in calibration can directly help improve the calibration results, as well as making the process simpler because the predicted RWSP changes are always used operationally.

In an attempt to improve the fit determined slopes, the timing of unloads was reviewed to search for any dependence of optimal delta-V on the time of day of the unload. None was found. It is conjectured that self-shadowing of a TDRS on its thrusters may affect momentum unload burn efficiency.

Two cases that were initially included in the study, –Z1A/–Z3A thruster pair momentum unloads on July 8 and 11, 1998, were later discarded. These two cases occurred during a brief time when the BRTS Multiple Access (MA) and Single Access (SA) range data varied by several meters, early after the Guam station began supporting TDRS-3, and are not included in the eighty-three cases in Table 7. Seven other optimization cases were done successfully for the –Z1A/–Z3A thruster pair.

The calibrations of delta-V with RWSP were used to redo the optimization cases to achieve results as would have been done operationally, given the slopes above.

All Solution Results

Table 7 shows solution statistics of the estimated C_R, TT&C and BRTS biases in meters, the weighted root-mean-square (WRMS) of solutions, definitive (def.) overlap comparisons, and comparisons over 1, 1.5, and 2.5 days of predictions (pred.) with definitive ephemerides. The maximum comparison over 1.5 days was 63 meters, which is less than the Terra requirement; the average comparison was 23 meters. At 2.5-days, for cases with less than 1.5 days notice for a pitch unload, the largest maximum comparisons were 200 and 110 meters. Otherwise, the maximum 2.5-day comparison was 104 meters. The average of all 2.5-day comparisons was 40 meters. The definitive overlap comparisons were very similar to the 1-day predictions. The highest definitive overlap comparison (76 meters) was propagating the December 24, 1999 epoch backward through two momentum unloads. The forward propagation consistency was 34 meters. Removing this one point drops the average definitive overlap difference just below the mean 1-day prediction difference. Five 2.5-day prediction differences exceeded 97 meters. Two of these were from short notice of unloads mentioned above, and four of the five were associated with a combination of a pitch unload and a roll/yaw unload within a 17-hour period.

Table 7. TDRS-3 Solution Statistics

	C_R	TT&C Bias (meters)	BRTS Bias (meters)	WRMS	Def. Overlap (meters)	1-day Pred. (meters)	1.5-day Pred. (meters)	2.5-day Pred. (meters)
Average	**1.3980**	**–37.20**	**–11.79**	**0.156**	**17.9**	**17.0**	**23.0**	**39.8**
SD	0.0245	13.92	14.64	0.059	12.9	12.1	13.2	33.6
Maximum	**1.4567**	**4.83**	**25.31**	**0.405**	**76.4**	**52.0**	**62.9**	**200.2**
Minimum	1.3129	–44.37	–28.40	0.062	3.2	3.0	5.1	4.4
Cases	83	83	83	83	74	69	65	52

2.5-Day Predictions

Out of fifty-two study cases, there were five cases for which either type of momentum unload modeling was not sufficient to meet the 75-meter predicted accuracy requirement after 2.5 days.

1.5-Day Predictions

Out of the sixty-one study cases, there were at least two cases for which pitch modeling alone was not sufficient to meet the 75-meter predicted accuracy requirement after 1.5 days. After the roll/yaw momentum unloads were modeled in daily solutions, all 1.5-day predictions were within 75 meters of definitive ephemerides.

Performance for both Monday-Wednesday-Friday and Daily Solutions

Table 8 gives performance results for solutions performed on an M-W-F schedule, which includes both the 1.5- and the 2.5-day results. Results show that when solutions are performed on an M-W-F schedule, 4.4 percent of the solutions fail the 75-meter accuracy requirement. This result would meet a 2σ 75-meter requirement.

Table 8. Performance for Solutions Three Times per Week

Prediction Span	Cases	Failures	Percent Failure	Sigma Requirement Met
2.5 Days	52	5	9.6 %	1σ
1.5 Days	61	0	0.0 %	3σ
1.5 and 2.5 Days	113	5	4.4 %	2σ

Table 9 gives performance results for solutions performed on a daily schedule, which includes only the 1.5-day results. Again, both pitch and roll/yaw unloads were modeled. Based on comparisons with definitive solutions, all daily solutions met the 75-meter 3σ accuracy requirement.

Table 9. Performance for Solutions Once per Day

Prediction Span	Cases	Failures	Percent Failure	Sigma Requirement Met
1.5 Days	61	0	0.0 %	3σ

4. SUMMARY AND RECOMMENDATIONS

This paper presents performance in meeting Terra requirements for TDRS-3 by modeling momentum unloads and estimating range biases. These 3σ requirements are 75 meters total positional accuracy and 5.5 millimeters per second total velocity accuracy predicted over 1 day onboard. Simultaneous orbit solutions were essential in estimating accurate range biases for TDRS-3. Modeling both pitch and roll/yaw momentum unloads is also required to meet the 75-meter requirement.

For a daily schedule, study results met the Terra requirement by at least 12 meters, compared with definitive ephemerides. Maintaining this performance operationally may require monthly assessment of range biases with simultaneous solutions, as well as closely monitoring for bias changes and for changes in roll/yaw momentum unloads.

To calibrate roll/yaw momentum unloads, we used actual reaction wheel speed (RWSP) changes for optimizing and predicted RWSP changes for calibrating, if both RWSP changes were available. Often only the actual values were available. For future analysis, it is recommended that the predictions be used for both optimizing and calibrating, so that if any bias exists in the predictions, this technique will be self-correcting. Moving the solution epoch before the unload aided at times both in optimization and in operational performance in modeling momentum unloads.

Study results did not always meet the Terra requirement when generating solutions on an M-W-F schedule. A major contributor to these failures was that there were often both pitch and roll/yaw unloads close together. Finer momentum unload modeling is required before the support may be done less often than daily. More precise simultaneous orbit determination may be achievable with the Ocean Topography Experiment (TOPEX)/Poseidon or Aqua (Earth Observing System PM-1) spacecraft and other TDRSs, perhaps with the user satellite as a constant in a TDRS-3 GTDS solution.

REFERENCES

1. National Aeronautics and Space Administration, Goddard Space Flight Center, *Earth Observing System AM-1 Detailed Mission Requirements*, November 1996
2. Computer Sciences Corporation, *TDRS Combined Pitch and Roll/Yaw Momentum Unload Modeling for Terra (EOS AM-1)*, memorandum from H. Offerman, D. Ward and L. Baxter, Code 453.2, to Mr. R. Caldwell, AlliedSignal Tech. Corporation, December 9, 1999
3. Computer Sciences Corporation, 6320-28221-320-01, *Improvement of BRTS Range Refraction Corrections for TDRS Orbit Accuracy*, M. Radomski, March 1998
4. Computer Sciences Corporation, *TDRS Pitch Momentum Unload Modeling for EOS AM-1*, memorandum from D. Ward and T. Thompson, Code 453.2, to Mr. R. Caldwell, AlliedSignal Tech. Corporation, December 23, 1998
5. J. Teles, M. Samii, and C. Doll, *Overview of TDRSS*, (paper presented at 30th COSPAR Scientific Assembly, Hamburg, Germany, July 11-21, 1994)
6. Computer Sciences Corporation, CSC-27434-40, *Summary of Tracking and Data Relay Satellite Orbit Determination and Prediction Accuracy Analyses*, W. Forcey et al., June 1997
7. Computer Sciences Corporation, 27434-43, *Orbital Determination Error Analysis for the Zone-of-Exclusion Tracking and Data Relay Satellite (TDRS-Z)*, D. Kelbel, T. Lee, M. MacWilliams, September 1997
8. White Sands Complex Spacecraft Engineering Group, *TDRS Momentum Unload Summary* (http://nmsp.gsfc.nasa.gov/WSC_SEG/momsumnow.htm), D. Perry, March 7, 2001
9. Computer Sciences Corporation, 27434-55, *Preliminary Assessment of the Accuracy of TDRS Orbit Determination Using Tracking from the K-Band Telemetry, Tracking, and Command System in Place of White Sands BRTS Tracking*, H. Offerman, D. Ward and W. Forcey, September 1997
10. D. Oza, D. Bolvin, J. Lorah, T. Lee, and C. Doll, *Accurate Orbit Determination Strategies for the Tracking and Data Relay Satellites*, (paper presented at the Flight Mechanics/Estimation Theory Symposium, Greenbelt, Maryland, May 16-18, 1995)
11. P. R. Bevington and D. K. Robinson, Data Reduction and Error Analysis for the Physical Sciences, 1969, First Edition, McGraw-Hill, Inc., New York, pp.58-59.
12. S. Wallace, *FDF COTS Infusion Operational Readiness Review*, [Consolidated Space Operations Contract (CSOC) presentation, Greenbelt, Maryland, March 30, 2000]

Preliminary Results of NASA's First Autonomous Formation Flying Experiment: Earth Observing-1 (EO-1)

David Folta
NASA/Goddard Space Flight Center
Greenbelt, MD

Albin Hawkins
a.i.-solutions
Lanham, MD

ABSTRACT

NASA's first autonomous formation flying mission is completing a primary goal of demonstrating an advanced technology called enhanced formation flying. To enable this technology, the Guidance, Navigation, and Control center at the Goddard Space Flight Center has implemented an autonomous universal 3-axis formation flying algorithm in executive flight code onboard the New Mellenium Program's (NMP) Earth Observing-1 (EO-1) spacecraft. This paper describes the mathematical background of the autonomous formation flying algorithm and the onboard design and presents the preliminary validation results of this unique system. Results from functionality assessment and autonomous maneuver control are presented as comparisons between the onboard EO-1 operational autonomous control system called AutoCon™, its ground-based predecessor, and a standalone algorithm.

INTRODUCTION

With the launch of NASA's Earth Observer-1 satellite (EO-1), the Goddard Space Flight Center is demonstrating the capability of satellites to fly in formation, to react to each other, and maintain a close proximity without human intervention. This advancement allows satellites to autonomously respond to each other's orbit changes quickly and more efficiently. It permits scientist to obtain unique measurements by combining data from several satellites rather than flying all the instruments on one costly satellite. It also enables the collection of different types of scientific data unavailable from a single satellite, such as stereo views or simultaneously collecting data of the same ground scene at different angles.

Figure 1. EO-1 Formation Flying Behind Landsat-7

The need for an innovative technical approach to autonomously achieve and maintain formations of spacecraft is essential as scientific objectives become more ambitious.[1,2] The development of small low-cost spacecraft and new scientific research such as large scale interferometry has led many programs to recognize the advantage of flying multiple spacecraft in formation to achieve correlated instrument measurements. Advances in automation and technology by the Guidance Navigation and Control (GN&C) center at the Goddard Space Flight Center (GSFC) has resulted in the development and demonstration of an autonomous system to meet these new guidelines.

The EO-1 technology incorporates the Folta-Quinn (FQ) 3-axis universal algorithm for formation control. This system can be used by single spacecraft or spacecraft in constellations and formations. It can also be applied to Low Earth Orbits, Highly elliptical orbits, and non-keplerian trajectories such as libration orbits. The system allows the burden in maneuver planning and execution to be placed onboard the spacecraft, mitigating some of the associated operational concerns while increasing autonomy.

The EO-1 formation flying requirements on our technology are to demonstrate the capability of EO-1 to fly over the same groundtrack as Landsat-7 within +/-3 kilometers at the equator while autonomously maintaining the formation for extended periods to enable paired scene comparisons between the two satellites. The required relative separation is 1 minute in mean motion, equivalent to 450km. The tolerance on this separation to meet the ground track is +/- 6 seconds, or roughly 42 km.

This paper presents preliminary validation results of formation flying of the NMP EO-1 spacecraft with respect to the Landsat-7 spacecraft. Results are presented as comparisons between the onboard autonomous formation flying control system and two ground systems. Both the onboard and the prime ground systems use AutoCon™ a high fidelity modeling package which incorporates the FQ Algorithm. This overall NMP autonomous control experiment is called Enhance Formation Flying (EFF).

FORMATION FLYING

Formation flying involves position maintenance of multiple spacecraft relative to measured separation errors. For EO-1, this relative separation between the EO-1 and Landsat-7 spacecraft is required to allow co-scene comparisons. An overview of the EO-1 formation flying using a two spacecraft differential drag example is presented here.

Mechanics Using Differential Drag

If two spacecraft are placed in similar orbital planes and similar altitudes with a small initial separation angle they will be equally affected by the Geopotential field of the Earth and by atmospheric drag provided that they have identical ballistic properties. As long as the separation angle is small enough that atmospheric density and gravitational perturbations can be considered constant, the relative separation will remain the same. If the spacecraft are separated in the radial direction, and the respective ballistic properties are different, their orbit velocities are also different, and one spacecraft (the EO-1 / chase spacecraft) will appear to drift relative to the other (Landsat-7 / control spacecraft). The drifting is most apparent in the along-track (orbital velocity) direction. The radial separation can be operationally planned or induced by differential decay rates caused by environmental perturbations. The concept of formation flying for EO-1 is based on the constructive use of the differential decay rates as a direct function of differential ballistic properties between a reference and a free-flying spacecraft.

EO-1 Example

An example of the orbit dynamics of EO-1 and Landsat-7 formation flying is shown in Figure 2. In the figure, EO-1 starts a formation at the red dot location, behind Landsat-7 by 450 kilometers and above by ~50 meters. Due to the differences in the drag accelerations from the atmosphere, the EO-1 orbit decays slightly faster. While above Landsat-7, EO-1 is drifting away from Landsat-7 since the average orbital velocity is less than that of Landsat-7. After several days of orbital decay due to atmospheric drag EO-1 will be below Landsat-7 and will drift towards

Landsat-7 since the average orbital velocity is now higher than that of Landsat-7. When EO-1 is outside the required tolerance box or if Landsat-7 has maneuvered, EO-1 will autonomously compute and perform a maneuver to reposition it to an initial condition to repeat the relative motion.

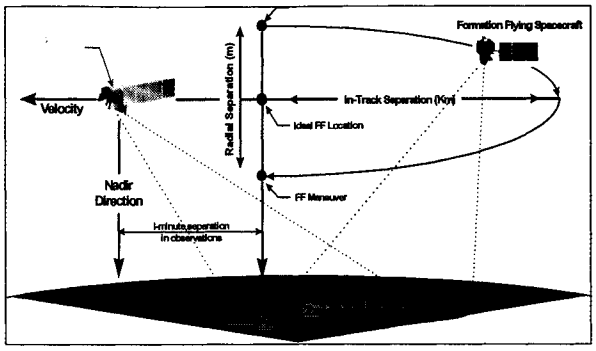


Figure 2. EO-1 Formation Flying Using Differential Drag

FORMATION FLYING ALGORITHM

The Folta-Quinn (FQ) algorithm is a new technology that is based on mathematics derived by Battin and adapted to the formation flying problem.[3,4] A patent application has been submitted to the GSFC Office of Patent Counsel for the application of Autonomous Closed Loop 3-Axis Navigation Control Of Spacecraft[5]. This patent-pending technology will allow full closed-loop maneuver autonomy onboard any spacecraft rather than the tedious and costly operational activity historically associated with ground based operations and control. The application to other missions is unlimited and can therefore be used to more fully explore the NASA mandate of faster, better, cheaper spacecraft.

FQ Algorithm Description

The FQ algorithm for formation flying solves the position maintenance problem by combining a modified Lambert's two point boundary value problem and Battin's 'C*' matrix with an autonomous system developed by a.i.-solution, Inc. of Lanham MD. called AutoCon™. The algorithm enables the spacecraft to execute complex three axis orbital maneuvers autonomously. Figure 3 illustrates the basic sets of information required for the EO-1 formation targeting as it is incorporated into AutoCon™. The FQ algorithm well is suited for multiple three axis burn scenarios but is more easily explained using a two-burn, co-planar example for clarity.

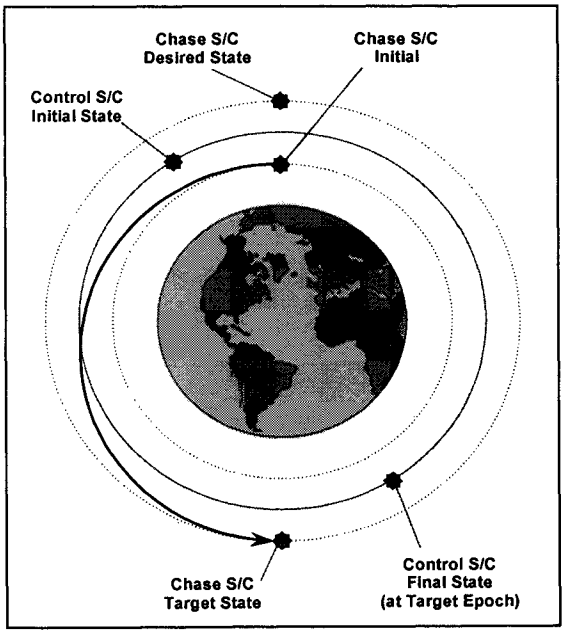


Figure 3. FQ Algorithm Inputs for EO-1 Formation Flying

The formation flying problem in this example involves two spacecraft orbiting the Earth. Landsat-7, the control spacecraft, orbits without performing any formation flying maneuvers. EO-1, the chase spacecraft monitors the control spacecraft, and performs maneuvers designed to maintain the relative position imposed by the formation requirements. In this example, the goal of the formation flying algorithm is for EO-1 to perform maneuvers which cause it to move along a specific transfer orbit. The transfer orbit is established by determining a path (in this case a Keplerian path) which will carry the EO-1 spacecraft from some initial state, $(\mathbf{r}_0, \mathbf{v}_0)$, at a given time, t_0, to a target state, $(\mathbf{r}_t, \mathbf{v}_t)$, at a later time, t_t. The target state is found to be one which will place EO-1 in a location

relative to Landsat-7 so as to maintain the formation. A desired state is also computed. This is accomplished by back propagating the target state to find the initial state that EO-1 would need at time t_0 for it to achieve the target state at time t_t without executing a maneuver. This back propagation of the target state gives rise to the desired state, ($\mathbf{r}_d$, $\mathbf{v}_d$) at time t_0. The initial state can now be differenced from the desired state to find:

$$\begin{pmatrix} \delta \mathrm{r} \\ \delta \mathrm{v} \end{pmatrix} = \begin{pmatrix} \mathrm{r}_0 - \mathrm{r}_d \\ \mathrm{v}_0 - \mathrm{v}_d \end{pmatrix}$$

STM Formulation

The FQ Algorithm uses state transition matrices, described below, for the calculation of the maneuver ΔV. Selecting initial conditions prescribed at a time t_0 so that the state at this time has all zero components except the jth term which is unity, a state transition matrix, $\Phi(t_1,t_0)$, can be constructed such that it will be a function of both t and t_0 and satisfies matrix differential equation relationships[5]. The initial conditions of $\Phi(t_1,t_0)$ are the identity matrix.

Having partitioned the state transition matrix, $\Phi(t_1,t_0)$ for time $t_0 < t_1$,

$$\phi(t_1,t_0) \equiv \begin{bmatrix} \phi_1(t_1,t_0)\phi_2(t_1,t_0) \\ \phi_3(t_1,t_0)\phi_4(t_1,t_0) \end{bmatrix}$$

We find the inverse may be directly obtained by employing symplectic properties

$$\phi^{-1}(t_1,t_0) \equiv \phi(t_0,t_1) \equiv \begin{bmatrix} \phi_1(t_0,t_1)\phi_2(t_0,t_1) \\ \phi_3(t_0,t_1)\phi_4(t_0,t_1) \end{bmatrix}$$

$$\phi^{-1}(t_1,t_0) \equiv \begin{bmatrix} \phi^T{}_4(t_1,t_0)\phi^T{}_2(t_1,t_0) \\ \phi^T{}_3(t_1,t_0)\phi^T{}_1(t_1,t_0) \end{bmatrix}$$

Where the matrix $\Phi(t_0,t_1)$ is based on a propagation forward in time from t_0 to t_1 and is sometimes referred to as the navigation matrix, and $\Phi(t_1,t_0)$ is based on a propagation backward in time from t_1 to t_0, and is sometimes referred to as the guidance matrix. We can further define the transition matrix partitions as follows:

$$\begin{aligned} \tilde{R}^*(t_0) &\equiv \phi_1(t_0,t_1) \quad \tilde{R}\ (t_1) \equiv \phi_1(t_1,t_0) \\ R^*(t_0) &\equiv \phi_2(t_0,t_1) \quad R\ (t_1) \equiv \phi_2(t_1,t_0) \\ \tilde{V}^*(t_0) &\equiv \phi_3(t_0,t_1) \quad \tilde{V}\ (t_1) \equiv \phi_3(t_1,t_0) \\ V^*(t_0) &\equiv \phi_4(t_0,t_1) \quad V\ (t_1) \equiv \phi_4(t_1,t_0) \end{aligned}$$

Substituting yields the following useful identities:

$$\begin{bmatrix} \tilde{R}^*(t_0) & R^*(t_0) \\ \tilde{V}^*(t_0) & V^*(t_0) \end{bmatrix} = \begin{bmatrix} V^T(t_1) & -R(t_1) \\ -\tilde{V}^T(t_1) & \tilde{R}(t_1) \end{bmatrix}$$

Where the starred quantities are based upon a guidance matrix and unstarred quantities are based on a navigation matrix. If a reversible Keplerian path is assumed between the two states, one should expect the forward projection of the state from t_0 to t_1 to be related to the backward projection of the state from t_1 to t_0. When the fundamental matrices C and C* are defined as

$$\tilde{C}^* \equiv \tilde{V}^* \tilde{R}^{*-1} \text{ and } C^* \equiv V^* R^{*-1}$$

We find the following:

$$\tilde{C}^* \equiv \left.\frac{\partial v_0}{\partial r_0}\right|_{v1=constant} \text{ and } C^* \equiv \left.\frac{\partial v_0}{\partial r_0}\right|_{r1=constant}$$

so that $C\delta r = \delta v_0$ becomes the velocity deviation required at time t_0 (as a function of the measured position error δr at time t_0) if the spacecraft is to arrive at the reference position r_1 at time t_1 (with arbitrary velocity). Recalling that the starred quantities were obtained based on the guidance matrix, the sympletic property allows them to be computed based on a navigation projection. It can therefore be shown that

$$\left[C^*(t_0)\right] = \left[V^*(t_0)\right]\left[R^*(t_0)\right]^{-1} = \left[\tilde{R}^T(t_1)\right]\left[-R^T(t_1)\right]^{-1}$$

Applying a universal variable formulation of the closed-form state transition matrix, the relevant state transition matrix submatrices are computed.[4,5] The expressions for F, G, F_t and G_t are derived from the Gauss problem of planar motion; K is a quantity derived from the Universal Variable (U) formulation.[5] These variables are dependent upon each other in their formulation, i.e. U(6) is dependent upon U(4) and on intermediate variables related to the classic f and g series. The target and desired states, r_d ,v_d, r_t , and v_t are computed from the propagated states. μ is the universal gravitational constant. R and R are then defined from the target and desired states as:

$$\mathbf{R}(t_t) = \frac{|\mathbf{r}_d|}{\mu}(1-F)\left[(\mathbf{r}_t - \mathbf{r}_d)\mathbf{v}_d^{\mathrm{T}} - (\mathbf{v}_t - \mathbf{v}_d)\mathbf{r}_d^{\mathrm{T}}\right] + \frac{K}{\mu}\left[\mathbf{v}_t\mathbf{v}_d^{\mathrm{T}}\right] + G[\mathbf{I}]$$

$$\tilde{\mathbf{R}}(t_t) = \frac{|\mathbf{r}_t|}{\mu}\left[(\mathbf{v}_t - \mathbf{v}_d)(\mathbf{v}_t - \mathbf{v}_d)^{\mathrm{T}}\right] + \frac{1}{|\mathbf{r}_t^3|}\left[\mathbf{r}_t|(1-F)\mathbf{r}_t\mathbf{r}_d^{\mathrm{T}} + K\mathbf{v}_t\mathbf{r}_d^{\mathrm{T}}\right] + F[\mathbf{I}]$$

From these variables and sub-matrices, the C* matrix is computed as follows:

$$R^*(t_0) = -R^T(t_t)$$
$$V^*(t_0) = \tilde{R}^T(t_t)$$
$$C^*(t_0) = V^*(t_0)\left[R^*(t_0)\right]^{-1}$$

The expression for the impulsive maneuver follows immediately:

$$\Delta V = \left[C^*(t_0)\right]\delta r_0 - \delta v_0$$

Keplerian and Non-Keplerian Transfer Orbits

Having established both actual and desired states of a spacecraft's location using standard GSFC propagators, all that is needed is a means of autonomously zeroing the difference between the two states. Given two Keplerian trajectories and a chronologically defined maneuver window, an arbitrary (possibly non-Keplerian) reference trajectory may be determined which will smoothly transport the spacecraft from its position on the first Keplerian path at the beginning of the maneuver window to a desired position on the second Keplerian path at the conclusion of the maneuver window. Control points on the reference trajectory in Figure 4 are calculated at regular time intervals consistent with the ability of the spacecraft to receive and process position data, fire its thrusters, and account for the effects of each firing.

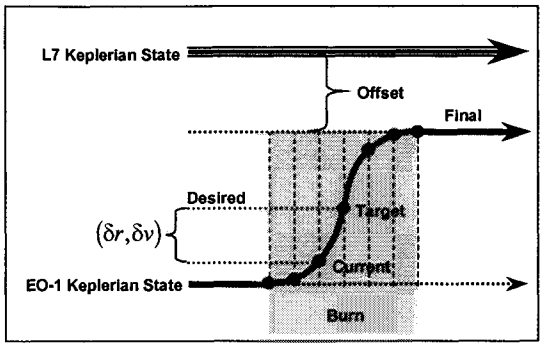


Figure 4. FQ Algorithm Transfer Trajectory

At each step in the process, the next control point on the reference path is examined and back-propagated along a Keplerian path to determine small differences between spacecraft position and velocity on the reference path and determine which Keplerian path would intersect the reference path at the next control point. These differences are then fed into the propagator via the state transition matrices to determine the incremental ΔV required to get the spacecraft to the next control position on the reference trajectory. At the conclusion of the maneuver window, a final burn is required to match the velocity required to maintain the new Keplerian trajectory. One can use single or multiple maneuvers to achieve the target condition.

For EO-1's orbit a long, iterative window requiring many small burns is not necessary and ΔV maneuvers resemble a Hohmann transfer. The generalized approach does not, however, require that the two Keplerian (pre- and post- window) paths intersect, nor does it require that the intervening reference trajectory be Keplerian. The approach is therefore a means of executing any smooth non-Keplerian trajectory that will get a spacecraft from anywhere to anywhere along any desired path, limited only by time, fuel, and spacecraft capabilities.

EO-1 Enhanced Formation Flying (EFF) Technology Description

This EFF is part of a new autonomous onboard technology, which features flight software that is capable of autonomously planning, executing, and calibrating routine spacecraft orbital maneuvers[5]. The autonomous formation flying control software AutoCon™ builds on GSFC GN&C existing capability for the maneuver planning, calibration, and evaluation tasks.[6] AutoCon™ can also use a fuzzy control engine, ideal for this application because it can easily handle conflicting constraints between spacecraft subsystems.

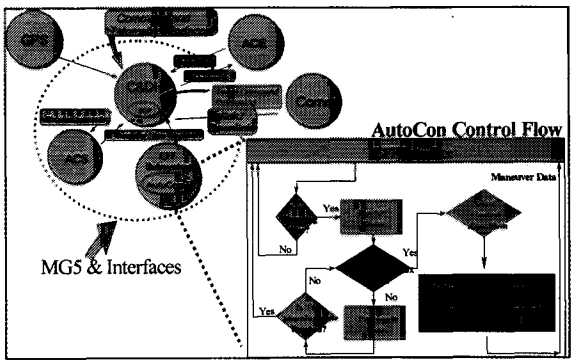


Figure 5. EFF Onboard Layout

Formation Flying Control

The AutoCon™ flight control system ingest data from EO-1 sensors and subsystems such as propulsion, navigation, and attitude data. It then autonomously generates, analyzes, and executes the maneuvers required to initialize and maintain the formation between Landsat-7 and EO-1. Figure 5 shows a functional diagram of EFF and AutoCon™ system. Because these calculations and decisions are performed onboard the spacecraft, the lengthy period of ground-based planning currently required prior to maneuver execution will be eliminated. The system is general and modular so that it can be easily extended to future missions. Furthermore, the AutoCon™ flight control system is designed to be compatible with various onboard navigation systems (i.e. GPS, or an uploaded ground-based ephemeris). The AutoCon™ system is embedded in the Mongoose-5 EO-1 spacecraft computer. Interfaces are handled with one interface to the C&DH system. This is used for the ingest of GPS states information, AutoCon™ commanding, EFF telemetry, and maneuver commands for EO-1 as well. The FQ algorithm needs input data for the current EO-1 state, the target state, and the desired state. These data are provided by AutoCon™. AutoCon™ takes the current EO-1 and uploaded Landsat-7 states and then propagates these states for a user-specified fraction of the period. Autonomous orbit control of a single spacecraft requires that a known control regime be established by the ground which is consistent with mission parameters. That data must then be provided to the spacecraft. When orbital perturbations carry the spacecraft close to any of the established boundaries, the spacecraft reacts (via maneuver) to maintain itself within its error box. The system is currently set to check the tolerance requirements every 12 hours. Form this point AutoCon™ propagates the states for 48 hours (a commandable setting) and will execute a maneuver plan if needed.

Algorithm Modes

There are five EFF maneuver control modes onboard EO-1 as shown in Figure 6. The control modes verified during this preliminary validation process are modes 1, 2 and a partial of 3. These modes were established to allow a incremental validation of the system performance, data interfaces, and maneuver computations before commands were generated onboard for an executable maneuver.

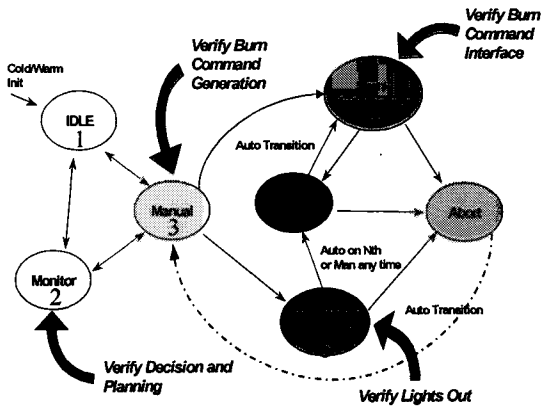


Figure 6 EFF Maneuver Modes

ΔV Computations and Quantized Maneuvers

The computation of the EO-1 maneuver ΔVs is performed using a sequence of two methods. The first method uses the FQ algorithm for the calculation of the maneuver to reach the targeted position relative to Landsat-7. Subsequently, a velocity-matching maneuver is then performed once the targeted position is attained. The FQ algorithm could also be used, but in an effort to conserve onboard resources a velocity matching method is employed. This velocity matching is computed from the difference in the velocity of the EO-1 transfer orbit and the targeted state

The EO-1 spacecraft propulsion system was designed so that the minimum maneuver duration is one second with larger burns selectable at one-second increments. This means that commands generated either onboard or on the ground will undergo a rounding of the maneuver duration based on the computed ΔV. For example if a maneuver is such that the computed maneuver duration is 5.49 seconds, the commanded maneuver will actually be 5 seconds, and a 5.51 second

duration would become 6 seconds. This results in a quantized maneuver duration for each maneuver and thus the achieved Keplerian trajectory will differ slightly from the targeted trajectory. To compensate for this effect the final ΔV is adjusted. The velocity match is perturbed slightly to compensate for the position error resulting from the prior maneuver's quantized burn duration. This allows the targeted orbit's SMA to be achieved with a trivial sacrifice of eccentricity.

PRELIMINARY VALIDATION RESULTS

On January 12, 2001, the Enhanced Formation Flying (EFF) Experiment onboard EO-1 became operational. EFF was started in the modes 1 and 2 whereby GPS data would flow though the C&DH interface into the AutoCon™ executable and maneuvers were computed continuously. Scripts and data uploaded via tables were enabled though the execution of EFF. With this data maneuvers were calculated at specified intervals. The overall computational interval was approximately 3 hours in duration and began with the ingest of a single GPS EO-1 state. This state, along with an uploaded Landsat-7 State, was then propagated onboard for durations of 12 hours, 24 hours, and 48 hours. Maneuvers were computed at the 12, 24, and 48 hour epoch marks. After the last maneuver was computed, a new GPS EO-1 state was ingested and the process began again. This enabled the continuous computation of maneuvers while verifying the ingest and data interfaces and propagation of states onboard EO-1.

Validation Results and Period of Performance

This EFF script ran over a several week period, Jan 12 through February 10, and generated over 530 maneuver plans. These maneuvers were planned in sets of three based on the three propagation durations. GPS data was ingested 177 times while tables were uploaded approximalety 30 times for script control, Landsat-7 data, and environmental data updates. The preliminary validation was accomplished by looking at several events and computations. These included:

- EO-1 GPS and Landsat-7 state ingest
- EO-1 and Landsat-7 Propagation Events (Target and Desired States)
- Folta-Quinn Targeting Algorithm Output
  - Quantized Maneuver ΔV
  - 3-D maneuver ΔV
  - Internal Calculations (Matrices, Variables, States)

EO-1 Relative Motion

The following results are comparisons taken directly from the EO-1 playback telemetry which provides the output from the onboard EFF AutoCon™ flight code to the output of using the playback states as input to the PC AutoCon™ ground system and the original MATLAB FQ algorithm. The Landsat-7 initial orbit conditions where taken from the playback telemetry. The Landsat-7 states uploaded for the test were obtained from the Landsat-7 project. The results from two comparisons show the general formation flying evolution and the effect on the mission groundtrack requirements. The evolution differences are due to the changing EO-1 state computed by the GPS receiver and Landsat-7 updates. Evolution of the ground track and the formation alongtrack, radial, and crosstrack are presented in a Landsat-7 centered rotating coordinate system with the radial direction (ordinate) being the difference in radius magnitude and the alongtrack direction (abscissa) being the arc between the position vectors. Crosstrack is a

direct measurement of cross track separation of the spacecraft which is a function the orbital plane separations necessary to meet the ground track requirement. Figures 7 and 8 present alongtrack, crosstrack, and radial, separations for two maneuver scenarios. In these plots, EO-1's initial position is located on the right side of the figure at approximately 456km and 487 km alontrack separation. Figure 9 presents the ground tracks for these maneuver scenarios.

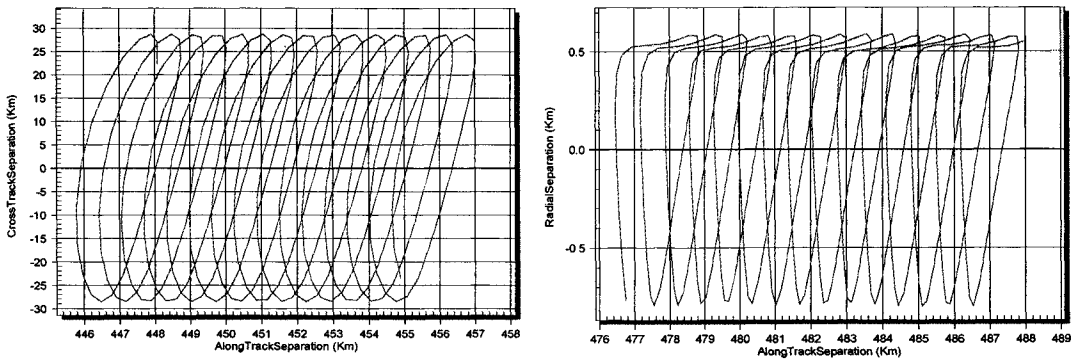


Figure 7. EO-1 Alongtrack vs. Radial Differences in a Rotating Coordinate System

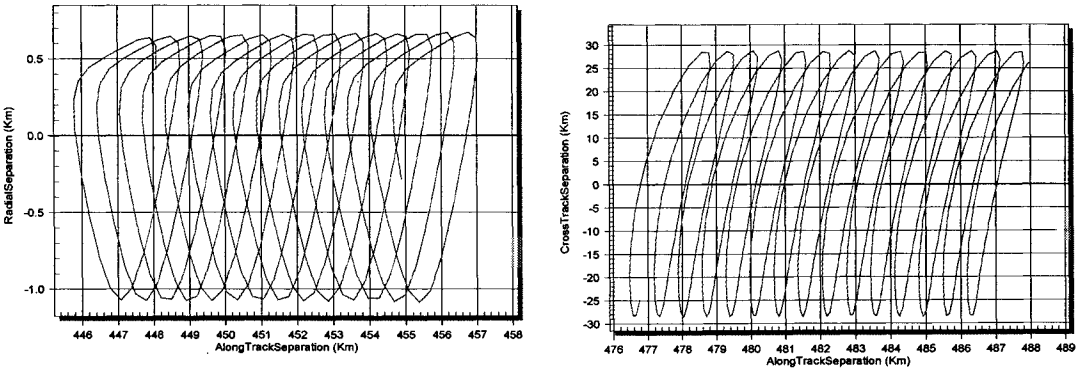


Figure 8. EO-1 Crosstrack versus Alongtrack Differences in a Rotating Coordinate System

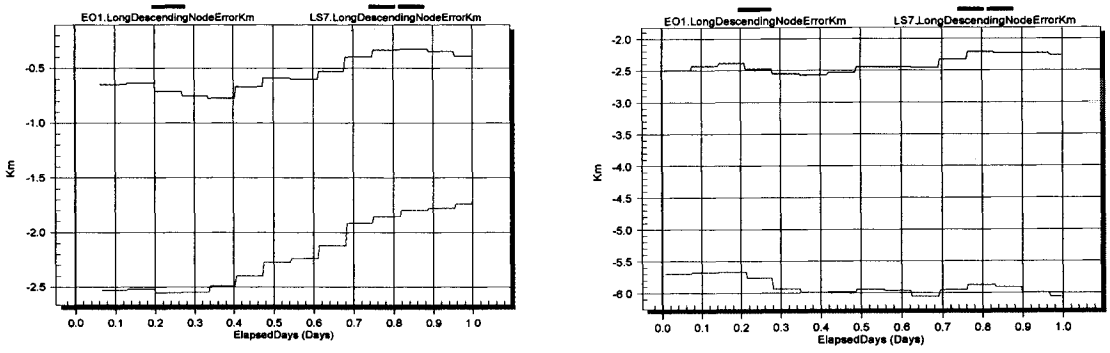


Figure 9. EO-1 and LS7 GroundTrack

Maneuver Comparisons

This section presents onboard and ground comparison results in terms of the absolute difference in the computed ΔV (cm/s) and the related percentage error for several maneuver scenarios. A total of 12 scenarios consisting of 3 maneuver sets (two maneuvers per set) for a total of 36 combined maneuvers were verified. The locations and epochs of these maneuvers were chosen randomly at approximately one per day over a three-week span. Figures 10 and 11 present the overall performance of each quantized maneuver as an absolute difference in the ΔV magnitude

and its percent error. The mean value of the quantized difference is 0.0001890cm/s with a standard deviation of 0.000133 cm/s. These data show that there is excellent agreement between the onboard system and ground validation system. The larger residual in figure 10 is due to a 1-second quantization of a velocity-matching maneuver. This difference is due to the onboard system yielding a maneuver duration near the mid point which rounded down while the ground system rounded up. The difference is still small at 1.4%. The next figures, 12 and 13, present maneuver comparisons for the 3-D computation. This provides the comparisons for the total ΔV required to align EO-1 directly behind Landsat-7 and involves all three ΔV components of radial, alongtrack, and crosstrack.

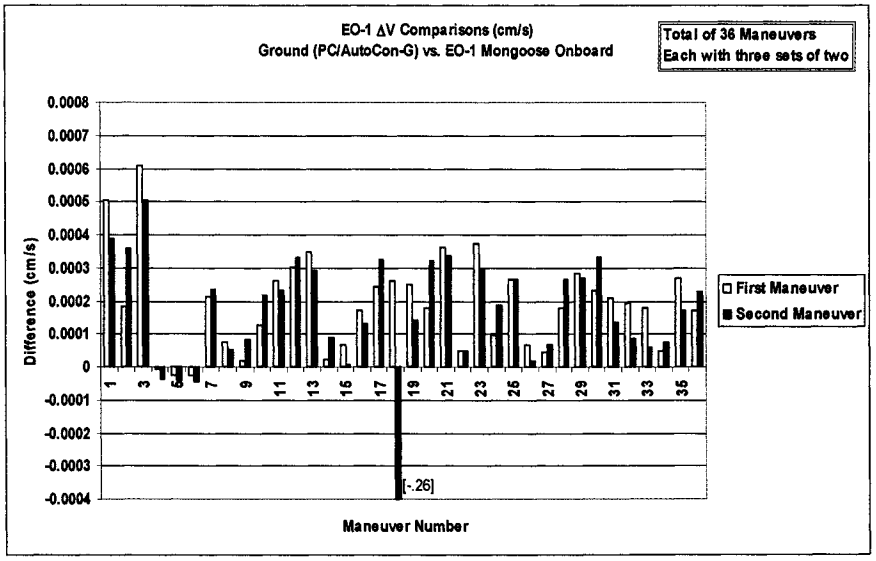


Figure 10. Difference in EO-1 Onboard and Ground Absolute ΔVs

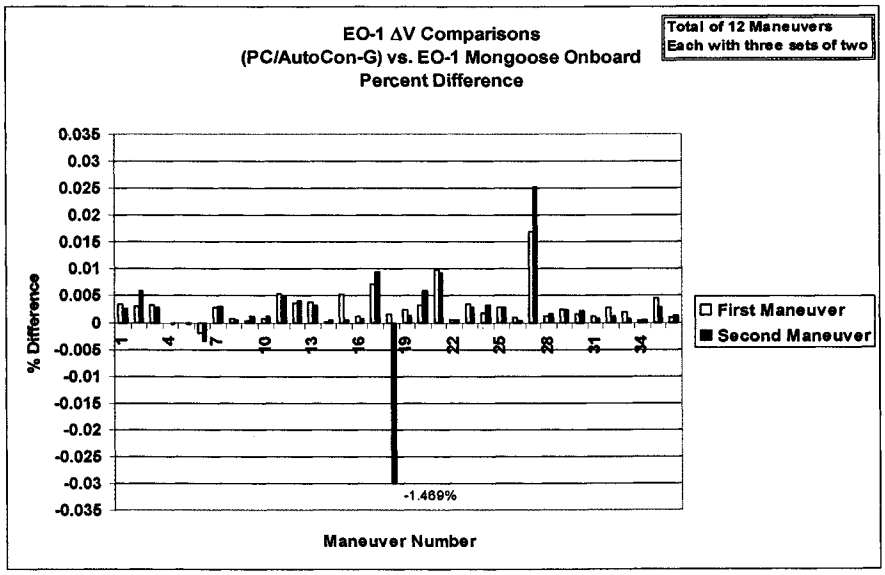


Figure 11. Percentage Difference in EO-1 Onboard and Ground Absolute ΔVs

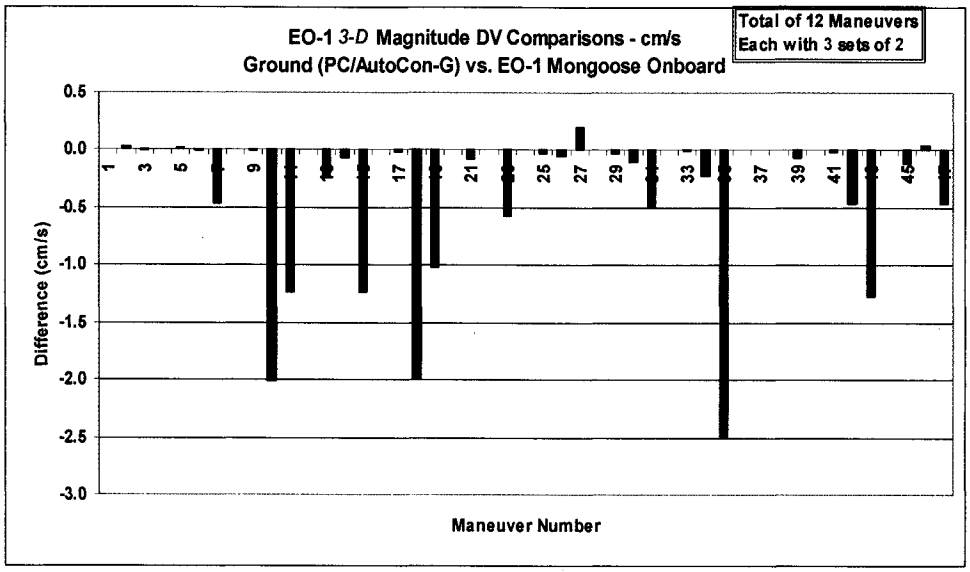


Figure 12. Absolute Difference in 3-D Onboard and Ground ΔVs

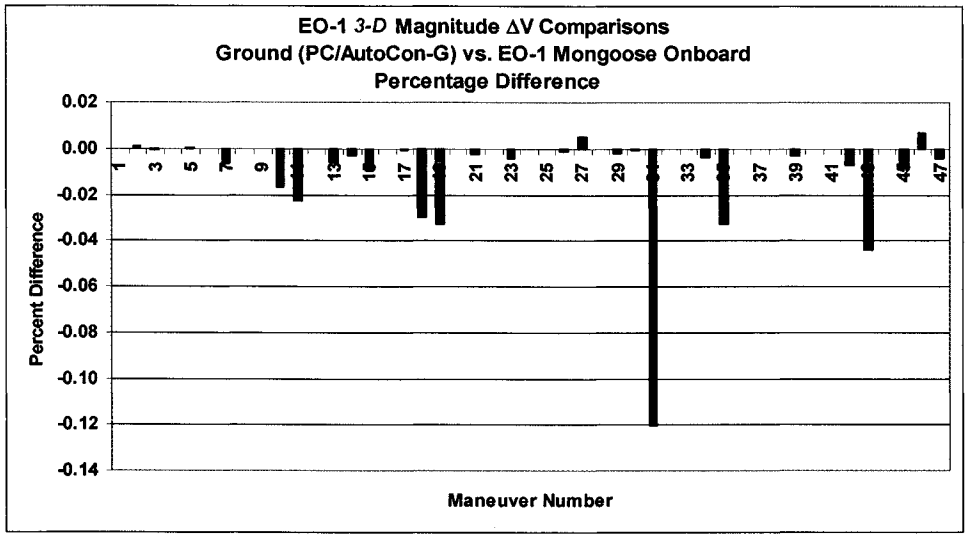


Figure 13. Percentage Difference in 3-D Onboard and Ground ΔVs

Obviously the crosstrack component is the driver with the largest magnitude. The comparisons show only the total ΔV magnitude, as this is the only information available in EO-1 playback telemetry.

With the comparisons between the ground and operational onboard version of the EFF completed, a comparison to the original FQ algorithm code was then performed. This comparison was done only for the first FQ targeted maneuver of each maneuver scenario. The state data from the playback telemetry was input into a MATLAB™ script with the FQ algorithm computing the maneuver without any propagation.[3,7] Figures 14 and 15 show the difference in cm/s and as a percentage respectively for the 3-D ΔV and an alongtrack ΔV. The alongtrack ΔV was represented in the MATLAB™ script by using a local-vertical local horizontal coordinate system based on the input states which is comparable to the EO-1 nominal attitude for maneuvers. The resulting ΔV difference gives a mean of 0.0727 cm/s and a standard deviation of 0.348058 for the 3-D and gives a mean of -0.03997 cm/s and a standard deviation of 0.278402 for the alongtrack. The mean percentage difference was 0.003 for the 3-D and 0.006 for the alongtrack. These results show excellent comparisons.

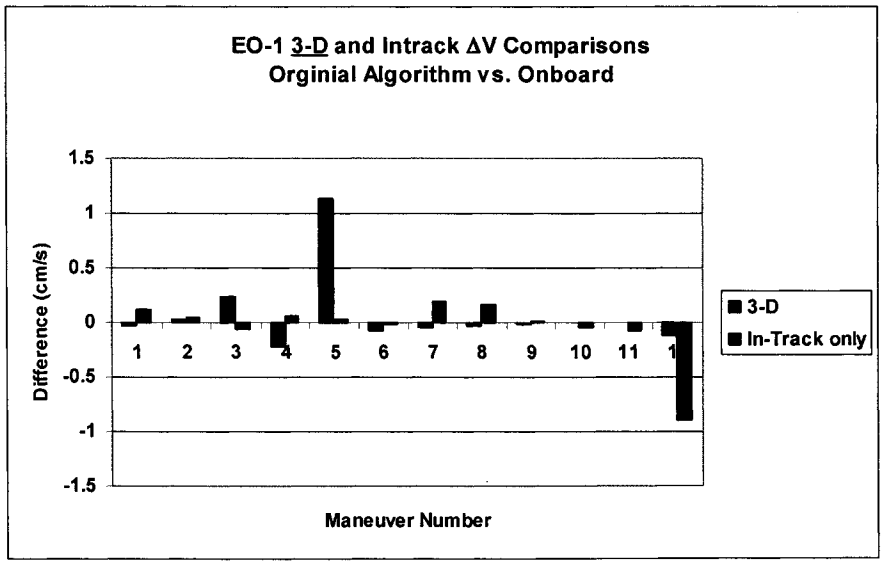


Figure 14. ΔV Difference in Original Algorithm and Onboard

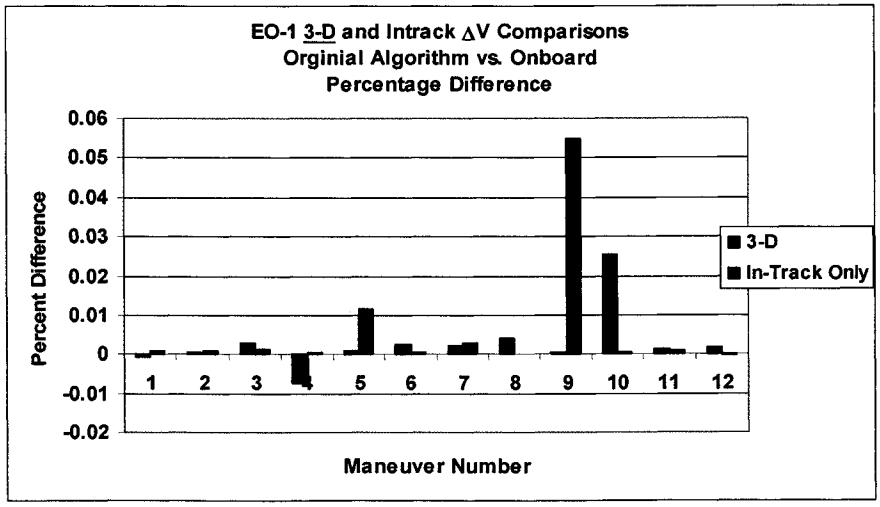


Figure 15. Percentage Difference in Original Algorithm and Onboard

Propagation Comparisons

The FQ Algorithm is dependent upon the generation of the target and desired states. These states are propagated onboard using a Runge-Kutta 4/5 with an 8x8 Geopotential model and a Jacchia-Roberts atmospheric drag model. The accuracy of the computed ΔV is dependent upon the accuracy of these propagated states. For EO-1, the states are propagated forward 1 and ½ orbits to compute the target state and then propagated 1 and ½ orbits backward to compute the desired state. As the desired state incorporates the longest propagation duration with a restart, a comparison was made in the onboard and ground states. The comparison results are shown below in figures 16 and 17. Figure 16 shows the position component and magnitude differences for six maneuver plans. Figure 17 shows the velocity differences. The maximum difference observed was 1.35 meters in the y-component of position and 1.4 cm/s in the velocity z-component. These small differences are still being investigated, but are believed to be the due to the integration into and performance of the EO-1 computer. The mean and standard deviations for position are listed in table 1.

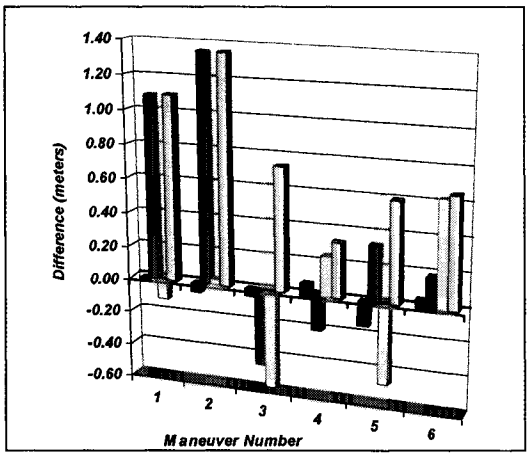


Figure 16. 1.5 Orbit Propagation Position Difference

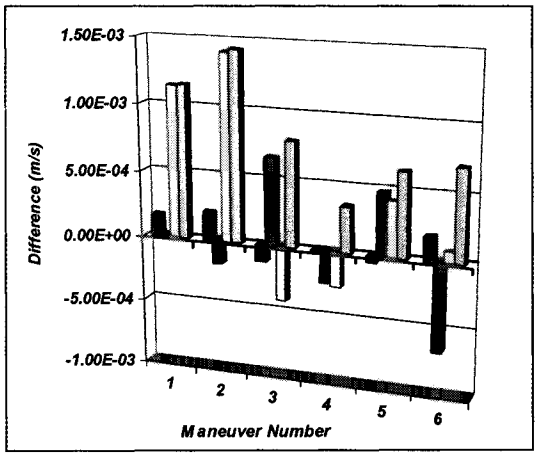


Figure 17. 1.5 Orbit Propagation Velocity Difference

Table-1. Propagation Mean and Standard Deviation for Desired State Computation

	X	Y	Z	Magnitude
Position Mean (m)	-0.02279	0.38221	-0.04550	0.79088
Position StDev (m)	0.07676	0.70684	0.45024	0.36886
Velocity Mean (m/s)	0.00007	0.00001	0.00040	0.00084
Velocity StDev (m/s)	0.00014	0.00049	0.00074	0.00039

SUMMARY

Using the formation flying algorithms developed by the Guidance, Navigation, and Control center of GSFC, onboard validation has shown that the EO-1 formation flying requirements can be easily met. To ensure the accuracy of the onboard FQ algorithm, several comparisons were performed against both original analytical calculations and ground based FQ numerical computations using AutoCon™ for given initial onboard-generated states. The FQ algorithm was validated by direct inputs of the initial taken from the onboard system. The ΔV results agree to millimeters/sec level for the numerical tests which include the effects of propagation. The Matlab simulations agree to the sub-cm/sec as well, due to the differences in PC and Mongoose applications.

CONCLUSIONS

The GSFC GNCC's Folta-Quinn formation flying algorithm is a innovative technology that can be used in a closed-loop design to meet science and mission requirements of all low Earth orbiting formation flying missions. The algorithm is very robust in that it supports not only benign groundtrack control and relative separation control, but also demanding 3-D control for inclination and non-Keplerian transfers. To best meet the NMP requirements, this innovative technology is flying onboard the EO-1 spacecraft. The algorithm was successfully integrated into AutoCon™ for ground support validation, closed-loop onboard autonomy, as well as operational support. The application of this algorithm and the AutoCon™ system to other NASA programs is unlimited, as it applies to any orbit about any planet and can be used to fully explore the NASA mandate of faster, better, cheaper spacecraft.

REFERENCES

1. Frank Bauer, David Quinn, Kathy Hartman, David Folta, and John Bristow "Enhanced Formation Flying Experiments For The New Millennium Program Earth Orbiter (EO)-1 Mission - Challenging Technology Program Management", AIAA, 5/97

2. John Bristow, David Folta, Kate Hartman, and Jessie Leitner, "A Formation Flying Technology Vision", AIAA-2000-5194, February 2000

3. David Folta and David Quinn , "A Universal 3-D Method for Controlling the Relative Motion of Multiple Spacecraft in Any Orbit," Proceedings of the AIAA/AAS Astrodynamics Specialists Conference, August 10-12, Boston, MA.

4. Quinn, D.A. and D. C. Folta (1996) *Patent Rights Application and Derivations of AutonomousClosed Loop 3-Axis Navigation Control Of EO-1.*

5. Battin, R. (1987) *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, Chapters 9 and 11.

6 . Sperling, R. (1997) *AutoCon User's Guide Version 3.0, February 1998,* AI Solutions, Inc., Greenbelt, MD. 20770

7. Matlab, The Math Works, Inc, Users Guide, 1995

THE INTEGRATION, TESTING AND FLIGHT OF THE EO-1 GPS

David A. Quinn
NASA/Goddard Space Flight Center

Paul A. Sanneman
SWALES Aerospace

Seth E. Shulman
Computer Sciences Corporation

Jennifer A. Sager
Honeywell

ABSTRACT

The Global Positioning System has long been hailed as the wave of the future for autonomous on-board navigation of low Earth orbiting spacecraft despite the fact that relatively few spacecraft have actually employed it for this purpose. While several missions operated out of the Goddard Space Flight Center have flown GPS receivers on board, the New Millenium Program (NMP) Earth Orbiting-1 (EO-1) spacecraft is the first to employ GPS for active, autonomous on-board navigation. Since EO-1 was designed to employ GPS as its primary source of the navigation ephemeris, special care had to be taken during the integration phase of spacecraft construction to assure proper performance. This paper is a discussion of that process: a brief overview of how the GPS works, how it fits into the design of the EO-1 Attitude Control System (ACS), the steps taken to integrate the system into the EO-1 spacecraft, the ultimate on-orbit performance during launch and early operations of the EO-1 mission and the performance of the on-board GPS ephemeris versus the ground based ephemeris. Conclusions will include a discussion of the lessons learned.

INTRODUCTION

The Global Positioning System (GPS) is not a new idea. Neither is using the navigation capability of GPS for Low Earth Orbiting (LEO) spacecraft a new idea. Although the idea has been around for quite some time and many papers have been written regarding the uses of GPS on orbit, only a handful of spacecraft have actually used GPS receivers on-board LEO spacecraft for primary real time navigation control. After more than a decade of experiments, the Earth Observing –1 (EO-1) spacecraft has become the first satellite designed and built at NASA's Goddard Space Flight Center to baseline GPS as the primary means of performing on-board real time navigation control.

GPS On-Orbit History

The first GPS receivers to fly on orbit were aboard NASAs Earth resources satellites Landsat-4 (launched in July of 1982) and Landsat-5 (launched in March of 1984). While both of these satellites possessed the capability to use GPS data as their primary source of ephemeris state information, at that time (the early 1980's) less than half of the GPS constellation was flying, so reliable navigation fixes were simply not possible over the prolonged times required by the Landsat mission. For these two spacecraft, uploaded ephemeris files were the main means of navigation and so GPS amounted to little more than an experimental footnote. Ten years later and although the GPS constellation was complete it was still not being used for active on-board navigation and attitude reference frame computations. Since that time, a host of spacecraft bearing GPS receivers have followed, most of which included GPS for experimental purposes only. **Table 1** [Ref 1] shows an extensive but by no means exhaustive list of such spacecraft.

Launch Target	Spacecraft	GPS Receiver	GPS Function
7-1982	Landsat-4	Motorola GPSPAC	Constellation incomplete - Experimental
3-1984	Landsat-5	Motorola GPSPAC	Constellation incomplete - Experimental
7-1991	ORBCOMM-X	--	Launch Failure
6-1992	EUVE	Motorola GPSDR	GPS Orbit Determination Experiment
8-1992	TOPEX/Poseidon	Motorola GPSDR	GPS Orbit Determination Experiment
multiple	Space Shuttle	multiple	Nav, Relative Nav & Att Experiments
6-1993	RADCAL	TANS Quadrex	Post Flight Attitude Experiment

7-1993	ORFEUS-SPAS-1	Alcatel/SEL	
9-1993	PoSat-1	TANS	Orbit Determination
2-1994	OREX	GPSDR	
3-1994	DARPASAT	AST-V	Orbit Determination
5-1994	TAOS/STEP-0	AST-V	Orbit Determination
5-1994	STEP-2	AST-V	Orbit Determination
11-1994	CRISTA-SPAS	TANS Vector	Attitude & Orbit Determination
'92 & '94	COMET	Ashtech SB24	
1-1995	Faisat-1	--	
3-1995	SFU	GPSR	
4-1995	ORBCOMM-FM1	TANS II	
4-1995	ORBCOMM-FM2	TANS II	
4-1995	OrbView-1	TurboStar	
9-1995	Wake Shield Facility-02	TurboStar & Tensor	Attitude & Orbit Determination
1-1996	GADACS / SPARTAN	Two TANS Vectors	Receiver Software Error
5-1996	GANE / STS-77	TANS Vector	Real-time Nav & Att Experiments
5-1996	MSTI-3	Viceroy	
5-1996	MOMS-2P	Viceroy	
11-1996	HETE	SEXTANT	
11-1996	Wake Shield Facility-03	multiple	Navigation Experiment
11-1996	ORFEUS-SPAS	Tensor	Relative Nav Experiment
2-1997	HALCA	GPS	
3-1997	Zeya	GPS and GLONASS	
8-1997	OrbView-2	redundant Viceroys	
8-1997	SSTI Lewis	Two Tensors	S/C Failed shortly after launch
9-1997	Faisat-2v	--	
9-1997	IRS-1D	--	
10-1997	Falcon Gold	TIDGET	
10-1997	YES	TANS II	Orbit Determination above GPS Constellation
11-1997	ETS-VII	--	Rendezvous and Docking Experiment
12-1997	Equator-S	Viceroy	Orbit Determination above GPS Constellation
12-1997	EarlyBird	Vector and Viceroy	
2-1998	GFO	Four TurboStars	Precise Orbit Determination
2-1998	Globalstar	Tensor	Orbit Determination and Time
2-1998	SNOE	MicroGPS (Bitgrabber)	
7-1998	FASat-Bravo	TANS II	
7-1998	TMSat-1	SGR-10	
10-1998	SEDSat-1	Ashtech G12	
10-1998	ARD	--	
11-1998	International Space Station	SIGI	Operational Navigation Attitude & Time
12-1998	SAC-A	TANS Vector	Orbit Determination & Attitude Experiment
1-1999	ORSTED	TANS, TurboStar	
2-1999	ARGOS	--	
2-1999	SUNSAT	TurboStar	Orbit Determination
4-1999	UoSAT-12	SGR-20	Orbit Determination
4-1999	Ikonos-1	Rockwell C/A code	
5-1999	IRS-P4 (OceanSat)	--	
6-1999	QuikSCAT	2 Viceroys	
9-1999	SRTM	Blackjack	
9-1999	JAWSAT	TANS Vector	
late 1999	AMSAT Phase 3D	Two TANS Vectors	Orbit & Attitude Det. above GPS Constellation
1999	STRV-C	Blackjack	
1998	ARISTOTELES	--	
1997	EOS-A	--	
1998	EOS-B	--	
1998	TSX-5	two TANS Vectors	Orbit Determination
1999	SAC-C	Lagrange, Tensor, Bitgrabber	
1999	QuickBird	2 Viceroys	
(late '90s)	European Polar Platform	--	
(late '90s)	RAMOS	--	
(1999)	Gravity Probe B	2 Vectors	Has yet to launch
(12-1999)	CHAMP	Blackjack	
(1999)	OSEM	Tensor or TANS Vector	

(5-2000)	Jason-1	Blackjack	Has yet to launch (summer 2001?)
(Aug 2000)	VCL	Blackjack	Has yet to launch (2002 ?)
(2000)	BIRD	Rockwell Collins	

Table 1: An incomplete history of GPS receivers in space.

...and there are plenty more to follow. In 1995 NASA Management Instructions directed that the lowest cost navigation system should be implemented on all NASA spacecraft. Since GPS was seen as a low cost alternative to standard practices, this instruction (which has since been superceded by the NASA Operational Directive Information System) was promptly interpreted in some quarters as to mean that all LEO spacecraft should include GPS receivers. Such was the origin of a GPS capability on NASA the New Millenium Program Earth Observing spacecraft EO-1. This spacecraft was to be a testbed for demonstrating new concepts in autonomy and advanced technologies and thus was a natural choice for showcasing GPS capabilities on orbit.

EO-1 ACS DESIGN

The EO-1 Attitude Control Subsystem establishes and maintains the pointing of the EO-1 spacecraft science instruments at desired targets while allowing for the proper orientation and operation of other spacecraft subsystems such as power, thermal and communications. Attitude control functions are performed by a suite of on-board attitude sensors, control actuators and mathematical control laws residing in a Mongoose 5 (M5) main spacecraft computer. Safe Hold Mode (SHM) controllers and electrical interfaces to most components is provided by the Attitude Control Electronics (ACE).

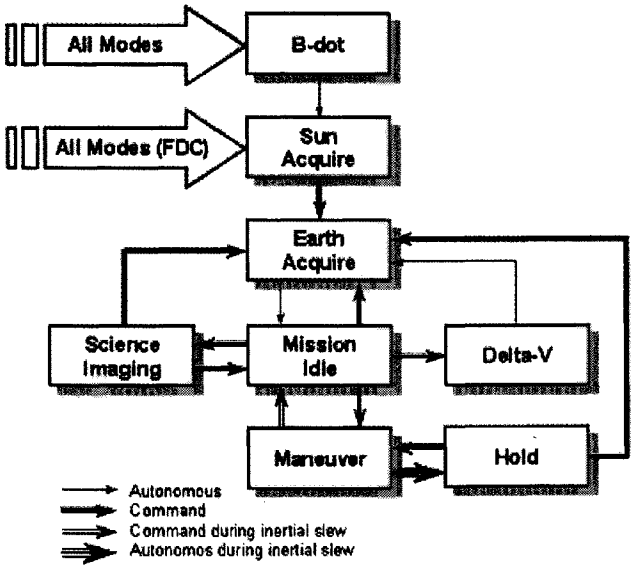


Figure 1: EO-1 ACS Mode Transitions

The EO-1 ACS flight software was derived from the GSFC Tropical Rainfall Measurement Mission (TRMM) design. For EO-1, the main ACS software on the M5 processor was to have attitude determination and closed-loop control modes for: magnetic de-spin following separation from the Delta launch vehicle, initial stabilization and sun acquisition, nadir-pointed science data collection and downlink, thruster maneuvers (for delta-V), and sun/moon slew/scan maneuvers for instrument calibrations.

The ACS flight software mode transitions are illustrated in **Figure 1**. Following separation from the launch vehicle, the ACS was to null the tip-off rates via a B-dot magnetic control law and stabilize the spacecraft before, during and after solar array deployment. During initial sun acquisition and "safe-hold" operations, the spacecraft would maintain an inertially fixed, solar-pointing attitude with the instruments away from the sun. During normal operations, the body-fixed science instruments point toward the earth with the center of their FOVs along the nadir axis. During sun and moon calibrations, the instruments point toward those bodies.

During launch, the GPS receiver would be powered off. A predicted ephemeris would be uploaded prior to launch as part of the pre-launch Absolute Time Sequence (ATS). This pre-launch orbit initialization was not a requirement for launch since orbit knowledge is not required for either B-dot or Sun Acquire modes. The initial predicted orbit would require further updating once the spacecraft was on orbit to account for any deviation in the launch vehicle trajectory. Shortly after launch and separation, the GPS would be powered on, configured and initialized for orbital operations. Until the GPS could be determined to be functioning on orbit, ephemeris uploads from the ground would be the method of choice for computing navigation and attitude reference solutions. Once outputs from the GPS were proven to be accurate, its navigation solutions would replace the ground loaded ephemeris in the active control loop.

GPS AS PART OF EO-1 ACS

Part of the charter of the New Millenium Program (NMP), of which EO-1 is a part, is to demonstrate new technologies that can be used to improve performance or lower the cost of future spacecraft . While GPS is hardly a new technology, until EO-1, no spacecraft operated by the Goddard Space Flight Center had ever been baselined to employ an on-board GPS receiver as the primary method of ephemeris estimation. The on board ephemeris estimation would be used to determine the attitude reference frame against which attitude measurements are compared. EO-1 would be the first spacecraft to employ GPS in this operational capacity. The space-qualified Space Systems/Loral Tensor™ was the unit selected to fly on EO-1. It is capable of providing 1 Hz position and velocity data in Earth Centered Inertial (ECI-J2000) coordinates, along with an absolute time reference. With internal Kalman filter software enhancements, the 3-sigma accuracies for position and velocity are ±150 meters and ±0.55 meter/second respectively. The one Pulse Per Second (PPS) coordinated universal time (UTC) reference output is accurate to ±1 microsecond.

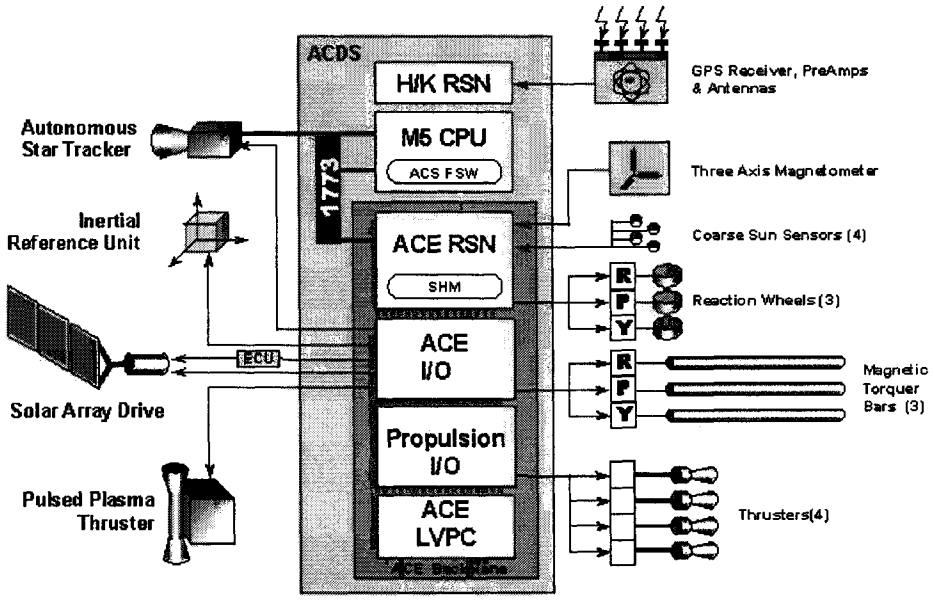


Figure 2: EO-1 Attitude Control System Components

The EO-1 design called for GPS data to be transferred via serial digital link to the Housekeeping RSN, and then read by the main ACS M5 processor over the 1773 bus shown in **Figure 2** Because the Tensor™ receiver was designed for a different power bus, a separate Power Conditioning Unit (PCU) was needed to provide the receiver 29±3 VDC from the 28±7 VDC EO-1main power bus. The components of the EO-1 GPS are shown in Figure 3.

The Tensor™ is a nine channel dual receiver. That is to say, the off the shelf unit consists of two redundant GPS receivers each of which can accumulate and process data from up to nine GPS SVs simultaneously. EO-1 required only one side of this receiver (as shown in Figure 3) so that the Tensor™ (as delivered by Space Systems/Loral) was split into two single string units.

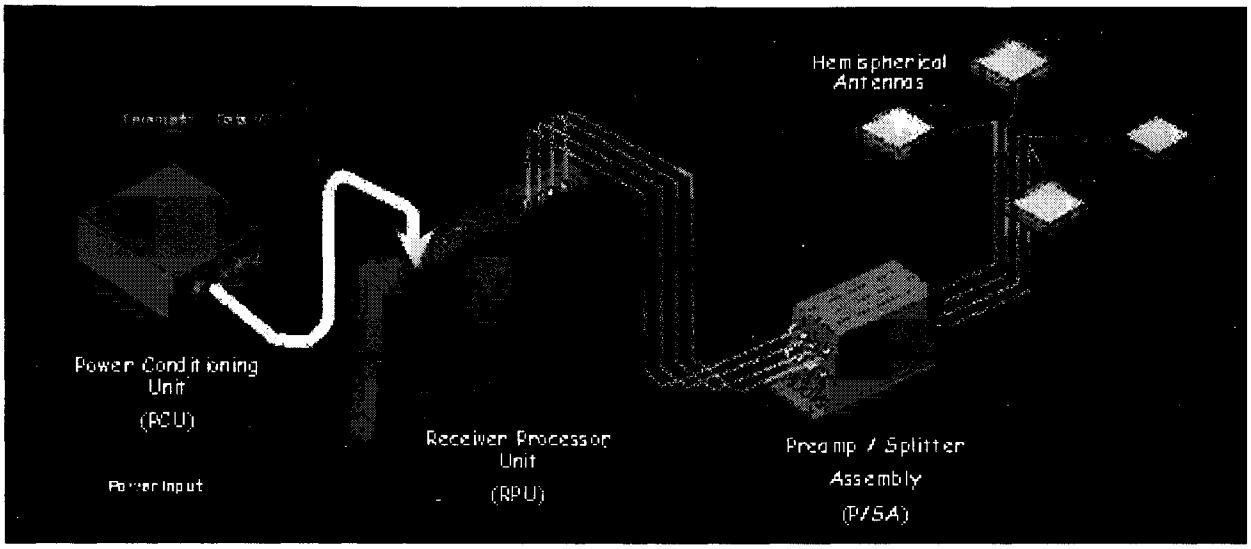


Figure 3: EO-1 Global Positioning System Components

When in operation, the GPS is primarily solving for x-y-z position and time. This is why at least four GPS Space Vehicles (SVs) must be in view providing four independent distance measurements to the receiver before an navigation solution can be attempted. If four SVs are in view, the receiver can determine the distance between the user spacecraft and each of the four GPS SVs at the broadcast time (all GPS SVs broadcast the current time). Since the locations of the GPS SVs are extremely well known, the user's position can be determined to a high degree of accuracy. This position and time data are then employed to resolve the orbital velocity. Since it is a derived quantity, velocity tends to be somewhat noisier than either time or position estimates.

Before any GPS receiver can begin computing navigation solutions, it must know three things. First it must know what time it is, it must know where it is and it must know where all the GPS SVs are as well. Only then can it determine which GPS SVs are visible and of those which are the best to use. The current time is simply that: the GPS time which is specified to be within ±1 nanosecond of UTC proper and is continually broadcast by all the SVs in the GPS constellation – if the receiver can acquire signals from even one GPS SV, it can get the time. The location of all the GPS SVs in orbit is contained in the GPS Almanac which specifies the orbital elements of all (up to 32) of the GPS SVs at a known epoch. This data is also simultaneously broadcast by all SVs in the GPS constellation and repeats every 12.5 minutes. If the receiver can lock up on even one SV for that long, it can acquire the position of all the other SVs in the constellation. Almanac files are also archived on the world wide web by the U.S. Coast Guard. The user location data is referred as the User Ephemeris data. Providing the receiver initial estimates of Time, Almanac and Ephemeris data allows the receiver to begin computing solutions within five minutes and is referred to as a 'warm start'. Waiting for the receiver to collect the time and Almanac data and then find four SVs on its own to begin computing navigation solutions is called a 'cold start' and can take considerably longer to accomplish.

There are two switches that must be configured before GPS outputs would be permitted to control the EO-1 spacecraft on orbit. The first enables ACS processing of on-board GPS ephemeris data so it can be monitored in telemetry, but would not enable the use of the GPS solution for orbit determination. The orbit solution propagated from the uploaded ground ephemeris data would continue to be used for on board orbit determination. The second switch would enable the use of the GPS solution for orbit determination so long as the solution is deemed valid. A GPS solution is considered valid if the GPS data packet containing the solution is valid, if the GPS packet checksum is good and if the health status is good and has been for at least five cycles. If all these conditions are met and the GPS propagated data passes the continuity checks, then and only then could the GPS solution be used for active on-board orbit determination.

When GPS is being used for on-board orbit determination, it provides ephemeris updates at 1 Hz. Each GPS data packet must pass all validity checks and the ephemeris data, propagated to the current cycle time, must pass the

same continuity checks described above for the ephemeris upload command. In this case, the position continuity tolerance is set to 50 km and the velocity continuity tolerance to 0.5 km/sec. If the GPS packet is invalid, the GPS packet checksum is bad, the GPS health status word is bad or the GPS propagated data does not pass the continuity checks, a GPS failure counter is incremented. The last valid ground based orbit solution is always propagated to the current time, and this solution is used instead of the GPS solution when the GPS solution fails at least one of the validity checks. The next valid GPS solution will clear the failure counter to zero, but if the GPS failure counter reaches the limit of 60 counts (this was later changed to 120 counts), GPS processing would be disabled.

EO-1 GPS INTEGRATION & TEST

Since the December 31, 1997 delivery of the GPS flight hardware from Space Systems / Loral to the GSFC Guidance Navigation & Control Center (and the subsequent acceptance testing performed at Goddard by Space Systems / Loral prior to formal acceptance), the unit successfully completed an extensive battery of tests designed to validate system compliance with design specifications and mission requirements as well as to characterize system performance under various circumstances which were expected to be encountered by the EO-1 spacecraft. This testing was accomplished in five phases over a period of several months and executed in order of importance to the project subject to resource availability

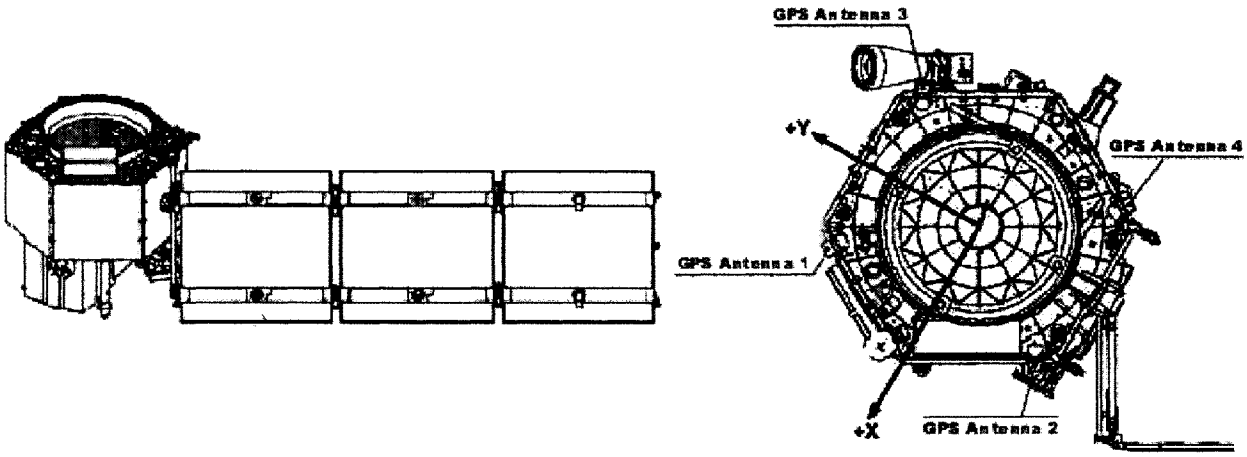


Figure 4: EO-1 Spacecraft and the Locations of the GPS Antennas on the Zenith Deck.

Phase I: Requirements Compliance (Performance) Testing:

The first sequence of tests on the EO-1 GPS Flight Receiver employed the GPS Simulator. For these tests, a single 6-hour simulator scenario was assembled which modeled the expected EO-1 orbital characteristics:

SemiMajor Axis:	7083.0 km	Longitude of A.N:	61.2 deg
Eccentricty:	0.00054	Argument of Perigee:	-37.9 deg (322.1 deg)
Inclination:	98.2 deg	Mean Anomaly:	0.00 deg

Epoch: 20-Mar-1998 00:00:00 UTC

The spacecraft characteristics were modeled to conform with those expected of the EO-1 spacecraft, the most significant of which is the non-nominal antenna pointing directions. Unlike standard GPS antenna orientations, where all antennas point in the expected zenith direction, the EO-1 GPS antennas were all canted 45° away from zenith so as to avoid interference and obscuration by the payload attach fitting (PAF) also mounted to the zenith deck of the spacecraft a seen in **Figure 4**. The PAF was present to attach EO-1 to the Delta launch vehicle. To model the EO-1 GPS antennas, the following locations and orientations for the antenna centers were observed:

	x-center	y-center	z-center	Azimuth	Elevation
Antenna-1:	0.3955 m	0.5927 m	-0.5040 m	30 deg	-45 deg
Antenna-2:	0.3955 m	0.5173 m	-0.5040 m	330 deg	-45 deg
Antenna-3:	-0.4705 m	0.5927 m	-0.5040 m	150 deg	-45 deg
Antenna-4:	-0.4705 m	-0.5173 m	-0.5040 m	210 deg	-45 deg

The x-y-z centers locate the center of the antenna in the spacecraft navigation frame (origin at the spacecraft center of mass) and where azimuth is measured in the navigation x-y plane with the x-axis as a zero reference and is positive towards the y-axis (counter-clockwise about the +z-axis). Elevation is measured off the x-y plane and is positive towards the positive z-axis.

A zero attitude error Earth pointing simulation scenario beginning at the design epoch and lasting for 6-hours was constructed. All subsequent Earth pointing tests in Phase I were executed using this scenario. For each test of the flight receivers (both sides were tested in parallel even though only one side would ultimately fly), they were powered up and placed in the same 'test specific' configuration. Almanac data was then provided to the receivers via laptop computer, along with initial estimations of ephemeris and time, thus allowing the unit to execute what will be henceforth referred to as a 'warm start'. This provides the receiver with enough information to attain an immediate (< 5mins) lock on the GPS constellation and begin providing navigation solutions, thus streamlining the acquisition process by avoiding the time consuming 'cold start' process which starts the receiver from an unknown navigation state and requires the receiver to initiate a random search for GPS satellites. The point of this first series of tests was performance evaluation, not cold start characterization, which was to be the subject of later testing.

Configuration Test 1: Default

For the first test, the default mode was tested. In the default mode, the receiver is free to switch back and forth between the Single Point Solution (SPS) and the embedded Kalman Filter outputs. This configuration assumes that the attitude function of the receiver should be switched on, even though attitude outputs were not to be employed by EO-1. In fact, since the attitude portion of the receiver software assumes that all antennas are pointed directly toward zenith, the canted configuration of the EO-1 antennas renders the attitude computations of the receiver virtually useless.

Configuration Test 2: Attitude Off

For the second test, attitude computations were switched off, demonstrating that there was no difference in navigation performance of the receiver whether the attitude computations are left on or switched off.

Configuration Test 3: Force SPS

For the third run, the embedded Kalman filter was not employed, yielding SPS data only for navigation. While in this mode, the receiver collects data from the GPS constellation and computes one navigation solution at a time (thus the name Single Point Solution). Each solution is independent of the last so solutions can be generated in this mode only when four or more GPS Space Vehicles (GPS-SVs) are visible. During those occasional periods when geometry is not sufficient to compute a single point solution, no data is output and the receiver is said to experience a 'navigation outage'. This test revealed outages to be a fairly frequent occurrence averaging five or more brief outages per orbit.

Configuration Test 4: Force Filter Solutions

For the last run, solutions from the filter was were output exclusively. As might be expected performance with the filter is considerably better than exclusive use of SPS data. Not only does the filter smooth the SPS data reducing overall errors, but the memory inherent in a Kalman filter allows the receiver to 'flywheel' through those outage periods were SPS solutions are unavailable. Only when SPS solutions are unavailable for an extended period of time does the filter show any signs of diverging. This can be seen by comparing the SPS plots (which show the SPS navigation outages) with the filtered output.

Conclusion of Phase I:

This first series of tests demonstrated that the Loral Tensor™ met the EO-1 navigation requirement and would meet the mission objectives [Ref 3].

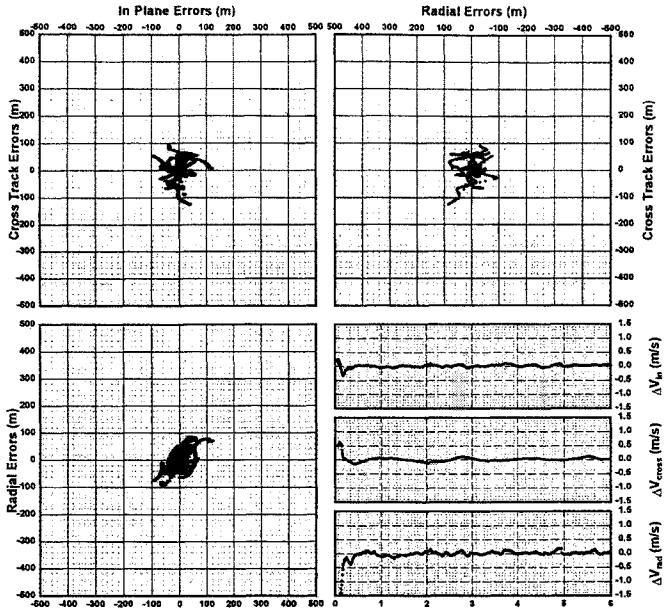


Figure 5: 3D Filtered Position & Velocity Errors from the Phase I simulation study.

	In-Track	Cross-Track	Radial	
Position:	100	130	100	m, 3σ
Velocity:	0.15	0.25	0.25	m/s, 3σ
Stability:	15.0	3.0	5.0	m over 1 sec intervals, 3σ

Table 2: Tabulated Results from the Phase I simulation study.

Phase II: Integration & Testing Plan Development:

The second sequence of tests validated the Ground Support Equipment (GSE) and procedures to be employed during Integration and Test (I&T) process. Each of these tests were first run in the lab in an effort to assure that all necessary procedures had been correctly developed, that all GSE hardware was available and functional, and to pre-quantify what to expect during the actual I&T process.

Test 1: Component Throughput / Aliveness:

The purpose of this set of tests was to develop an ability to verify that proper Radio Frequency (RF) power levels reach the receiver. Replicating the conditions expected during the real I&T process required a GPS antenna on the roof of the building (with a clear line of sight to the real GPS constellation) with enough RF cable to allow this rooftop antenna to act as an RF source for the flight receiver once mounted to the spacecraft (in the I&T test facility). This was expected to require no more than 120 feet of RF cable. Compensating for losses in such a long cable (which is not representative of the flight configuration) required the inclusion of a pair of powered RF amplifiers to boost the signal. The resulting GSE setup employed a rooftop GPS antenna followed by a 3 meter run of RG-142 coaxial cable. This led to a pair of powered amplifiers whose output was sent to the test chamber via approximately 200 ft of Belden 9913 low noise cable. The end of this cable was then connected to the spacecraft Preamp / Splitter Assembly (P/SA), providing signal to the receiver. This setup would allow the testing of P/SA and RPU and flight cable throughput, bypassing only the flight antennas.

Test 2: System Throughput / Aliveness:

The purpose of this test set was to develop an ability to verify proper RF power levels reach the receiver when connected to the P/SA via the flight antennas. Replicating the conditions expected during the real I&T test required the same GSE set-up as the first test, with one important addition. Here again, RF signals are captured via a rooftop antenna, amplified, passed through the long GSE cable and ultimately fed into the flight system. In this case, the flight antennas are not bypassed, but instead the signal is fed into a non-flight source antenna which re-radiates RF signal into one of the four flight antennas. The non flight source antenna is held steady about 1 inch from the flight antennas by a mechanical fixture hereafter referred to as a 'hat coupler' (four of which were constructed, one for each of the four flight antennas). The flight antennas are then connected to the Preamp via the flight cables. In this way, throughput of the complete flight system can be verified against the real GPS constellation.

The test was virtually identical to the one above, with the exception that the physical set-up was as described above, and the GSE hardware had already been constructed and its performance quantified as part of the first tests. The first objective of these tests was to demonstrate that the antenna-to-antenna RF 're-radiation' concept provided sufficient RF input to the GPS receiver for it to properly perform its navigation function. The second objective was to establish any sensitivity difference between this method of input to the GPS receiver and the method employed by the first set of tests (which by-passed the antennas). The third (and primary) objective was to verify proper RF throughput of the complete EO-1 Flight GPS, that is through all four antennas. All of these objectives were successfully accomplished.

Test 3: System Functional Performance:

As a final check, orbital performance was tested using the GPS RF simulator employing the same scenario developed for the Phase-I study. For this performance test, the simulator output is to be connected to the four GSE hat couplers, and the generated signal re-radiated into the four flight antennas in a manner identical to that employed during the previous test. Here, the configuration was the same as before, only now the simulator provided RF signals to all four antennas simultaneously. Based on this work, it was determined that the I&T procedure previously developed could be used to warm start the receiver for performance verification of EO-1 GPS during I&T.

Phase III: Timing Performance Testing:

The third sequence of tests on the EO-1 GPS Flight Receiver involved testing the synchronization Pulse Per Second (PPS) output of the receiver. Specification documentation [Ref 2] from Loral states the following:

> "An RS-422 pulse shall be output within ±1.0 μsec of every GPS second"

Later, the same document states:

> " The PPS leading edge occurs sometime in the 999th millisecond of an integer clock time. The countdown process to generate the leading edge is performed in that last millisecond. The Pulse Per Second, PPS, signal is output once per second with an accuracy within one microsecond with respect to GPS time."

Which means there are two quantities which needed to be verified, one being the offset of the receiver PPS from the GPS-SV time (which is itself specified to be within 100 nanoseconds of the theoretical 'absolute' UTC) and the time difference between consecutive PPS pulses. Not all of this could be tested with the equipment residing in the GSFC GPS facility at the time. What was tested was the time difference between the trailing edge of a plus PPS output and the leading edge of the corresponding minus PPS output which was measured for every other pulse pair. In the past, taking measurements exclusively off of either plus or minus PPS outputs allowed measurements to occur only for every forth pulse, so this approach is seen as an improvement over previously employed techniques. Since the same physical oscillator drives all four PPS outputs of a given receiver (plus-1, minus-1, plus-2, minus-2), one six hour test was executed on both the A and B receivers (both of which were warm-started per the procedure developed during Phase II).

Timing tests of the two sides of the GPS Tensor™ receiver were completed to the level possible with the test equipment present in the GPS facility. The characteristics of the PPS alignment to absolute time was not measurable

in the GPS Test facility, while the PPS period was. The characteristics of the PPS period for the A and B sides of the Tensor™ were found to be as follows:

	Pulse Width	Period	Mean Error (μ)	Stand. Dev (σ)
Side-A:	0.978 μsec	1 sec ± 1 msec	0.01423 μsec	0.3791 μsec
Side-B:	0.978 μsec	1 sec ± 1 msec	0.01416 μsec	0.4152 μsec

Table 3: Tabulated Results from the Phase III timing tests.

Phase IV: Other Simulation studies:

While no requirements were levied on the performance of the GPS receiver during non-nominal operational modes, it was considered wise to characterize the performance during any such modes that could be expected to be encountered during the run of the EO-1 mission.

Characterization Test 1 - Sunpoint

For this simulation study, the same scenario from Phase I was employed. In it, the theoretical spacecraft was held inertialy fixed, (attitude such that the Sun remained at a fixed position in the EO-1 body frame). This was accomplished by starting the spacecraft at the Earth Pointing attitude at the beginning of the six hour run (using a warm start to quickly lock up on a navigation solution) only this time commanding the simulator to hold the Sun fixed in the body frame representing a transition from Earth Point to Sun Point. The results of this run demonstrated that the receiver in the nominal EO-1 flight configuration should be able to maintain navigation solutions with the following general characteristics:

	In-Track	Cross Track	Radial	
Position	320	120	250	m, 3σ
Velocity:	0.35	0.20	0.40	m/s, 3σ

Table 4: Tabulated Results from the Phase IV characterization tests.

Again, since there were no requirements levied on the receiver performance while in the Sun Pointing safehold attitude, these numbers were provided for the purpose of characterization only.

Characterization Test 2 - J2000 Confirmation

For this study, the theoretical spacecraft was run in the same Earth pointing mode as it was during the performance simulation studies. The key difference here being that both sides of the receiver were configured to output the navigation data in the ECI-J2000 (Earth Centered Inertial) coordinate frame rather than the default ECI-True-of-Date frame. The results of this test demonstrated that the receiver performance while configured to output navigation solutions in the J2000 coordinate frame was comparable to the True-of-Date solutions cited earlier to validate performance. This indicated that the proper rotations from Earth Centered Earth Fixed (ECEF) to True-of-Date to ECI-J2000 was occurring in both sides of the receiver, thereby validating the receiver performance in all coordinate frames.

Characterization Test 3 - Spacecraft Slews

For this study, the same scenario was again employed however, both sides of the receiver were warm started, and a set of slews executed at every descending node. For each of four descending nodes a different slew profile was executed which characterizes all of the different slew profiles likely to be encountered by the EO-1 spacecraft. To allow sufficient recovery time after the last slew set, the run was extended to seven hours in duration instead of the usual six. The four slew sets executed were as follows:

Slew Set 1: At (DN-1)-22.5 minutes, execute a 6 min +90° Pitch (up).
At (DN-1)+12.5 minutes, execute a 6 min -90° Pitch (recovery).

Slew Set 2: At (DN-1)-22.5 minutes, execute a 6 min -90° Pitch (down).
At (DN-1)+12.5 minutes, execute a 6 min +90° Pitch (recovery).

Slew Set 3: At (DN-1)-22.5 minutes, execute a 6 min +90° Roll (right).
At (DN-1)+12.5 minutes, execute a 6 min -90° Roll (recovery).

Slew Set 4: At (DN-1)-22.5 minutes, execute a 6 min -90° Roll (left).
At (DN-1)+12.5 minutes, execute a 6 min +90° Roll (recovery).

While the receiver did not have time for the filter to fully converge before the first slew was executed, it became clear that the receiver in the flight configuration was quite capable of outputting valid navigation solutions during and after all the slew profiles even though there was no requirement for the receiver to be able to do so (in fact the baseline plan was to disable the use of GPS navigation data during the slews).

Characterization Tests 4 & 5 - Cold Start: Earth Point / Sun Point

The point of this pair of tests was to measure the time required to cold start the receiver in both the Earth pointing and Sun pointing attitudes. Limitations of the GPS Simulator prevented the execution of these final two tests before delivery of the flight unit to the EO-1 project. The GPS Simulator can simulate RF signals from only ten of the 32 GPS Space Vehicles (SVs) at any one time. The selection of which SVs the simulator provides is based upon the simulators knowledge of the users location and the assumption that the user antennas are aligned along the zenith direction. In the case of the canted orientations of the EO-1 antennas, this assumption is invalid. From a cold start, the receiver employs only one antenna at a time in a random search for nine of the GPS SVs. Since the antennas for EO-1 are not pointed in the zenith direction, the simulator is modeling SVs that the search antenna cannot see. Combine that with the random nature of the search employed by the receiver, and it becomes clear that the receiver may be looking for SVs that are not being modeled while many of the SVs being modeled cannot be seen. Any statements on the cold start performance of the receiver under these conditions would not be a fair assessment of the receivers capabilities for either the Earth pointing or Sun pointing cases.

Phase V: RPU/PCU Electrical Characterization:

The final phase of the testing involved confirming that the GPS receiver functions properly while obtaining its power from the Power Conditioning Unit (PCU) designed to provided proper power to the system while in flight. Up until now, all testing has been done with power provided to the GPS receiver from a GSE power supply. Since the spacecraft main power bus was deemed incompatible with the specifications of the GPS receiver, a power conditioner had to be designed to bridge the gap. This last set of tests were designed to ensure that not only does the PCU perform as expected, but that the GPS receiver performs in the presence of the PCU and to demonstrate the in-rush characteristics of the integrated GPS. The test setup is shown in **Figure 6**.

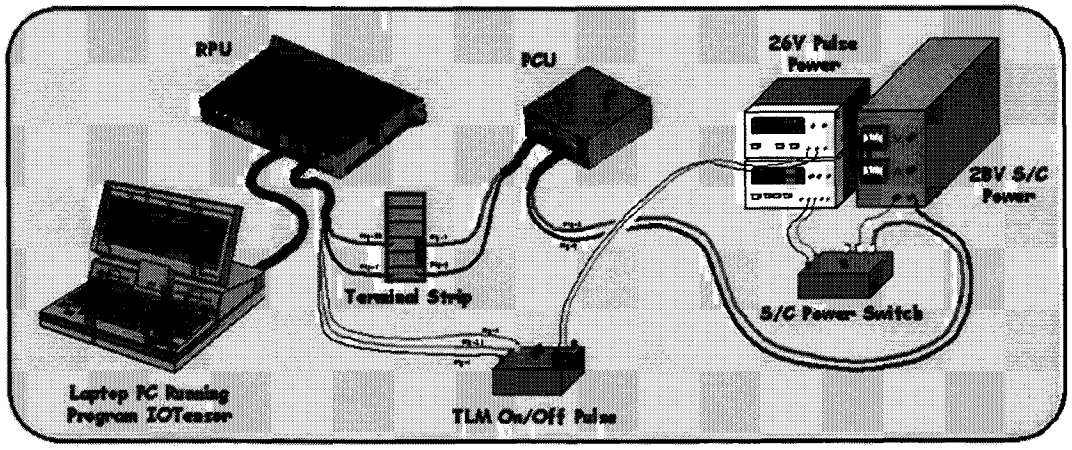


Figure 6: Test setup used to characterize in-rush and check PCU performance.

The simplest analogy to draw upon here is a simple desk lamp, which has a plug and a switch. To turn the light on, it is possible to turn the switch on and then plug the light in, or to plug the light in (with the switch off) and then turn the switch on. In the case of the GPS receiver, the connection to the 28 VDC main power bus is the 'plug', and the relay internal to the RPU is the 'switch'. Therefore, testing all the possibilities involves three separate tests:

1. RPU Relay = ON; S/C Power = OFF; S/C Power from OFF to ON (plug it in with the switch on).
2. RPU Relay = OFF; S/C Power = OFF; S/C Power from OFF to ON (plug it in with the switch off).
3. RPU Relay = OFF; S/C Power = ON; RPU Relay from OFF to ON (switch it on while plugged in).

Each of which were performed three times on each RPU. The first set was run on the RPU alone, so as to distinguish between characteristics of the RPU and those imposed by the presence of the PCU. The second set was run on the RPU/PCU combination with the probe located at the output of the PCU. The third set was run on the RPU/PCU combination with the probe located at the input of the PCU. During each test, a digital oscilloscope was employed to capture the curves which characterize the behavior of the system and were then included in a set of drawings which show the test set-up and resulting output of each test. The full test set was run first on the Engineering Test Unit (ETU) in combination with an early PCU breadboard many months before the delivery of either flight unit. These same tests were then performed on the RPU Side-A (the flight unit), and finally on the RPU Side-B (the flight spare) in combination with the flight PCU.

The ETU test set was run not only to test the PCU design (in breadboard), but to be sure the test was sufficient to fully characterize the flight unit, as well as to provide a set of baseline expectations. Comparison between tests yields a few notable differences between the ETU tests and the flight unit tests. For each Test-1, an oscillation is present on the plots of the flight unit that is not seen in the corresponding ETU test. This is due to the nature of the switch used to apply the 28 VDC spacecraft power. For the ETU tests, a solid state switch was used. This switch provided a clean start-up but introduced series resistance which would slightly lower the in-rush current. For the flight unit tests, a mercury switch was employed which has a noisier turn on characteristic, but would not limit the current as much as the solid state switch.

Other differences are best characterized by direct comparison of corresponding Test-3's which shows the relay applying power to the receiver. In the case of the ETU, a high frequency oscillation is present near the end of the power on cycle which is not present in either of the flight unit tests. This simply means that the relays in the flight unit are much cleaner (and therefore better) than those present in the ETU.

Conclusion of Phase V:

The in-rush tests performed fully characterized the power characteristics of the Receiver/PCU combination intended for flight on the EO-1 spacecraft. Up until this point, the two sides of the receiver were bolted to one another in a single fully redundant unit. At this time, the two sides of the receiver were separated from one another and a set of validation tests demonstrated that the separation was successful into dividing the formerly redundant dual receiver into two single string receivers dubbed side-A (the flight unit) and side-B (the flight spare). The GPS flight hardware (side-A) was then delivered to the EO-1 project on June 22, 1998 and subsequently integrated into the spacecraft. Once integration of the GPS hardware was complete, the three Phase II tests were repeated this time using the spacecraft power as a power source to the PCU and the GPS receiver. This repeat proved that the Receiver/PCU combination would operate as expected and that the GPS hardware had been successfully integrated into the EO-1 spacecraft.

Since all simulator testing to date employed the same scenario developed earlier, it was mutually agreed by all involved that checking the system against another epoch would be a wise precaution and so another scenario was developed. This new scenario was run using the flight spare receiver (so-called Tensor B) before the simulator was moved over to the spacecraft staging area. In this way, not only could the scenario be checked, but the performance of the flight receiver anticipated. This scenario would accomplish several goals in a single stoke: it would allow the receiver to demonstrate its performance at a time other than the original test time, it would verify receiver performance in an epoch after the GPS system rollover (which occurred the morning of August 22, 1999), it would demonstrate the ability of the receiver to fly through the transition to the year 2000 (the so-called Y2K issue) and finally, it would provide the ground its first opportunity to command, configure and warm start the receiver without help from any GSE whatsoever.

Having completed these tests, if was determined that the integrated EO-1 GPS was functioning in a manner consistent with the performance seen prior to delivery, that is to say the integrated EO-1 flight GPS system was functioning properly and the ground system had demonstrated its ability to correctly interface with the GPS to command and configure the system as well as receive and employ navigation from the system. The EO-1 GPS was now ready for spacecraft level testing in the thermal vacuum chamber.

Thermal Vacuum Test 1:

During the first run at thermal vacuum testing (November 1999), the GPS was run for prolonged periods (+24 hours) of time at both hot and cold plateaus. The system performed well, when it was provided sufficient signal from the GSE setup used to relay signals from the roof of the building into the test chamber. Several dropouts were seen and a problem report generated.

GSE Upgrade:

As a result of unrelated piece part problems with the EO-1 spacecraft, thermal vacuum testing would need to be run again. It was decided that this time a lesson would be learned from the first set of tests and the RF setup used to carry GPS constellation signals from the roof to the test chamber would be modified so as to minimize the possibility of interruptions or dropouts. The GSE setup used during TV-1 employed a rooftop GPS antenna followed by a 3 meter run of RG-142 coaxial cable. This led to a pair of power amplifiers whose output was sent to the test chamber via approximately 200 ft of Belden 9913 low noise cable. The end of this cable was then connected to one of four inputs to the test chamber. Four 6 meter runs of RG-142 cable connected the inputs to the chamber to the flight antennas via four hat couplers spaced approximately 6" from the flight antenna.

RF experts examined the TV-1 setup and found plenty of room for improvement. It was discovered that the signal losses associated with the RG-142 cable were significant over long distances, so it would be eliminated wherever possible. This was done by constructing a weather resistant box which housed the rooftop antenna externally and contained a low noise amplifier (Down East Microwave model 520LNA20WP) and a gain amplifier (Mini-Circuits model ZEL-1217LN), thus eliminating the 3 meter run of RG-142 originally used to keep the powered amplifiers from being exposed to the weather. This cable it was found, reduced the input signal by 3dB before it ever got to the first amplifier, so it's elimination was crucial to improved performance. The 200 ft run of Belden 9913 low noise cable was retained since 10dB lost in the run was more than made up for by the 40dB of combined amplifier gain provided by the previous change. Although they could not be eliminated altogether, the RG-142 cables in the test chamber itself were minimized to 3 meter runs. Finally, the gaps between the GSE antennas and flight antennas maintained by the four hat couplers was reduced to less than 0.5". The modified setup was then checked against the flight spare receiver before connecting it to the flight system. The net result of these changes were strong clear signal to the GPS receiver which allowed for consistently quick warm starts (~10 mins) and continuous lock on the GPS constellation. The setup was declared both safe and functional and was then connected to the flight system where behavior identical to that of the backup receiver was noted.

Thermal Vacuum 2:

For TV-2 four data sets would be collected. The first set (June 28, 2000) would baseline the behavior of the GPS connected to the new GSE setup while the spacecraft was outside the test chamber. The receiver was warm started and signal strength data was collected using a laptop PC receiving telemetry directly from the receiver via the EO-1 skin connector. A minimum of 10 minutes of data was collected along each of the four antenna paths into the flight receiver. The constellation provided uninterrupted navigation signals the entire time by maintaining constant lock through all four paths. Needing signals from only four GPS satellites at any time to provide a navigation estimate, the flight GPS maintained constant lock on no less than six (most commonly seven) throughout the test.

The second data set (July 7, 2000) was collected during a cold plateau phase of TV-2. Because the time was available, an hour's worth of data was collected through each antenna path even though a minimum of 10 minutes is all that is required. During this test as well, strong signal strengths from no less than six (most commonly seven or eight) GPS satellites were observed with no signal dropouts or interruptions to navigation.

The third data set (July 10, 2000) was collected during a hot plateau phase of TV-2. This time, there was insufficient time for the luxury of collecting a full hour from each antenna path, so an hour was collected from antennas 1 & 3 and 30 minutes from antennas 2 & 4. Here again, strong signal strengths from no less than seven (most commonly eight or nine – the maximum the receiver can process) GPS SVs were observed with no signal dropouts or interruptions to navigation.

Finally, the last data set was collected (July 21, 2000) after the spacecraft was removed from the thermal vacuum chamber. A minimum of 10 minutes worth of data was collected from each of the four antennas to replicate the pre-Thermal Vacuum test data and verify that there were no observable changes in the behavior of the system as a result of the thermal vacuum tests. The success of this test meant that the EO-1 GPS pre-launch testing was complete and ready for the spacecraft to be shipped to the Launch facility for the expected launch on October 17, 2000.

EO-1 LAUNCH & EARLY OPERATIONS

The EO-1 spacecraft was launched with its companion payload, SAC-C, aboard a Delta II 7320 ELV on Tuesday, November 21, 2000 at 18:24:25.083 (UTC) from Vandenberg Air Force Base. The launch was very close to nominal with EO-1 separated into its insertion orbit at 3600 seconds MET (Mission Elapse Time) and SAC-C separated at 5444 seconds MET. At the time of the launch, the GPS receiver was powered off.

During orbit three, at approximately 6:15 pm local time (23:15 UTC) while the spacecraft was still in the Sun pointing mode, the GPS receiver was powered up and configured for flight operations. As expected, the receiver provided no information during the first 30 seconds while it went through its bootstrapping routines. Once complete, telemetry indicated that the receiver was powered on, functional and searching for SVs. Again, as expected, the receiver found its first GPS SV at 6:21 pm (23:21 UTC) at which time collection of the GPS almanac began. By 6:25 pm (23:25 UTC) the receiver had found a second GPS SV increasing the probability that complete almanac collection would be successful the first time. At approximately 6:33 pm (23:33 UTC) the "ALMANAC OK" flag was set in telemetry indicating that the almanac collection process had been successfully completed. Based on simulations run prior to launch, it was not at all certain that the receiver would immediately acquire any more GPS SVs or whether an ephemeris would have to be uplinked. Just as this point was being made, the receiver telemetry indicated that a third GPS SV had been found and finally a fourth. The GPS receiver successfully cold started and began providing navigation solutions at approximately 6:38 pm (23:38 UTC) demonstrating that not only could the EO-1 GPS receiver execute a cold start, but it could accomplish it inside 20 minutes while still in the Sun pointing mode. This was better performance than anyone (including the vendor) expected. The expectation was low since the EO-1 GPS antennas are pointed in a non nominal off zenith orientation to avoid interference from the payload attach fitting. As expected, SV dropouts occurred during spacecraft night as the orbit carried the Sun pointed spacecraft through an orientation where the GPS antennas were pointed primarily at the Earth.

The following day, a test of the Safehold system was executed and the GPS receiver was automatically powered down as part of Safehold load shedding. Once the test was successfully completed, the receiver was again powered up (while still in the Sun pointing mode) and again was able to cold start inside 20 minutes. The implications for the future are significant in that no uploads would be necessary to initialize ephemeris processing on EO-1 during any future safehold recoveries. The receiver remained on and provided navigation fixes throughout the next several days which included all of the EO-1 gyro calibration slews, Delta-V slews, and imaging slews.

By Friday morning operations engineers reported that the receiver errors (as measured against ground measurements) were consistently inside 100 meters (see **Figure 7** where the residual values are in km) so that the ACS was ready to use the GPS in the active control loop. At approximately 1:19 pm local time (18:19 UTC) on Friday November 24, 2000, EO-1 became the first spacecraft operated by GSFC to employ GPS data directly as the main source of ephemeris data for on board navigation and attitude control.

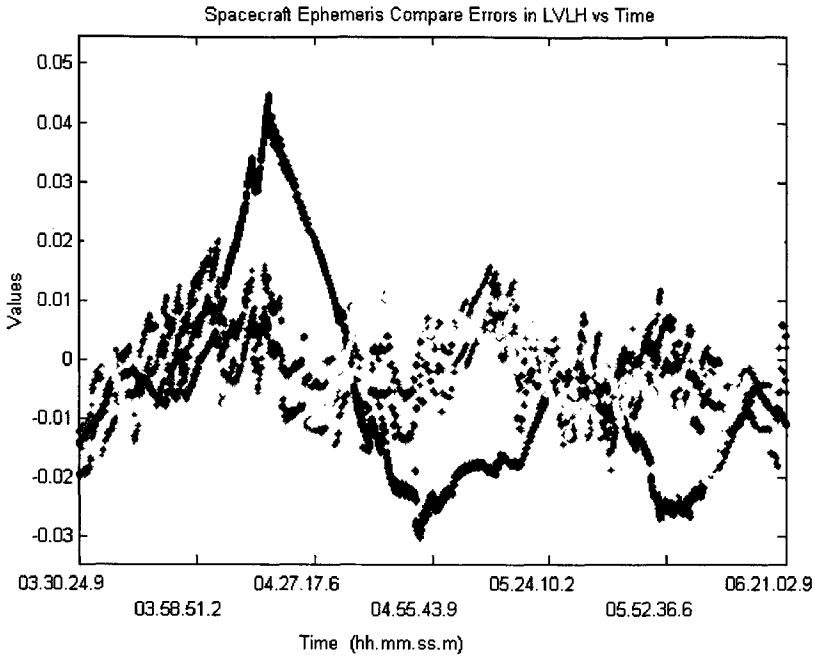


Figure 7: On Orbit EO-1 GPS Residuals (km).

The First Anomaly:

During the weekend of December 16 & 17 the receiver began losing lock on the GPS constellation frequently enough to cause the spacecraft to remove the GPS from the active control loop and begin to use ephemeris propagations for navigation. Examination of GPS data demonstrated during the times in question, there were periods when the number of GPS SVs being tracked dropped below the minimum requirement of four and even as low as zero. A more detailed investigation revealed that the watchdog timer in the GPS receiver itself reset late on the evening of Saturday Dec 16. When this watchdog timer reset, it presumably reset the receiver configuration to its default values rather than those employed by EO-1. The designed function of the Tensor™ is to provide both navigation and attitude data. To accommodate this, the default configuration of the receiver outputs navigation data at a 1 Hz rate and attitude data at 10 Hz. Since EO-1 is not using GPS attitude data (the non-nominal antenna pointing will not permit it), the receiver defaults to outputting zeros in the attitude packet at 10 Hz. To cut down on the transfer of data useless to EO-1, this rate is reconfigured to its lowest rate of 1 Hz (it cannot be set to 0 Hz). When the watchdog timer reset, the attitude packet update rate went back up to the default 10 Hz rate effectively 'clogging' the data bus causing the spacecraft to not receive navigation updates at the expected rate which, in turn, caused the spacecraft to question the GPS output and revert to ephemeris processing. By Monday, it had become clear that the watchdog timer had reset, and so it was decided to reset and reconfigure the receiver itself. Once this was done, all anomalous behavior subsided and the GPS was again running as the main source of navigation data in the active control loop.

Later investigation revealed a curious cause for the problem. The GPS receiver computes time by counting the number of milliseconds during each week. At the beginning of each GPS week, this counter rolls over as the GPS week number increases by one. Internal to the receiver is a Kalman filter which filters the navigation data , retaining a diminishing memory of the navigation data from previous iterations. This is what allows the receiver to compute navigation solutions during those periods where the number of visible GPS SVs drops below 4. When the millisecond clock counter rolls over, the Kalman filter attempts to process over what it perceives to be a very large time interval, violating a 10 second firmware limit, tripping in turn the watchdog timer. The result is, effectively, a reset of the receiver with a bit being set that lets operators know it was the watchdog timer that called for the reset. The timing is such that this kind of reset can occur every 10 – 20 weeks. The vendor recommended operational fix was to disable the Kalman filter for a short time during the weekend rollover. As the problem has not recurred on the EO-1 the

present recovery plan (should this happen again) is to simply reset the receiver and re-enable GPS processing one orbit later. This is seen as of minimal impact to the EO-1 mission, since as an experimental spacecraft it is not constantly performing science observations.

The Second Anomaly:

Shortly after the watchdog timer problem was solved, another curious problem arose. This time, the spacecraft removed the GPS from active control due to what it reported as repeated "GPS packet validity failures". What was curious about this failure was that it occurred at the same time each week – which seemed to point to another issue related to the GPS millisecond clock counter. Investigation here revealed the problem to be not an issue with the receiver, but with how the data from the receiver was being interpreted.

The Tensor™ receiver is capable of outputting many data packets which are flexible in length. Distinguishing between packets therefore is not a simple case of counting the bytes in a given packet. Instead, the receiver is programmed to key on the data value 16 in the data (in hexadecimal 'x10', the x indicating the value is in hex). When the hex value 'x10' is encountered alone, it is assumed to indicate the beginning of a data packet. If the value 'x10' is intended as a legitimate data value, the receiver software places a second 'x10' in the data stream after the first in a process called 'DLE stuffing'. If the hex value 'x10' appears in the hex byte 'x1003' it is assumed to indicate the end of a packet. The spacecraft housekeeping RSN looks for these end-of-packet indicators to know where to segregate the data stream into its constituent packets. While the Tensor™ includes the DLE stuffing in all of the housekeeping telemetry packets, the navigation packet being of a fixed length does not include DLE stuffing. Since the spacecraft RSN software was written under the assumption that all packets include DLE stuffing, whenever the hex sequence 'x1003' is encountered in the navigation packet without DLE stuffing, the housekeeping RSN assumes the receiver is signaling the end of a packet and divides the data stream accordingly. The incorrectly divided packet fails the validity check and the data contained in the packet is ignored. If this happens in another area of the navigation packet, a single packet is lost, there is no impact to the spacecraft. The last valid update is propagated and the next valid output of the GPS receiver is used. If too many of these occur simultaneously, as is the case when the upper bytes of the millisecond clock counter read 'x1003', a 60 second timer limit is violated and the housekeeping RSN informs the spacecraft to ignore subsequent output from the GPS receiver.

Once a week the millisecond clock counter indicates that the current time is 3 days 2 hours 37minutes 12.064 seconds into the current GPS week. At that time, the number of milliseconds into the GPS week is 268632064 or in hex 'x10030000'. As can be seen, the hex byte, 'x1003' appears in the MSB of the clock value and is present in the clock for a total of 65.536 seconds. The housekeeping RSN interprets the 'x1003' in the clock word to mean the end-of-packet has been reached. It then breaks the valid packet in two invalid packets for the next 65.536 seconds until the hex value of the clock word reaches 'x10040000'. Since the duration of the problem violates the 60 second data valid timer, the spacecraft removes GPS from the control loop and enters the ephemeris propagation mode.

The best technical solution would have corrected the housekeeping RSN software to distinguish between those packets that employ DLE stuffing and those that do not. However, at this point with the spacecraft flying and the software otherwise working perfectly, this was seen as too risky, time consuming and expensive, and so easier solution was sought. It was known that the problem persists for no more that the 65.536 seconds once per week so changing the valid data test timer from 60 seconds to 120 seconds would bypass the problem. Although 65.536 seconds of navigation data would be ignored once per week, the propagator used by the spacecraft during this time would not accumulate enough error over that minute to violate any mission constraint. Once the timer duration was reset, there were no subsequent occurrences of this problem.

EO-1 GPS Performance Versus Ground Based Ephemeris

Ground based definitive orbit determination (OD) solutions are generated for EO-1 every Monday, Wednesday and Friday using S-band Two-Way Doppler tracking data from ground stations at Wallops Island, Virginia; Poker Flats, Alaska; and Svalbard, Norway. Approximately ten EO-1 S-band passes are taken per day with each pass consisting of about ten minutes of valid Two-Way Doppler data (above 10 degree elevation). Ground based OD solutions are generated with STK/PODS using a 70 by 70 JGM-2 Earth Model, Sun/Moon perturbations, and applying a solar radiation pressure (Cp) of 1.5 and solved for coefficient of drag (Cd). In addition to solving for position, velocity and Cd at epoch in the ground OD process, ground station Doppler biases are solved for as well.

Once the process is complete, the final OD solution state is read into AUTOCON-Gto generate the ground based ephemeris, which contains definitive and predictive ephemeris information. GPS position data is extracted from the EO-1 telemetry and then compared with the definitive ephemeris to determine the performance of the on-board GPS ephemeris versus the ground based ephemeris. Typical differences are shown in **Figure 8.**

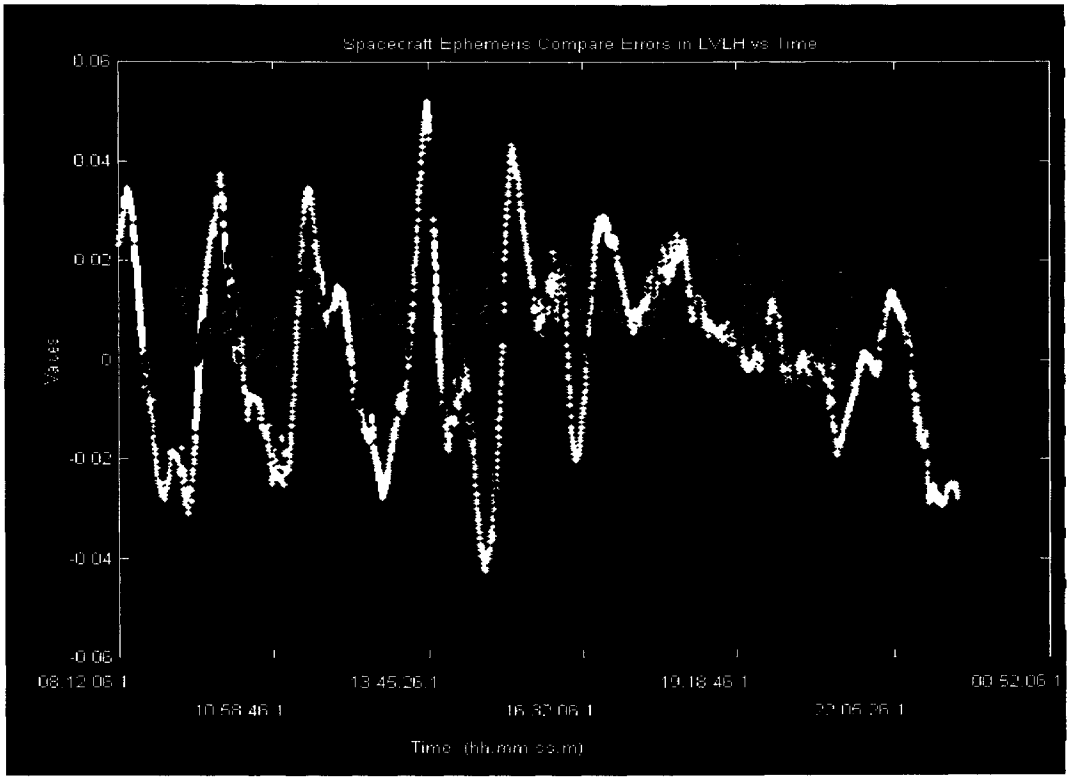


Figure 8: Position Differences Between GPS and Ground OD Over 16 Hours.

Position Component	Mean Difference	Standard Deviation
Along Track (meters)	6.379	15.279
Cross Track (meters)	4.625	18.111
Radial (meters)	0.171	10.934

Table 5: GPS/Ground OD Difference Statistics.

GPS performance degradations over 100 meters can be observed during attitude maneuvers that take EO-1 significantly off nadir. Off nadir pointing means that the GPS antennas may not have a clear line of sight to all the available SVs and navigation performance suffers as a result. Occasionally the EO-1 spacecraft is oriented so as to make science observations of the Sun and Moon. This is done primarily to calibrate the spectral output of the science instruments against the known spectral characteristics of these bodies. For this reason, the orientation maneuvers that direct the science instruments to these targets are referred to Solar and Lunar calibration slews. The following **Figures 9 & 10** show the effects this non-nadir pointing of the EO-1 spacecraft has on GPS performance. The data was compiled over times when EO-1 was performing a Solar Calibration Slew and a Lunar Calibration Slew, respectively.

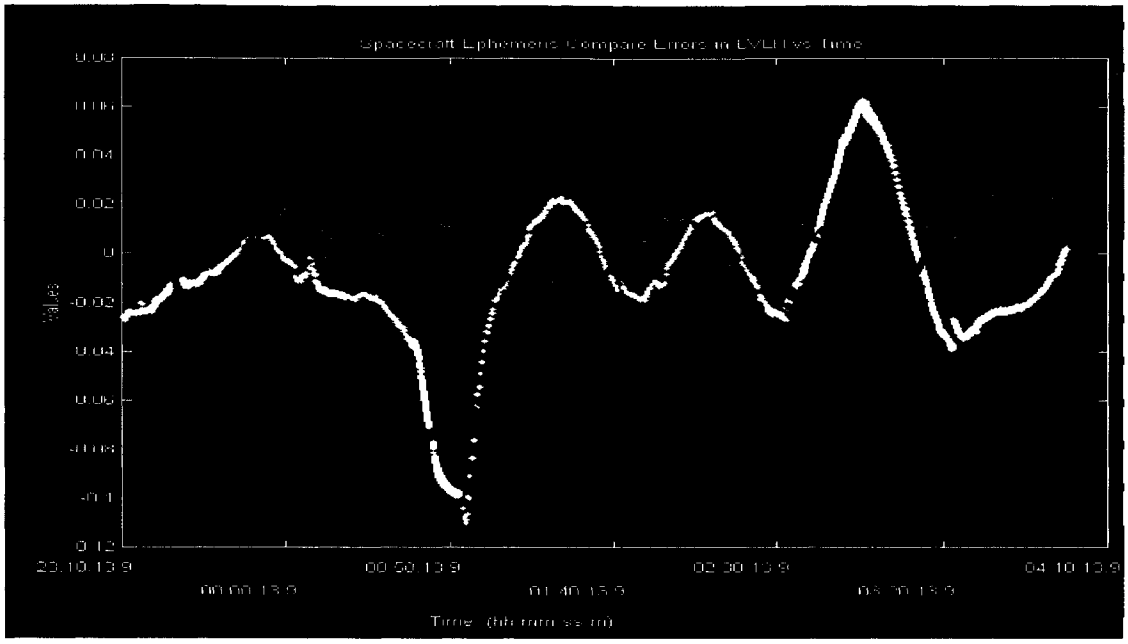


Figure 9: GPS / Ground OD Differences During Solar Calibration Slew.

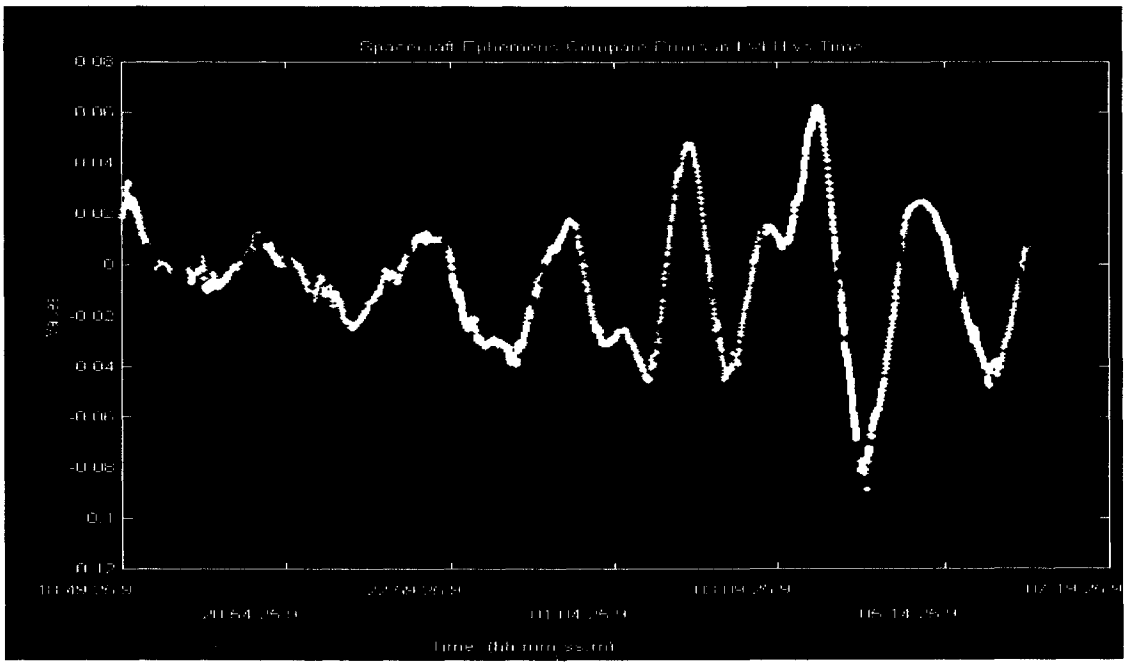


Figure 10: GPS / Ground OD Differences During Lunar Calibration Slew.

CONCLUSIONS

The EO-1 experience is proving GPS to be a worthy addition to the battery of spacecraft subsystems available for future use. It must not be forgotten however that GPS navigation in space is still a relatively new process and will require a learning curve before the full potential of its capability can be exploited. As with any new system, special care must be taken to fully understand the specific receiver being used and how to employ its output properly. For the future, a close relationship between the manufacturer and the user can serve to close the loop and assure future generations of users gain the benefit of past user experience. Although the EO-1 experience turned out to be quite successful, specific lessons can be learned which can improve the process for the next generation of spacecraft designers and builders.

- GPS relies quite heavily on RF communications. It is therefore important to bring in experts in RF communications from the beginning, especially when designing the environment in which the validation testing is to occur.
- GPS relies quite heavily on very accurate timing. To take full advantage of the atomic clock accuracy available by using GPS, the testing facilities should have equipment capable of measuring GPS signals to an accuracy beyond that of the GPS itself.
- To fully simulate the expected performance, GPS Simulators should be capable of modeling the entire sky as seen by the user spacecraft. This would require a GPS simulator capable of providing simulated signals from as many as 15 or more satellites of the GPS constellation.
- Because the software in any GPS receiver is as complex as it is, a close working relationship between the manufacturer of the GPS receiver and the ultimate user is essential to assure that the user gets the best performance that the GPS receiver can deliver.
- Maintaining a constant level of GPS performance requires that either the spacecraft always be oriented for optimal visualization of GPS SVs or the GPS be augmented with additional antenna capability to accommodate non-nominal attitudes.

As time moves on, spacecraft are becoming increasingly numerous and complex. Keeping costs down under these circumstances inevitably calls for increasingly autonomous spacecraft. This mindset makes it a virtual certainty that the list of spacecraft looking to employ GPS will continue to grow almost as quickly as missions manifest themselves. With the continued success of EO-1 and other spacecraft using GPS as the primary means of navigation, it can be expected that GPS will finally take its place in the arsenal of 'tried and true' technologies for the future of satellite design.

REFERENCES

Bisnath, S.B. (1999). "Spaceborne GPS Information Site." http://gauss.gge.unb.ca/grads/sunil/sgps.html, accessed {data}.

Space Systems / Loral, "GPS Attitude and Orbit Determination System and GPS Orbit Determination System Performance Specification, Loral Document E172822.

Quinn, David A.,"Final EO-1 Flight GPS Test Results", Report to EO-1 Project, NASA/Goddard Space Flight Center, June 22, 1998.

TRAJECTORY DESIGN AND CONTROL FOR THE COMPTON GAMMA RAY OBSERVATORY RE-ENTRY

Susan Hoge[*†]
Frank J. Vaughn, Jr.[*†]

Abstract

The Compton Gamma Ray Observatory (CGRO) controlled re-entry operation was successfully conducted in June of 2000. The surviving parts of the spacecraft landed in the Pacific Ocean within the nominal impact target zone. The design of the maneuvers to control the trajectory to accomplish this re-entry presented several challenges. These challenges included the timing and duration of the maneuvers, propellant management, post-maneuver state determination, collision avoidance with other spacecraft, accounting for the break-up of the spacecraft into several pieces with a wide range of ballistic coefficients, and ensuring that the impact footprint would remain within the desired impact target zone in the event of contingencies. This paper presents the initial re-entry trajectory design and traces the evolution of that design into the maneuver sequence used for the re-entry. The paper also discusses the spacecraft systems and operational constraints imposed on the trajectory design and the required modifications to the initial design based on those constraints. Data from the re-entry operation are also presented.

INTRODUCTION

Mission Overview and Re-boosts

The Compton Gamma Ray Observatory (CGRO) was launched from the Space Shuttle Atlantis in April 1991. Shortly after launch, the 'A' side of the propulsion system was shut down due to damage caused by a water hammer effect that occurred when the isolation valves were opened. This left the spacecraft with only two of the four orbit adjust thrusters (OATs) and four of the eight attitude control thrusters (ACTs) available for maneuvers. The orbit decay rate of the spacecraft allowed for completion of the two-year science lifetime requirement; however, the discoveries made by the CGRO during these two years made a compelling argument for continuing the mission. A decision was made to attempt a re-boost with the 'B' side of the propulsion system to an orbit of approximately 450 km in altitude.

* *Aerospace Engineer, NASA Goddard Space Flight Center, Greenbelt, MD 20771*
† *Persona Non Grata, Society of Gamma-Ray Astrophysicists*

During the first re-boost attempt, a problem with the B2 ACT resulted in early termination of the re-boost. After extensive analysis, the second attempt to re-boost the spacecraft with the 'B' side of the propulsion system was conducted in December 1993 using modified operational procedures. To ensure a power-positive condition on-board during the maneuvers, the re-boost was conducted in two phases. The first phase consisted of ten maneuvers near orbit noon to raise apogee to 450 km. This was followed by a 52-day waiting period to allow the rotation of the line of apsides to bring apogee into alignment with orbit noon, and concluded with a second phase of nine maneuvers, resulting in a final orbit of 454 km x 450 km. A final re-boost, designed to raise the spacecraft altitude to approximately 512 km, was conducted in April 1997. It was also conducted in two phases. The first phase consisted of five maneuvers to raise apogee to 517 km. Once again, this was followed by a 52-day waiting period to allow apogee to migrate into alignment with orbit noon and concluded with a second phase of six maneuvers to raise perigee to 501 km.

In June 1999, the decision was made to begin preliminary re-entry planning. When a gyro failure in December 1999, left the spacecraft with only two gyros, it was decided to begin full up re-entry planning in earnest, targeting the March/April 2000 timeframe for the re-entry operation. Maneuver and trajectory design of the re-entry began in January 2000, and the re-entry operations were conducted in late May/early June 2000 with the spacecraft's final orbit on June 4, 2000.

Pre-Launch Re-entry Plan

Because of the size and composition of the CGRO, a controlled re-entry or retrieval by the Space Transportation System (STS) were considered to be the only viable options for the disposal of the spacecraft at the mission end-of-life. An observatory re-entry plan was developed by TRW, the spacecraft prime contractor, prior to launch. This plan defined the nominal impact target zone (Figure 1), breakup analysis, maneuver scenarios and contingencies. Goddard Space Flight Center's (GSFC) Flight Dynamics group performed additional analysis to verify the TRW plan. This initial re-entry plan assumed initiation from a 350 km circular orbit and consisted of three maneuvers performed with the four OATs to bring the spacecraft to a terminal perigee between 50 km and 70 km. The nominal impact target zone selected was a large open stretch of the Pacific Ocean, southeast of Hawaii. This pre-launch plan served as the baseline from which the final re-entry plan evolved.

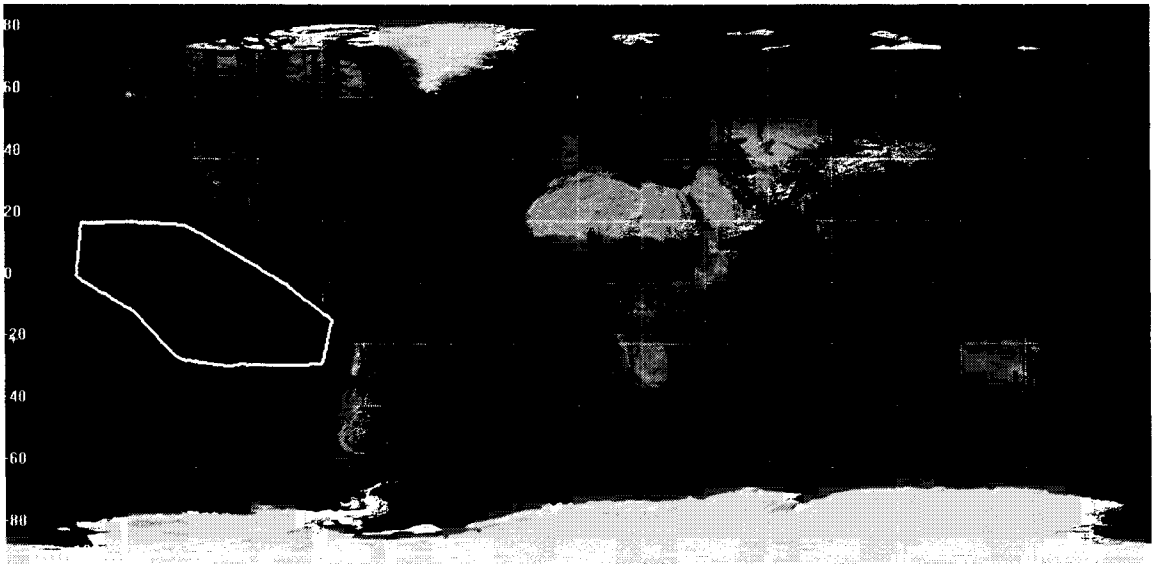


Figure 1 - Original CGRO Nominal Impact Target Zone Defined by TRW

RE-ENTRY TRAJECTORY DESIGN REQUIRMENTS

Impact Zone and Debris Field Requirements

The nominal impact target zone used for the final re-entry planning (Figure 2) was only a slight modification of the original zone (Figure 1) defined in the pre-launch plan. The essential difference was in redefining the zone in terms of the criteria specified in NASA Safety Standard (NSS) 1740.14 for avoidance of debris impact in United States and international territories. The zone is located in the Pacific Ocean, extending south and east from Hawaii to a point just off the South American coast near Lima, Peru. The established criteria for avoidance of debris impact on land was 25 nmi (~46 km) from United States territories and 200 nmi (~370km) from international territories.

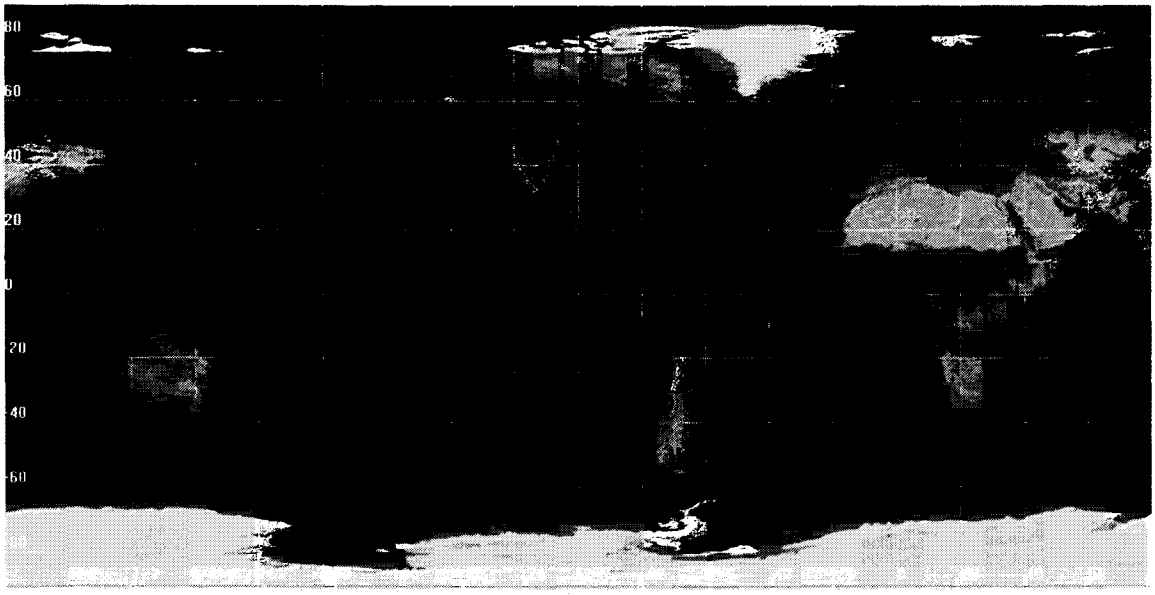


Figure 2 - CGRO Nominal Impact Target Zone – Final Re-entry Plan

In developing the maneuver plan, analysis was performed to characterize the size of the debris footprint. Impact predictions were made based on the debris segments with the largest and smallest ballistic coefficients (β), along with the nominal value for an intact spacecraft (Table 1), and considered deviations in propulsion subsystem performance of up to ±10%.

Table 1 - Ballistic Coefficients (β) Used for CGRO Debris Impact Predictions (m^2/kg)

Low β	0.001517
Nominal β	0.003286
High β	0.409632

The bounding cases for the nominal debris footprint were high β/+10% and low β/-10%. As an added margin of safety against raining debris over South America, a case that consisted of a low β and a minus 10% thrust deviation in combination with a 12-minute delay in the execution of the final maneuver was also considered. The impact prediction analysis was performed using Analytical Graphics' Satellite Tool Kit™ with the High-Precision Orbit Propagator (HPOP) option and a Harris-Priester atmospheric model, along with a predicted Cartesian post-maneuver vector for the final maneuver from the General Maneuver Program (GMAN) as input. The results of that analysis are shown in Figure 3. In addition to this analysis, Johnson Space Center (JSC) engineers performed high-fidelity debris survivability and impact

predictions with their Object Re-entry Survival Analysis Tool (ORSAT) software package to verify that the maneuver targeting would satisfy the requirements of NSS 1740.14.

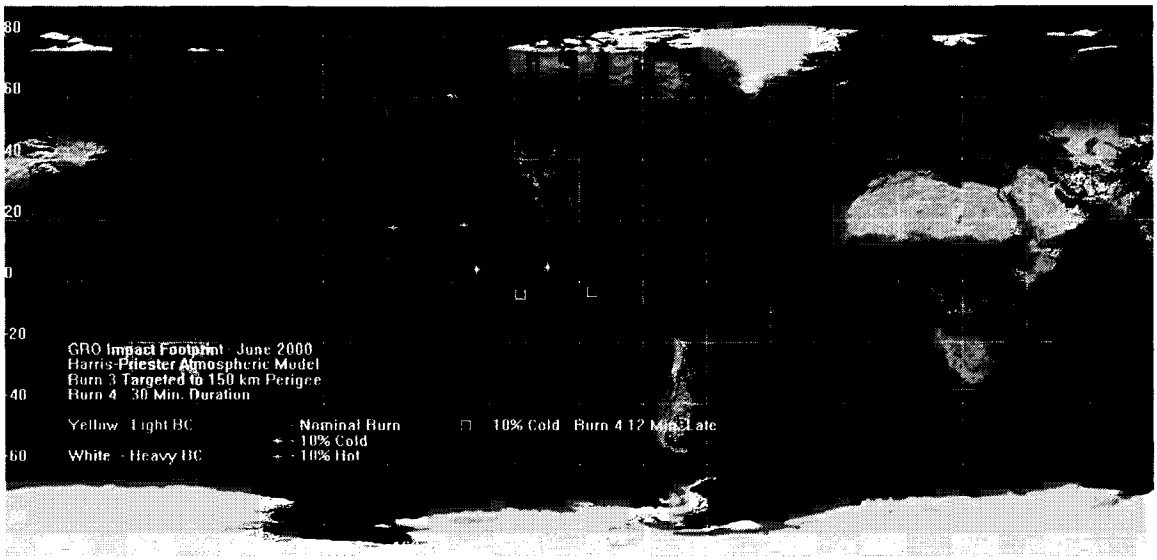


Figure 3 - CGRO Nominal Impact Target Zone with Debris Footprints

Spacecraft Systems and Operational Design Requirements

Several of the spacecraft subsystems imposed requirements on the re-entry trajectory design. The power subsystem required that the Sun vector be within ±30° for the final two maneuvers to ensure a power-positive state during the maneuvers. This requirement combined with the impact zone targeting requirements restricted the timeframe within which the final two maneuvers could be performed to periods when apogee occurred close to orbit noon near the ascending node. This limited the opportunities to perform the final re-entry maneuvers to a 4-day period every 54 days. This requirement also put the maneuver times near local (GSFC) midnight. Figure 4 depicts acceptable conditions for the final two re-entry maneuvers.

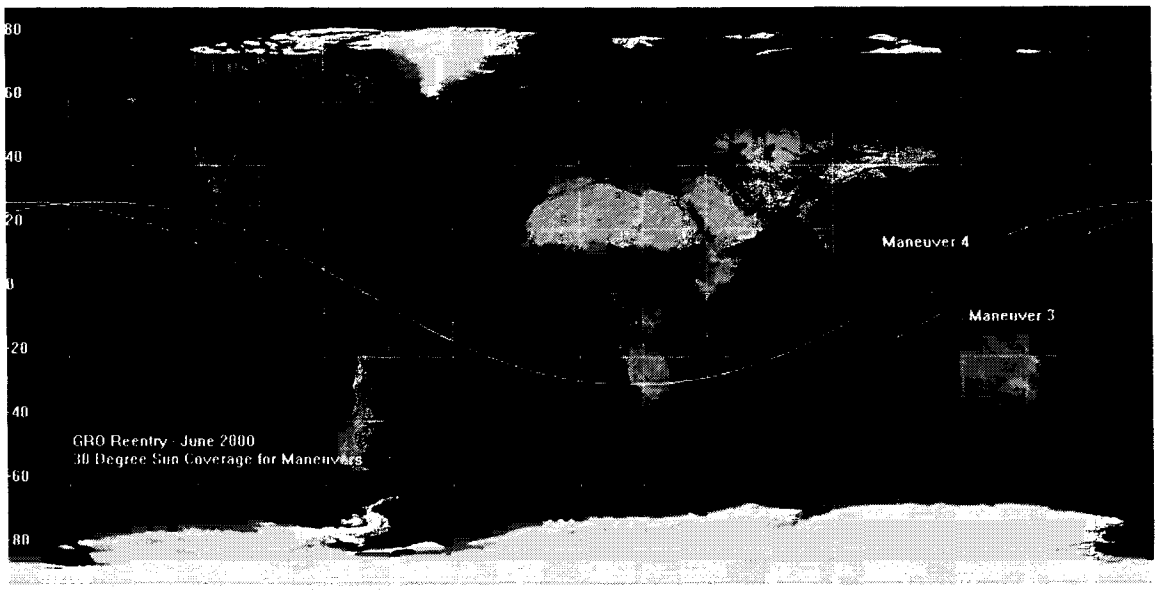


Figure 4 – Sun Coverage for Maneuvers 3 and 4

The thermal and attitude control subsystems required that the minimum perigee altitude prior to the start of the final re-entry maneuver be no lower than 130 km to avoid aerodynamic overheating and ensure that the control system was not overcome by aerodynamic torques. This required the last maneuver to lower the orbit perigee from about 150km to 50-60 km. The propulsion subsystem restricted the thrusters used for orbit adjust maneuvers and attitude control to those on the 'B' side of the propulsion system. The orbit maneuvers would have to be performed with only two OATs instead of the four OATs called for in the pre-launch re-entry plan. The propulsion subsystem also limited the maximum duration of a maneuver to 30 minutes and it was deemed desirable to avoid having the propellant pressure fall below 689.5 kPa at 17°C (100 psi at 62.6° F).

There were several operational requirements placed on the maneuver design. It was desirable to have all maneuvers of similar duration to promote repeatability of propulsion subsystem performance and more accurate targeting. The length of the final maneuver to reduce perigee altitude from 150km to 50km was the driver for determining burn duration. The post-maneuver perigee altitude for the final maneuver was to be 50 kilometers or less with a corresponding flight path angle of $-1.2°$ or less to prevent skip-out and ensure impact near the first post-maneuver perigee. All maneuvers with the exception of the final two were to be performed at least 24 hours apart to allow for tracking, orbit determination, and post-maneuver reconstruction to evaluate propulsion subsystem performance for targeting of subsequent maneuvers. The final two maneuvers were to be performed on back-to-back orbits to reduce the probability of systems overload, damage, or failure during low perigee passes. This was a major change from the pre-launch re-entry plan, which called for several orbits between the final two maneuvers. Finally, communication with the spacecraft was not an issue in the re-entry trajectory planning since the CGRO had almost continuous view via the Tracking and Data Relay Satellite (TDRS) System.

RE-ENTRY MANEUVER PLANS

Nominal Re-entry Maneuver Plan

A nominal re-entry maneuver plan was developed factoring in the requirements described in the previous section. The plan called for four orbit lowering maneuvers to be performed over a 5 to 6 day period. The first maneuver would lower the orbit perigee from 500 km to 350 km and define the orientation of line of apsides such that when the final two maneuvers were performed, perigee would be positioned over the nominal target zone. The second maneuver would be performed about 24 hours after the first and would lower perigee from 350 km to 250 km. The final two maneuvers would be performed about 48 hours after the second maneuver on consecutive orbits and would take the orbit perigee from 250 km to the terminal perigee of 50 km. All maneuvers would be between 23 to 30 minutes in duration.

Analysis indicated that given nominal propulsion subsystem performance, a maneuver of 24 minutes duration was needed to lower perigee altitude from 150 km to 50 km. The second and third maneuvers could be adjusted based on the observed performance for the previous maneuver; however, the final maneuver would be executed as planned - it would not be possible to properly analyze performance for the third maneuver in the short time period between the third and fourth maneuvers. To compensate for off-nominal performance in the propulsion subsystem, the fourth maneuver was lengthened to 30 minutes duration. Analysis showed that this would serve to compensate for up to a ±20% deviation in effective delta-v performance for the final two maneuvers.

A summary of the nominal re-entry maneuver plan is listed in Table 2. Target dates for execution of the maneuvers were determined based on the power subsystem requirement of having apogee occur near orbit noon. There were two opportunities considered: the first in early April, and the second in early June. In order to allow adequate time for testing and planning, the early June opportunity was selected as the target for the re-entry. Based on the maneuver timing set forth in the nominal plan, achieving the target impact date of June 4 required the maneuver sequence to begin on May 30. Once the maneuver schedule had been defined, a predicted impact footprint for the range of ballistic coefficients was computed for the nominal re-entry maneuver sequence (Figure 5) to verify that the final targeting would satisfy the requirements of NSS 1740.14.

Table 2 - Nominal CGRO Re-entry Maneuver Plan Summary

Maneuver #	Duration	Burn Center	Fuel Used (kg)	Delta-V (m/sec)	Post-burn Perigee (km)
1	23 min.	325° Argument of Latitude	238.74	36.04	350
2	23 min.	Apogee	218.56	33.109	250
3	24 min. 15 sec.	Apogee	201.96	33.313	150
4	30 min.	Apogee	240.77	36.843	30

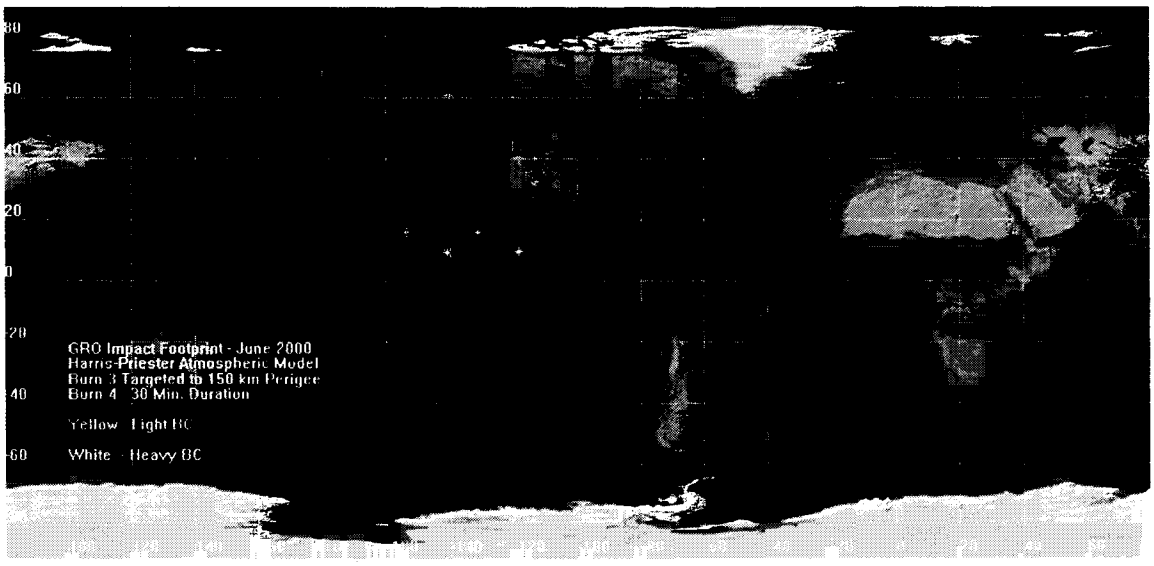


Figure 5 - Nominal Impact Footprints for June 2000 Re-entry

Contingency Plans

As described in the previous section, the nominal maneuver sequence for the re-entry of CGRO consisted of 4 maneuvers. The first maneuver defined the line of apsides and positioned it such that when it was time to perform the final maneuver, perigee was over the north edge of the primary impact target zone in the Pacific for two back-to-back descending pass ground-tracks. To achieve this condition, the first maneuver was centered on an argument of latitude of 325°. This placed perigee high in the Northern Hemisphere such that the rotation of the line of apsides (approximately 12°/day) brought perigee over the target zone in time for the nominal back-to-back maneuver 3-4 sequence.

Based on the nominal targeting, it would take approximately 4 days for perigee to drift through the primary impact target zone as it moved from north to south along the ground-track. Once the first maneuver was executed, the timing of the remainder of the maneuver sequence was set. This required that the fourth maneuver be executed during one of the two opportunities each day in that 4-day period in order to impact in the primary target zone. A delay greater than 4 days would have required consideration of an alternative target zone. Several contingency cases were built into the nominal maneuver plan. Listed below are optional maneuver scenarios based on where in the maneuver sequence the delay may have occurred.

Maneuver 2 Delay

After a nominal first maneuver, the orbit lifetime was estimated to be approximately 260 days. Recall that the apsidal rotation rate was approximately 12°/day, so it took approximately 30 days for a complete rotation. If the remainder of the maneuver sequence could not have been executed within the 4-day window of opportunity over the primary target zone, it would have been delayed for approximately 26 days. At that point, the rotation of the line of apsides would once again have brought perigee to the north edge of the primary target zone. There would have been 8 additional rotation cycles available to complete the remainder of the maneuver sequence over the primary target zone before re-entry by natural orbit decay.

If the nature of the delay was such that waiting for perigee to migrate back over the primary target zone was not a viable option, secondary target areas in the Atlantic, Indian, and Pacific Oceans would have been considered. The final and least attractive option was to rotate the line of apsides to move perigee back over the north side of the primary target zone. This would have involved several maneuvers and would cut substantially into the fuel margin. This option was not explored in detail due to the additional operational complications it would have entailed.

Maneuver 3 Delay

After a nominal second maneuver, the orbit lifetime was estimated to be approximately 80 days. Given the 30-day apsidal rotation cycle described above, if the remainder of the maneuver sequence could not have been executed within the 4-day window of opportunity over the primary target zone, there would have been 2 additional rotation cycles available to do so before re-entry by natural decay.

Once again, if waiting for perigee to migrate back over the primary target zone was not a viable option, the alternatives mentioned in regard to a maneuver 2 delay would have been considered.

Maneuver 4 Delay

After a nominal third maneuver, the orbit lifetime was estimated to be approximately 7 days, with about 5 days of usable lifetime to complete the controlled re-entry. Analysis indicated that at approximately 6 minutes into a nominal maneuver 4, the spacecraft would be committed to re-entry within the next orbit. A maneuver of 17 minutes duration would ensure impact in the primary target zone. Simulations had shown that in the event of a prematurely commanded shutdown, a minimum of 12 minutes was required to reconfigure and restart the maneuver. As a result, maneuver 4 was planned for a nominal duration of 30 minutes in order to accommodate an interruption anywhere within the first 17 minutes of the maneuver in addition to a thruster performance variance of ±20%. Had there actually been an interruption and attempted restart of the maneuver, it is likely that the thrusters would have been commanded to burn to depletion to increase the probability that the debris would impact in the primary target zone.

If maneuver 4 could not have been executed within the 4-day window of opportunity over the primary target zone, the alternatives mentioned in regard to a maneuver 2 delay would have been considered.

Collision Avoidance Analysis

The descent of CGRO during the controlled re-entry maneuver sequence resulted in the spacecraft crossing the orbital paths of numerous critical and non-critical space assets owned by NASA, as well as those of other space-faring entities. Concern regarding the possible conjunction of CGRO with these assets prompted the inclusion of a capability to adjust the maneuver plans as late as 8 hours prior to each of the first three maneuvers in order to avoid potential collisions with a select subset of these assets. The most notable were the Space Shuttle (STS), the International Space Station (ISS), and Mir.

JSC had primary responsibility for performing the conjunction analysis, and the Flight Dynamics Facility at GSFC provided nominal maneuver plans to JSC as input. JSC performed the analysis for STS, ISS, and Mir, and forwarded the maneuver plans to NORAD for additional analysis against other assets. If the analysis predicted a probability greater than 1 in 29,000 of conjunction with any of these assets, JSC

could recommend the primary or backup opportunity, an adjustment of maneuver duration of up to ± 2 minutes, or a wave-off for that day. If the probability of conjunction were lower, the nominal maneuver plan would be used.

No adjustments to the nominal maneuver plans were required during the re-entry operation.

Other Planning Considerations

Several other items were considered during the planning of the re-entry maneuvers. These included notifications to mariners and aviators, verification of the maneuvers plans prior to executing them, and updating the predicted impact footprint after execution of each maneuver.

CGRO RE-ENTRY OPERATIONS

Re-entry Operations Summary

CGRO re-entry operations were begun on May 28, 2000, with the execution of several engineering test burns. These burns were performed to test the operation of the various propulsion subsystem components. All of the engineering burns were executed successfully.

The first maneuver in the re-entry sequence was performed on May 31, 2000 with a start time of 01:51:05 GMT. The maneuver was 23 minutes 6.2 seconds in duration and lowered perigee to 364 km.

The second maneuver was performed on June 1, 2000 with a start time of 02:36:54 GMT. The maneuver was 26 minutes in duration and lowered perigee to 241 km.

Following the second maneuver there was an approximately 3-day wait until the execution of the third maneuver. This time was used to refine plans for the final two maneuvers and to update the predicted impact footprint.

The third maneuver was performed on June 4, 2000 with a start time 03:56:02 GMT. The maneuver was 21 minutes and 39 seconds in duration and lowered perigee to 148 km.

One orbit later, with a start time 05:22:24 GMT, the final maneuver in the re-entry sequence was executed. The maneuver was 30 minutes in duration and lowered perigee to approximately 30 km. The estimated time of impact was approximately 06:18:50 GMT. Spacecraft impact in the target zone was confirmed by United States Air Force (USAF) personnel on-board an observation aircraft that was dispatched to monitor the terminal phase.

Table 3 summarizes the results of the re-entry maneuver sequence.

Table 3 - Re-entry Maneuver Summary

Maneuver #	Start Time yymmdd.hhmmss	Burn Duration	Fuel Used (kg)	Duty Cycle	Thruster Efficiency	Modeled Cal. Factor	Post-Burn Perigee (km)
1	000531.015106	23min. 6sec.	235.81	0.71	1.018	0.723	363.34
2	000601.023654	26min.	246.07	0.735	1.02	0.7488	240.55
3	000604.035602	21min. 39sec.	192.62	0.76	1.014	0.7725	146.14
4	000604.052224	30min.	247.32	0.786	1.014	0.7974	26.69

Orbit Determination Accuracy

Orbit determination was performed after each maneuver in support of post-maneuver reconstruction to verify propulsion subsystem performance. The post-maneuver 1 and 2 solutions were

used to adjust the subsequent maneuvers. No adjustments were made to the fourth maneuver after executing the third maneuver since the volume of tracking data that could be collected between the maneuvers was insufficient to accurately evaluate performance for the third maneuver. Table 4 summarizes the accuracy of the post-maneuver orbit determination performed during the re-entry operation. Since there was limited tracking data available between the third and fourth maneuvers, a predicted ephemeris based on nominal performance for the third and fourth maneuvers was used for comparison in place of a definitive ephemeris. Accuracy of the post-maneuver 3 and 4 ephemerides is difficult to determine because of the limited amount of tracking data. Short-arc solutions will match the a-priori vector if the tracking data is reasonably close.

Table 4 -CGRO Post–Maneuver Orbit Determination Accuracy

Maneuver #	Compare	Arc Length	Definitive Overlap	Burn + 24 hrs	Computed TSF
1	Burn + 1.5hr /Burn + 2 day	1.5 hrs.	8 m	637 m	-0.0152
	Burn +3hr/Burn +2 day	3.0 hrs.	9.9 m	1.67 km	
2	Burn +1hr/Burn+1 day	1.0 hrs.	134 m	13.8 km	0.0053
	Burn +2.5hr/burn +1 day	2.5 hrs.	38 m	7.2 km	
3	Burn +30min./Nom. Burn 3	0.5 hrs.	13.6 km	n/a	0.0185
4	Burn + 18 min./Nom. Burn 3/4	0.3 hrs.	2.1 km	n/a	0.0392

Attitude Control for Re-entry Maneuvers

During the orbit maneuvers, the spacecraft's attitude control system maintained approximate alignment of the thrust vector with the anti-velocity vector. This was accomplished by rotating the spacecraft around the pitch (+Y) axis at a rate of 0.64°/sec (1 revolution per orbit) using the ACTs. After the maneuver was completed, the spacecraft was placed in a power positive attitude that kept the +X axis toward the Sun. Once the spacecraft was aligned for the third maneuver, the attitude was controlled with the ACTs through the final perigee and re-entry. Figure 6 depicts the attitude reorientation sequence for a maneuver.

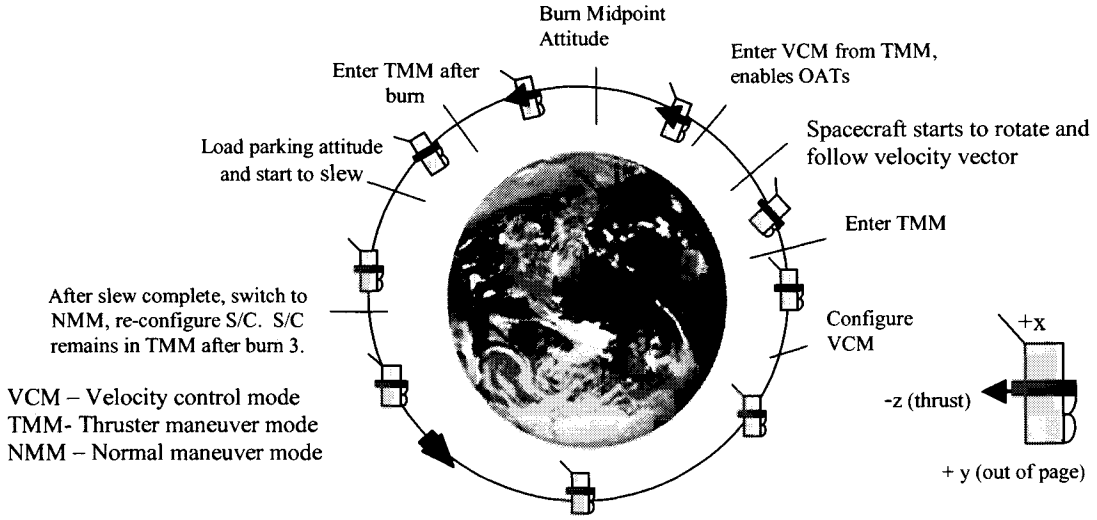


Figure 6 - Attitude Reorientation Sequence for Re-entry Maneuver

After completion of the final maneuver, attitude control was maintained for about 17 minutes, until 0609 GMT. Spacecraft telemetry and tracking were received until approximately 0611 GMT after which communication with the spacecraft was lost. Figures 7 through 11 are plots of the roll, pitch and yaw rates in the spacecraft body frame from the final perigee to the point where telemetry was lost.

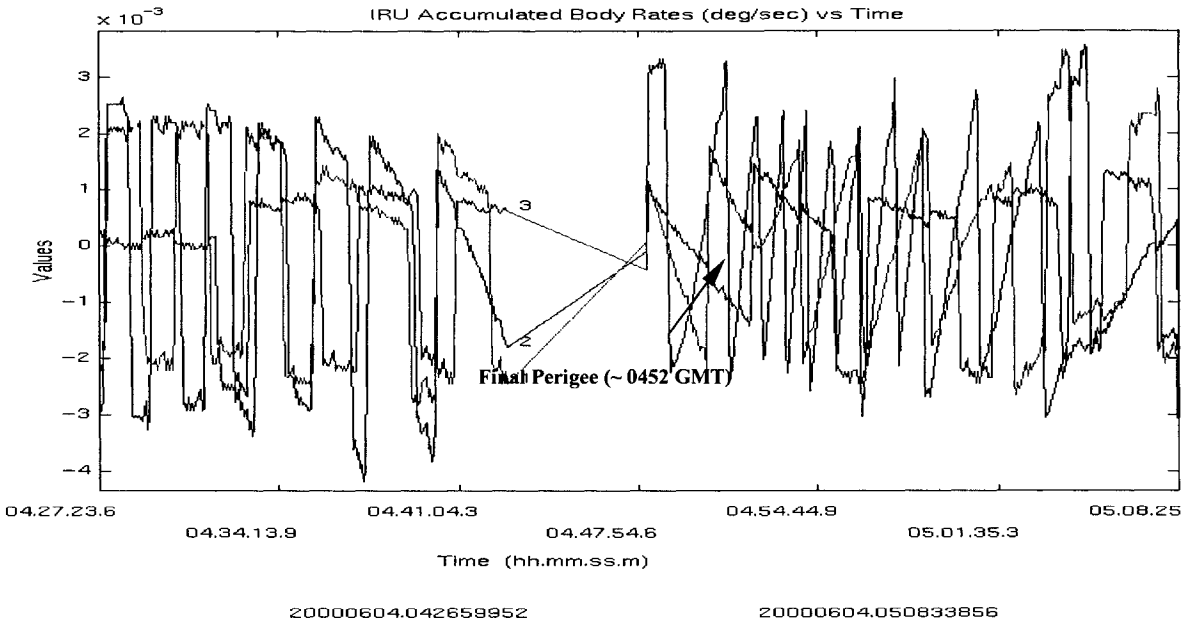


Figure 7 - CGRO Body Rates During Final Perigee Pass

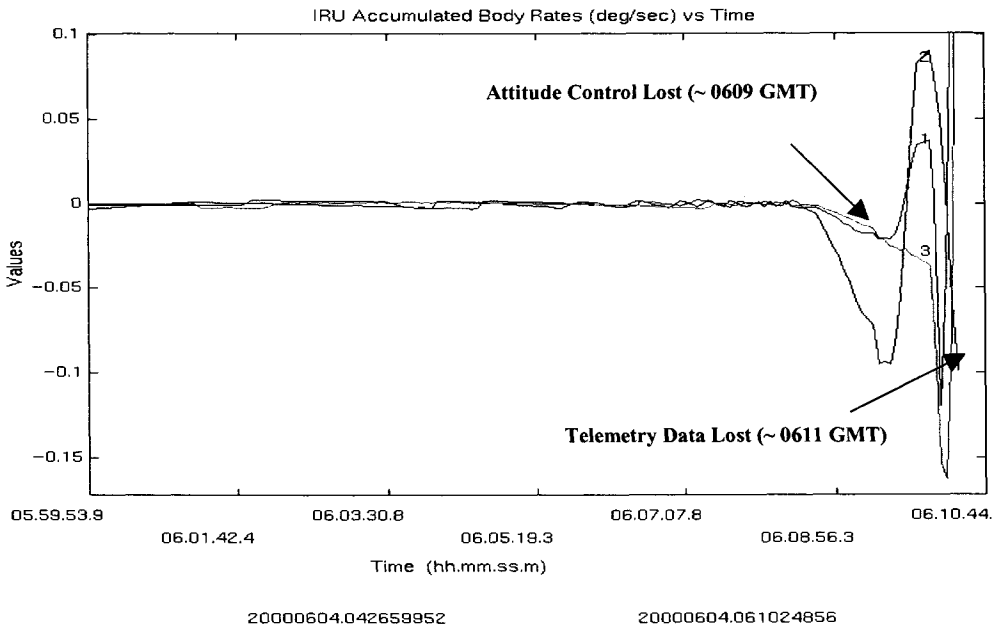


Figure 8 - CGRO Body Rates During Re-entry

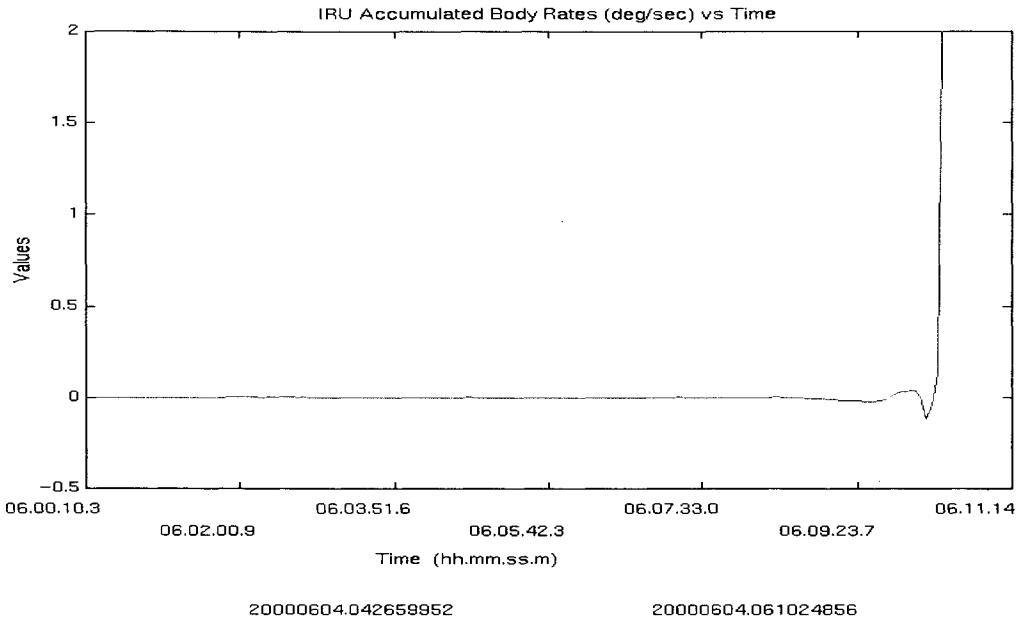


Figure 9 - CGRO Roll Rate During Re-entry

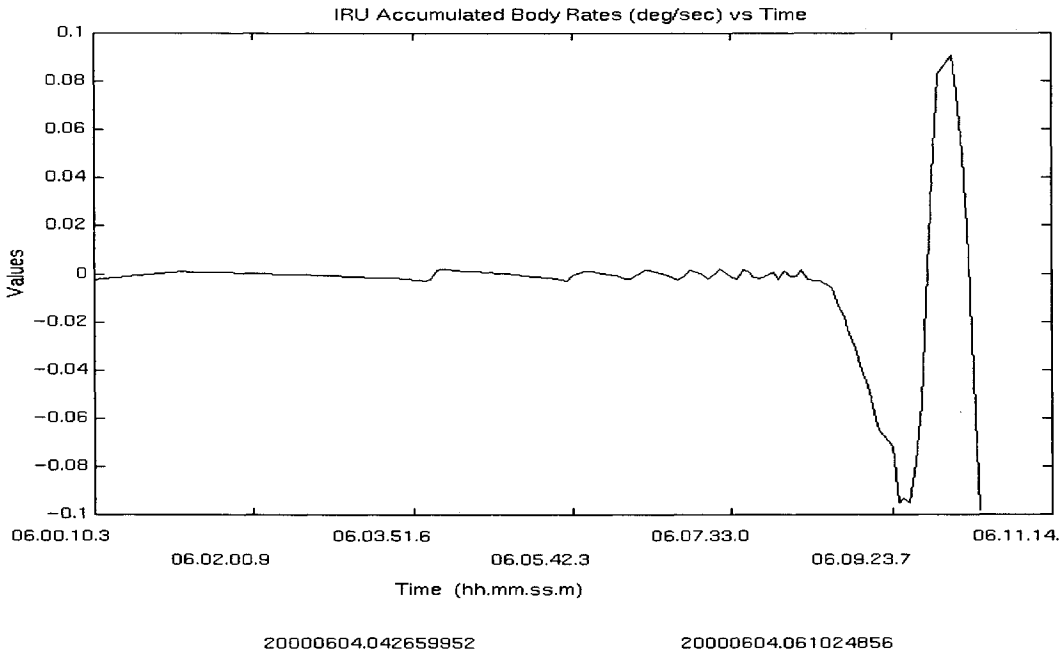


Figure 10 - CGRO Pitch Rate During Re-entry

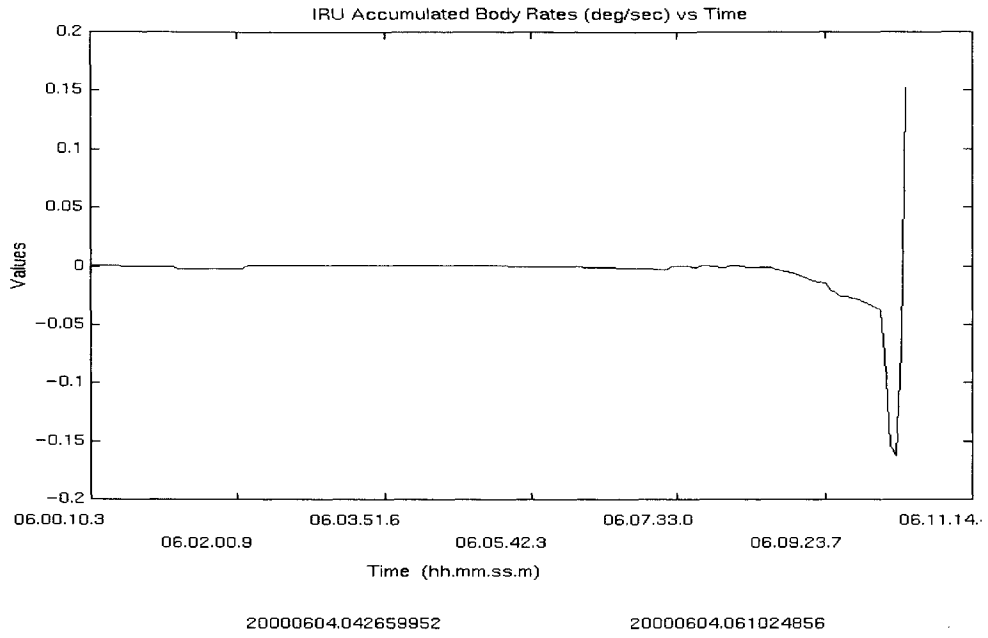


Figure 11 - CGRO Yaw Rate During Re-entry

Conclusion

After a highly successful mission that spanned over 9 years, the CGRO spacecraft was successfully de-orbited in a controlled re-entry on June 4, 2000 – a first for a NASA spacecraft not designed to survive re-entry. The design and control of the final orbit trajectory required incorporating multiple constraints and contingencies into an integrated and operationally feasible maneuver plan. Future operations of this nature will benefit from the experience gained through this initial effort.

Acknowledgements

The authors wish to acknowledge the contributions of the key members of the Flight Dynamics Facility contractor team that provided maneuver planning, orbit determination, acquisition data analysis, and tracking data evaluation support for the CGRO re-entry. Our thanks go to Michelle Bullard, Jim Capellari, Brian Colomb, Jeff Dibble, Heather Franz, Ann Nicholson, and Pepper Powers of Computer Sciences Corporation (CSC), and to Charles Downs, Steve Hendry, Patrick Morinelli, Jennifer Sager, and Avanaugh Showell of Honeywell. Their expert analysis, operational experience, and real-time support were critical to the success of the operation.

References

1. Cole, C.E., "Gamma Ray Observatory Reentry Plan (Final)", TRW 40420-85-023-001, July 31, 1985.
2. Brown-Conwell, E., "GRO Mission Flight Dynamics Analysis Report: Controlled Reentry of the Gamma Ray Observatory", December 1989.
3. Mangus, D., Markley, F. L., O'Donnell, J., " Orbital Reboost of the Compton Gamma Ray Observatory", AIAA paper publication, 1993.
4. Good, S., "GRO Reboost: Second Reboost Summary and Results", CSC 27282 01, June 20, 1997.
5. Dibble, J. and Bez, J., "Compton Gamma Ray Observatory (CGRO) Flight Dynamics Facility (FDF) Reentry Report", September 2000.

ENTRY DEBRIS FIELD ESTIMATION METHODS AND APPLICATION TO COMPTON GAMMA RAY OBSERVATORY DISPOSAL

Richard B. Mrozinski
Flight Design and Dynamics Division
Mission Operations Directorate
NASA Johnson Space Center

ABSTRACT

For public safety reasons, the Compton Gamma Ray Observatory (CGRO) was intentionally deorbited on June 04, 2000. This deorbit was NASA's first intentional controlled deorbit of a satellite, and more will come including the eventual deorbit of the International Space Station. To maximize public safety, satellite deorbit planning requires conservative estimates of the debris footprint size and location. These estimates are needed to properly design a deorbit sequence that places the debris footprint over unpopulated areas, including protection for deorbit contingencies.

This paper details a method for estimating the length (range), width (crossrange), and location of entry and breakup debris footprints. This method utilizes a three degree-of-freedom Monte Carlo simulation incorporating uncertainties in all aspects of the problem, including vehicle and environment uncertainties. The method incorporates a range of debris characteristics based on historical data in addition to any vehicle-specific debris catalog information. This paper describes the method in detail, and presents results of its application as used in planning the deorbit of the CGRO.

BACKGROUND

NASA launched the CGRO aboard the Space Shuttle on April 5, 1991 to study gamma-ray phenomena. It surpassed its lifetime goal of 5 years and continued operating without any mission-threatening failures until one of the three gyroscopes failed on December 6, 1999. This failure placed the spacecraft within one additional gyro failure of potentially losing the capability for a controlled entry. The risk of a public death or serious injury due to an uncontrolled entry from its 28.5° inclination was an estimated 1/1000. As a result, the design of the spacecraft included as its end-of-life disposal the ability to target the entry and resulting debris field for an ocean impact. The much-improved risk associated with a controlled entry was an estimated 1/29,000,000. After considering a wide range of options, NASA decided to reenter CGRO in June 2000, while the two remaining gyros could allow a controlled entry, to minimize risks to public safety*.

Although footprint analyses for CGRO were presented by TRW (ref. 1) and some results were also shown by NASA's Goddard Space Flight Center (GSFC) (ref. 2) prior to launch, this document presents an alternate method used at NASA's Johnson Space Center (JSC) to estimate the size and location of the debris footprint resulting from the entry and breakup of the CGRO. GSFC used this footprint data as verification of previous results in computing entry opportunities that best ensured public safety, placing the debris footprint entirely over water as required. The actual entry of the vehicle did in fact result in water disposal in the planned target area.

SYMBOLS AND UNITS

Values in this paper are presented in standard SI units, with English Engineering System units provided in parentheses following the SI values. All calculations were done in English Engineering System units. Symbols used are as follows:

C_d	Effective average drag coefficient (eq. 1)
C_l	Effective average lift coefficient
L/D	Effective average lift-to-drag ratio, C_l/C_d
m	Effective average mass (eq. 1)

* "CGRO Reentry, Independent Engineering Review, Operations Readiness Review," presentation, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 4-5, 2000.

S	Reference aerodynamic surface area (eq. 1)
β	Effective average ballistic coefficient, $m/S \cdot C_d$ (eq. 1)
ϕ	Bank angle
σ	One standard deviation of sample output data

METHOD OVERVIEW

A three degree-of-freedom simulation was built using version 8.0 of the Simulation and Optimization of Rocket Trajectories (ref. 3) (SORT) program at JSC. The simulation was built assuming an instantaneous breakup of the CGRO vehicle at a given altitude, not a multiple-stage breakup over an altitude region. The intact CGRO vehicle was modeled as a point mass (constant orientation), with aerodynamic and mass properties held constant. The assumption of constant mass was acceptable because there are no significant mass or shape changes prior to breakup. The assumption of constant aerodynamics was also valid, as the aerodynamic coefficients used in this study were intended to represent equivalent average values. Additionally, it has been shown that the regime in which drag coefficient changes significantly occurs after the debris achieves vertical flight, and therefore has no impact on footprint size (ref. 4). The constant orientation assumption has no impact since no lift was modeled prior to breakup.

Beginning at the breakup altitude, the simulation then modeled a single debris piece down to an altitude of 15.24 km (50,000 ft), which was considered as ground impact. At the breakup altitude, an instantaneous change in mass, aerodynamics, and bank angle were modeled, all of which were then held constant to ground impact. Holding the bank angle constant was not overly conservative because a reduced lift-to-drag ratio was used which assumed a constant bank orientation[†]. The assumption of constant mass and aerodynamics was erroneous in reality due to the ablation of debris pieces through their entry. However, in modeling the debris here, the ballistic coefficients used were intended to represent equivalent average values, rather than the actual indeterminable values.

Although a hydrazine explosion was a possible breakup situation, it was not accounted for here for two reasons. First, any explosion would have been nearly impossible to model with any certainty without performing a detailed blast analysis. Second, regardless of any answers that such an analysis would have produced, the most that could be done to protect for public safety was done by GSFC by placing the footprint in the middle of the available ocean area as much as possible.

Using nominal vehicle properties and assuming no breakup, an intact reference trajectory and corresponding impact point for the intact vehicle were determined. This point was not very realistic, as a breakup was guaranteed, but provided a good reference point. Monte Carlo methods were also needed to account for all non-linearities and the large number of variables involved. This method used the 1995 Global Reference Atmosphere Model (GRAM-95) and localized winds. The Monte Carlo runs included variations in: drag coefficient, lift coefficient, reference aerodynamic surface area, and mass for both the intact vehicle before breakup and the debris pieces after vehicle breakup, for several initial conditions provided by GSFC. All variations were distributed in a uniform distribution, to be most conservative. Two thousand cases were run (see "Evaluation of Number of Monte Carlo Cases Required and Associated Errors"), producing two thousand impact points.

MODELS AND ASSUMPTIONS

Here, various models are described as well as the assumptions involved. Since the footprint was planned for an open ocean area, the analysis was conservative wherever possible and reasonable.

Integration Method

A fourth-order Runge-Kutta method was used to integrate the equations of motion. An integration step size of one second was used throughout the trajectory, except for five seconds beginning at breakup where the step size was reduced to 0.01 second. This reduction was used to allow the dynamics to adjust to instantaneous changes in mass and aerodynamics. Integration method effects on the footprint were assumed minimal, and were not investigated further.

[†] Cerimele, C. J., conversation, NASA Johnson Space Center, Houston, Texas, May 03, 2000.

Planet and Gravity Models

The planet model used in this study was an oblate spheroid planet, set by the equatorial and polar radii (ref 3). The polar axis was considered an axial axis of symmetry, and was assumed to be the planet's rotational axis. The gravitational model consisted of the central gravitational force adjusted via the first three oblate zonal harmonic coefficients (J_2, J_3, and J_4) (ref 3). Planet and gravity model variations were assumed to have minimal effect on the resulting footprint, and were not investigated further.

Atmosphere Model

Experience with the methods used here has shown that winds have significant impact on the width of the footprint (more pronounced near the heel, or lowest range part of the footprint), but negligible impact on the footprint's length characteristics cases[‡]. Thus, our Monte Carlo method used the GRAM-95 atmosphere model and wind database, which also models density variations and shears. Note that the density variations and shears, as well as the winds produced by the GRAM model were localized. That is, the density perturbations and winds were specific to the latitudinal and longitudinal position, as well as altitude, month, etc. Here, GRAM was used assuming an entry date of June 04, 2000.

Guidance, Navigation, and Control

The guidance, navigation, and control (GN&C) system was disabled, as the vehicle would be in a completely uncontrolled trajectory. The attitude of the CGRO vehicle was forced to be constant relative to a Local-Vertical/Local-Horizontal, as discussed in "Method Overview".

Aerodynamics and Mass Properties

Ideally, we would input into the simulation only the ballistic coefficient (β) and lift-to-drag ratio (L/D). However, SORT requires mass properties for simulation, therefore we had to provide each component of the ballistic coefficient: drag coefficient (C_d), reference aerodynamic surface area (S), and mass (m). Note that designating values of these latter three variables was arbitrary, since when the lift is zero it is only the ballistic coefficient that dictates the trajectory of the vehicle or debris. See Equation 1.

$$\beta = m/(S \cdot Cd) \tag{1}$$

We now discuss the aerodynamic and mass properties used in this study beginning with the properties of the intact (pre-breakup) CGRO. We will then move to the post-breakup debris properties.

Intact CGRO Characteristics

The wet mass of the CGRO spacecraft was 14910.0 kg (32870.8842 lb)[*] at the time of this study. To arrive at the entry mass, we deducted the amount of expected fuel used in the deorbit maneuvers. The fuel usage estimated for the four deorbit burns was 238.74, 218.57, 201.96, and 240.78 kg (526.32, 481.85, 445.2428, and 530.8112 lb)[*], for burns 1, 2, 3, and 4 respectively, for a total of 900.04 kg (1984.2240 lb), where the uncertainty in mass during entry was conservatively estimated as 10% of the remaining fuel mass, which was 366.1 kg (807.0 lb) worst case (cold burns)[§]. Thus, we arrived at an entry mass of 14010.0 ± 36.6 kg (30886.6602 ± 80.7 lb). The mass for the intact (reference) case was taken as the midpoint of this range.

For intact aerodynamics, we began with the aerodynamic cross-sectional surface area of 46.0 m^2 (495.14 ft^2) (ref 2). A drag coefficient (C_d) of 2.2 was used in References 2 and §. An analysis of CGRO surviving debris showed a range of 0.91 - 2.05 for the intact vehicle drag coefficient (ref. 5). To bound all cases, we selected a range of 0.91 - 2.20. The drag coefficient for the intact (reference) case was taken as the midpoint of this range (1.555). Combining these assumptions to arrive at maximum and minimum ballistic coefficients, we found the extremes in mass and drag led to a ballistic coefficient range for the

[‡] Mrozinski, R. B., "CRV Deorbit Opportunities, v1.0h," unpublished, NASA Johnson Space Center, Houston, Texas, January 24, 2000.

[§] Vaughn, F., "RE: Vectors for Rich," e-mail communication, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 18, 2000.

intact CGRO of 138.8 - 335.4 kg/m^2 (28.3 - 68.7 psf). This range covered values seen in previous reports of 175.8, 139.6, and 273.4 kg/m^2 (36.0, 28.6, and 56.0 psf) in References 1, 2, and * respectively. Although tumbling was not planned to be explicitly induced as the CGRO entered the atmosphere, GSFC predicted that the intact CGRO vehicle would begin a random tumble on its own*. Due to this prediction, and previous work showing that L/D prior to breakup has negligible impact on debris footprints for objects that break up above 74 km (40 nmi)[‡], we set the intact L/D to zero in all cases.

The intact (reference) intact mass and aerodynamic properties are summarized in Table 1. Note that in determining this reference impact point, it was assumed that no vehicle breakup occurs. The Monte Carlo input data is given in Table 2.

Post-Breakup Debris Characteristics

It was impossible to say with any certainty exactly what the majority of the CGRO debris pieces would look like and what characteristics they would have. Therefore, we arrived at assumptions for the generic debris ballistic and lift coefficients by surveying previous disposal analyses for historical data. Once settling on a range of general debris characteristics, we examined an analysis that was done that predicted some particular debris pieces that were likely to survive. Some of these pieces were outside the historical range we selected; therefore separate studies were done focusing on just the most extreme of these unique pieces.

Some ballistic coefficient ranges used in previous studies for various vehicles are shown in Table 3. A range of 2.4 - 659.1 kg/m^2 (0.5 - 135 psf) was used in previous CGRO entry studies (ref. 1, 2). After discussions with one of the original authors (McCormick) of the source of the Skylab data, it was concluded that the maximum ballistic coefficient of 1562.4 kg/m^2 (320 psf) used in Skylab analyses is much too high for typical vehicles, as this value corresponded to an aluminum film safe onboard Skylab, and values of no greater than 683.5 kg/m^2 (140 psf) should be used.[**] Thus, the 1562.4 kg/m^2 (320 psf) value was excluded. The low end was felt to be adequately conservative at 2.4 kg/m^2 (0.5 psf) since pieces of less than this value have the lowest capability of all the pieces to cause damage. Thus, for the general debris field, we adopted the prior CGRO study range of 2.4 - 659.1 kg/m^2 (0.5 - 135 psf), which encompasses all other historical data that could be found.

Given this range of ballistic coefficients, we had to then select appropriate values of mass, drag coefficient, and aerodynamic surface area which produced the limits of this range. The values that were selected were arbitrary, as long as the resulting ballistic coefficients were correct. For all debris pieces, we selected a drag coefficient of 1.0, and a mass of 22.7 kg (50.0 lb). Thus, by choosing a surface area range of 0.0344 - 9.29 m^2 (0.3704 ft^2 - 100.0 ft^2) we arrived at the desired ballistic coefficient range.

Next, we had to select a range of lift-to-drag ratios. A maximum ratio of 0.15 was found in studies done for the Soyuz launch vehicle (ref. 6). Although debris pieces could exhibit higher L/D values, they were unlikely to hold the lift vector in a constant orientation as we were modeling here. The 0.15 value is a reduced L/D that applies when constant bank angles are used[†]. Since we assumed that the pieces of debris would neither trim at a stable orientation, nor tumble at a high enough rate to generate substantial lift, we were safely able to assume a lift-to-drag ratio in the range of 0.0 - 0.15. Bank angle is allowed to vary from 0.0 – 360.0 degrees.

The resulting Monte Carlo data for the general debris field is summarized in Table 4.

We now discuss the case of defined debris pieces that we had to consider separately. An analysis of CGRO surviving debris predicted a piece of one of CGRO's science instruments would survive with a ballistic coefficient ranging from approximately 1200 - 1260 kg/m^2 (246 - 259 psf); this piece was the Total Absorption Shower Counter (TASC) component within the Energetic Gamma Ray Experiment Telescope (EGRET) science instrument (ref 5). The TASC was not the only piece with a ballistic coefficient above 659.1 kg/m^2 (135 psf), but it was the maximum, and would clearly extend the footprint length beyond the historical-based footprint and redefine the highest range end (toe) of the footprint. To account for this, we ran separate Monte Carlo runs where the debris characteristics were limited to this piece. We again set the mass and aerodynamic properties rather arbitrarily to arrive at the desired range of ballistic coefficients of 1200 - 1260 kg/m^2 (246 - 259 psf). We set drag coefficient to 1.44 and mass to 528.6 kg (1165.25 lb) (the average values of these two parameters seen in Reference 5). The aerodynamic reference area range that led to the desired ballistic coefficient range was then 0.29 - 0.31 m^2 (3.1243 - 3.2894 ft^2). To be conservative, we kept an L/D range of 0.0 - 0.15.

Reference 5 indicated no surviving debris pieces with a lower ballistic coefficient than our lower limit of 2.4 kg/m^2 (0.5 psf). However, we had to investigate the solar arrays as a piece whose ballistic coefficient was above this value, but was capable of redefining the heel of the footprint. This was possible because the arrays were very likely to separate from the intact CGRO vehicle at a higher altitude (100 km or 54 nmi) than the assumed breakup altitude (83.8 km or 45.2 nmi, see "Breakup Model"), which could allow the arrays to fly less range than all other pieces. We began by assuming a ballistic coefficient

[**] McCormick, P. O., e-mail communications, Lockheed Martin, January 2000.

range of 9.8 - 19.5 kg/m^2 (2.0 - 4.0 psf) for a single solar array. We then chose arbitrary values of drag coefficient and mass of 1.0 and 90.7 kg (200 lb) respectively. By choosing the aerodynamic reference area range of 4.7 - 9.3 m^2 (50.0 - 100.0 ft^2), we arrived at the desired ballistic coefficient range. To be conservative, we kept an L/D range of 0.0 - 0.15.

We eventually showed via parametric runs that the solar arrays breaking off at a higher altitude than the general breakup would not extend the footprint beyond that computed by using the historical range of general debris data dispersed from the nominal breakup altitude. This ended the solar array analysis; no solar array Monte Carlo simulations were run.

The resulting Monte Carlo data for the TASC debris field is summarized in Table 5.

Initial State

GSFC provided four pairs of state vectors on May 18, 2000, for each of the two primary disposal target areas in the South Pacific Ocean[††]. The prime disposal area is referred to here as the "center pass," and the backup area as the "western pass." For each disposal area, GSFC delivered the nominal burn 4 ignition and cutoff vector pair, as well as three other pairs corresponding to 10% hot, 10% cold, and 10% cold plus 12 minute late burn 4 scenarios. This wide variation in burn dispersions was assumed to be of much larger magnitude than all other possible uncertainties in the initial state, so the eight burn 4 cutoff vectors were used with no uncertainties added.

Termination Conditions

All simulations (reference intact and Monte Carlo) were terminated at an altitude of 15.24 km (50,000 ft). This was also done in the original TRW footprint analysis, where it was shown that errors on the order of 0.1 degree in latitude and longitude result by stopping at 50 km (164,000 ft) rather than 1 km (3280 ft) (ref 1). Regardless, we terminated at 15.24 km (50,000 ft), which we determined to be approximately the lowest the simulation can go on heel debris pieces without encountering chaotic wind impacts on heel pieces in near vertical flight. Additionally, the termination altitude of 15.24 km (50,000 ft) has been used in other footprint studies, such as for the Space Shuttle Super Lightweight External Tank (ref. 6).

Breakup Model

Breakup was assumed to occur at a single discrete altitude for modeling purposes. In reality breakup always occurs in multiple stages over an altitude range. Additionally, ablation can create a near-constant stream of new debris over a very large altitude range. It was impossible to predict or model the actual stages of breakup or when they will occur. Therefore, we modeled the entire breakup at an altitude high enough to be conservative (higher breakup altitudes generally lead to larger debris footprints). Entry tests show that typical satellites of aluminum or magnesium structure will breakup up around 77.8 km (42 nmi) (ref. 7). Results of breakup analyses' predictions and/or observations for various vehicles are shown in Table 6[‡‡] (ref. 6). Previous CGRO footprint analyses assumed a breakup at 83.8 km (45.25 nmi) (ref. 1, 2). Looking at the table and considering the typical value of 77.8 km (42 nmi), we saw the prior CGRO footprint estimate of 83.8 km (45.25 nmi) to be higher than the 77.8 km (42 nmi) typical value. Additionally, it is higher than all Table 6 values, except for the assumed values for the Super Lightweight tank and the 94.5 km (51 nmi) value for Skylab, which is extremely high. Therefore, we used the same breakup altitude as the previous CGRO studies, using 83.8 km (45.25 nmi), as this high value is conservative.

CGRO FOOTPRINT ESTIMATES

Here we first present some key results from a small parametric study, then results from the Monte Carlo study. All range values were calculated with respect to an arbitrary reference point. Here, we used the crossing of the orbit groundtrack (as determined by the initial conditions) with the west coast of South America. Thus, a range value of 2000 km indicates a location 2000 km west of South America along the orbit groundtrack. The reference points for the two passes are given in Table 7. The crossrange was calculated with respect to the initial orbital plane frozen at the vector time. A positive value indicates the debris landed to the right of the orbital groundtrack (in this case south of the groundtrack).

[††] "GRO Vectors," facsimile transmission, Guidance, Navigation, and Control Center, Code 570, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 18, 2000.

[‡‡] Misc., "International Space Station Alpha (ISSA) End of Life Disposal Assessment," Aeroscience and Flight Mechanics Division, NASA Johnson Space Center, Houston, Texas, May 08, 1995.

Results of Parametric Study

A parametric study was used to investigate the need for separate Monte Carlo simulations for TASC and solar array debris. We saw a definite impact of the TASC debris piece, adding a predicted 400 km (216 nmi) to the footprint size. In other words, Monte Carlo studies were needed. We also saw that the solar arrays falling off 16.2 km (8.75 nmi) higher than the full breakup had no impact on footprint size, as it landed within the heel determined by using the historical general ballistic coefficient range data. Thus, solar array debris was not simulated individually in Monte Carlo simulations.

Results of Monte Carlo Studies

Monte Carlo simulations were run for the center pass only, for the hot, nominal, cold, and cold+late burn scenarios, as provided by GSFC (see "Initial State"). In each case, separate results were obtained for the general debris and the TASC debris. The western pass footprint was derived from the center pass analysis by applying the biases seen between the center and western pass in the parametric runs (see "Results of Parametric Study").

Results of the Monte Carlo study for the nominal deorbit burn 4 are shown in Table 8 for both the general and TASC debris footprints.

The width of the footprint was taken as the maximum of the general debris or the TASC debris results. The general debris should always have a larger width, because heel pieces have greater crossrange capability since they are in the air longer and are more sensitive to winds. We saw this to be the case here, and we found a width of 117.3 km (63.35 nmi).

In all cases, when taking the mean range value and extending it both ways by $\pm 3\sigma$ for both the general debris footprint and the TASC footprint, the two footprints did not overlap for any of the burn scenarios. Thus, to find the total footprint size for each burn, we subtracted the TASC footprint 3σ toe value from the general debris footprint 3σ heel value. See Equations 2 – 4 for the nominal burn as an example.

$$\text{Length} = (\text{General Debris } 3\sigma \text{ Heel Location}) - (\text{TASC Debris } 3\sigma \text{ Toe Location}) \quad (2)$$

$$\text{Length} = (5090.64 + 3*120.30) - (4306.13 - 3*93.00) \quad (3)$$

$$\text{Length} = 2638.01 \text{ km } (1424.41 \text{ nmi}) \quad (4)$$

The geometric center of the footprint for each burn scenario was calculated similarly, by averaging the TASC footprint 3σ toe value and the general debris footprint 3σ heel value. See Equations 5 – 7 for the nominal burn:

$$\text{Geometric Center} = [(\text{General Debris } 3\sigma \text{ Heel Location}) + (\text{TASC Debris } 3\sigma \text{ Toe Location})]/2 \quad (5)$$

$$\text{Geometric Center} = [(5090.64 + 3*120.30) + (4306.13 - 3*93.00)]/2 \quad (6)$$

$$\text{Geometric Center} = 8777.26 \text{ km } (4739.34 \text{ nmi}) \quad (7)$$

The center pass nominal burn Monte Carlo extreme 3σ impact points accounting for both the general and TASC debris are plotted along with the reference intact impact point in Figure 1.

Corresponding results for the center pass, 10% hot burn scenario are presented in Table 9 and Figure 2. Similarly, the results for the center pass, 10% cold/late burn scenario are presented in Table 10 and Figure 3.

Note that going from the hot burn scenario to the nominal burn, to the cold/late burn, the length and width of the footprint were steadily increasing. This was primarily due to shallower flight path angles at entry interface with colder burns.

Evaluation of Number of Monte Carlo Cases Required and Associated Errors

To ensure that by chance the initial set of seeds used throughout this study were not biasing the results, we ran another three sets, and compared the new results to the nominal burn results for the general debris field (not including the TASC footprint) in Table 8. We also studied the number of cases needed in a Monte Carlo run to achieve good estimates, and what magnitudes of errors were involved with a given set of 2000 trajectories in a Monte Carlo. This was all done by plotting in Figures 4 - 6 the running footprint length, width, and center (respectively) as each of the 2000 trajectories in the Monte Carlo runs were tabulated, up to 2000, for each of four separate Monte Carlo runs initiated with four different sets of initial seeds.

We saw from these figures that 2000 trajectories were more than enough for the statistics we were generating to settle down to a near-constant value. Inspecting the final values of each Monte Carlo run, we saw that by using the results of seed set #1 (as we did throughout), we may have had about an 148.2 km (80 nmi) error in length, less than 18.5 km (10 nmi) error in width, and 18.5 km (10 nmi) error in position of the geometric center.

Overall Footprint Estimates

Results of the Monte Carlo studies for the three burn scenarios for the center pass are summarized again in the top half of Table 11. The bottom half of this table shows the results when the three individual burn statistics were combined to achieve an overall footprint area to protect, as depicted in Figure 7. Note: the intact (reference) impact point ("INTACT") in this figure corresponds to the intact entry parametric case, not the geometric center of the footprint, which is labeled as "GEOMETRIC_CENTER".

No Monte Carlo cases were run for the Western Pass, only an intact reference impact trajectory. We approximated the corresponding extreme 3σ impact points by applying the same uprange/downrange deltas from the intact reference point as we saw for the Center Pass. Doing so provided a graphical estimate of the Western Pass footprint in Figure 8, and estimates of length and range position in Table 12 (note that crossrange position, length, and width do not change from the Center Pass, due to the approximation).

Footprint Proximity to Land Masses

Considering the proximity to the west coast of South America, we found that the closest a piece can come, with 3σ probability, was 4260 km (2300 nmi) in range (the toe of the TASC footprint for the cold/late burn) for the center pass, and 6465 km (3490 nmi) in range for the western pass.

After looking at various islands in the Pacific Ocean, we considered Figure 9 which plots the approximate centerlines for the overall footprint area to protect (covering all burn scenarios), for both the center and western pass (note that the center footprint extends beyond the right-hand-side of the page). To show the proximity to islands, we drew circles around those that were closest to the footprint of radius equal to half the footprint width (whole width was 67.5 km, or 36 nmi), plus either 46.3 km (25 nmi) if United States soil, or 370.4 km (200 nmi) if foreign soil. The 46.3 km (25 nmi) and 370.4 km (200 nmi) values are landmass miss distance guidelines suggested in NASA Safety Standard 1740.14 (ref. 8). By adding half the footprint width to the radius of these circles, the figures are interpreted as: the 3σ edge of the footprint was closer than the corresponding land-miss guideline (46.3 or 370.4 km) if the groundtrack went through a circle. Thus, we saw that the center pass footprint completely met NSS 1740 guidelines, while the western pass did not. However, for both passes, predictions showed that no island could hit by debris within a 3σ probability. For the center pass, the closest a piece could come to land, with 3σ probability, was 250 km (135 nmi) in crossrange to the Hawaiian Islands. For the western pass, the closest a piece could come to land, with 3σ probability, was 119 km (64 nmi) in crossrange to Palmyra Island. Other values are presented in Table 13.

CONCLUSIONS

This document presented results of Monte Carlo simulations of the CGRO entry, assuming breakup at 83.8 km (45.25 nmi), and debris flight to ground impact. Three burn scenarios were considered: nominal burn #4, 10% hot burn #4, and 10% cold and late burn #4. These scenarios were studied for both a prime, or "central pass" debris target zone, as well as a backup, or "western pass" zone. In addition to size and location results, island proximity was studied. These results showed that the Central Pass had no safety issues associated with it, and although the western pass did not satisfy land miss distance guidelines in NASA Safety Standard 1740.14, it was also seen that no debris should hit any land mass within a 3σ probability for either pass. All primary results are summarized in Table 14.

MODELING CHANGES SINCE CGRO

Since the successful deorbit and safe disposal of CGRO, the methods detailed herein have been continuously improved. These improvements include improved statistical handling of non-Gaussian output, leading to valid results for any number of Monte Carlo cases, with appropriate protection and confidence levels. Capabilities to model explosions during breakup, and breakup over an altitude range are also in the process of being added. The crossrange calculation is also being changed to be relative to the entry groundtrack, not the orbital groundtrack frozen at the initial state vector time.

Table 1: Intact (Reference) Mass and Aerodynamic Data, Intact CGRO Vehicle

Variable	Intact (Reference) Value
Mass	14010.0 kg (30886.7 lb)
Drag Coefficient	1.555
Aerodynamic Surface Area	46.0 m^2 (495.14 ft^2)
Ballistic Coefficient	195.8 kg/m^2 (40.1 psf)
L/D Ratio	0.0
Lift Coefficient	0.0

Table 2: Monte Carlo Mass and Aerodynamic Data, Intact CGRO Vehicle

Variable	Mean Value	Dispersion Limit (Uniform Distribution)
Mass	14010.0 kg (30886.7 lb)	36.6 kg (80.7 lb)
Drag Coefficient	1.555	0.645
Aerodynamic Surface Area	46.0 m^2 (495.14 ft^2)	0.0 m^2 (0.0 ft^2)
L/D Ratio	0.0	0.0

Table 3: Debris Ballistic Coefficient Ranges for Previous Disposal Analyses (ref. 6)

Entry Vehicle	Ballistic Coefficient Range
Space Shuttle External Tank	13.7 - 283.2 kg/m^2 (2.8 - 58 psf)
Apollo Service Module	2.4 - 463.8 kg/m^2 (0.5 - 95 psf)
Soyuz Service Module	6.8 - 566.4 kg/m^2 (1.4 - 116 psf)
Skylab	4.9 - 1562.4 kg/m^2 (1 - 320 psf)

Table 4: Monte Carlo Mass and Aerodynamic Data, General Debris

Variable	Mean Value	Dispersion Limit (Uniform Distribution)
Mass	22.7 kg (50.0 lb)	0.0 kg (0.0 lb)
Drag Coefficient	1.0	0.0
Aerodynamic Surface Area	4.66 m^2 (50.1852 ft^2)	4.63 m^2 (49.8148 ft^2)
L/D Ratio	0.075	0.075
Bank Angle	0.0 deg	180.0 deg

Table 5: Monte Carlo Mass and Aerodynamic Data, TASC Debris Only

Variable	Mean Value	Dispersion Limit (Uniform Distribution)
Mass	528.6 kg (1165.25 lb)	0.0 kg (0.0 lb)
Drag Coefficient	1.44	0.0
Aerodynamic Surface Area	0.30 m^2 (3.20685 ft^2)	0.01 m^2 (0.08255 ft^2)
L/D Ratio	0.075	0.075
Bank Angle	0.0 deg	180.0 deg

Table 6: Breakup Altitudes Found in Previous Disposal Analyses (ref. 6)

Entry Vehicle	Breakup Altitude
Space Shuttle External Tank (Actual Results)	61.1 - 83.3 km (33 - 45 nmi)
Space Shuttle Super Lightweight Tank (Assumed)	61.1 - 88.9 km (33 - 48 nmi)
Soyuz Service Module	64.8 - 83.3 km (35 - 45 nmi)
VAST/VASP	77.8 km (42 nmi)
Apollo Service Module	83.3 km (45 nmi)
Skylab	77.8 - 94.5 km (42 - 51 nmi)

Table 7: Reference Points for Range Values

Case	Center Pass	Western Pass
Geodetic Latitude	25.15° South	28.43° South
Longitude	70.46° West	71.21° West

Table 8: Nominal Burn 4 Monte Carlo Results, Center Pass

Statistic	Uprange from South America West Coast (km)	Crossrange from Orbit (km)
General Debris Footprint		
Maximum	9720.22	+80.40
Minimum	8227.70	-30.45
Average	9427.87	+24.72
Standard Deviation	222.80	19.56
Footprint Size	1336.77	117.32
TASC Footprint		
Maximum	8309.74	+71.01
Minimum	7365.03	-25.85
Average	7974.95	+21.84
Standard Deviation	172.24	18.21
Footprint Size	1033.42	109.29
Overall Footprint Geometric Center	8777.26 (4739.34 nmi)	+23.28 (12.57 nmi)
Overall Footprint Size	2638.01 (1424.41 nmi)	117.3 (63.35 nmi)

Table 9: 10% Hot Burn 4 Monte Carlo Results, Center Pass

Statistic	Uprange from South America West Coast (km)	Crossrange from Orbit (km)
General Debris Footprint		
Maximum	10406.39	+76.34
Minimum	9060.91	-33.21
Average	10132.83	+22.11
Standard Deviation	204.52	18.98
Footprint Size	1227.10	113.84
TASC Footprint		
Maximum	9132.55	+66.84
Minimum	8300.66	-26.54
Average	8842.56	+19.13
Standard Deviation	154.22	17.52
Footprint Size	925.30	105.08
Overall Footprint Geometric Center	9563.15 (5163.69 nmi)	+20.63 (11.14 nmi)
Overall Footprint Size	2366.47 (1277.79 nmi)	113.84 (61.47 nmi)

Table 10: 10% Cold/Late Burn 4 Monte Carlo Results, Center Pass

Statistic	Uprange from South America West Coast (km)	Crossrange from Orbit (km)
General Debris Footprint		
Maximum	7233.17	+70.14
Minimum	5278.76	-46.58
Average	6886.18	+12.74
Standard Deviation	272.32	20.65
Footprint Size	1633.91	123.94
TASC Footprint		
Maximum	5437.29	+60.50
Minimum	4079.40	-44.15
Average	4958.01	+8.72
Standard Deviation	231.02	19.59
Footprint Size	1386.11	117.58
Overall Footprint Geometric Center	5984.05 (3231.13 nmi)	+10.74 (5.80 nmi)
Overall Footprint Size	3438.18 (1856.47 nmi)	123.94 (66.92 nmi)

Table 11: Overall Monte Carlo Footprint Results for Center Pass

Statistic	10% Hot Burn 4	Nominal Burn 4	10% Cold/Late Burn 4
Crossrange Position of Geometric Center	+20.6 km (11.1 nmi)	+23.3 km (12.6 nmi)	+10.7 km (5.9 nmi)
Range Position of Geometric Center	9565 km (5165 nmi)	8780 km (4740 nmi)	5985 km (3230 nmi)
Length	2370 km (1280 nmi)	2640 km (1425 nmi)	3440 km (1855 nmi)
Width	114 km (61.5 nmi)	118 km (63.5 nmi)	124 km (66.9 nmi)
Overall Area to Protect:			
Crossrange Position of Geometric Center	15.6 km (8.4 nmi)		
Range Position of Geometric Center	7510 km (4055 nmi)		
Length	6480 km (3500 nmi)		
Width	135 km (72 nmi)		

Table 12: Overall Monte Carlo Derived Footprint Results for Western Pass

Statistic	Overall Values
Crossrange Position of Geometric Center	15.6 km (8.4 nmi)
Range Position of Geometric Center	9685 km (5230 nmi)
Length	6480 km (3500 nmi)
Width	135 km (72 nmi)

Table 13: 3σ Footprint Closest Approach Distances to Various Land Masses

Land Mass	**3σ Closest Approach Distance**
Center Pass	
West Coast of South America	4260 km (2300 nmi) uprange
Hawaii, United States	250 km (135 nmi) crossrange
Western Pass	
West Coast of South America	6465 km (3490 nmi) uprange
Palmyra Island, United States	119 km (64 nmi) crossrange
Washington Island, Kiribati	144 km (78 nmi) crossrange
Fanning Island, Kiribati	165 km (89 nmi) crossrange
Christmas Island, Kiribati	236 km (127 nmi) crossrange
Marquesas Islands, French Polynesia	252 km (136 nmi) crossrange

Table 14: Summary of GRO Debris Footprint Results

Statistic	**Nominal Burn Scenario**	**Covering all Burn Scenarios**
Length	2640 km (1425 nmi)	Monte Carlo: 6480 km (3500 nmi)
Width	118 km (63.5 nmi)	135 km (72 nmi)
Geometric Center Range from South America (in-plane)	Center Pass: 8780 km (4740 nmi) Western Pass: 10960 km (5915 nmi)	Center Pass: 7510 km (4055 nmi) Western Pass: 9685 km (5230 nmi)
Geometric Center Crossrange from Initial Orbit Groundtrack	23.3 km (12.6 nmi)	15.6 km (8.4 nmi)
Closest Approach to South America (uprange, 3σ)	Center Pass: 4260 km (2300 nmi) Western Pass: 6465 km (3490 nmi)	
Closest Approach to an Island (crossrange, 3σ)	Center Pass: 250 km (135 nmi) (Hawaii, United States) Western Pass: 119 km (64 nmi) (Palmyra, United States) 144 km (78 nmi) (Washington Island, Kiribati)	

Figure 1: Center Pass Nominal Burn 4 MC Extreme 3σ Impact Points - General Debris + TASC

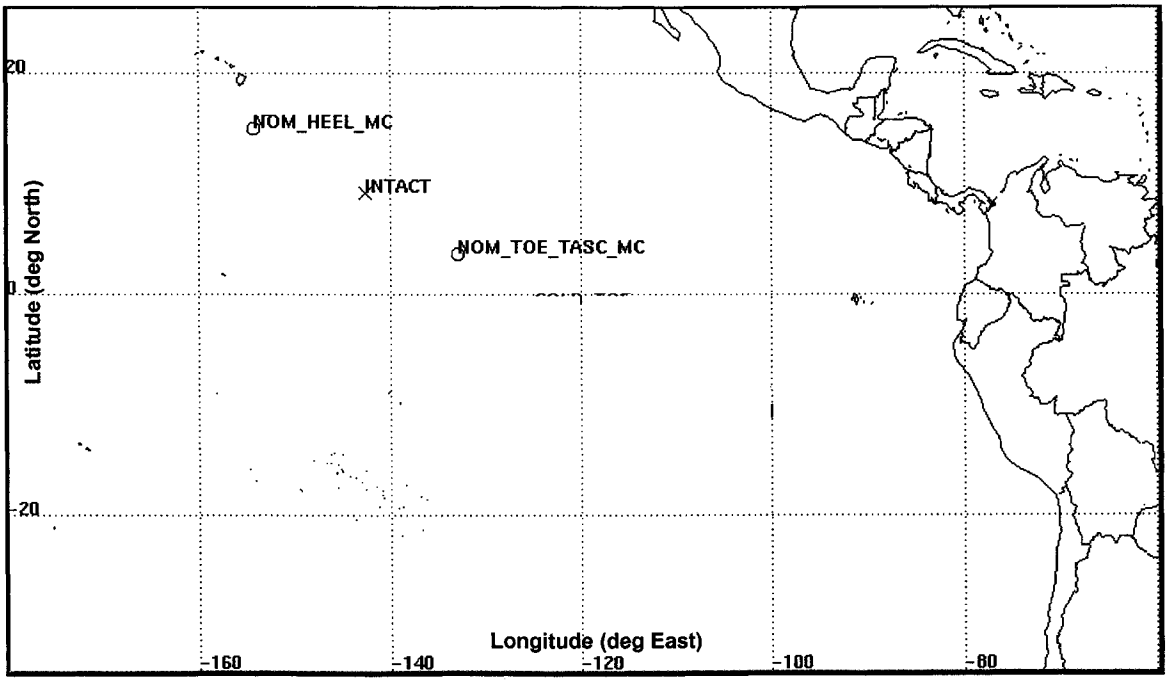


Figure 2: Center Pass 10% Hot Burn 4 Monte Carlo Extreme 3σ Impact Points - General Debris + TASC

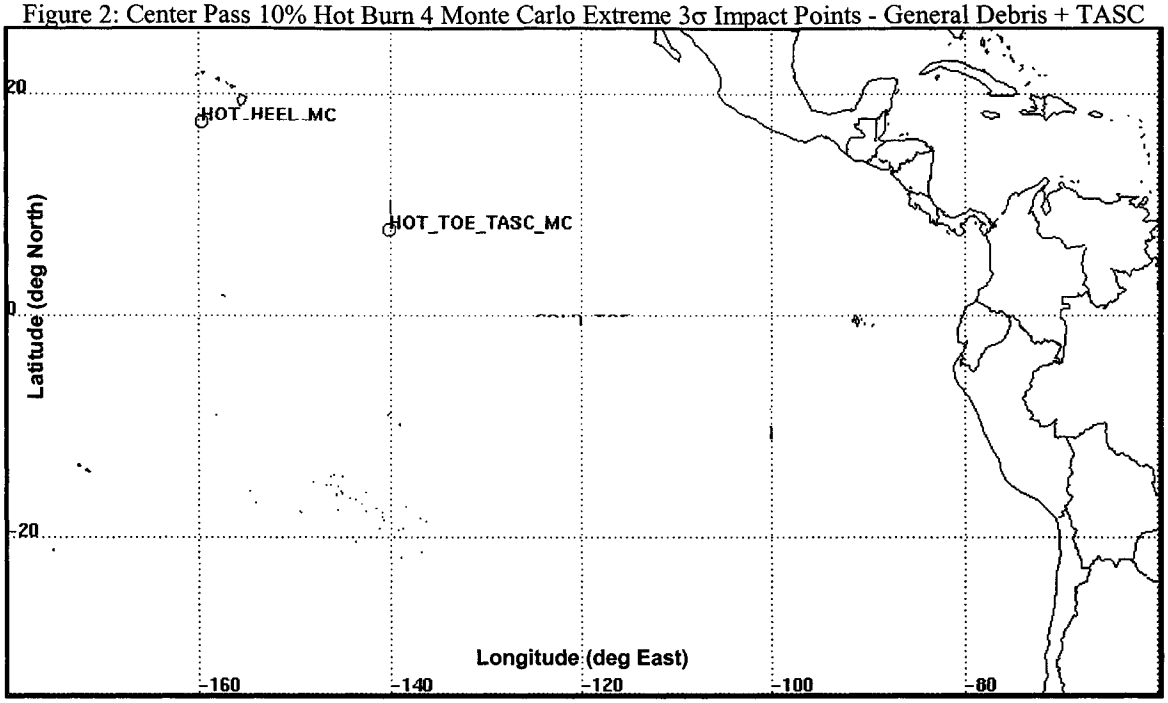


Figure 3: Center Pass 10% Cold/Late Burn 4 MC Extreme 3σ Impact Points - General Debris + TASC

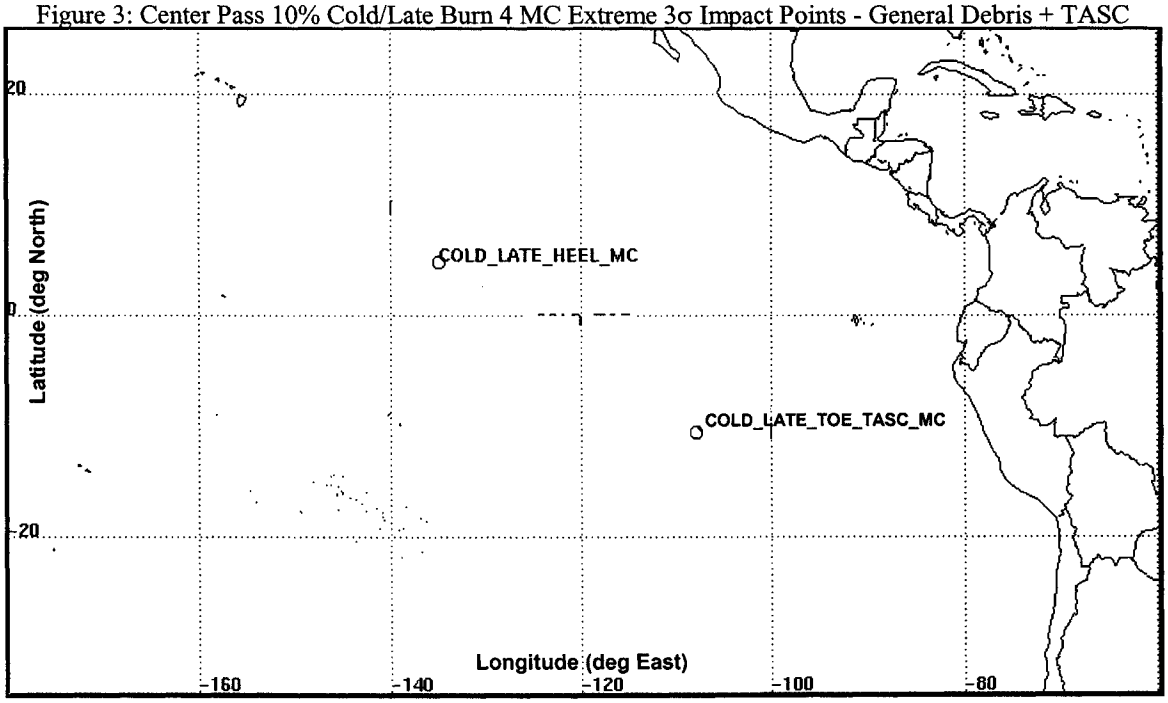


Figure 4: Footprint Length (General Debris) vs. Number of Monte Carlo Runs

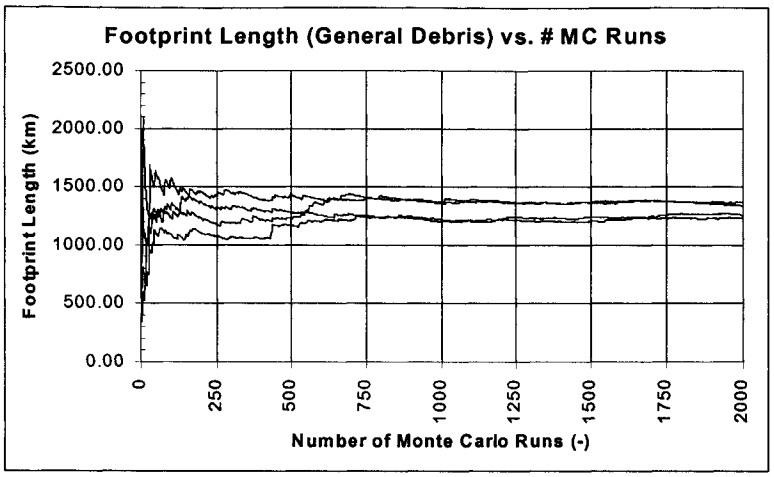


Figure 5: Footprint Width (General Debris) vs. Number of Monte Carlo Runs

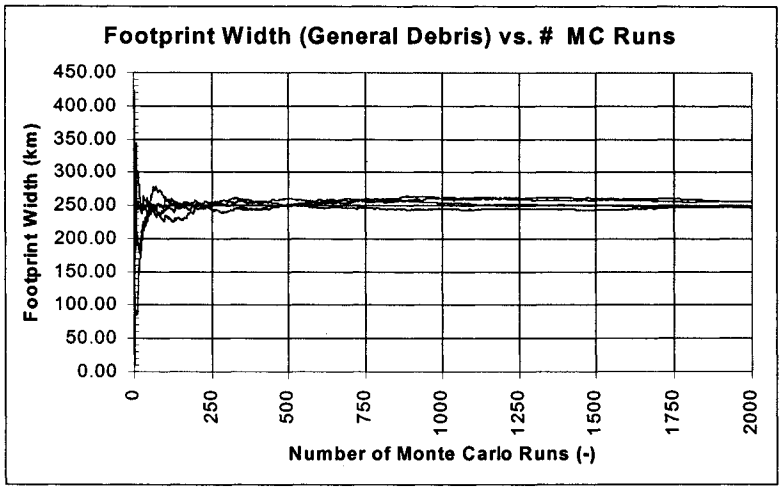


Figure 6: Footprint Center (General Debris) vs. Number of Monte Carlo Runs

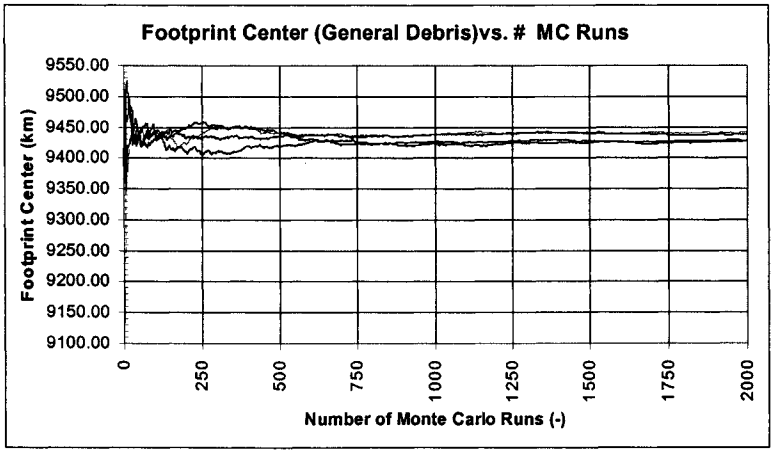


Figure 7: Center Pass Monte Carlo Extreme 3σ Impact Points, All Burns - General Debris + TASC

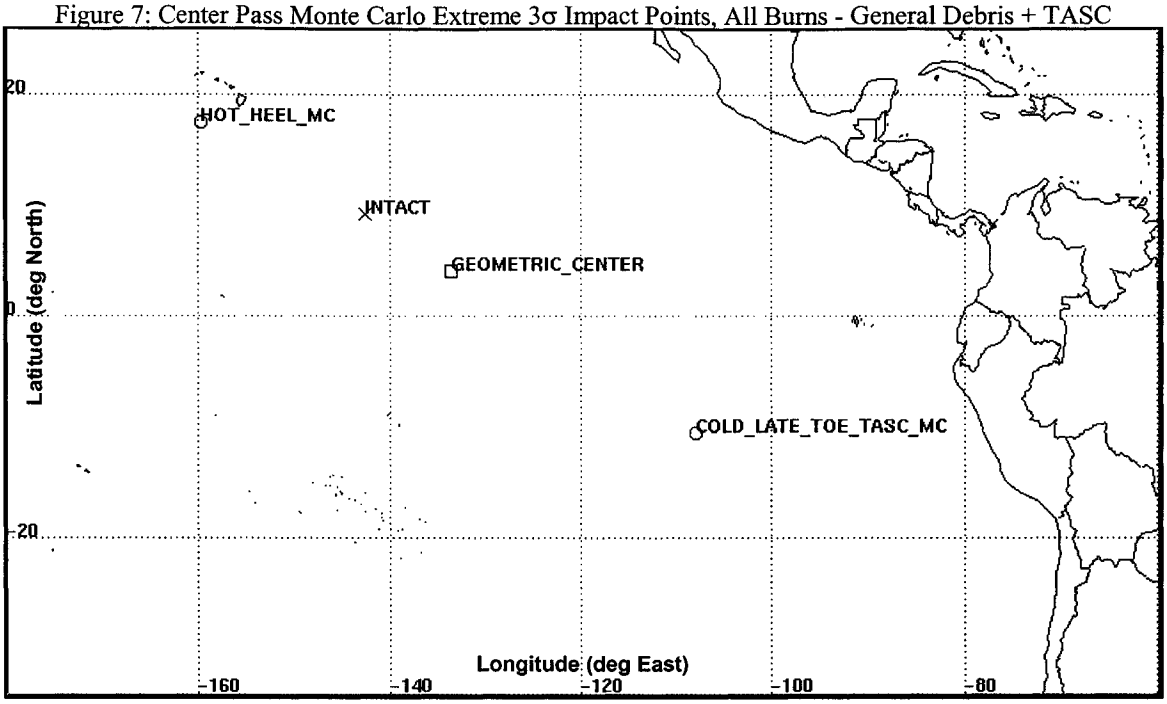


Figure 8: Derived Western Pass Monte Carlo Extreme 3σ Impact Points, All Burns - General Debris + TASC

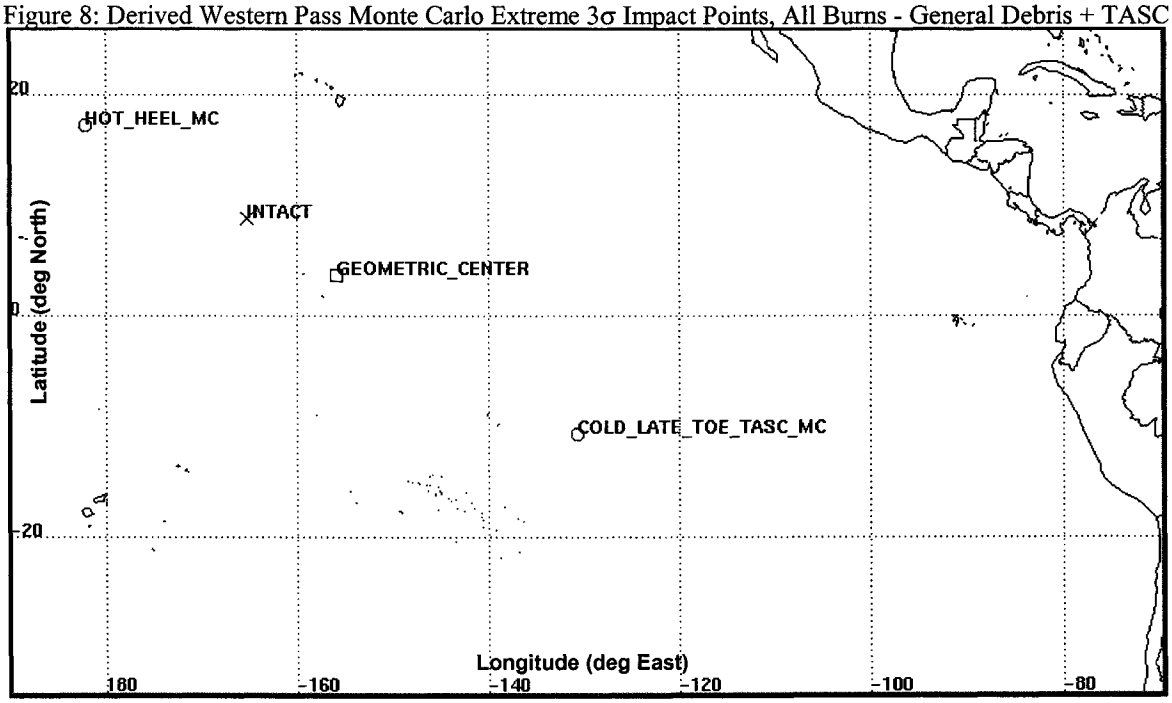
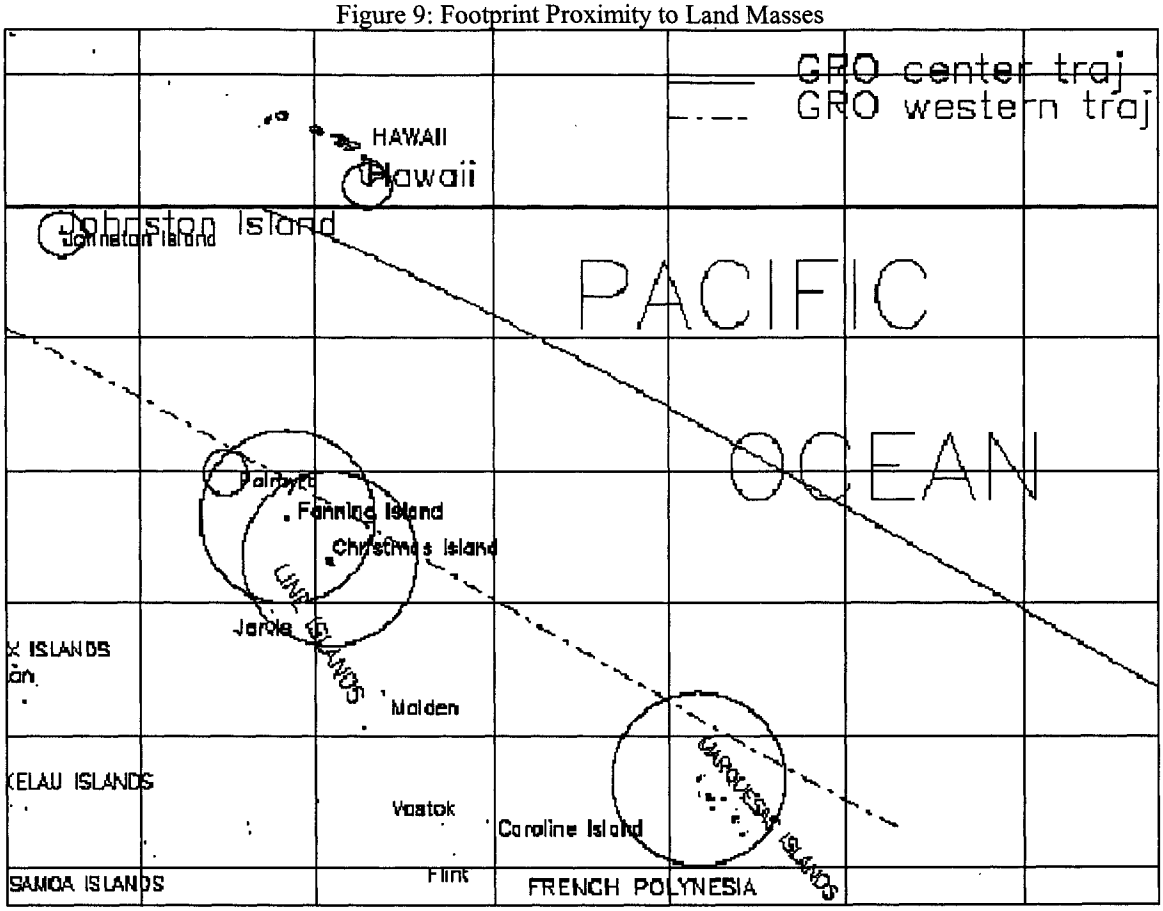


Figure 9: Footprint Proximity to Land Masses



REFERENCES

1 Cole, C. E., "Gamma Ray Observatory Mission Contract, Observatory Reentry Plan (Final)," DRL 023, 40420-85-023-001, TRW Space & Technology Group, Federal Systems Division, July 31, 1985.

2 Brown-Conwell, E. R., "GRO Mission Flight Dynamics Analysis Report: Controlled Reentry of the Gamma Ray Observatory," CSC/TM-90/6001, Mission Report 90001, Flight Dynamics Division, NASA Goddard Space Flight Center, Greenbelt, Maryland, November 1989.

3 Berning, M. J., Sagis, K. D., "User's Guide for the Simulation and Optimization of Rocket Trajectories (SORT) Program, Version 7," NAS9-17900, Lockheed Engineering & Sciences Company, Houston, Texas, October, 1992.

4 Rao, P. P., Woeste, M. A., "Monte Carlo Analysis of Satellite Debris Footprint Dispersion," AIAA 79-1628, 1979.

5 Smith, R. N., Rochelle, W. C., "Reentry Survivability Analysis of Compton Gamma Ray Observatory (CGRO)," JSC-28929, Lockheed Martin Space Operations, NASA Johnson Space Center, Houston, Texas, March 2000.

6 Herdrich, R. J., and Nguyen, P. D., "Super Lightweight Tank (SLWT) Footprint Analysis: Technical Report," JSC-27712, NASA Johnson Space Center, Houston, Texas, March 01, 1997.

7 Refling, O., Stern, R., and Potz, C., "Review of Orbital Reentry Risk Predictions," Aerospace Report No. ATR-92(2835)-1, Programs Group, The Aerospace Corporation, El Segundo, California, July 15, 1992.

8 Gregory, F. D., "Guidelines and Assessment Procedures for Limiting Orbital Debris," NASA Safety Standard 1740.14, August 1995.

ORBIT DETERMINATION FOR THE QUIKSCAT SPACECRAFT*

Blair F. Thompson, Matthew C. Meek, Daniel G. Kubitschek**,
Kenn L. Gold, Penina Axelrad, George H. Born

Colorado Center for Astrodynamics Research
Department of Aerospace Engineering Sciences
University of Colorado
Boulder, CO 80309

ABSTRACT

The QuikSCAT satellite was launched on June 19, 1999. The primary mission payload is the SeaWinds scatterometer designed to measure wind speed and direction near the ocean's surface. The Colorado Center for Astrodynamics Research (CCAR) was contracted to provide an operational orbit determination (OD) system capable of providing 100 meter (3σ) orbit accuracy using the navigation solutions computed by the onboard GPS receivers, and to investigate two alternative OD strategies. Based on several sample solutions, the operational OD system is shown to produce root-mean-square (RMS) position errors between 28 and 31 meters, seven-day overlapping arc position errors between 15 and 25 meters (RSS), and one-day arc overlaps between 5 and 6 meters (RSS). Periodically, ten minute bursts of pseudorange and carrier phase data are telemetered to the ground. We have investigated the performance of an OD system based on these observations. Orbits generated with the smoothed navigation solutions differ from single-differenced pseudorange solutions by less than 10 meters (RSS). A third investigation considers the feasibility of a back-up system using antenna azimuth and elevation angles from three ground tracking stations. Results of processing these three data types are presented: 1) GPS navigation solutions, 2) GPS pseudorange burst data, and 3) ground-based azimuth and elevation data.

INTRODUCTION

The QuikSCAT spacecraft was launched from Vandenburg Air Force Base on June 19, 1999. It carries the SeaWinds scatterometer for measuring wind speed and direction near the ocean's surface. The operational orbit and spacecraft characteristics are listed in Table 1.

Table 1. QuikSCAT Spacecraft Characteristics and Mean Orbital Elements (4 April 2001)

Orbit Type	Sun-Synchronous
Period	101 min
Altitude (perigee x apogee)	807 x 827 km
Inclination	98.63 degrees
Eccentricity	0.0001187
Argument of Perigee	62.426 degrees
Spacecraft Mass	820.48 kg
Cross-sectional Area (average)	3.2 m^2

* work supported under contract 97BSM00005 by Ball Aerospace Systems and University of Colorado's Laboratory for Atmospheric and Space Physics (LASP)
** currently with the Navigation and Mission Design Section, Jet Propulsion Laboratory, Pasadena, CA

The QuikSCAT spacecraft was built by Ball Aerospace Systems Division, and is being operated by the University of Colorado's Laboratory for Atmospheric and Space Physics (LASP). The program is managed by NASA's Jet Propulsion Laboratory (JPL). Prior to launch, the Colorado Center for Astrodynamics Research (CCAR) was contracted to provide an operational orbit determination system capable of providing 100 meter (3σ) orbit accuracy using the GPS navigation solutions computed onboard and brought down in the telemetry stream.

The QuikSCAT orbit determination system consists of an advanced user interface built around the MicroCosm orbit determination software system[1]. The user interface was designed to take care of several tasks required to routinely generate daily orbit ephemerides at 30-second intervals. The orbit ephemerides are computed in the inertial true-of-date coordinate system. The observations consist of once-per-second Motorola Viceroy GPS receiver NAV positions. Because the velocity solutions are not computed to better than 1 meter per second, orbit arcs are fit only to the navigation position solutions.

MicroCosm uses a high fidelity dynamic model consisting of the following[2]:

- JGM-3 gravity model complete to degree and order 70 (70x70)
- Jacchia-71 atmospheric density model
- Tabular data consisting of the latest solar flux and geomagnetic data (updated on a weekly basis)

Shortly after launch, it was determined that the NAV solutions have a consistent timing offset of exactly 1 second. The QuikSCAT operational orbit determination system accounts for the timing error in a FORTRAN routine. This routine is used to convert the NAV position and velocity information to the proper format for use by the MicroCosm tracking data formatter (TDF).

ORBIT DETERMINATION USING GPS NAVIGATION SOLUTIONS

In order to assess the orbit accuracy, six 7-day arcs were selected for the period immediately following the QuikSCAT orbit raising campaign. The 7-day length was selected in order to ensure a good estimate of the orbit and an accurate orbit prediction. Each arc overlaps its adjacent arcs by 24 hours. Onboard computed GPS navigation solutions (NAV positions) were used as data for each of the 7-day orbit arcs. Spacecraft position and velocity at reference epoch and a single drag coefficient (C_d) were estimated for each arc using the first NAV position in each data file as the initial estimate. Table 2 shows the root-mean-square (RMS) of fit to the NAV positions and the estimated drag coefficient for each of the 6 arcs computed.

Table 2. Orbit determination results for 7-day arcs

Arc	RMS of fit (m)	Estimated C_d
1 15-22 July 1999	28.8	4.1
2 21-28 July 1999	29.1	3.1
3 27 July – 3 August 1999	30.7	3.3
4 2-9 August 1999	29.1	3.0
5 8-15 August 1999	29.4	1.7
6 14-21 August 1999	28.8	1.8

Orbit determination consistency was assessed through overlap comparisons of successive arcs. Orbit comparisons (differences) were computed at one minute intervals for the 6 arcs listed above. Table 3

shows the mean and RMS of differences in the radial, alongtrack, and crosstrack directions. Also included in Table 3 is the root-sum-square (RSS) of the RMS of difference in each direction; this is representative of the total error in position.

Table 3. 24-hour overlap comparisons for 7-day arcs

Overlapping Arcs	Mean (m) Radial	Alongtrack	Crosstrack	RMS (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1 and 2	0.17	9.1	-0.04	6.73	17.5	3.57	19.1
2 and 3	0.22	18.9	0.04	6.05	24.1	2.80	25.0
3 and 4	-0.24	15.4	0.00	5.53	21.2	2.43	22.0
4 and 5	0.21	11.2	0.00	3.83	15.6	3.02	16.3
5 and 6	-0.35	3.51	-0.03	2.94	15.0	2.49	15.5

Based on overlap comparisons for the 6 arcs shown, the orbit accuracy (total position) for the QuikSCAT spacecraft is estimated to be 25 m (1σ) or better. This accuracy is within the 100 m (3σ) requirement. However, the estimates of the drag coefficient are not consistent from arc to arc. The RMS of fit to the tracking data seems to be smallest for extreme values of the drag coefficient. The variation from arc to arc could be caused by errors in the gravity and atmospheric drag model for QuikSCAT. The alongtrack differences seen in the overlap comparisons dominate, further suggesting inconsistent modeling of forces acting in the alongtrack direction or systematic timing errors that manifest themselves in the alongtrack direction.

The Department of Defense turned off selective availability (S/A) on the GPS signals on May 1, 2000. An orbit determination study was conducted using data collected after this date in order to determine the effects of S/A on the QuikSCAT orbit accuracy. Five orbit arcs were computed after S/A was removed from the GPS signals. As before, each arc was seven days long, and 24 hour overlaps were used to compare the arcs to each other. The results are shown below in Tables 4 and 5.

Table 4. Post S/A orbit determination results

Arc	RMS of fit (m)	Estimated C_d
1 3-10 May 2000	10.97	3.34
2 10-17 May 2000	12.3	3.42
3 17-24 May 2000	11.51	3.19
4 24-31 May 2000	12.31	2.53
5 31 May – 7 June 2000	12.63	2.58

Table 5. 24-hour post S/A overlap comparisons for 7-day arcs

Overlapping Arcs	Mean (m) Radial	Alongtrack	Crosstrack	RMS (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1 and 2	0.55	14.93	-0.06	4.69	25.13	3.11	25.27
2 and 3	-0.32	17.70	0.01	5.53	22.99	3.13	23.85
3 and 4	0.36	19.79	-0.05	5.96	25.27	3.78	26.23
4 and 5	-0.16	27.62	0.03	6.72	30.86	2.21	31.66

These results show that the RMS of fit is better than the results obtained using data with S/A on the GPS signals. Because S/A is high frequency relative to the dominant orbit perturbation frequencies, the total position RSS for overlapping arcs is about the same magnitude as the arcs with S/A present.

An analysis was also conducted using 1-day arcs for the period July 15-18, 1999. This is the same data set used for the burst data analysis in the next section of this paper. The results of the NAV orbit fits are show in Tables 6 and 7. The fact that the 1-day arcs are more consistent than 7-days arcs is attributed to errors in the dynamical model, mostly gravity and drag i.e., these errors do not affect the orbit as much for the shorter arcs.

Table 6. Orbit determination results for 1-day arcs

Arc	RMS of fit (m)	Estimated C_d
1 15-16 July 1999	29.26	0.67
2 16-17 July 1999	28.51	0.63
3 17-18 July 1999	29.44	0.56

Table 7. 3-hour overlap comparisons for 1-day arcs

Overlapping Arcs	Mean (m) Radial	Alongtrack	Crosstrack	RMS (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1 and 2	0.03	-2.48	0.04	1.44	4.13	2.10	4.85
2 and 3	0.03	5.25	-0.15	1.48	5.91	1.37	6.25

ORBIT DETERMINATION USING GPS PSEUDORANGE BURST DATA

In standard operation, the GPS receiver onboard QuikSCAT does not send raw pseudorange and carrier phase data to the ground. The standard downloaded GPS data consists of onboard computed state vector solutions. However, the receiver, and satellite communication sub-systems can be configured to send short sets of raw GPS data (pseudorange and carrier phase), along with other telemetry information. These short sets consist of approximately ten minutes of contiguous data gathered three to four times per day. These data sets are referred to as "burst data."

Processing with Gipsy/OASIS II Software

Processing with Gipsy/OASIS II software (GOA) began by fitting an orbit through the NAV positions to generate a reference trajectory. The reference orbit converged to approximately 30 meters RMS for the NAV positions for all three days, which is consistent with the performance of the MicroCosm fit. The reference orbit was then used as the nominal orbit in the GPS solution with the burst data, and was the basis of comparison for orbit results.

The GOA solutions estimate satellite position and velocity, as well as a stochastic timing bias for the onboard receiver's clock. GPS orbits and clocks were held fixed to solutions from the global network which are available online from JPL. Both pseudorange and carrier phase were used to produce the GOA results.

Results are shown in Table 8 for the three days for which burst data was available. In general, the GPS orbit solutions are consistent with expectations during periods when burst data observations are present. The GOA processing consists of forward filtering with a batch-sequential filter, followed by a backwards smoothing of the data. When the solution orbit is mapped, orbit trajectory points are produced for every epoch in the arc. The results show that the estimated orbit diverges from the reference trajectory during periods between burst data segments. The estimated orbit returns to the reference trajectory when burst data is available.

Table 8. GPS burst data orbit determination results using Gipsy/OASIS II

Comparison to NAV solutions only where burst data is present	RMS of fit (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1 15 July 1999	15.3	18.8	16.6	29.4
2 16 July 1999	16.2	17.4	15.3	28.3
3 17 July 1999	17.2	19.5	18.5	31.9
Comparison of burst data solution to entire NAV solution				
1	62.8	93.0	60.2	127.3
2	65.1	96.1	62.2	131.7
3	77.2	93.2	58.5	134.4

These results show that it is possible to process the GPS burst data, and that a "good" orbit can be produced for periods where data is available. Without data, the solution orbit is seen to quickly diverge away from that produced with the on-board solution. Selective Availability (S/A) was active during the period in which this data was collected.

Processing with MicroCosm Software

Like the GOA processing, processing with MicroCosm[2] begins by fitting an orbit to the NAV positions to generate a reference orbit. This fit is used to provide initial conditions for the solution based on the GPS burst data, and is also used as a basis for comparison of the generated orbits. The MicroCosm results described here are based on processing only the pseudorange data.

In order to process the pseudorange data, it is necessary to know the positions of the GPS satellites. A fit by MicroCosm to the precise ephemeris provided by the International GPS Service (IGS) was used for this purpose. This fit was used to provide MicroCosm with the precise positions of the GPS satellites during processing.

Early processing runs suggested that single day arcs may be too short for processing the burst-data. The value of C_d and the overall orbit accuracy varied widely between single day arcs. For the solution results given below, a three-day arc was used. This longer arc gives a much more reasonable estimate for C_d and compares well to a solution based on smoothing the NAV positions.

Four solution estimates were produced using different combinations of the available data and estimated parameters (see Table 9). All solutions estimated satellite position, velocity and C_d. Solutions involving one-way pseudorange data also required the estimation of receiver clock biases.

Single-difference pseudorange data involves using two one-way pseudorange observations for two different GPS satellites taken at the same epoch. These two observations are subtracted yielding an observation data type which is free of receiver clock biases.

Table 9. Observation data types for burst data solution sets

Solution	Observation Data type	Additional estimated parameters
1	One-way pseudorange	receiver clock bias and drift
2	Single-difference pseudorange	no clock estimation
3	NAV data combined with one-way pseudorange	receiver clock bias and drift
4	NAV data combined with single-difference pseudorange	no clock estimation

The results shown in Table 10 are the differences of positions from the different solutions shown in Table 9 and the reference trajectory based on smoothing the NAV solutions. Note that the estimates of drag coefficients shown in Table 11, although small, are very consistent.

Table 10. GPS burst data orbit determination results using MicroCosm compared to smoothed NAV solutions

Solution	Radial (m)	Alongtrack (m)	Crosstrack (m)	Total Position (m) (RSS)
1	2.5836	8.0280	1.2377	8.5238
2	3.5945	8.1131	1.4876	8.9975
3	1.5190	3.3833	1.0464	3.8535
4	1.5516	3.3514	1.0141	3.8298

Table 11. Drag coefficient estimates from the burst data solutions

Solution	C_d
NAV fit	0.6703
1	0.7376
2	0.6969
3	0.6629
4	0.6662

These results show that it is possible to process the pseudorange burst data and generate a good orbit solution. It is believed that the pseudorange burst data can be also be used to augment the onboard navigation solution. The results above indicate that this is possible. At this time we did not have access to enough contiguous burst data to compute orbit overlap statistics. Future work will include multiple arcs of burst data, both pseudorange and carrier phase, so that overlap comparisons can be performed.

ORBIT DETERMINATION USING AZIMUTH AND ELEVATION TRACKING DATA

The final investigation was for a backup OD system that processes azimuth and elevation angles collected from ground-based tracking stations. There are three tracking stations used to command the

QuikSCAT satellite and download telemetry. As QuikSCAT passes over each station, azimuth and elevation angle observations are collected from the station antennas. A QuikSCAT pass over a tracking station typically lasts about ten minutes. Observations are collected every ten seconds. Approximately five observation sets are collected in a 24-hour period from the three stations combined. The tracking station locations were considered fixed; they were not estimated in the OD process (see Table 12).

Table 12. Tracking Station Data

Station	ID #	Location	Geodetic Latitude	East Longitude	Elevation
AGS	29	Alaska	65.1167331° 65° 7' 0.239''	212.5384511° 212° 32' 18.424''	430.34 m
SGS	30	Norway	78.2303125° 78° 13' 49.125''	15.3928411° 15° 23' 34.228''	455.0 m
WAPS	64	Virginia	37.9249256° 37° 55' 29.732''	284.5234775° 284° 31' 24.519''	-20.10 m

The data used in this analysis were collected during the month of August 2000. From this data, a five-day subset (15 August through 19 August) was selected and assumed to be representative of a typical five-day period. The observation uncertainties (1σ noise levels) of the azimuth and elevation angles were not known for the three tracking stations. By analyzing the pre-fit residuals of the five days of data (nearly 5,000 azimuth and 5,000 elevation observations), the approximate data noise levels were conservatively estimated (1σ) to be one degree for both azimuth and elevation.

Figures 1 and 2 are plots of the azimuth residuals before any processing. The MicroCosm software processes both azimuth and elevation data, but only produces azimuth residual information in a format suitable for plotting. Tabular results of azimuth and elevation residuals indicate that the two data types have similar noise levels. Figure 1 shows all of the azimuth residuals over the five-day period. The residual mean is –0.32 degrees, and the RMS is 6.86 degrees. Figure 2 is a plot of the same data with all points having a residual greater than 3σ (or 3 x 6.86 degrees) discarded. Even after removing the 3σ outliers, the raw data has a relatively high level of noise. Figures 3, 4, and 5 show the unprocessed azimuth residuals for each of the three individual tracking stations. It appears that tracking station #29 (Alaska) provides observations best suited for the OD process. Stations #30 and #64 have large discontinuities and sharp, acute changes in the data.

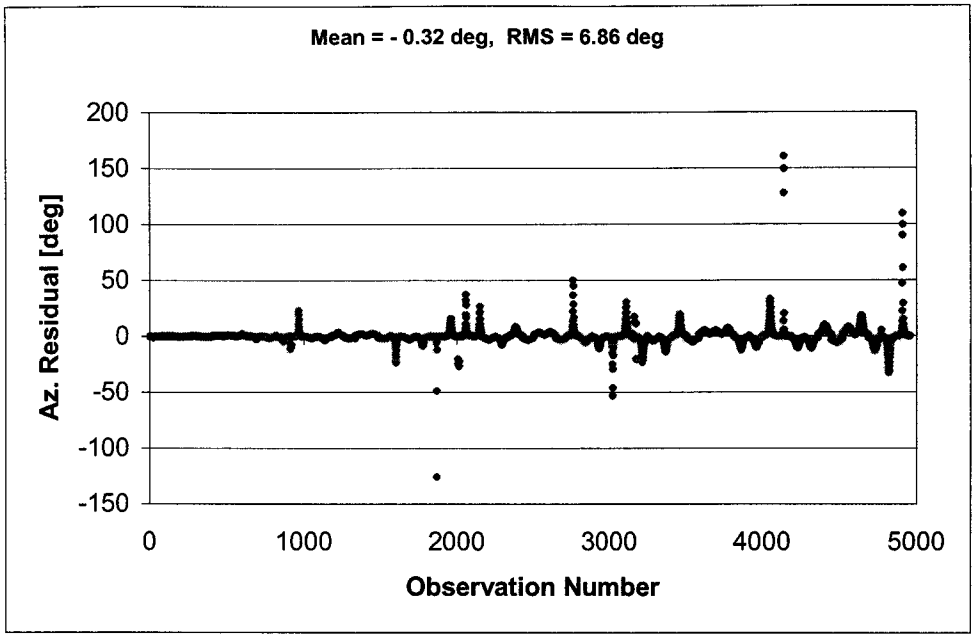


Figure 1. Pre-fit azimuth residuals (all stations)

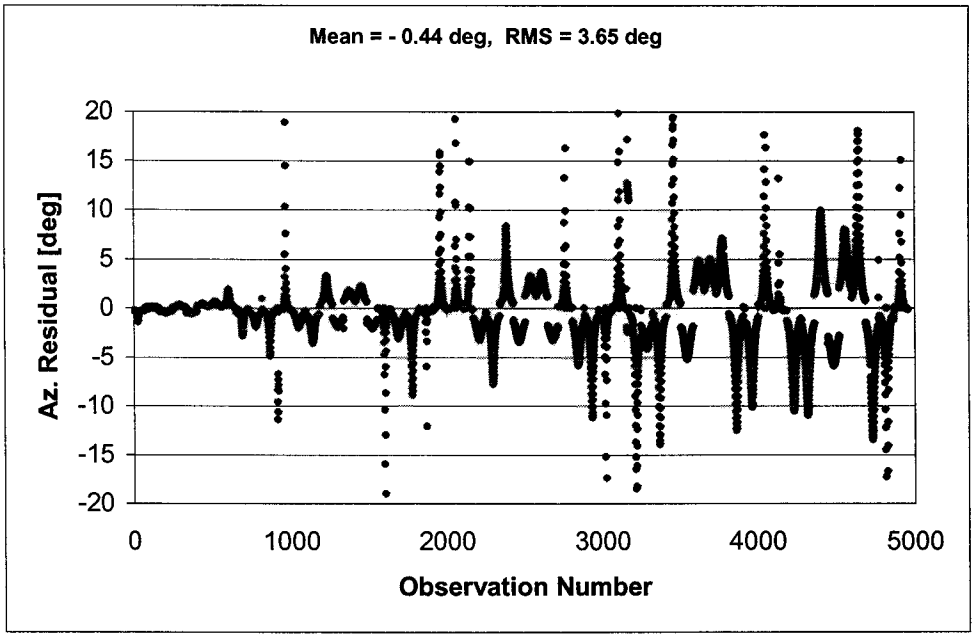


Figure 2. 3σ edited pre-fit azimuth residuals (all stations)

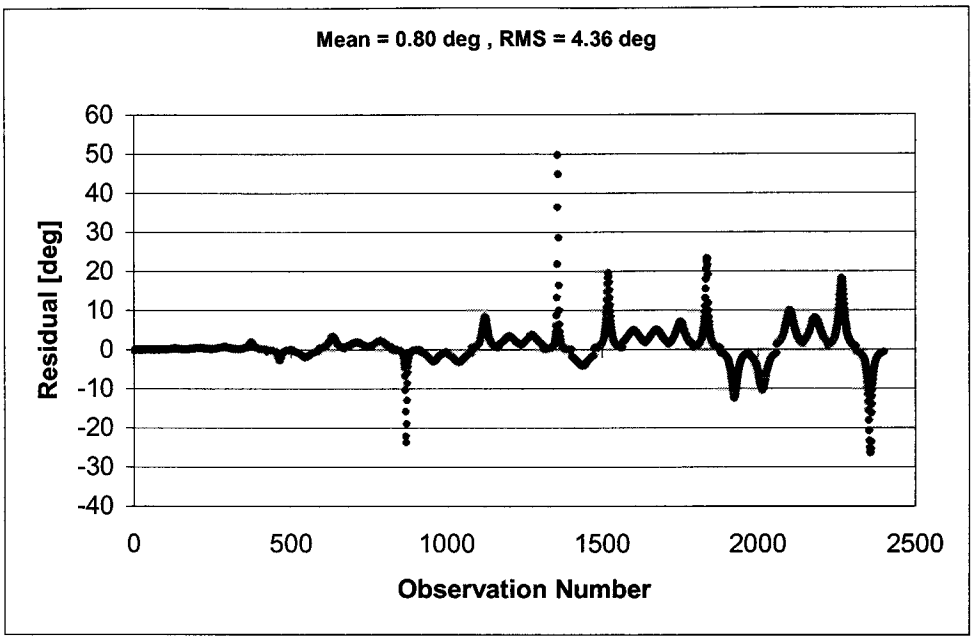


Figure 3. Station #29 (Alaska) pre-fit residuals

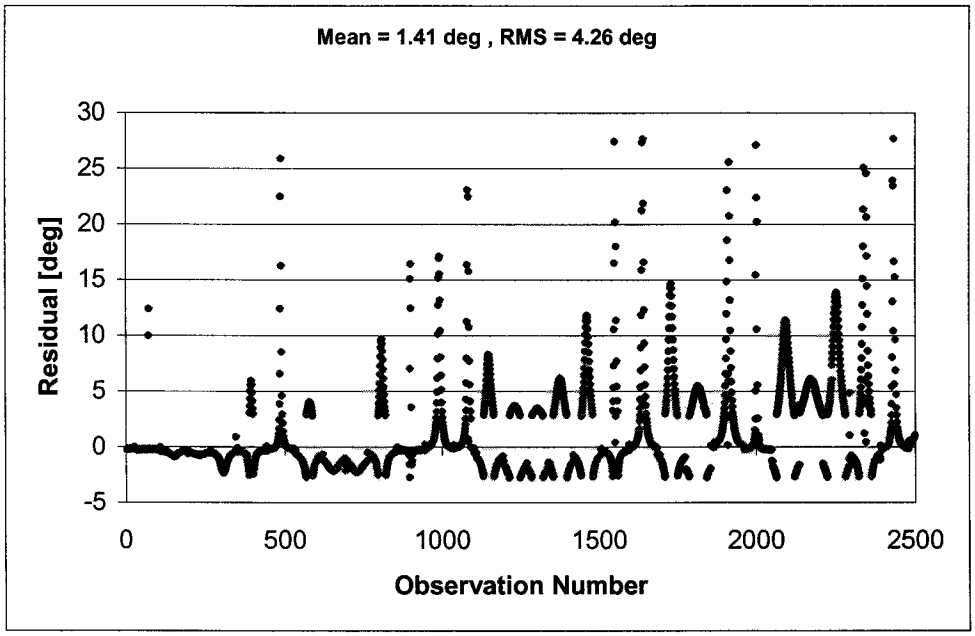


Figure 4. Station #30 (Norway) pre-fit residuals

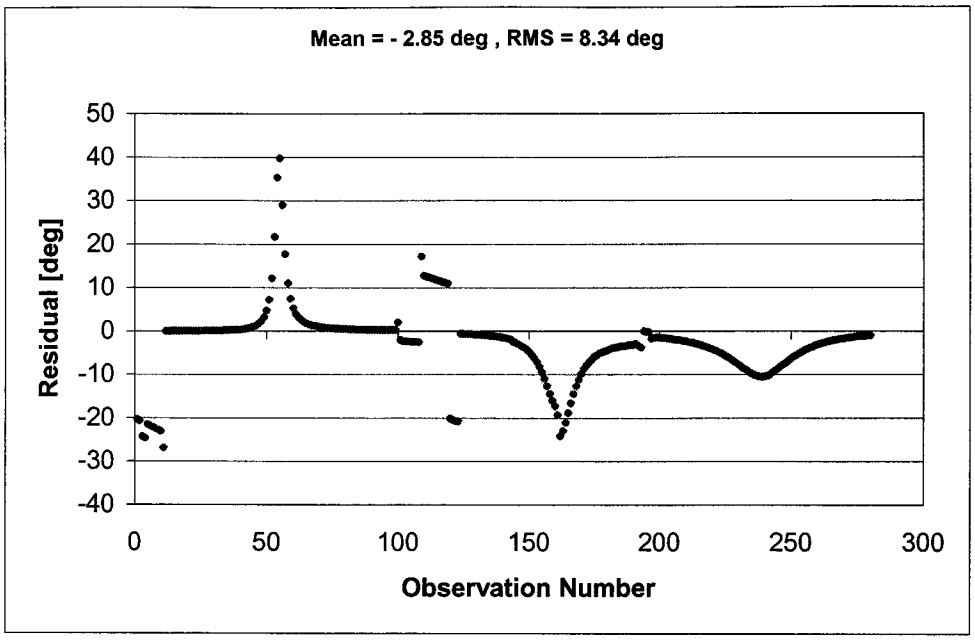


Figure 5. Station #64 (Virginia) pre-fit residuals

The azimuth and elevation angles data from the ground tracking stations is stored in binary files in the Universal Tracking Data Format (UTDF)[3]. The data file name contains the station number and date the observations were made. However, it was discovered that some files had names that did not correspond to the actual data within the file. A program was written in C++ to convert the binary angles data into the GEOS-C format, which is suitable as input to MicroCosm. The program extracts the station, date and time, azimuth, and elevation of each observation. The program reformats the data and produces three text files. Each file contains all of the observations in the proper format for each of the three tracking stations.

The first assessment of orbit accuracy was to compare solutions of the primary, GPS-based OD system with solutions computed by using the backup azimuth and elevation data system. For this analysis, the GPS solution was considered to be the "truth." The truth orbits were computed using five days worth of GPS NAV data. By experimenting, it was found that at least five days of angle data are required to converge on an OD solution reasonably close to that of the truth. The angles-only orbits were computed using angles data gathered over the same five days used for the primary OD system. The a priori state for the first arc of the angles-only system was computed using Gauss's method for angle-only orbit determination[4, 5]. Subsequent a priori states were taken from the ephemerides generated by the angles-only system.

To ensure the best possible comparison between orbits, the dynamic model of the angles-only system was matched as closely as possible to the primary GPS OD system dynamic model. The gravity model, atmospheric model, and integration controls were set identical to the GPS OD model. The elevation angle cutoff was set to ten degrees to attempt to remove atmospheric induced errors encountered at low elevation angles. The a priori variances of the satellite position and velocity were 1.0×10^7 meters^2 and 10.0 meters^2/second^2.

Using only the data from station #29, four orbit arcs were generated with the angles-only OD system (see Table 13 for the dates of these arcs). These arcs were each generated using four separate, five-day sets of data. The post processing azimuth residuals for the first arc are shown in Figure 6. Approximately one-third of the raw data were found to be outlying data points and were edited by

MicroCosm. As expected, it was found that better results (compared to the primary system) were obtained by only using tracking data from station #29. Table 13 contains the estimated values of C_d, which are all consistent. The four arcs were compared to four arcs generated by the primary GPS OD system. Differences in the radial, alongtrack, and crosstrack directions were taken every 60 seconds over a five-day period. Statistic on the differences from each arc comparison can be found in Table 14.

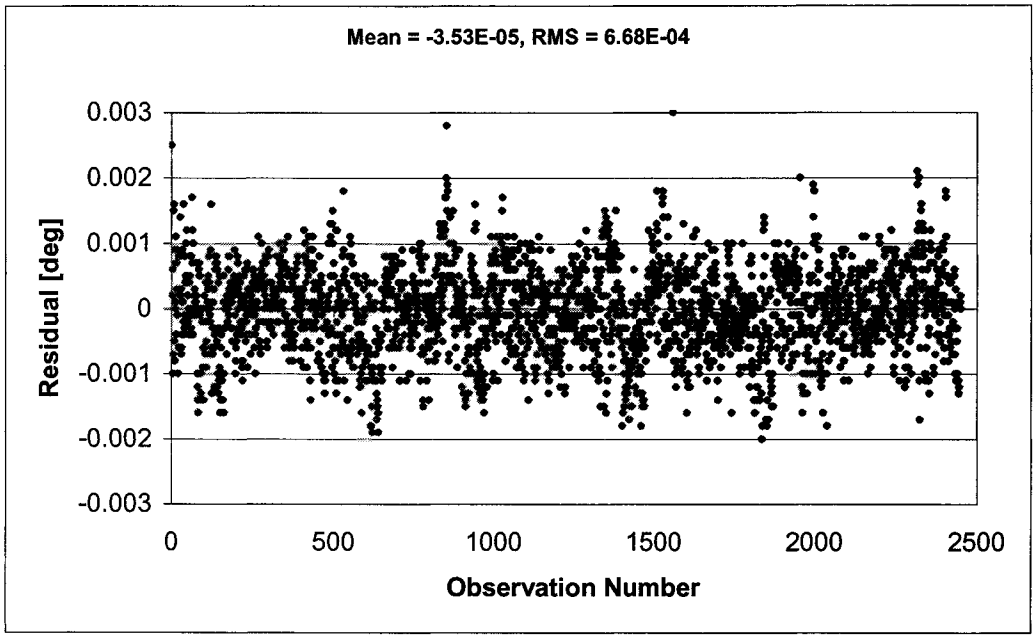


Figure 6. Station #29 (Alaska) post-fit azimuth residuals

Table 13. Drag coefficient estimation results for 5-day arcs

Arc		Estimated C_d
1	14-19 Aug 2000	5.47
2	16-21 Aug 2000	5.50
3	18-23 Aug 2000	5.53
4	20-25 Aug 2000	5.48

Table 14. Angles-only arcs compared with GPS NAV arcs

Arc	Mean (m) Radial	Alongtrack	Crosstrack	RMS (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1	-0.35	-243.52	-0.69	203.58	496.09	266.04	598.60
2	0.42	-477.73	1.62	250.34	693.25	284.41	790.03
3	-1.11	-411.65	-0.55	232.84	637.95	227.09	716.08
4	-0.87	105.48	1.04	159.36	354.54	185.50	430.70

A second means of accuracy assessment was to compare overlapping arcs of solutions using only the angles data from separate sets of data. Three arc overlap comparisons were made. Each arc comparison was made by differencing the overlapping arcs every 60 seconds for a twenty-four hour period. The results of these comparisons are shown in Table 15.

Table 15. 24-hour overlap comparisons for 5-day arcs

Overlapping Arcs	Mean (m) Radial	Alongtrack	Crosstrack	RMS (m) Radial	Alongtrack	Crosstrack	Total Position (RSS)
1 and 2	4.89	67.93	1.26	122.22	261.98	60.33	295.32
2 and 3	1.15	-204.89	0.60	24.06	211.96	26.95	215.02
3 and 4	-2.06	-587.00	-0.12	163.96	673.66	27.99	693.89

The results shown above indicate that the QuikSCAT satellite orbit can be determined using only azimuth and elevation angles from one station, but the desired level of operational accuracy cannot be achieved. This is primarily due to the quality (noise) and systematic errors in the tracking data. In order to determine the typical noise levels on azimuth and elevation that would result in orbits with the desired accuracy, a simulation was conducted. Simulated azimuth and elevation tracking data were generated from a GPS based QuikSCAT ephemeris. Increasing amounts of normally distributed noise (zero mean) were added to the tracking data. The effects of the noise on the orbit accuracy were determined by comparing the orbit computed using the simulated data with the original orbit. The results are shown in Table 16 and Figure 7.

Table 16. Total (3-D) position error

Observation Noise (degrees)	Position Error (m) 1 Station	Position Error (m) 2 Stations	Position Error (m) 3 Stations
0.00	0	0	0
0.05	117	142	99
0.10	234	284	199
0.20	467	568	397
0.40	935	1139	796
0.80	1866	2215	1552
1.00	2330	2867	1999
2.00	5040	5595	4040
4.00	10142	11269	8060
5.00	12636	14361	10182

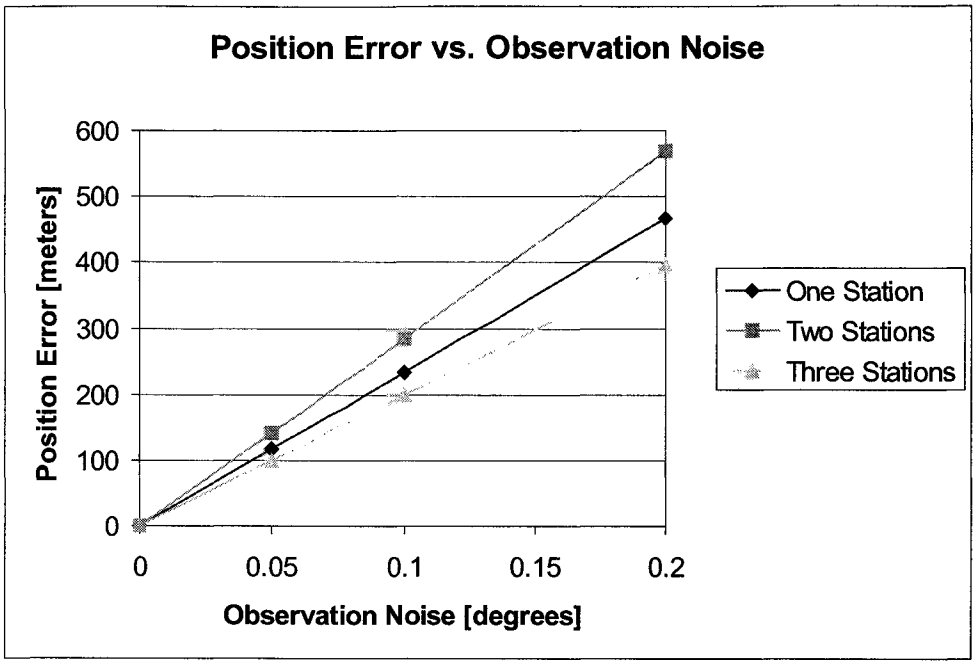


Figure 7.

The data in Table 16 indicates that with three stations and a 1σ noise level no greater than approximately 0.05 degrees on the azimuth and elevation data, assuming no systematic errors, would produce the desired operational orbit accuracy of 100 meters (3σ).

CONCLUSION

We have shown that the operational orbit determination system, using on-orbit GPS NAV solutions as observations, produced RMS position errors between 28 and 31 meters, and seven-day overlapping arc position errors between 15 and 25 meters (RSS). This system fulfills the mission requirement of 100-meter (3σ) position accuracy.

Comparison of the pseudorange burst data solutions to the GPS NAV solutions produced RMS position errors between 4 and 9 meters for 3-day arcs. This, combined with the RMS errors of the NAV solution, indicates that pseudorange burst data is also capable of achieving mission position accuracy requirements. The best approach for incorporating burst data is to use the MicroCosm batch filter, along with both NAV solutions and burst data as observations.

The investigation of using azimuth and elevation data as a back-up OD system show that this type of data could be used to track QuikSCAT in order to command the satellite and keep it operational. The test results show that the back-up OD system computed solutions with a mean 3-D difference from the primary GPS NAV solutions between 100 and 480 meters, and a 3-D RMS difference between 430 and 790 meters over a five-day period. Twenty-four hour overlapping arcs showed a mean 3-D difference between 68 and 587 meters, and a 3-D RMS difference between 215 and 693 meters. Hence the 100-meter (3σ) position accuracy requirement cannot be met using this back-up OD system. If data from all three tracking stations were comparable to that from station #29, we could possibly meet the accuracy requirements. However, it is still possible to operate the payload with the understanding that the position of the satellite will not be as well determined as it would be using the primary GPS OD system.

REFERENCES

1. Davis, George W., et al, "A Low Cost, High Accuracy Automated GPS-Based Orbit Determination System for Low Earth Satellites," *ION GPS Proceedings*, Institute of Navigation, 1997, p. 723.

2. *MicroCosm* software manuals, version 1999, Van Martin Systems, Inc., Rockville, Maryland, November 2000.

3. "Tracking and Acquisition Handbook for the Spaceflight Tracking and Data Network", 450-TAH-STDN (formerly STDN No. 724), Mission Operations and Data Systems Directorate, Goddard Space Flight Center, Greenbelt, Maryland, October 1994.

4. Escobal, Pedro R., Methods of Orbit Determination, Krieger Publishing Company, Malabar, Florida, 1965.

5. Vallado, David A., Fundamentals of Astrodynamics and Applications, The McGraw-Hill Companies, Inc., New York, 1997.

Session 6: In-Flight Attitude Experience

CALIBRATION OF HUBBLE SPACE TELESCOPE FOCAL-LENGTH VARIATIONS USING THE EMBEDDING TECHNIQUE

Lee Barford, Nicholas Tufillaro and Daniel Usikov
Agilent Laboratories, Systems and Solutions Lab

Leonid Marochnik and Robert McCutcheon
Computer Sciences Corporation / Space Telescope Science Institute

ABSTRACT

A modeling method that allows one to rapidly build data-driven models of nonlinear components is discussed. The models are constructed from input/output time domain data and their "embeddings". The notion of models built from embedded data is described in the Taken's Embedding Theorem and has been extensively explored for modeling systems in the physics literature. The authors from Agilent Laboratories are developing practical methods to extend these results to non-autonomous systems by creating tools that allow engineers to rapidly build models for driven nonlinear components. These models can be used in simulation, process control, diagnostics, and sensor calibration.

Using these methods a "black-box" data-driven model is generated to calibrate Hubble Space Telescope (HST) focal-length changes on a 5-minute time grid for the period from 1995-1999. These models are built using a program, CHAOS, developed by Agilent Laboratories. The data-driven model predicts the focus for the measured points about 36 percent better than the Full-Temperature Model (FTM) constructed from a detailed knowledge of the telescope structure. As demonstrated by this HST focal-length calibration, data-driven models, such as those generated with the CHAOS package, have great potential for application to a wide spectrum of HST/Next Generation Space Telescope (NGST) calibration problems. In particular, for sensor calibration applications, black-box nonlinear models can be generated rapidly, which have similar or better performance than models built from a detailed understanding of the system structure.

1. INTRODUCTION

This paper describes a nonlinear data-driven "black-box" model of HST focus behavior at 5-minute intervals for the period 1995-1999. The model uses techniques originally developed in "chaos theory" for describing the behavior of nonlinear dynamic systems based only on observed data. A short description of this model can be found at the top of each quarter-year file on the HST website:

http://www.stsci.edu/instruments/observatory/focus/ephem.html

This paper contains no information about the HST framework and optical telescope assembly as related to thermal effects on the telescope focal length. A comprehensive description of these systems and thermal effects can be found in the paper [1], which also describes three physics based models – the Four-Temperature Model, the Attitude Model, and the Full-Temperature Model – created by John Hershey[1]. We do not discuss these models here, nor do we discuss their comparative pros and cons with respect to the black-box data-driven model from the physical point of view. We do, however, discuss the comparative fits of the various models to the focus measurement data, and we compare model prediction characteristics at times between focus measurements.

We use the CHAOS Program (CP) to rapidly develop a nonlinear black-box model. We refer to this black box model as the Chaos Program Model (CPM) when comparing this model to previously developed models such as FTM. The general approach and method used to build these data-driven models is described in Section 2. Section 3 provides a brief description of what the CP is and how it works. Section 4 of this paper provides additional details about data files used and the models constructed.

One of the goals of this paper is to demonstrate the applicability of black-box modeling using embedding techniques to spacecraft sensor calibration problems. By *black box*, we mean a model that requires no deep information about the device (i.e., science instrument, fixed-head star tracker, etc., to be calibrated) be known before model building. The nonlinear model extraction behind our methods is based on the so-called *dynamic-reconstruction theory*, which is described in Section 2. This theory of nonlinear system identification is based on the Taken's Embedding Theorem, which very roughly states that the evolution of points in the "reconstructed" state space of our data-driven model can be related by a change of variables to that of actual dynamics (a first principles model) believed to be determining the systems motion. Thus, the embedding theorem opens the way toward a general solution for extracting black-box models for nonlinear devices directly from time-domain measurements. We selected the HST focus calibration task for this initial demonstration project because of easy access to focus data and the existence of focus models developed using more traditional methods.[1]

We found that across the period from 1995 through early 1999, the black-box model fits the observed focus data better than the existing Four-Temperature, Attitude, and Full-Temperature models. The root-mean-square (RMS) residuals for the CPM show 42 percent, 48 percent, and 36 percent improvements over the residuals for the three existing models (see details below).

Figure 1a shows focus observations along with FTM and CPM focus predictions from 1995 through the first quarter of 1997 [mostly before HST servicing mission 2 (SM2)]. Figure 1b plots the same data from the second quarter of 1997 through the first quarter of 1999. These two figures demonstrate how well the FTM and the CPM predict the observations.

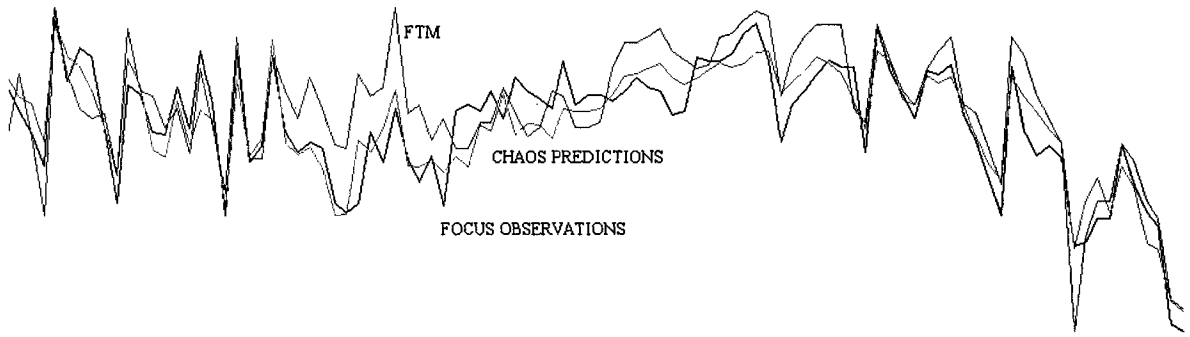


Figure 1a. Focus Observations and FTM and CPM Focus Estimates From 1995 Through First Quarter of 1997 (Mostly Before SM2). The Chaos Program prediction is 29.3% better than FTM.

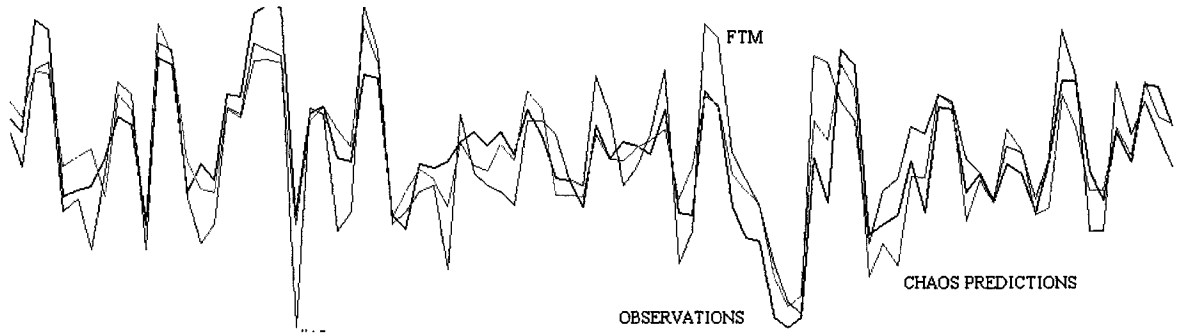


Figure 1b. Focus Observations and FTM and CPM Focus Estimates From Second Quarter of 1997 Through First Quarter of 1999 (Mostly After SM2). The Chaos Program prediction is 36% better than FTM.

---

[1]An HST web site presents the focus data for these models at 5-minute intervals for each quarter of each year from 1994 through 1999. John Hershey provided us with raw temperature data for the same time period.

2. "DYNAMIC RECONSTRUCTION THEORY" for SENSOR and SYSTEM CALIBRATION

If a system is well described by linear theory, then a transfer function can be used for building data driven models for the system [2]. Nonlinear systems, especially those that are not weakly nonlinear, require a different approach. Some methods that have been developed to model and calibrate nonlinear systems include Volterra Series, neural nets, and cluster weighted models to name a few [3]. Here we are concerned with models that are constructed only from input/output data, and which require very little system specific information. The advantage of such data driven models and calibration systems is that they can be developed much more rapidly than detailed first principle models. The disadvantage of such data driven models is that they may require extensive data sets for model training and they typically will have poor extrapolation properties (as an attempt to predict the system's behavior out of the training box of inputs).

A dynamical systems approach to "black box" modeling was first suggested by Casdagli [4] (see [5]). Using time domain input/output data, an attempt is made to embed the original data in a higher dimensional space, built from transforms of the original data of sufficient dimension so that the determinism of the dynamical system is recovered. This approach to nonlinear system identification is sometimes called "Dynamic Reconstruction Theory" [6] and begins with a state space representation of the form

$$\dot{x} = f(x(t), u(t)); \quad y(t) = h(x(t)); \tag{1}$$

or a numerical version of difference equations,

$$x_{n+1} = f(x_n, u_n). \tag{2}$$

In these equations f, x, u are typically vectors and $u(t)$ is the input, drive, or stimulus, $x(t)$ is the state, and $h(t)$ is a measurement function.

Attempts to build the state space models appear hard on at least two counts: first, without any specific form for a model the relevant dynamical variables, x, appear to be unknown, and second, even if one knows what variables are needed to be included, they still may not be accessible to experimental measurements. Both of these issues, essentially the nonlinear order, or dimension of the model, and model selection and calibration are discussed below.

A simple approach to nonlinear modeling in the time domain could begin by plotting the input and output on a graph. Next we could create a function from the stimulus, $u(t)$, to the response $y(t)$,

$$y(t) = F[u(t)]; \tag{3}$$

but this function might not be unique. As seen in Figure 2, for example, two different inputs can have the same output. However, the slope at each of these input points is different. So an "embedded'' input variable created from a vector formed by the input and its derivative does have a unique output in this example.

A key insight of dynamic reconstruction is thus to embed the measured variables to resolve any indeterminacy by building a function not just with $y(t)$, but also transforms of y, for example its numerical derivatives. In more general terms, an "embedding" is a map that places an "m" dimensional manifold, in this case a one-dimensional curve, in a higher "n" dimensional space. We then attempt to build input/output models not based on the measured scalar data, but rather maps on vectors of embedded data.

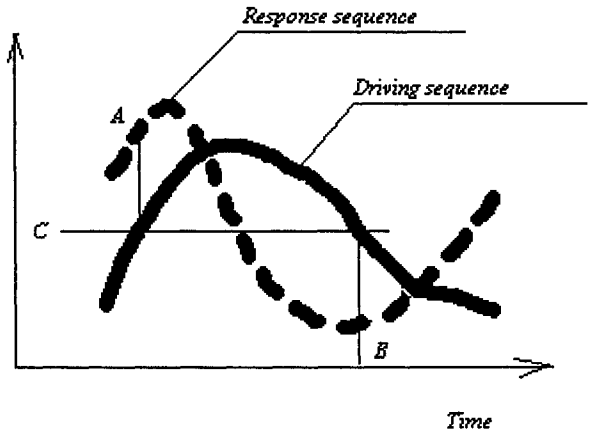


Figure 2. Embedding: Effect of Driving Sequence on Response Sequence

One simple class of models would be a polynomial model that could be of the form,

$$y(t) = a_0 + a_1 u + a_2 \dot{u} + a_3 u\dot{u} + a_4 u^2 + a_5 \dot{u}^2 ...; \tag{4}$$

The unknown coefficients ($a_0, a_1, a_2, ...$) can be determined by least squares. Plotting the embedded trajectory in the enlarged phase space can untangle and remove the indeterminacy. This idea can also be applied to difference equations as well,

$$f(n+1) = F[f(n)] \tag{5}$$

and in effect create a numerical approximation for the differential equations generating the flow.

Due to a theorem of Taken's (with an extension to the driven case by Stark [5]) these embedded models are diffeomorphic to the dynamics of the original system. This means that there is a continuous and differentiable map from the original system trajectory to the new embedded system (created from the measured variables) of sufficient dimension. In particular, deterministic prediction is possible from an embedded model that will mimic the actual dynamics.

Thus, embedding opens the way toward a general solution of extracting black box models for the observable dynamics of nonlinear systems directly from input/output time series data. It can solve the fundamental existence problem, however, the gulf between these theoretical results and practical implementation is wide.

In most of the research physics literature, the components' behavior is described by embedding both the inputs and outputs in the form

$$z(t) = G[y(t-\tau), y(t-2\tau), ..., y(t-l\tau), u(t), u(t-\tau), ..., u(t-(k-1)\tau)] \tag{6}$$

where G is fitted to the data using nonlinear modeling methods such as global polynomials, neural nets, or radial basis functions [3]. The form of the equations shows a "lag" embedding with a time delay, τ, input lag dimension, k, and output lag dimension, l, though in practice one can find that better quality models can often be built using other embeddings such as linear transforms, integral and differential transforms, and wavelets to help bring out the salient dynamical features in the data.

Given the above model form, the problem now reduces to a number of technical issues such as: Determination of the dimensions (k and l), determination of lags (τ) or other forms of embeddings and embedding parameters,

determination of model class G, and fitting the model parameters, model validation, and design of excitation signals (where possible) for a given model/signal class.

It might be helpful to point out that this relation between a continuous dynamical system and an embedded model built from time-delayed input/output signals can be made explicit in the case of linear systems. The details for an algorithm are described by Franklin [7], which shows how to go from the linear system with matrices A, B, and C to a model based only on delayed variables. Unfortunately, no explicit constructive proof exists for nonlinear systems.

For embeddings built from a time delay lag, τ,

$$y(t) = G[y(t-\tau),...,y(t-l\tau),u(t),u(t),...,u(t-(k-1)\tau)] \tag{7}$$

we can use an extension of the algorithm for the theory of embedded autonomous systems known as ``False Nearest neighbors" [8], [9]. The smallest "k" and "l" can be found by creating a statistic that checks if vectors close in a delay space are also close in a delay space of greater dimension. If they are not, they are the false neighbors and G is not single valued. This diagnostic is independent of G.

For models built from time delays, one should estimate the time lag value, τ. Again, a diagnostic can be used from the theory of autonomous systems. The mutual information or the first zero of the autocorrelation function [8] work well. In cases where there is a single dominant frequency band, both of these diagnostics often turn out to be about one-quarter of the dominant frequency. In other words, τ is chosen so that the delay variables are decorrelated as much as possible. In practice, though, these diagnostics are not nearly so useful as software tools such as the Chaos Program that allows one to rapidly build and test models with different combinations of embedding functions and parameters.

Once a suitable embedding is found, the next task is to find the function approximation of G. We try to keep things here as simple as possible. First we usually try a global polynomial fitted by least squares. For other applications we have had some success with radial basis functions

$$y = \alpha + \beta u + \sum_{i=1}^{N} \omega_i \phi(\| c_i - u \|) \tag{8}$$

that use optimization algorithms which can automatically determine the number and placement of the basis functions [10]. We found that the neural net approach sometimes works better on extrapolation than the polynomial-based methods, but the neural net requires much more time for building and verifying the created models.

With software tools such as the CP, one can rapidly test out different basis functions and model structures. For validation, i.e. testing the models on overfitting and underfitting, simple cross validation methods are sufficient in the examples we consider here. That is, models are only built from a subset of the available data (the training set), and then tested on the remaining data.

3. CHAOS PROGRAM

The CP uses embedded signals to build models as described in Section 2. The program provides a graphical user interface with tools that assist the user with importing, inspecting, and analyzing data. Next the program allows the user to create various data embeddings, and finally to fit and evaluate models based on the embedded data. For the models considered here the initial input data can be described as a time series of drive and response variables, and embeddings created from the drive variables, such as their numerical derivatives.

To create models, the input and output data are fitted with a polynomial, as in the following equation:

$$f_{response} = x_0 + x_1 f_{drive} + x_2 \dot{f}_{drive} + x_3 f_{drive} \dot{f}_{drive} + x_4 \left(\dot{f}_{drive}\right)^2 + ... \tag{9}$$

The ``embedded'' derivative variables are calculated numerically from the measured drive data. The unknown coefficients $(x_0, x_1, x_2, x_3, x_4, \ldots)$ of the polynomial are found by a least squares methods (LSM) on the training sequence and are then used for prediction.

Thus, the problem solved is a construction of a black box behavioral model for an object with nonlinear behavior (a polynomial of degree higher than one) describable by a functional or differential relation among the state variables. Complete documentation for the program and more details about the CP can be found in the Chaos User's Manual by Usikov [11].

The accuracy of any data-driven model is restricted by the amount of noise in the source data and the accuracy of the computation. The latter can be mitigated by improved computation technique (i.e., by double precision, robust algorithms, etc.). However, the experimental accuracy often cannot be improved. This is one reason why the embedding technique has inherent limitations. Also, these techniques are limited to systems that can be effectively described by only a few dimensions, i.e., embedding variables. This limitation is not as harsh as might first appear since dissipation in even a very complex system can greatly reduce the effective degrees of freedom, the number of modes actually excited. Additional difficulties arise when the data to be modeled have many different time-scales. For these systems additional embedding strategies, such as using wavelet embedding, are very useful. A wavelet embedding algorithm is implemented in the Chaos program and can be explained using the following example. Consider a wavelet with width = 3. The embedding value derived from a source variable is calculated on six sequential points in time: $I_6 = f_6 + f_5 + f_4 - f_3 - f_2 - f_1$. It is a kind of averaged first derivative (i.e., the derivative, which is smoothed over a span of points).

By averaging the data in this way we can create new embedded variables that are more in tune with longer time constants.

4. CHAOS PROGRAM MODEL

To build an HST focus model, we could use any of the data used to create the three already existing models. The first of these, the Four-Temperature Model, is based on four temperature sensors in the light shield near the secondary mirror spider. The second model, the Attitude Model, is based on HST attitude information from the mission scheduler files (Sun angles, off-nominal rolls, occultations, and day/night satellite positions). The third model, the Full-Temperature Model (FTM), is based on a large number of temperature sensors throughout the telescope, including the four sensors in the first model. Because the FTM represents the observations somewhat better than the other two models do, we chose to use the raw data used to create the FTM to build the CPM.

Table 1 shows an extraction from the typical raw data file (from the first quarter of 1997) used to build the CPM.

Table 1. Sample From Raw Data File

Year	Day	Hour	Min.	Julian Date	Temperature Functions (degrees Celsius)						Observed Focus Position (microns)
					T1	T2	T3	T4	T5	T6	
1997	9	22	20	50457.92969	28.8	-9	-24.1	-8.4	-37.3	16.8	
1997	9	22	25	50457.93359	28.6	-8.8	-23	-8.6	-36	16.8	
1997	9	22	30	50457.93750	28.8	-9	-22.7	-8.5	-34	16.8	
1997	9	22	35	50457.94141	28.8	-9	-21.7	-8.1	-32.7	16.8	
1997	9	22	40	50457.94531	28.7	-9.2	-21.2	-8.1	-31.5	16.8	
1997	9	22	45	50457.94922	28.8	-9	-20	-7.9	-30.5	16.8	
1997	9	22	50	50457.95313	28.8	-9	-19.3	-7.8	-30.4	16.8	3.00E-01
1997	9	22	55	50457.95703	28.8	-9	-19.1	-7.6	-30.6	16.7	9.00E-01
1997	9	23	0	50457.96094	28.7	-9.2	-19	-7.4	-30.9	16.8	
1997	9	23	5	50457.96094	28.7	-9	-19.7	-7.4	-31.5	16.8	
1997	9	23	10	50457.96484	28.8	-9	-20.5	-7.4	-32.5	16.8	

T1, T2, T3, T4, T5, and T6 are six functions of many telescope temperature sensors, including four light shield temperatures near the secondary mirror; each function is the mean of a number of sensor measurements extracted by the Observatory Monitoring Program:

- T1: Truss axial temperature difference function
- T2: Truss diametrical temperature difference function
- T3: Aft shroud temperature function
- T4: Forward shield temperature function
- T5: Light shield temperature function
- T6: Primary mirror temperature function

The last column of the table contains focus position observations; in most cases, the value shown in this column is blank, indicating that no observations were made for the corresponding time row. The focus observations are measured in units of secondary mirror microns and are relative to the best focus (zero microns) of the wide field planetary camera 2 (WFPC2)[2]

We have excluded from consideration the focus measurements of 1994 and four focus observations of 1995 and 1996 because no regular temperature data (T1-T6) exist for the appropriate times. Consequently, we are dealing with a typical calibration task: using existing focus observations and appropriate raw data (T1-T6) for the time of

[2] According to convention (see http://www.stsci.edu/instruments/observatory/focus/focus2.html), we express all focus changes in the science instrument-independent units of secondary mirror microns. A +1-micron change is equivalent to a physical movement of the secondary mirror by that amount away from the primary mirror. One micron of secondary mirror defocus translates to 110 microns at the focal plane. The focus zero point is defined as WFPC2 best focus as determined by point spread function (PSF)-fitting (phase retrieval) software for PSFs in the ~400-800 nanometer range.

observations to predict focus values for time rows with no observations. To do that, we must create the focus model using the existing set of observations.

We have divided all focus observations into two data sets. The first data set, mostly before SM2, includes the time period from 1995 to the first quarter of 1997 and contains 98 observations. The second data set, after SM2, includes the time period from the second quarter of 1997 to the first quarter of 1999 and contains 86 observations.

By experimenting, we have found one transfer function (embedding scheme) that applies equally well to both data sets. It uses the following terms: a constant offset; modified Julian date; temperature functions T1, T4, and T5; three-point wavelets of T1, T4, and T5; and the first derivative of T5:

$$F_E = C_0 + C_1 t + C_2 T1 + C_3 W(T1) + C_4 T4 + C_5 W(T4) + C_6 T5 + C_7 W(T5) + C_8 \frac{\Delta T_5}{\Delta t}$$

where

$$\begin{aligned} F_E &= \text{focus estimate (microns)} \\ t &= \text{modified julian date} \\ W(T) &= \text{three-point wavelet function of temperature function T} \\ \frac{\Delta T5}{\Delta t} &= \text{first derivative approximation for } T_5 \\ &= (T5(t_2) - T5(t_1))/(t_2 - t_1) \\ C_1 \cdots C_9 &= \text{polynomial coefficients} \end{aligned}$$

Calculations show that including T2, T3, and T6 in the transfer function does not improve the fit.

Note that the sensitivity to three-point wavelets means that the current value of the focus depends on the nearest time history of temperatures before the focus observation. The three-point wavelet uses six points of data; for a 5-minute time grid, this equals 25 minutes, about one-fourth of the orbital period for HST. The temperature inertia of the sensors might be responsible for this result.

We have computed two sets of transfer function coefficients (one for data set one and one for data set two), and we have used these to evaluate the fit to the observed focus measurements. Table 2 shows standard deviations from focus observations for the Four-Temperature, Attitude, Full-Temperature, and Chaos Program Models. In the first column, the notations 95_1, 97_2, etc., mean first quarter of 1995, second quarter of 1997, etc. One can see from Table 2 that the CPM fits the observed focus data points better than the existing Four-Temperature, Attitude, and Full-Temperature models.

Figure 3 gives an overview of data for the first quarter of 1997. It shows the behavior of all functions of interest with 5-minute time intervals. From this figure one can see the behavior of all three temperature functions (T1, T4 and T5), estimates of all four models, and observations. Note the drop in the CPM focus estimates at a time when there is a jump in several of the temperature functions.

Recall that only 184 observed focus measurements were available to construct the focus models. Fortunately, this is sufficient to conduct a statistical test to check inconsistencies in estimated focus amplitudes for the different models. One can plot a histogram of the number of measured foci in an interval of focus values vs. the focus value. Assuming that the focus measurement is a random process (in time) and that the typical time interval of focus variation is much smaller than the time between focus measurements, such a histogram, calculated on the set of measured points only, would be similar to the histogram calculated if exact focus values were known for all times. The model whose histogram is closest to the measured focus histogram should be the more statistically acceptable model.

Table 2. Standard Deviations (microns)

Year_ Quarter	Four-Temp. Model	Attitude Model	Full-Temp. Model	Chaos Program Model	Number of Focus Observations	Comments
95_1	1.312	1.310	1.068	0.584	6	
95_2	2.2	2.2	2.2	0.56	1	
95_3	1.444	0.850	1.196	0.974	4	
95_4	1.116	1.113	0.775	0.693	6	One observation has been removed
96_1	0.995	1.603	0.789	0.780	3	One observation has been removed
96_2	0.913	1.294	0.469	0.446	3	Two observations have been removed
96_3	1.522	2.033	1.570	0.877	8	
96_4	1.917	1.354	1.577	1.015	8	
97_1	1.222	1.134	0.848	0.726	59	
97_2	1.186	1.268	1.078	0.919	16	
97_3	1.822	1.952	1.647	1.003	10	
97_4	0.925	0.753	0.870	0.664	8	
98_1	1.074	1.312	0.832	0.664	11	
98_2	1.917	2.765	1.967	0.940	18	
98_3	1.600	1.512	1.304	0.780	10	
98_4	0.685	1.251	0.699	0.844	10	
99_1	0.519	1.173	1.266	0.472	3	
Average Standard Deviation:	1.316	1.463	1.186	0.761	184	

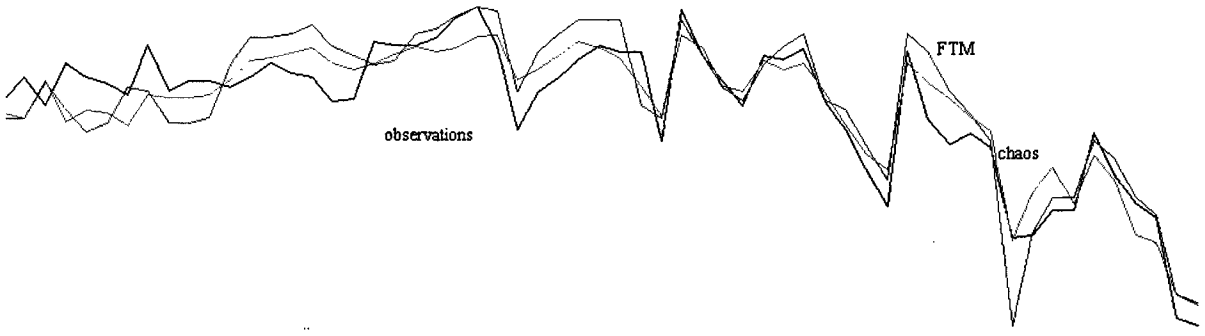


Figure 3. Overview of Focus Model Behavior for First Quarter of 1997. FTM and Chaos Program standard deviations are 0.848 and 0.720, respectively.

Figures 4 and 5 show the histograms. Figure 4 gives frequencies (absolute numbers) of amplitudes predicted by the CPM and the FTM that can be compared with the frequencies of amplitudes found for focus observations. This histogram takes into account all 184 focus measurements and the corresponding estimates from the CPM and the FTM. A comparison of the statistics shows that the CPM estimates are closer than the FTM estimates to the observed frequencies.

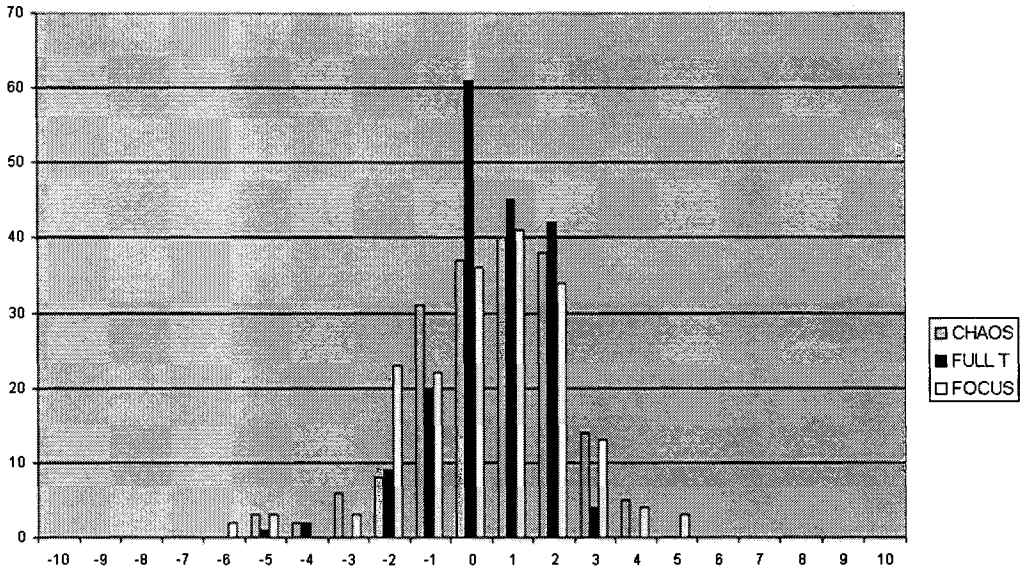


Figure 4. Frequencies of Amplitudes (in Microns) Predicted by CPM and FTM in Comparison With Frequency of Amplitudes Found for Focus Observations (Observational Points Only)

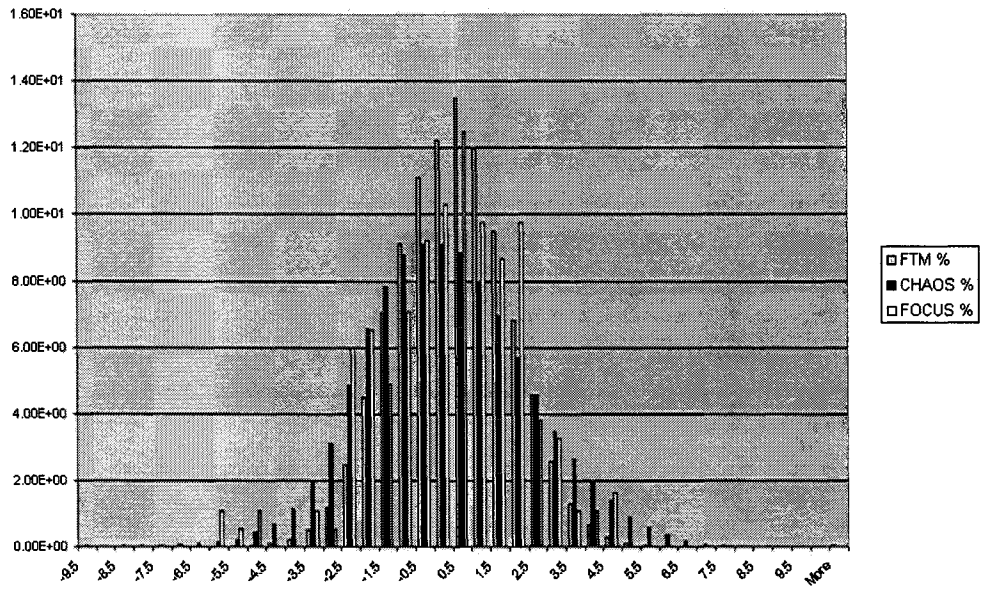


Figure 5. Frequencies of Amplitudes (in Microns) Predicted by CPM and FTM in Comparison With Frequency of Amplitudes Found for Focus Observations (All Points With 5-Minute Grid)

One can see that the FTM overestimates small amplitudes and underestimates large ones. As a result, when time profiles of focus estimates are overlapped, it appears (but only appears) as though the CPM is overestimating large amplitudes.

Figure 5 is similar to Figure 4. The only difference is that the estimates of focus amplitude from the CPM and the FTM are for all points (i.e., for all points on the 5-minute time grid, not just the 184 observation points). This figure confirms that the CPM is closer than the FTM to the observed focus distribution in the region of large amplitudes.

5. CONCLUSION

The CPM fits the focus observations somewhat better than existing physical models do in terms of both achieved LSM deviation on the measured points and the histogram statistics. We expect that the fit could be improved even more if the additional available data (i.e., the four light shield temperatures and the attitude parameters) are taken into account. Perhaps more importantly, the time to develop and verify a data-driven model can be significantly less than other approaches.

We expect also that the Chaos Program or similar software could be successfully used for the calibration of other HST/NGST science instruments and devices. Using embedding variables that capture the appropriate time scales of interest in the data before performing an LSM fit is a key step that allows the creation of accurate data-driven models.

REFERENCES

1. Hershey J., 1997. Space Telescope Science Institute (STScI) document SESD-97-01, http://www.stsci.edu/instruments/observatory/focus/focus2.html
2. Rugh W. J., 1996. Linear System Theory, Prentice Hall, NJ.
3. Gershenfeld Neil A. 1999. The Nature of Mathematical Modeling, Cambridge University Press.
4. Casdagli M. 1992. Dynamical systems approach to modeling input-output systems, Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity. In: Proc. Vol XII, Eds. M. Casdagli and S. Eubank, Addison-Wesley.
5. Stark J., 1999. Delay embeddings and forced systems, *J. Nonlinear Science*, 9, 255-332
6. Haykin S., Principe J., 1998. IEEE Signal Processing Magazine, p. 66-81.
7. Franklin G. F., 1994. Feedback control of dynamic systems, Addison-Wesley.
8. Kantz H., Schreiber T., 1997. Nonlinear time series analysis, Cambridge University Press.
9. Walker David M., Tufillaro Nicholas B. 1999. Phase-space reconstruction using input-output time series data, *Phys. Rev. E* 60 (4), 4008-4013.
10. Walker David M., Brown Reggie, Tufillaro Nicholas B. 1999, Constructing transportable behavioral models for nonlinear electronic devices, *Phys. Letts A* (255) 4-6, 236-242.
11. Usikov D., 1999. Chaos Program Users Manual, Agilent Technologies (unpublished)

IMAGE MISSION ATTITUDE SUPPORT EXPERIENCES[†]

N. Ottenstein[*], *M. Challa, A. Home*

Computer Sciences Corporation, 7700 Hubble Drive, Lanham-Seabrook, Maryland USA 20706

R. Harman, R. Burley

NASA Goddard Space Flight Center, Greenbelt, MD 20771

ABSTRACT

The spin-stabilized Imager for Magnetopause to Aurora Global Exploration (IMAGE) is the National Aeronautics and Space Administration's (NASA's) first Medium-class Explorer Mission (MIDEX). IMAGE was launched into a highly elliptical polar orbit on March 25, 2000 from Vandenberg Air Force Base, California, aboard a Boeing Delta II 7326 launch vehicle. This paper presents some of the observations of the flight dynamics analysts during the launch and in-orbit checkout period through May 18, 2000.

Three new algorithms - one algebraic and two differential correction - for computing the parameters of the coning motion of a spacecraft are described and evaluated using in-flight data from the autonomous star tracker (AST) on IMAGE. Other attitude aspects highlighted include: support for active damping consequent upon the failure of the passive nutation damper, performance evaluation of the AST, evaluation of the Sun sensor and magnetometer using AST data, and magnetometer calibration.

INTRODUCTION

This paper describes some of the major events during IMAGE in-orbit checkout (IOC) and also presents new algorithms that were developed for the use of attitude processing during this period. Attitude support was predominantly via an enhanced version of the Multimission Spin Axis Stabilized Spacecraft Attitude Determination System (MSASS), which is the controlled software used in the Flight Dynamics Facility at NASA GSFC. Note that, only highlights during IOC support are presented here; more detailed analyses and system descriptions are given in Refs. 1 and 2.

The octagon-shaped IMAGE spacecraft is 2.25 m in diameter by 1.52 m height and weighs 494 kg, including its instrument payload. There are four thin radioplasma imager (RPI) radial antennas, positioned 90 deg apart, which define the spacecraft X-Y plane while the two RPI axial antennas define the Z-axis. When fully extended, the axial antennas measure 20 m tip-to-tip, and the radial antennas 500 m.

[†] This paper was supported by the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under Contract: GS-35F-4381G, Task Order: S-43411-G. IMAGE mission support was performed under NASA GSFC Contract: GS-35F-4381G, Task Order: S-36490-G.

[*] E-mail: nottenst@cscmail.csc.com, Phone: (301) 794-2437

IMAGE was placed into an orbit that had an apogee altitude of about 45885 km (~ 7 Earth-radii), an inclination of 90.008 deg, and an eccentricity of 0.752624. It is spin-stabilized about its Z-axis, with closed-loop spin-rate control. Following separation from the launch vehicle the attitude determination and control subsystem was required to align the spin axis parallel to the nominal attitude within 1° of negative orbit normal and with a spin rate between 0.5 and 20 revolutions per min (rpm). Once attitude acquisition was accomplished and the antenna deployments were completed, the spin rate is to be maintained at 0.5 ± 0.01 rpm. During nominal mission (commencing approximately 40 days after separation) the attitude knowledge is to be within 0.1° for both spin phase angle and spin axis RA and DEC while maintaining the spin rate at 0.5 ± 0.01 rpm.

The IMAGE attitude hardware consists of: one Lockheed Martin ATC AST-201R autonomous star tracker (AST), one Adcole 44690 Sun sensor assembly (SSA), one MEDA model TAM-2A three-axis magnetometer (TAM), an Ithaco model 750UPR magnetic torque rod (MTR), and a nutation damper (ND). These are described in detail in Ref. 3.

EVENT TIME LINE

There were several major events during IOC. This time period consisted of several stages of interest.

The first stage was from launch on March 25, 2000 until April 2, 2000. This stage include the following events:

- The satellite was first checked out and it was discovered that the nutation was not damping.
- An initial procedure to decrease the spin rate of the satellite increased the nutation.
- After considerable analysis it was decided that the onboard passive nutation damper was malfunctioning, and that active nutation damping must be used to reduce the nutation.
- Several spin-downs were performed before active nutation damping could be attempted, even though these did increase the nutation.
- The first attempt to reduce the nutation on April 1, 2000 actually increased the nutation as it was done 180 deg out of phase. More details on the failure of the nutation damper and the actions to deal with it are discussed in the Ref. 4.

Figure 1 shows the Sun angle for the first full day of operation (March 26, 2000). The Sun angle is the angle between the spin axis vector (the Z-axis) and the Sun. The Sun sensor boresight is elevated 1.922 deg above the spin plane and the plots do not compensate for this bias. There is one point every spin period. The spin period at this time is about 19.4 seconds. The figure shows no noticeable damping of the nutation. It also appears that there was a slight gravity gradient effect at the first perigee and possibly a thermal effect from the eclipse, which adjusted the attitude slightly. The slope in the Sun angle for the two orbits shown in the figure is clearly different. Figure 2 shows the effects of two spin down procedures on the Sun angle. The spin period was changed from ~20.3 seconds to ~22 seconds and then to ~26.5 seconds. As seen, the nutation was slightly increased each time and the attitude was changed as is shown by the change in the mean Sun angle.

The second stage of interest was from April 3, 2000 through April 10, 2000. During this time the nutation was actively damped by use of the torquers. During the last few days of this period the spin rate was reduced to 0.75 rpm. Once the satellite was at this rate the AST was turned on and began tracking stars. Figure 3 shows the nutation in the Sun angle damped to less than 0.001 deg on April 9, 2000. It shows a clear gravity gradient effect in the Sun angle near perigee, as indicated by the change of slope in the Sun angle.

The third stage of interest was the reorientation of the attitude towards negative orbit normal. This was from April 10, 2000 through April 14, 2000. Figure 4 shows the AST quaternions during the first attitude maneuver towards negative orbit normal. The data gap towards the lower right of the figure is when the AST lost track due to Earth occultation. More discussion of the AST is presented later in this paper. The attitude circles have a radius that is of the nutation angle at the time.

The fourth stage was the spin-up in order to deploy the radial booms. This was from April 15, 2000 through April 19, 2000. The AST lost track shortly into the first spin-up at a point above 1.123 rpm as was be expected from the AST specifications

The final stage involved the RPI antenna deployments. The radial booms were deployed from April 20, 2000 through May 13, 2000. The Z-axis booms were deployed on May 14, 2000. The AST was turned on again and began tracking towards the end of the first deployment on April 20.

Figures 5 and 6 show the first 125 m of deployment of the radial booms, preceded by another segment of active nutation damping. A gravity gradient effect coming out of perigee is also seen. Figure 6 shows in more detail the effects of the deployment, which induced new frequencies in the attitude behavior.

IOC ended on May 18, 2000.

STAR TRACKER PERFORMANCE

The AST was able to calculate attitudes that could be verified with the Sun angle data.

During the spin-up after attitude acquisition the AST lost track due to rising spin rate at about GMT 415.0126. The spacecraft was in eclipse with the spin rates before and after eclipse being 1.123 and 1.434 rpm respectively.

Figure 7 compares predicted (using AST quaternions) and observed Sun angles for a full day. The known Sun angle bias is taken into account in this figure. The straight line is the predicted Sun angles calculated from the AST quaternion generated attitude. The mean difference is 0.098 deg with a standard deviation of 0.012 deg. At these Sun angles the specified error in the Sun sensor is 0.25 deg so the AST result is well within that.

As discussed previously, the AST did lose track during the first attitude maneuver, and during the early mission timeframe it lost track a few more times near perigee. The first occurrence was thought to have been caused by Earth occultation, but the subsequent dropouts did not have the Earth interfering with the AST. It is now believed that the dropouts near perigee were due to radiation, which interfered with the charge-coupled device readings. By default, the AST drops into a standby mode after three consecutive acquisition failures. The flight operations team (FOT) changed it from 3 to 256 attempts and this resulted in satisfactory "on-times" of the AST.

Figure 8 shows the AST losing track during a perigee and then regaining it later. The perigee was at Greenwich Mean Time (GMT) 0546. There was also an eclipse during GMT 0543-0551. It shows the RA versus time. Note that as the AST lost track the quaternions registered a constant default value, and resumed the oscillations (due to coning) when tracking was regained. At this time the angle between the spin axis and the Sun vector was about 144 deg, which is about only 26 deg away from the AST boresight. Later, after the IOC period the star tracker lost track at a closer Sun angle. This was believed to be due to contamination from stray sunlight. The problem disappeared when the Sun was further away from the AST.

MSASS SUN-ONLY SOLUTIONS

The differential correction (DC) algorithm (i. e., batch least-squares estimator) in MSASS solves for a constant spin-axis attitude from SSA data. These solutions used data over the course of 6 to 12 hours. However the attitude varied significantly during early mission due to nutation and smaller effects such as gravity-gradient torques. Thus the DC Sun-only solutions (when the algorithm converged) during early mission have larger errors, which are estimated to be about half the nutation angle.

The initial MSASS attitude solution was (203, 53) deg where, as in the rest of this paper, the first number is the right ascension (RA) and the second number is the declination (DEC) in the geocentric inertial frame (GCI). This was about 6.8 deg away from the expected a priori attitude of (213.7, 50.9) deg. As noted in the event timeline (and seen in Figures 2 and 5), the spin and attitude changed because of the active nutation damping. Accurate and reliable attitudes were available only after AST became operational on 4/10/00, the first AST attitude being (201.16, 61.77) deg.

The best procedure for obtaining Sun-only attitude would be by taking constant attitude data from successive apogees. Unfortunately, due to active damping occurring at perigees, there was rarely a case when the attitude was not changed between apogees. Thus data were taken only during a single orbit. A key assumption in getting Sun-only attitude solutions is that the attitude is a constant. Once the AST was turned on, it was noticed that there was a considerable drift in the attitude over an orbit. A drift of 0.2 deg in DEC over the coarse of an orbit was often seen. This changing DEC also contributed to the inability of obtaining a reliable Sun-only solution if the RA and DEC were allowed to vary from their a priori values.

MSASS PREDICTION AND ONBOARD PROCESSOR SUPPORT (POPS) UTILITY

The multi-mission nature of MSASS was helpful during the IOC when POPS (which is normally used to predict Earth cuts by horizon sensors) was used in three originally unanticipated situations. In one early situation it was used to verify antenna contact angle during a period when contact with the spacecraft was lost. Later, by using the AST parameters (instead of those of a horizon sensor), the utility was used to check for Earth-occultation of the AST during the first attitude maneuver. Finally, POPS was also useful in generating Sun angle predictions for various attitudes. This was used to validate the Sun sensor telemetry as well as to predict times when the Sun would be outside the Sun sensor field of view.

MSASS CONING ANGLE (CA) UTILITY

Theory

The CA utility computes the parameters of uniform coning by a spacecraft, and was included in MSASS as part of IMAGE attitude support (Ref. 5). Given a history of the spacecraft attitude, $\{\hat{S}(\alpha_i,\delta_i), i=1,2,...,N\}$, coning about an unknown cone axis, $\hat{S}_c(\alpha_c,\delta_c)$, this utility computes (α_c,δ_c) and the cone half-angle, θ. Here α and δ generally denote RA and DEC respectively. A sketch of the scenario is given in Figure 9.

The principal algorithm in the CA utility is the Batch-Cone algorithm, which is a DC algorithm that estimates the above cone parameters. To overcome possible convergence issues, two new "circle" algorithms - Triplet and Batch-Circle - were also implemented for computing the a priori Batch-Cone state vector. These circle algorithms are an approximation, for they compute the center and radius of a circle in RA-DEC space given the coordinates of circumference points, which are the tips of $\hat{S}$ in Figure 9. Such a circle is not an unrealistic approximation during the IMAGE nominal mission mode when the cone axis is expected to be the GCI y-axis, although it could be drastically incorrect for other mission geometries such as when coning is about the GCI z-axis. (See the CA evaluation below for a scenario where the circle approximation does not hold.) Of the two circle algorithms, Batch-Circle is a differential corrector whereas Triplet is a pure-algebraic algorithm without any convergence issues. The Triplet algorithm can be particularly useful when the measurements are available over only a small arc in RA-DEC space, for the algorithm is very general and will yield the center even if only 3 points are available. By default, CA uses all three algorithms in sequence: Triplet, whose results are input to Batch-Circle, whose results are input to Batch-Cone.

These three algorithms are described next. Note that, although we assume here that GCI is the input/output reference frame, the choice of frame is not essential to the computations - the components of the cone axis are in the same frame as the input vectors.

The Batch-Cone Algorithm

The state vector is the 3×1 vector, $\vec{X}$, defined via:

$$\vec{X} = [\alpha_c \;\; \delta_c \;\; \theta]^T, \tag{1}$$

where the superscript "T" denotes matrix transpose. As usual $\vec{X}^0$ denotes the a priori state while $\vec{X}$ denotes the current one, and S0 denotes the 3×3 diagonal matrix of user-specified weights associated with $\vec{X}^0$. Noting that

$\hat{S}_c = [\cos\delta_c \cos\alpha_c \quad \cos\delta_c \sin\alpha_c \quad \sin\delta_c]^T$, with a similar expression for the i-th attitude vector, $\hat{S}_i(\alpha_i, \delta_i)$, the angle θ_i between $\hat{S}_i$ and $\hat{S}_c$ is given by

$$\cos\theta_i = \cos\delta_i \cos\alpha_i \cos\delta_c \cos\alpha_c + \cos\delta_i \sin\alpha_i \cos\delta_c \sin\alpha_c + \sin\delta_i \sin\delta_c . \tag{2}$$

We seek an estimator which would make all the θ_i identical in the ideal noise-free scenario. Accordingly, the vector of residuals, $\vec{\rho} = [\rho_1 \ \rho_2 \ ... \ \rho_N]^T$ is constructed from:

$$\rho_i = \theta_i - \theta , \ i = 1, 2, ..., N. \tag{3}$$

The theoretical development is similar to that in Ref. 6 , except that, for simplicity, all the residuals were given the equal weight of $1/(\Delta\theta)^2$, where $\Delta\theta$ is user-specified. The loss function, J, is defined via: $J = \frac{1}{2}\frac{\rho^T\rho}{(\Delta\theta)^2} + \frac{1}{2}\left[\vec{X} - \vec{X}^0\right]^T S^0 \left[\vec{X} - \vec{X}^0\right]$. Minimizing J with respect to the components of $\vec{X}$ yields the following matrix equation:

$$G^T \vec{\rho} = S^0\left(\vec{X} - \vec{X}^0\right)(\Delta\theta)^2 , \tag{4}$$

where the elements of the $N\times3$ matrix G are given by:

$$G_{ik} = -\partial\rho_i / \partial X_k , \quad i = 1,2,...,N, \quad k = 1,2,3 , \tag{5}$$

Let $G^0 \equiv G(\vec{X}^0)$, i. e.,

$$G^0 \equiv \begin{bmatrix} -\left.\frac{\partial\theta_i}{\partial\alpha_c}\right|_{\vec{X}_0} & -\left.\frac{\partial\theta_i}{\partial\delta_c}\right|_{\vec{X}_0} & 1 \\ \cdots & \cdots & \cdots \\ -\left.\frac{\partial\theta_N}{\partial\alpha_c}\right|_{\vec{X}_0} & -\left.\frac{\partial\theta_N}{\partial\delta_c}\right|_{\vec{X}_0} & 1 \end{bmatrix}, \tag{6}$$

which can be readily computed using Eq. (2). We can compute ρ_i using ρ_i^0 and G^0 via:

$$\vec{\rho} \approx \vec{\rho}^0 - G^0\left(\vec{X} - \vec{X}^0\right). \tag{7}$$

Using Eq. (7) and approximating G by G^0 in Eq. (4) yields the leading order formula for the state innovation:

$$\vec{X} - \vec{X}^0 = \left[G^{0T} G^0 + S^0(\Delta\theta)^2\right]^{-1} G^{0T} \vec{\rho}^0 . \tag{8}$$

If $\vec{X} - \vec{X}^0$ is greater than a user-specified tolerance, the algorithm is repeated after setting $\vec{X}^0 = \vec{X}$.

The Batch-Circle Algorithm

It is assumed here that the attitude history $\{\alpha_i, \delta_i\}$ traces an arc of a circle whose center is at (α_c, δ_c) and whose radius is r. The 3×1 state vector is defined as: $\vec{X} = [\alpha_c \quad \delta_c \quad r]^T$. We now seek an estimator that would make all the radii identical in the ideal noise-free scenario. Accordingly, the residual vector is constructed from: $\rho_i = r_i - r$, $i = 1, 2, \ldots, N$, where

$$r_i = \sqrt{(\alpha_i - \alpha_c)^2 + (\delta_i - \delta_c)^2}\ . \tag{9}$$

The development is similar to that of the Batch-algorithm. As before, all the residuals have the equal weight of $1/(\Delta r)^2$, with G^0 now given by:

$$G^0 \equiv \begin{bmatrix} \dfrac{(\alpha_1 - \alpha_c^0)}{r_1^0} & \dfrac{(\delta_1 - \delta_c^0)}{r_1^0} & 1 \\ \cdots & \cdots & \cdots \\ \dfrac{(\alpha_N - \alpha_c^0)}{r_N^0} & \dfrac{(\delta_N - \delta_c^0)}{r_N^0} & 1 \end{bmatrix}, \tag{10}$$

and the state innovation given by: $\vec{X} - \vec{X}^0 = \left[G^{0T} G^0 + S^0 (\Delta r)^2\right]^{-1} G^{0T} \vec{\rho}^0$.

Denoting the final converged values of the state by the superscript "f", $\{\theta_i^f\}$ are computed using (α_c^f, δ_c^f) in Eq. (2). The cone parameters are then obtained using: $\hat{S}_c^f = [\cos\delta_c^f \cos\alpha_c^f \quad \cos\delta_c^f \sin\alpha_c^f \quad \sin\delta_c^f]^T$

and $\theta^f = \sum_{i=1}^{n} \theta_i^f / n$.

The Triplet Algorithm

This algorithm computes the center, (α_C, δ_C), and radius, r, of the circle given three points (α_1, δ_1), (α_2, δ_2), and (α_3, δ_3) on the circumference of a circle. Let r_{ic} be the distance between the i-th circumference point and the center, i. e., the i-th radius. Demanding $r_{1c} = r_{2c}$ and $r_{2c} = r_{3c}$ yields respectively:

$$2\alpha_c(\alpha_1 - \alpha_2) + 2\delta_c(\delta_1 - \delta_2) = (\alpha_1^2 - \alpha_2^2) + (\delta_1^2 - \delta_2^2), \tag{11}$$

$$2\alpha_c(\alpha_2 - \alpha_3) + 2\delta_c(\delta_2 - \delta_3) = (\alpha_2^2 - \alpha_3^2) + (\delta_2^2 - \delta_3^2). \tag{12}$$

Multiplying Eq. (11) by (δ_2 - δ_3), Eq. (12) by (δ_1 - δ_2), subtracting, and re-arranging the terms yields:

$$\alpha_c = \frac{1}{2}\left[\frac{(\alpha_1^2 + \delta_1^2)(\delta_2 - \delta_3) + (\alpha_2^2 + \delta_2^2)(\delta_3 - \delta_1) + (\alpha_3^2 + \delta_3^2)(\delta_1 - \delta_2)}{\alpha_1(\delta_2 - \delta_3) + \alpha_2(\delta_3 - \delta_1) + \alpha_3(\delta_1 - \delta_2)}\right], \tag{13}$$

$$\delta_c = -\frac{1}{2}\left[\frac{(\alpha_1^2 + \delta_1^2)(\alpha_2 - \alpha_3) + (\alpha_2^2 + \delta_2^2)(\alpha_3 - \alpha_1) + (\alpha_3^2 + \delta_3^2)(\alpha_1 - \alpha_2)}{\alpha_1(\delta_2 - \delta_3) + \alpha_2(\delta_3 - \delta_1) + \alpha_3(\delta_1 - \delta_2)}\right]. \tag{14}$$

To overcome noise, the averages (α_c^f, δ_c^f) over several triplets are computed. The cone parameters $\hat{S}_c^f$ and θ^f are then computed as in the Batch-Circle algorithm. A set of radii is computed using Eq. (9) and the average value is input to the Batch-Circle algorithm. The principal advantage of the Triplet algorithm is that it yields results even when only 3 points are available, although one must ensure that these data are well separated when processing noisy data.

Performance of the Coning Angle Utility

In practice, the utility was used as follows. An a priori attitude solution was obtained by using the triplet algorithm, which was fed to the Batch-Cone algorithm. This solution then was fed into the Batch-Cone algorithm for the final solution. Figure 10 is an output of the Coning Angle Utility using 6 min (720 points), i.e., 3 spin periods, of IMAGE in-flight data. Here circles represent the attitude history and an asterisk at the center of the circle represents the estimated cone axis. The Batch-Cone algorithm yielded (283.226, -0.24619) deg for (α_c, δ_c) and 0.196 deg for the cone angle. The Batch-Circle and Batch-Cone algorithms converged here in 2 and 1iterations respectively, and yielded very small residuals - less than 0.01 deg - thus indicating that the estimation was accurate. (Note that the circle approximation is not a drastic one here; in fact, since the "circle" is complete, a simple averaging of the (RA, DEC) data yielded reasonable values of (283.211, -0.246) deg for (α_c, δ_c).) When only 30 sec of the same data were input, the Batch-Cone algorithm yielded (283.202, -0.249) deg for (α_c, δ_c) and 0.170 deg for the cone angle, results not far from the definitive ones using 6 min of data. In general, data spanning 1/2 spin period yield results which are essentially no different from those obtained using 3 spin periods of data.

A dramatic example of the usefulness of the Batch-Cone algorithm is presented in Figure 11 which presents results using noise-free simulated data of coning close to the GCI z-axis. One cone period (240 points) of spin axis data were generated with truth model parameters of: $(\alpha_c, \delta_c) = (75, 80)$ deg, $\theta_c = 15$ deg. These are shown in Figure 11 as circles (points 1 - 144) as well as dashes (points 145 - 240). Note that the (RA, DEC) plot of the attitude history is not even a closed curve, let alone a circle. Thus intuition as well as Batch-Circle and Triplet algorithms are of no use here in determining the characteristics of the motion. The Batch-Cone algorithm was directly run, using a priori values of: $(\alpha_c^0, \delta_c^0) = (60, 64)$ deg, $\theta_c^0 = 12$ deg, *and only 60% of the attitude history* (points 1-144). The algorithm converged in 10 iterations to within 0.1 deg of all the parameters (see the legend in the figure). It should be emphasized that the Batch-Cone algorithm does not always perform as well; in fact, both batch algorithms usually do not converge if a priori errors are over 20%.

TAM AND MTR PERFORMANCE

The TAM biases were estimated by simply examining the field at apogee which is at 7 Earth radii. This yielded biases of (-12.05, +4.42, -17.25) mG in BCS. The MTR coupling coefficients were computed by Lockheed Martin Missiles and Space Division (LMMS) personnel separately via in-flight tests at apogee on 3/26/00, and were found to be changed only slightly from their pre-launch values. Numerous instances were seen, however, of differences between the TAM and MTR timetags. An example of this is shown in Figure 12, where TAM and MTR data are compared near an apogee on 3/26/00. The top plot of this figure shows that there is significant MTR current until 9318 sec. The middle plot, however, shows that the TAM measurements dropped significantly at about 9307 sec. (The bottom plot shows that the magnitudes are near zero during 9300 sec - 9307 sec after accounting for the MTR currents.)

Additional evidence of TAM timetag issues can be seen in Figure 13 where we see large differences between the predicted and measured Sun-magnetic field angle (which is attitude-independent) on 3/26/00. Note that timetag errors of a few seconds are sufficient to generate large angular errors since IMAGE was spinning at 3 rpm at that time. Reliable three-axis attitude and rates could not therefore be computed using SSA and TAM data; some interesting simulations of effects of timetag errors are given in Ref. 1.

MAGNETIC FIELD MODELS AND PREDICTIONS

A consequence of the failure of the nutation damper was that nutation damping had to be accomplished actively using the MTR to generate the appropriate magnetic control torques near the perigees in early April 2000. This required unplanned attitude support involving the generation of 3-hour spans of predicted magnetic fields (via non-MSASS utilities at NASA-GSFC called the Flight Dynamics Toolbox) at over 30 perigees. The input here consisted of an estimated spin axis attitude from the MSASS Sun-only solutions, and ephemeris data. The field predictions were generated during early April. The accuracy of the predictions could occasionally be evaluated only after the AST became operational after 4/10/00. Nevertheless, it was possible, sometimes, to evaluate the TAM after the AST was powered-on on 4/10/00. It was seen that the measured fields at apogee were within 5 mG of the predictions (which were used to estimate the extent of active damping). A particularly revealing example is provided in Figure 14, where the axis-wise predictions (using AST attitudes) are compared with measurements at an apogee of 4/26/00. The bias calibration yields good agreement even at 7 Earth radii. (Note that the plots do not include periods of MTR activity at which times large TAM residuals were again seen.)

SUMMARY

There were several lessons learned during the support.

Pre-launch simulations were useless from an attitude support point of view due to the lack of sensor data with any fidelity. Various telemetry problems were discovered early in the mission. Practically immediately it was discovered that the bits of the sun sensor telemetry were reversed from the specifications. This was quickly corrected in the software. Similarly, there were the previously mentioned timetag problems with the TAM/MTR. End-to-end tests are needed resolve data issues while the spacecraft is still on the ground.

Despite the above, many of the difficulties were quickly overcome due to constant communications between the attitude analysts and developers, the FOT, and LMMS personnel during early operations. The close contacts also facilitated quick responses to the nutation damper anomaly by, for example: (1) creating several different versions of the MSASS telemetry processor function to process different data formats, (2) predicting antenna contacts, (3) predicting perigee magnetic fields, and (4) keeping personnel schedules flexible.

The new MSASS Coning Angle utility worked as intended and was useful during and after the IOC period. The new algorithms introduced for estimating the cone parameters are very general and should be useful in future missions.

REFERENCES

1. M. Challa, N. Ottenstein, A. Home, and M. Marchowsky, *Imager for Magnetopause to Aurora Global Exploration (IMAGE) Attitude Support Post-Launch Report*, prepared by Computer Sciences Corporation (CSC) for NASA GSFC, Greenbelt, MD, June 30, 2000

2. A. Home, M. Marchowsky, N. Ottenstein, and M. Challa, *Imager for Magnetopause to Aurora Global Exploration (IMAGE) Multimission Spin Axis Stabilized Spacecraft (MSASS) Attitude Determination System - User's Guide Version 2.5*, prepared by CSC for NASA GSFC, Greenbelt, MD, June 30, 2000

3. LMMS/P458668A, *IMAGE Spacecraft/Observatory - Spacecraft Description and Operator's Manual CDRL A028 (FINAL)*, LMMS, Sunnyvale, CA 94088-3504, January 5, 1999

4. C. Hubert and D. Swanson , "Surface Tension Lockup in the IMAGE Nutation Damper - Anomaly and Recovery." *2001 Flight Mechanics Symposium*, NASA/ CP—2001-209986, 2001. (Paper 39 of this compilation.)

5. M. Challa and A. Home, *Coning Angle Utility - Functional Specifications*, prepared by CSC for NASA GSFC, Greenbelt, MD, January 14, 2000

6. J. S. Wertz (Editor), *Spacecraft Attitude Determination and Control*, D. Reidel Publishing Co., Dordrecht, Holland, 1978, pp. 449–451

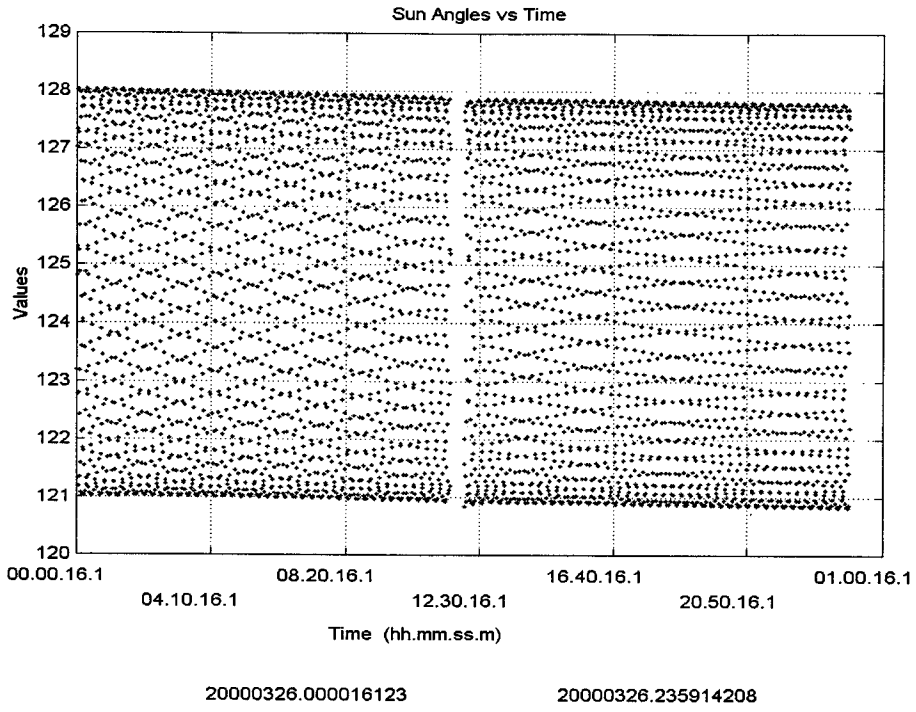


Figure 1 **Sun Angles (Deg) During the First Full Day of Operations**

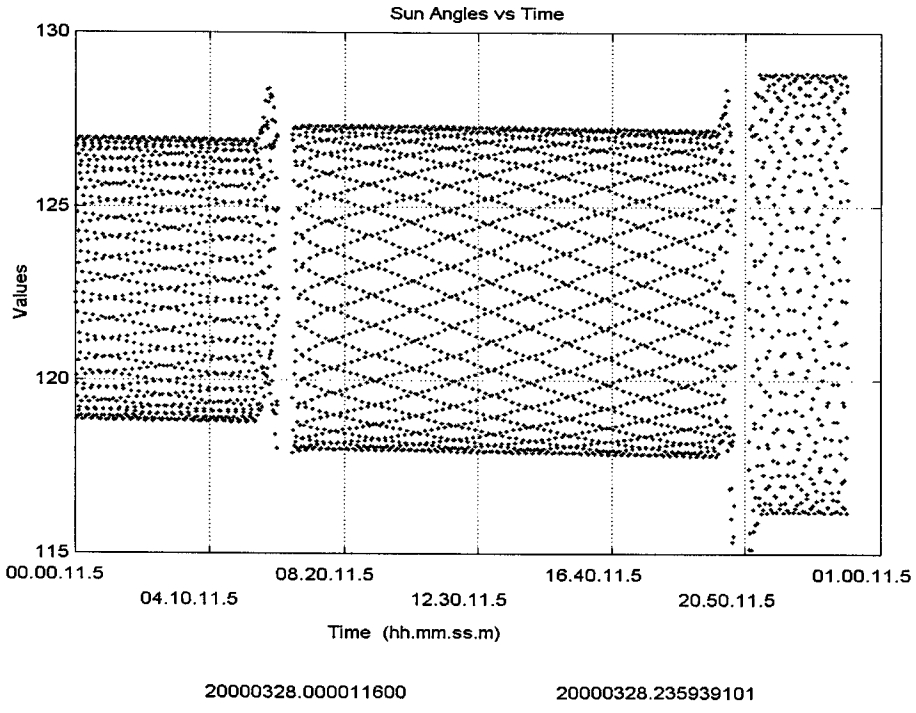


Figure 2 **Sun Angles (Deg) During the 2nd and 3rd Spin-Down Procedures**

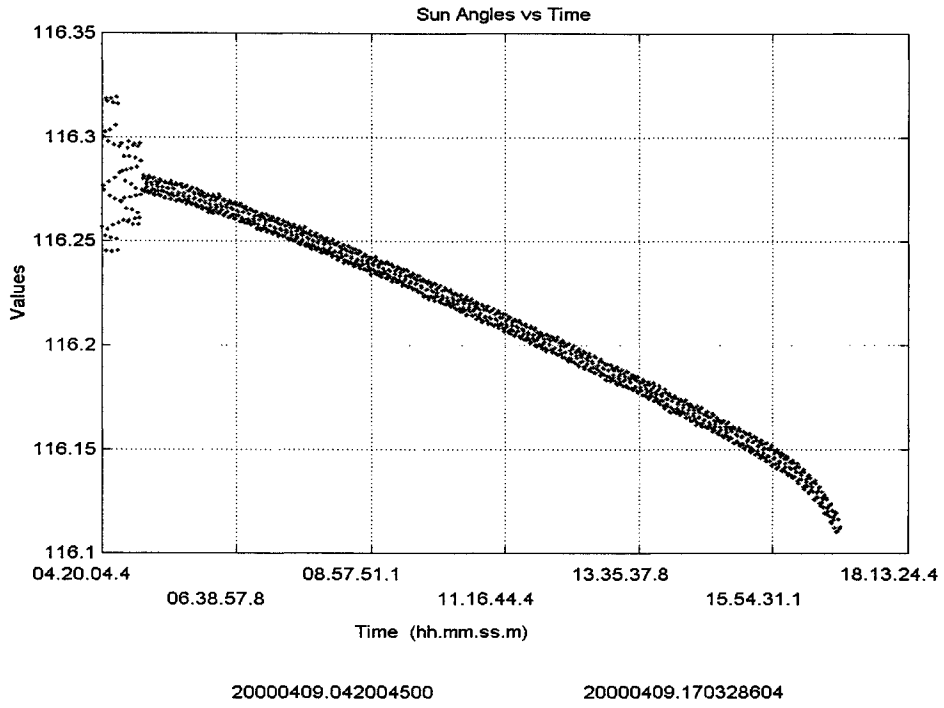


Figure 3 Sun Angle Data (Deg) Showing Gravity-Gradient Effects

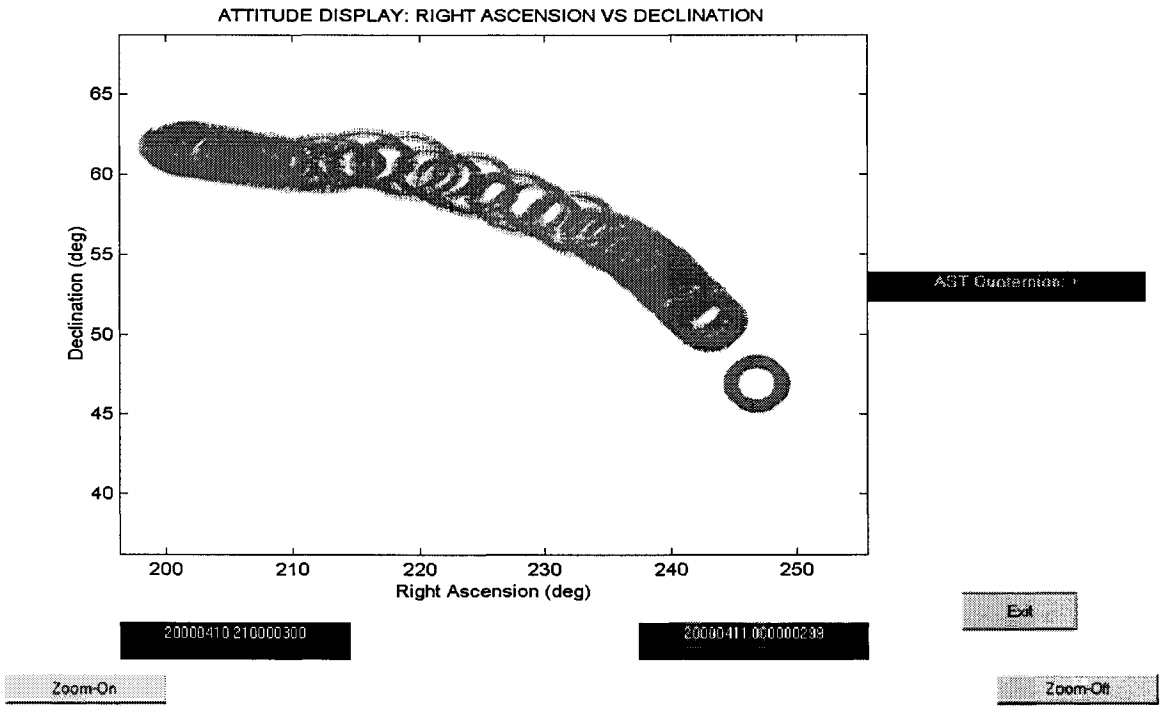


Figure 4 AST Attitude Data During the First Reorientation

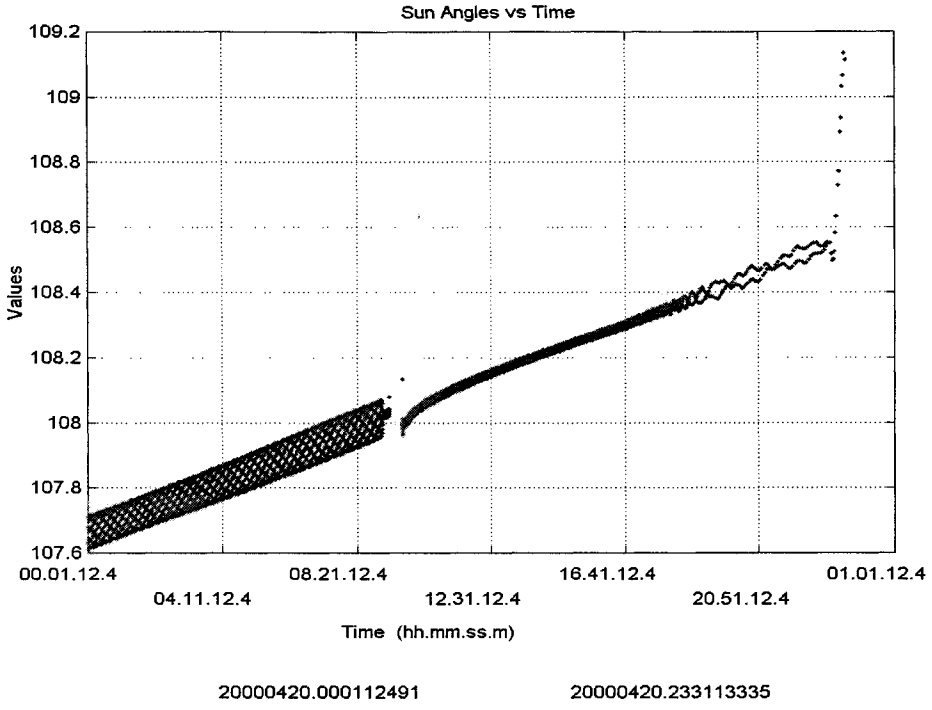


Figure 5 **Sun Angle Data (Deg) During Radial Boom Deployment to 125 m**

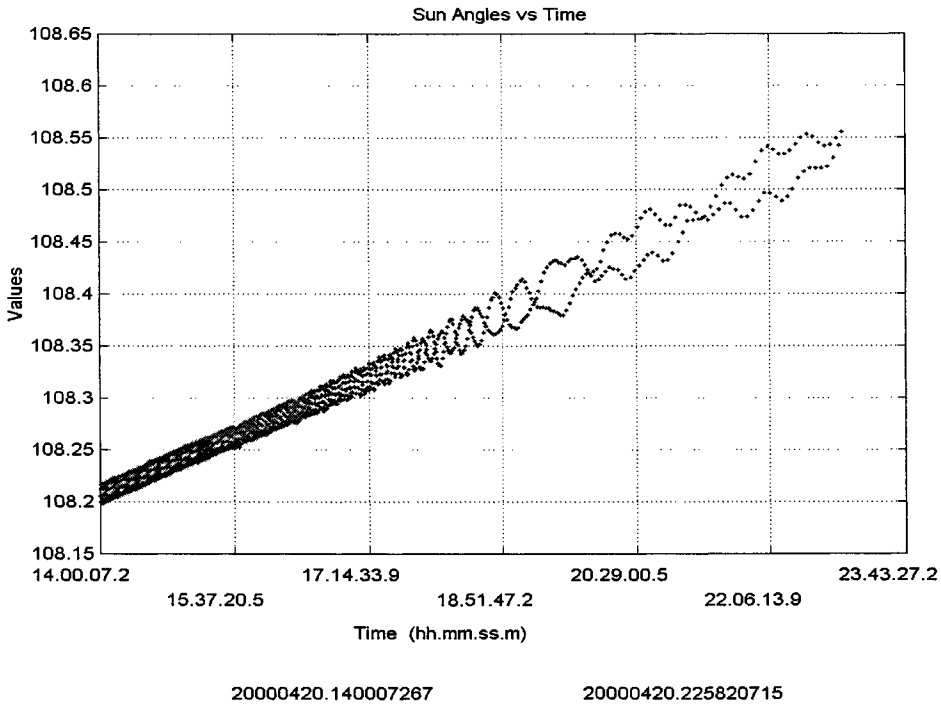


Figure 6 **A Close-Up View of Figure 5 - Sun Angles (Deg)**

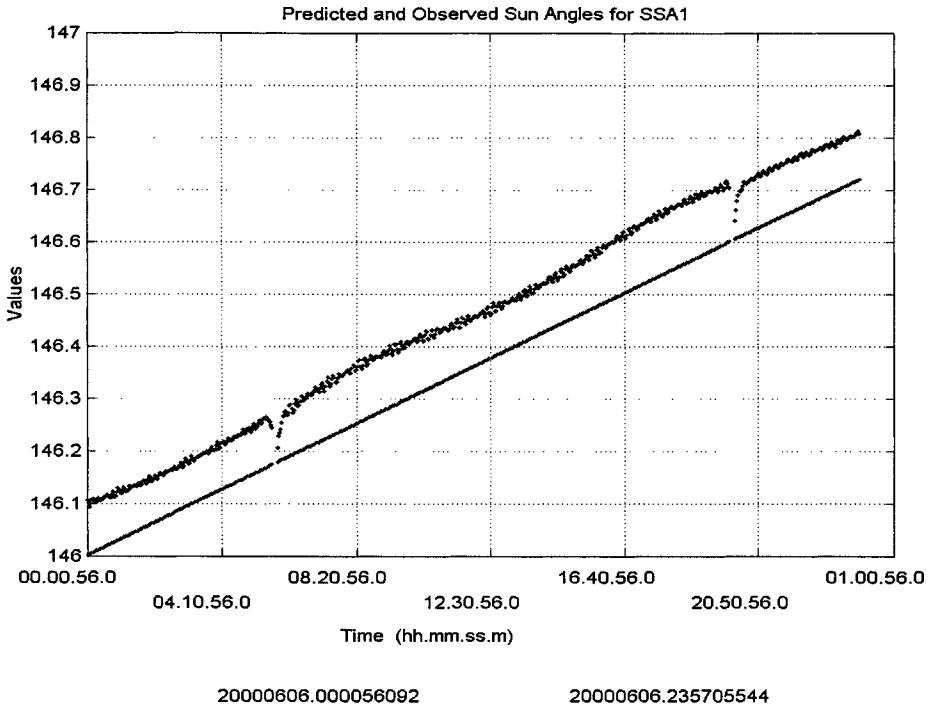


Figure 7 Comparison of AST-Predicted and Observed Sun Angles (Deg)

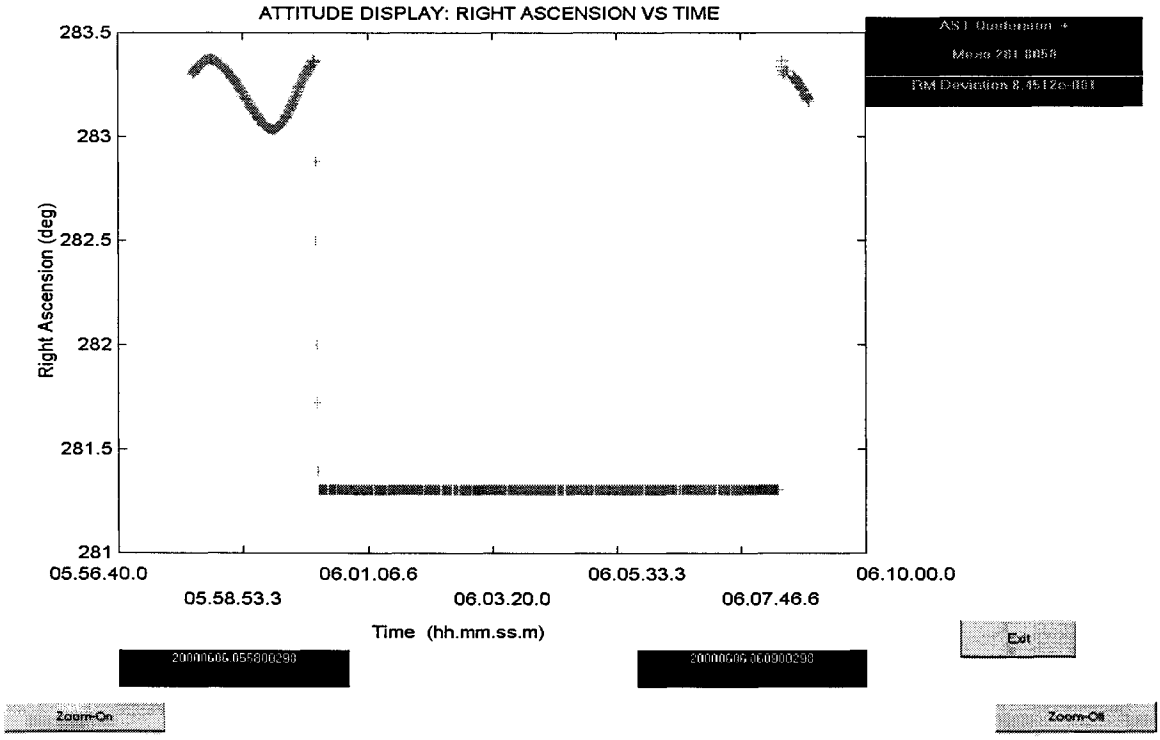


Figure 8 AST RA History Showing AST Dropout

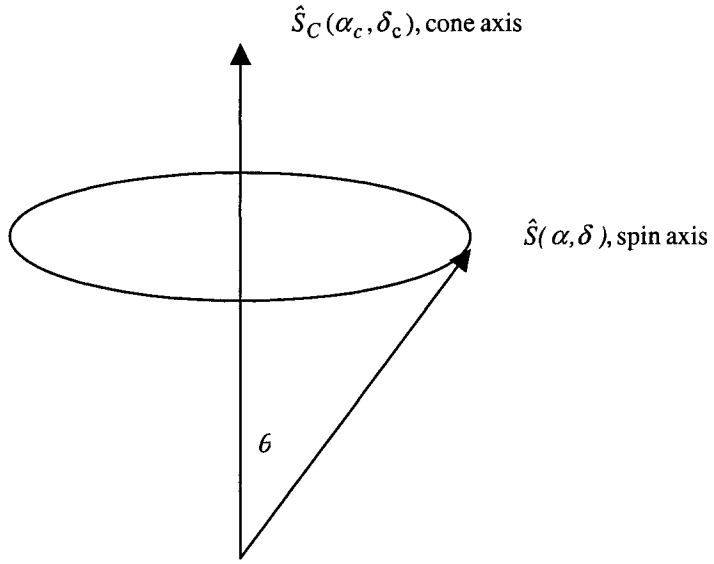


Figure 9 **Schematic of Spacecraft Coning**

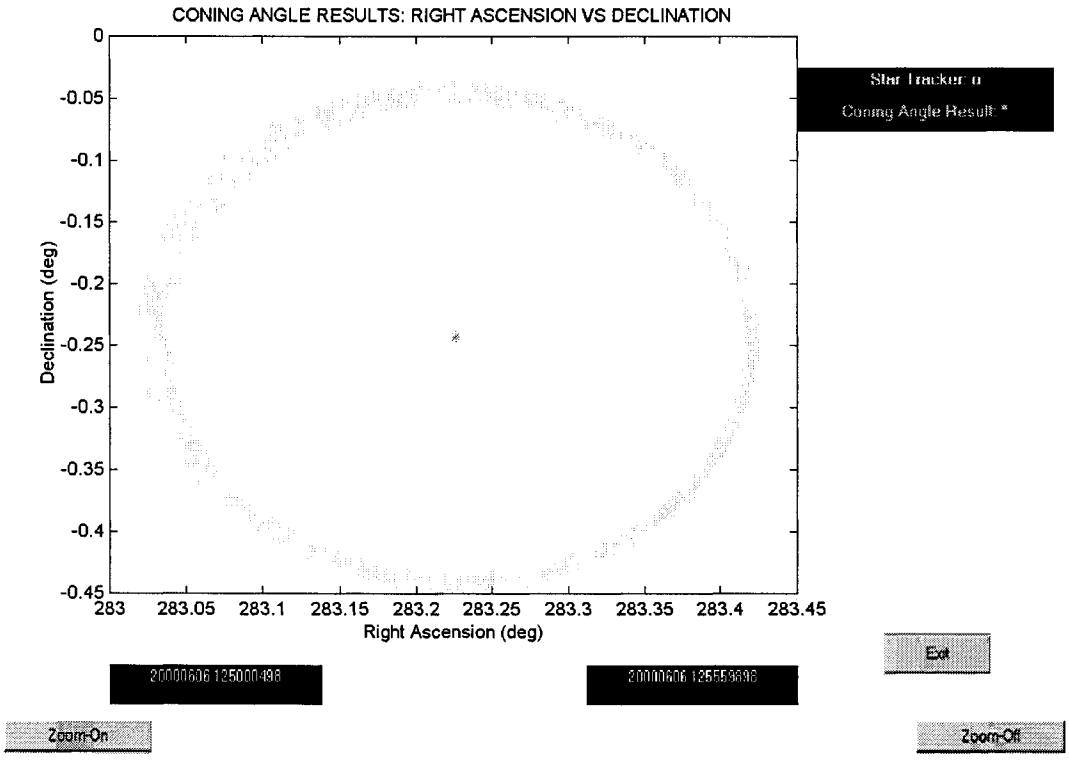


Figure 10 **DEC vs. RA plot of 3 Spin Periods of IMAGE Data with Coning Angle Solution (Center)**

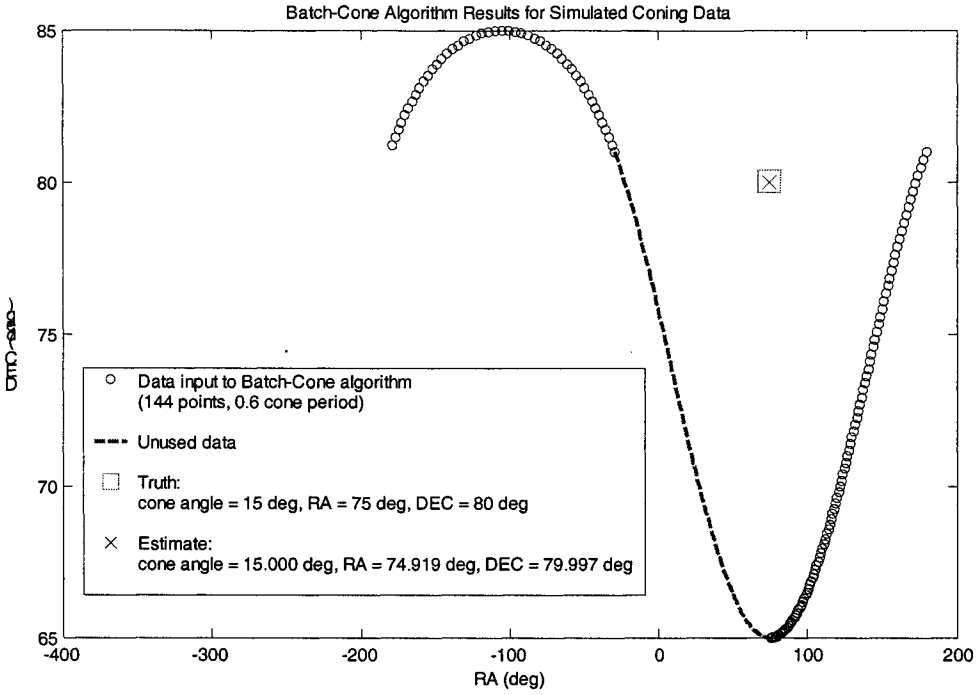


Figure 11 **Batch-Cone Algorithm Results for Simulated (non-IMAGE) Coning Data**

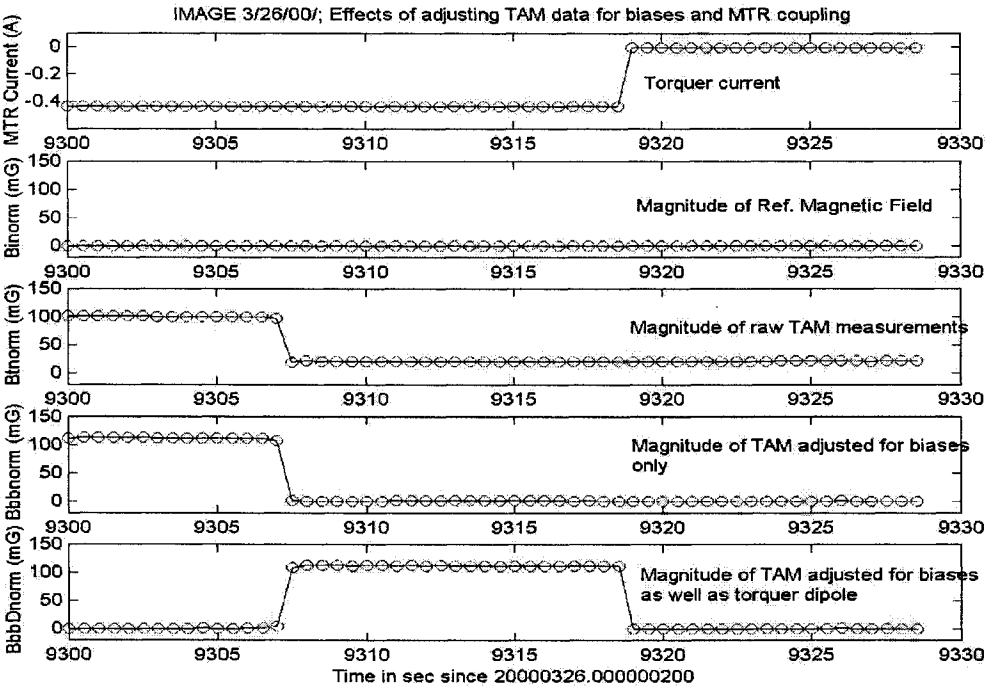


Figure 12 **Discrepancies between TAM and MTR Data at Apogee of 3/26/00**

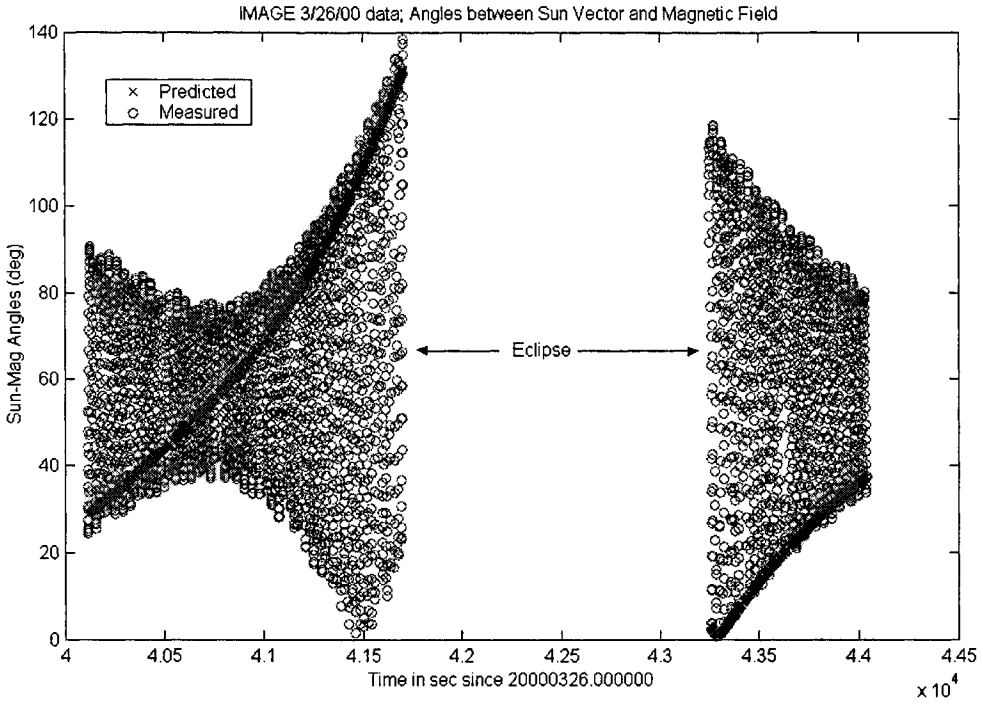


Figure 13 Sun-Mag Angles in BCS and GCI frames on 3/26/00

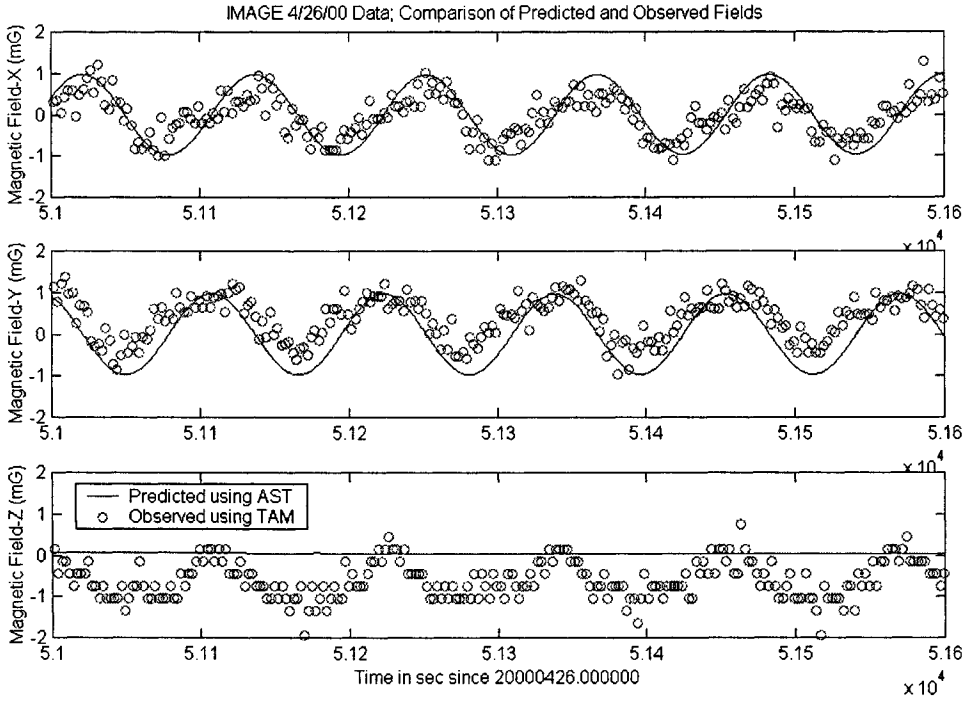


Figure 14 Comparison of Predicted and Observed Fields at Apogee on 4/26/00

ON-ORBIT POINTING PERFORMANCE FOR THE NEW MILLENNIUM EARTH OBSERVING-1 SPACECRAFT

Kathie Blackman
the Hammers Company

Seth Shulman
Computer Sciences Corporation

Dave Speer
Litton Advanced Systems

Teresa Hunt, Paul Sanneman
Swales Aerospace

ABSTRACT

The NASA New Millennium Program (NMP) Earth Observing-1 (EO-1) spacecraft was launched in November 2000 on its primary mission to validate advanced remote sensing instruments. This paper provides a summary of the early orbit checkout of the spacecraft, paying special attention to the initial acquisition sequence, science pointing performance, Kalman filter attitude determination, and system calibration. EO-1 pointing performance is meeting the 2-sigma pointing requirement as well as the 3-sigma goal for science imaging. The operational issues and estimated accuracy are presented for the Lockheed Martin AST-201 Star Tracker, the Litton G&C Space Inertial Reference Unit and the Loral Space Systems GPS Tensor receiver. This paper will also provide an overall summary of all attitude control modes, including the raster scan slew required for instrument lunar calibration.

INTRODUCTION

The goal of NASA's NMP is to enable 21st-century missions through identification, development, and flight validation of key breakthrough technologies so that future spacecraft can take advantage of them without assuming the risks inherent in their first use. The NMP technology development and validation process will also provide a significant return of valuable science data, so that immediate benefits of NMP flights are realized along with the steady stream of new technologies for future science missions. NASA's Jet Propulsion Laboratory is managing the NMP, and Goddard Space Flight Center is running the Earth Orbiter series of NMP missions. EO-1 is the first in this series of smaller, faster, cheaper Earth observing spacecraft[1, 2]. The Swales Aerospace EO-1 spacecraft was launched from Vandenberg Air Force Base on November 21, 2000 into a circular, sun-synchronous polar orbit at an altitude of 705 kilometers. EO-1 was co-manifested on the Boeing Delta II (7920-10C) launch vehicle with the SAC-C satellite developed by Argentina. The EO-1 orbital inclination (98.2 degrees) and descending nodal crossing time (10:01 am) puts it in "formation flight" with Landsat-7 and EOS AM-1. With all three satellites following the same ground track, EO-1 will be flying "behind" Landsat-7 but "ahead of" EOS-AM (Terra).

The primary EO-1 science instrument is the Advanced Land Imager (ALI), which incorporates wide field-of-view push-broom optics to achieve high-resolution and a 1°x3° field-of-view (FOV), with panchromatic detectors covering the visible/near infrared (VNIR) and short wave infrared (SWIR) spectral bands. The spatial resolution for the panchromatic detectors is 10 meters. The ALI has an open/close aperture cover and an internal stimulation lamp, but also utilizes new calibration methods. Using the sun and the moon, the ALI will demonstrate techniques for 5% absolute and 2% pixel-to-pixel relative radiometric calibration. The spacecraft Attitude Control System (ACS) performs slew-and-hold maneuvers to point the body-fixed instrument for sun calibrations on an approximately weekly basis, and performs a complex series of slew-and-scan maneuvers for lunar calibrations every month.

The pointing budget for EO-1 allows a total ground targeting pointing error of 132 asec in roll, 174 asec in pitch and 122 asec in yaw. The budget includes an ACS allocation of 54 asec in roll and yaw and 108 asec in pitch for attitude determination, and 30 asec in each axis for attitude control. The requirement is to meet these values as a 2σ variance and the goal for imaging is to meet the same values as a 3σ variance. These allocations were established in September 1998 following the addition of the Hyperion instrument. The mass properties growth associated with the

Hyperion instrument addition was a challenge to retaining attitude control performance. The launched spacecraft mass was 571 kg and the diagonal elements of the inertia tensor were estimated as [443 179 429] kg-m^2 with the solar array deployed at the 0 degree position. The EO-1 spacecraft is pictured in Figure 1.

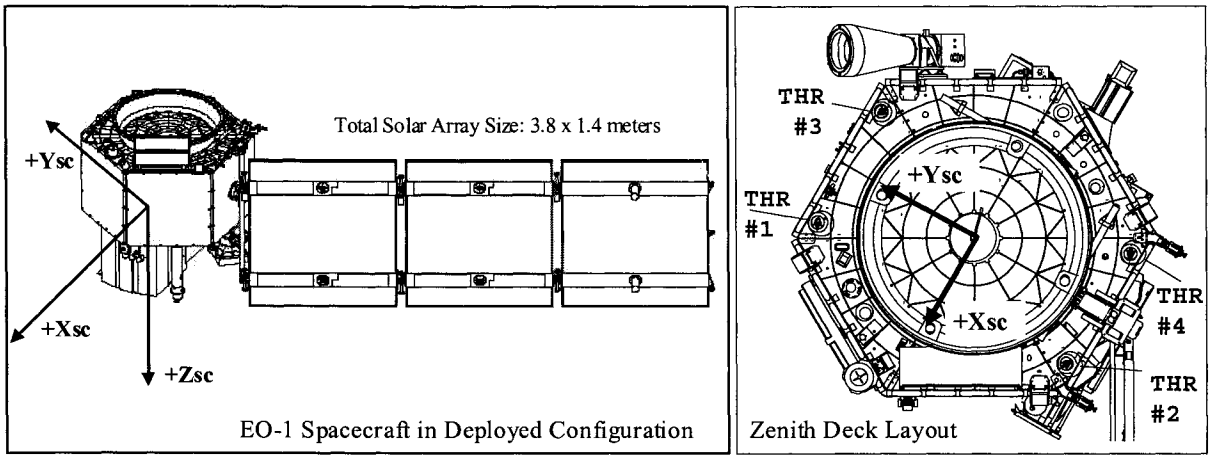


Figure 1 EO-1 Spacecraft

ACS DESIGN SUMMARY

The architecture and major components of the ACS are presented in Figure 2. All of the attitude control functions are performed within the Attitude Control and Data System (ACDS). The circuit cards of interest within the ACDS are along the center of the Figure. The primary attitude control software resides in the Mongoose 5 main spacecraft computer. The Attitude Control Electronics (ACE) provides electrical interfaces to most ACS components and hosts the Safe Hold Mode controller. Table 1 includes a description of each of the ACS components.

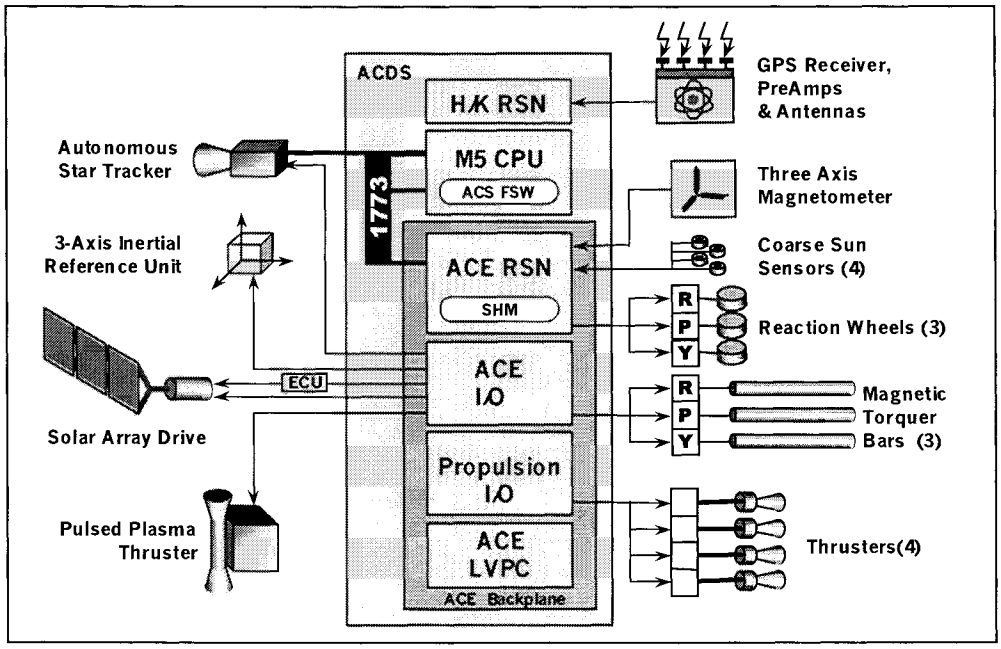


Figure 2 EO-1 ACS Components

Table 1 ACS Component Capabilities

ACS Component	Vendor & Model	Description
Reaction Wheels Assembly	Ithaco Space Systems Type A w/MDE	±4 N-m-s momentum at 5100 RPM, ±0.025Nm max torque; three wheels utilized with on-axis orthogonal mounting
Magnetic Torquer Bars	Ithaco Space Systems TR60CFR	±60 Am linear dipole moment with on-axis orthogonal mounting and linear coil drive capability
Three Axis Magnetometer	SAIC/Ideas and Nanotesla NT600s	±100.0 μTesla range on each of three axes; resolution of 0.05 μTesla
Inertial Reference Unit	Litton Guidance & Control SS-SIRU (with 3 HRG)	±10 degree/sec max rate; bias stability ≤0.015°/hr over 8 hours; ARW ≤0.001°/hr $^{1/2}$ (3σ); resolution of 0.05 arcsec
Autonomous Star Tracker	Lockheed Martin / ATC AST-201 with sun shade	8 x 8 deg FOV, 1 Hz ECI attitude quaternion output, with 5, 5, 25 arcsec, 1-sigma accuracy; up to 50 stars in solution
GPS Receiver	Space Systems Loral GPS Tensor with 4 antennas	1 Hz time, position to ±150 meters, 1-sigma and velocity to ±0.55 meter/second, 1-sigma
Coarse Sun Sensors	Adcole Corporation Model 29450	Peak output current of 650 micro-amps; four eyes provide full 4π steradian coverage

Attitude Control Electronics

The basic purpose of the ACE is to provide the low-level electrical interfaces and functions required for the ACS software to be "connected" to the physical hardware of the attitude control sensors and actuators. The ACE functions are done by a set of four circuit boards. A low-voltage power converter (ACE LVPC) board draws +28V power from the Power System and converts it to +5V logic power and ±15V analog power for all four ACE boards. The ACE LVPC also has ON-OFF switchable +28V outputs or "services" that are used to power many of the ACS sensors and actuators. The ACE RSN board has the rad-hard microprocessor, the 1773 fiber-optic data link, and direct electrical interfaces to the coarse sun sensors (CSS), three-axis magnetometer (TAM), and reaction wheel (RW) motor drive electronics (MDE) units. A general purpose input/output (ACE I/O) board interfaces with the inertial reference unit (IRU), magnetic torquer bars (MTB), autonomous star tracker (AST), solar array drive (SAD), and pulsed plasma thruster (PPT) experiment. A propulsion system input/output (Prop I/O) board handles the controls and telemetry to and from the thrusters, latch valve, and other components of the hydrazine propulsion system. A local power and signal backplane connects all of the ACE boards to each other. Figure 3 is a block diagram showing the four ACE boards and the backplane, and the signal interfaces to the various ACS sensors and actuators.

The Remote Services Node (RSN) board is the heart of the ACE. Its UT69R000 microprocessor executes the ACE software, which distributes ACS actuator commands and collects telemetry data from the ACS sensors and actuators. ACE software tasks are scheduled at either 1 Hz or 25 Hz, and the ACE software contains the Safe Hold algorithm. The ACE RSN has a digital-to-analog converter (DAC) that converts 12-bit digital data received from ACS into analog command voltages for the three RW torques and the three MTB dipole moments. The RSN also has a 12-bit analog-to-digital converter (ADC). In conjunction with a network of analog multiplexers, this ADC can digitize a wide variety of sensor, control, and telemetry signals. . The ACE RSN also has a Special Command Reset interface with the S-band transponder that allows a ground command to reset the ACE processor.

The ACE I/O board provides interfaces to the IRU for bi-directional serial data communications, digital status inputs and analog telemetry inputs. For the digital health, status, and inertial angle data, the IRU sends out 64-byte packets over a 422 kbaud RS-422 serial data link. The ACE I/O card does serial-to-parallel conversion of the 422 kbaud stream, and has a RAM buffer to store the packets. The ACE I/O board also interfaces to the SAD for controlling and measuring the angular rotation of the solar array (SA). The fastest SA motor stepping rate is 25 Hz (0.19°/sec) and the normal daylight rate of 8 Hz can be adjusted by ±1%. SA rotation can be positive/forward for "orbit day" or

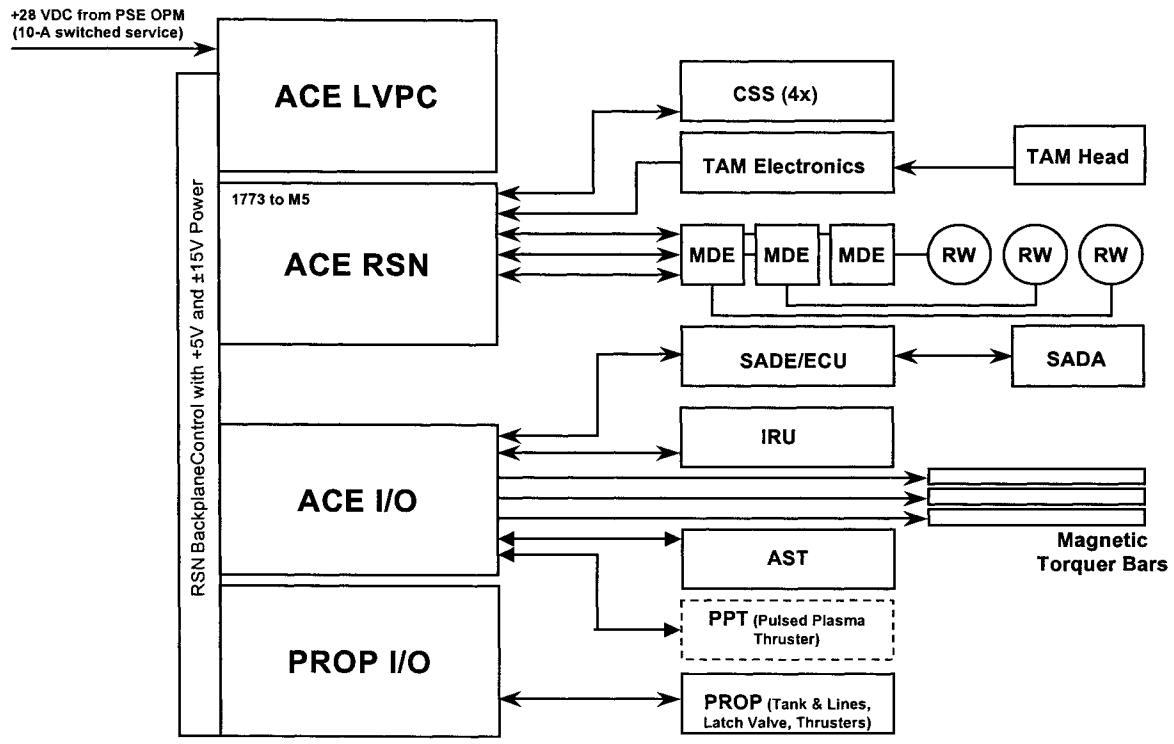


Figure 3 ACE Block Diagram

negative/reverse for rewinding during "eclipse". The solar array drive has two potentiometers that provide a coarse analog voltage proportional to the absolute angular position of the solar array (SA). The max range is –177° to +177°, and the normal operating range is about –120° to +120°. For each MTB, the ACE I/O board has a second-order low pass filter with a cut-off frequency at about 1 Hz, and a unity-gain power amplifier capable of driving the 20 ohm, 2.5 henry MTB load. A separate DC-to-DC converter on the ACE I/O board provides dedicated ±15V supplies for the MTB power amplifiers, and these ±15V supply voltages are digitized by the RSN. For the AST, the ACE I/O board provides a 1 Hz synchronization signal, a reset signal and a base-plate temperature measurement. For the PPT, the ACE I/O board has digital command lines for controlling charge time on the main capacitor and then firing one of the two spark plugs. The ACE I/O board also has analog telemetry input interfaces for PPT capacitor voltage and spark plug voltages, capacitor and transformer temperatures, and two "fuel gauge" potentiometers.

The Propulsion I/O board provides for control functions and telemetry measurements for the hydrazine Reaction Control System (RCS). Control functions include +28V pulses for opening or closing the latch valve, and +28V pulses for turning each of four thrusters on and off. The thruster firing circuits are designed with multiple logic and power stages that are electrically connected for ANDed operation. First, there is a separate Enable/Disable bit for each of the four thrusters. After the thrusters are enabled, they are activated with a two-step ARM and FIRE scheme. The two steps allow verification of the arm command before executing it to actually fire the thrusters. Due to the ACE RSN software 25 Hz cycles, the width of thruster firing pulses is always an integer multiple of 40 milliseconds. The Propulsion I/O board has redundant +28V driver circuits (primary and secondary) for opening the latch valve, and redundant circuits for closing the latch valve. Latch valve status (open or closed) is measured via an analog telemetry voltage, and the analog tank pressure signal has a resolution of about 3447 Pa (0.5 PSIA) per ADC count. Additional propulsion system telemetry includes two propellant tank temperatures, two propellant line temperatures, four thruster valve temperatures, and a Propulsion I/O board temperature.

ACS FSW & Control Modes

The EO-1 ACS flight software architecture is derived from the GSFC Tropical Rainfall Measurement Mission (TRMM). The EO-1 ACS software includes attitude determination and closed-loop control modes for magnetic de-spin following separation from the Delta II launch vehicle, initial stabilization and sun acquisition, nadir pointed science data collection and downlink, thruster maneuvers for delta-V, and solar/lunar slew/scan maneuvers for instrument calibrations.

The ACS flight software mode transitions are illustrated in Figure 4. Following separation from the Delta II launch vehicle, the ACS nulls the tip-off rates via a B-dot magnetic control law and stabilizes the spacecraft. During initial Sun Acquisition, the spacecraft maintains an inertially fixed, solar pointing attitude with the instruments facing away from the sun. During normal operations in Mission Idle, the body fixed science instruments point toward the earth as the spacecraft maintains a fixed attitude with respect to the orbit frame. Solar calibration requires a slew maneuver to point the instruments toward the sun, followed by an Inertial Hold. A series of transitions between slew maneuvers and holds is used to perform the lunar calibration raster scan that sweeps the moon across each of the instrument detectors. A transition to Delta-V is preceded by a slew maneuver to orient the spacecraft for the thruster burn. Transition back to Earth pointing from any attitude is achieved using an Earth Acquisition slew. The ACS Fault Detection and Correction (FDC) system includes actions that can force transitions to either ACS Sun Acquisition or ACE Safe Hold.

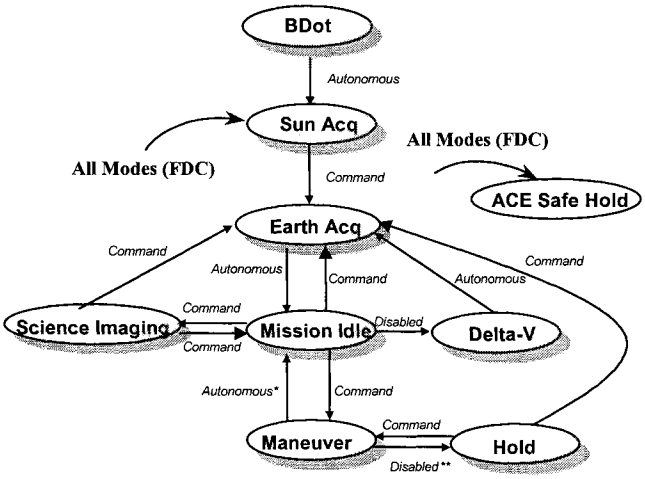


Figure 4 ACS Control Mode Diagram

Kalman Filter

EO-1 uses a six state Kalman filter that estimates the spacecraft inertial to body attitude quaternion and IRU drift bias. It uses AST quaternion measurements at 1 Hz intervals, and propagates the attitude using IRU body rates between valid AST measurement updates. The Kalman filter is used to compute spacecraft attitude in all modes other than B-dot and Sun Acquisition, which do not require attitude knowledge.

Safe Hold Mode

The independent Safe Hold algorithm software resides in the ACE RSN processor. If a fault condition is detected, this fully autonomous control capability will drive the solar array to the 0 degree reference position (orbit noon

position) and put the EO-1 spacecraft in a thermal and power safe attitude that is inertially fixed relative to the sun. Coarse sun sensors on the solar array and spacecraft main body will provide spacecraft attitude with respect to the sun with the IRU providing rate feedback. The control about the sunline will only be rate damped during orbit day, but all axes will be inertially fixed during eclipse. Continuous magnetic unloading of reaction wheel momentum will be performed using a cross-product control law. The thruster commands will be zeroed at Safe Hold entry and the propulsion interface circuit power will be removed to ensure against a failed open thruster solenoid.

LAUNCH & EARLY ORBIT PERFORMANCE

Launch Summary, Tip-off Rates and Despin

The EO-1 spacecraft separated form the launch vehicle with tip off rates of [3.61 0.42 –0.19] deg/sec. The rates decreased to [1.11 1.72 1.0] deg/sec following deployment of the solar array. The ACS B-Dot mode controller was able to decrease the spacecraft system momentum to below 4 Nms in all three axes in 78 minutes at which point the ACS transitioned to Sun Acquire mode. Sun Acquire mode was initiated at 00-326-20:42 just prior to the orbital eclipse period. The ACS controller performed angular rate damping and reaction wheel momentum dumping during the eclipse period. Upon entering orbit day, a pitch flip maneuver was performed to get the sun off of the backside of the solar array. Sun Acquisition was completed at 22:00, although Reaction Wheel momentum dumping did not complete until 23:00. The spacecraft body rates for this initial acquisition sequence are plotted in Figure 5.

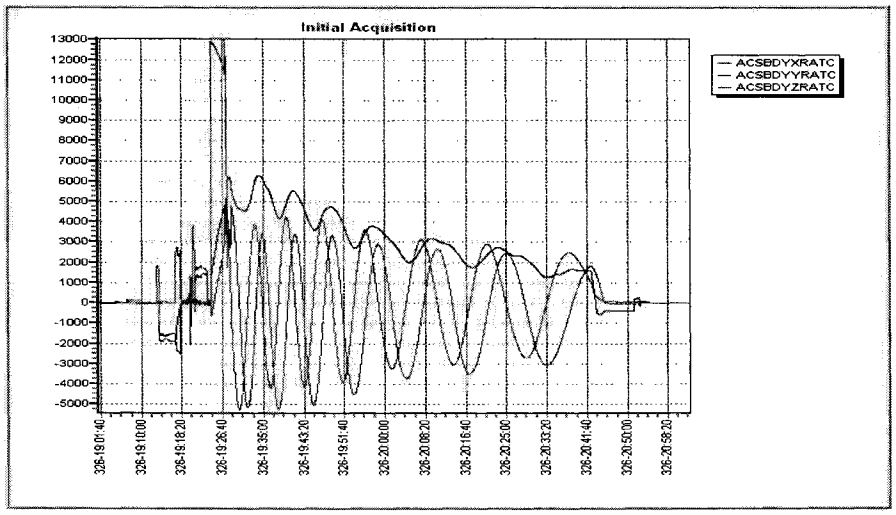


Figure 5 EO-1 Initial Acquisition Body Rates (Arcsecond/second)

Safe Hold Mode Test

A test of the ACE independent Safe Hold Mode was planned as part of Mission Day 2 activities. The power load shed commands executed and powered off the non-essential services as expected. The ACE performed an autonomous transition to the redundant side of the SAD/ECU upon entering Safe Hold mode, and the solar array rotation from +10 deg to 0 deg index was demonstrated successfully. The ACE Safe Hold mode attitude controller exhibited excellent performance for sun pointing the solar array. The solar array sun pointing specification was 25 degrees, 3σ and the Safe Hold mode sun pointing performance, shown in Figure 6, was 3 to 5 deg, 3σ including CSS albedo errors.

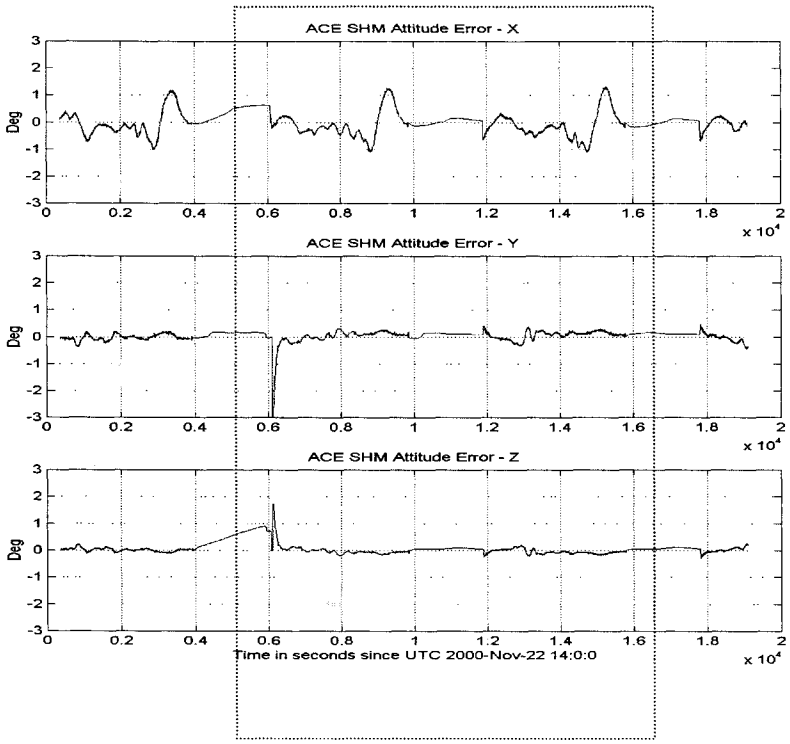


Figure 6 Safehold Mode Control Attitude Errors

Earth Pointing, Mission Idle

ACS pointing performance has been well within mission requirements as shown in Table 2. The EO-1 mission has been scheduling between 5 and 10 imaging and calibration maneuvers per day. This is significantly higher than the two to four images per day originally planned. The ACS has been capable of performing all of these images with the pointing performance listed in Table 2.

Table 2 Science Pointing Accuracy Assessment

	Requirement	On-Orbit Performance
Attitude Control	[30, 30, 30] asec, 3σ	X & Z consistently under 30 asec Y varies between 0 to 50 asec depending on settling time after image prep slew maneuver
Attitude Knowledge	[54, 108, 54] asec 3σ	All axes consistently under 36 asec, 3σ during normal nadir-pointed operations
Navigation Accuracy	[130 m Cross-Track, 100 m Along Track] 3σ	Cross-track 45m, Along track 55m, Radial 30m, 3 σ
Jitter/Rate Stability		Better than 0.5 asec/sec, 3σ during imaging

Figure 7 demonstrates that the typical attitude control accuracy achieved during an image observation meets the 30 asec, 3σ goal. The top position and rate error plots are over a 6 hour period, and the bottom plots zoom in on a 30 second image duration.

Figure 8 shows the position errors during a pitch reaction wheel zero crossing, which is the worst case of the three wheels. The reaction wheel zero crossings cause a peak transient of about 80 asec for the roll wheel, 120 asec for the pitch wheel and 50 asec for the yaw wheel, each with about a 2 minute width. The reactions wheels are momentum biased in order to avoid zero wheel crossings during an image.

Figure 9 shows the 300 asec peak transient in the pitch axis due to opening the ALI cover prior to an image. The opening of the cover is timed such that this transient is zeroed within 2 minutes prior to the start of an image. Figure 10 and Figure 11 show that the thermal snap transients as a result of night-to-day and day-to-night transitions have a 40 to 80 asec peak error per axis with a 2 minute width.

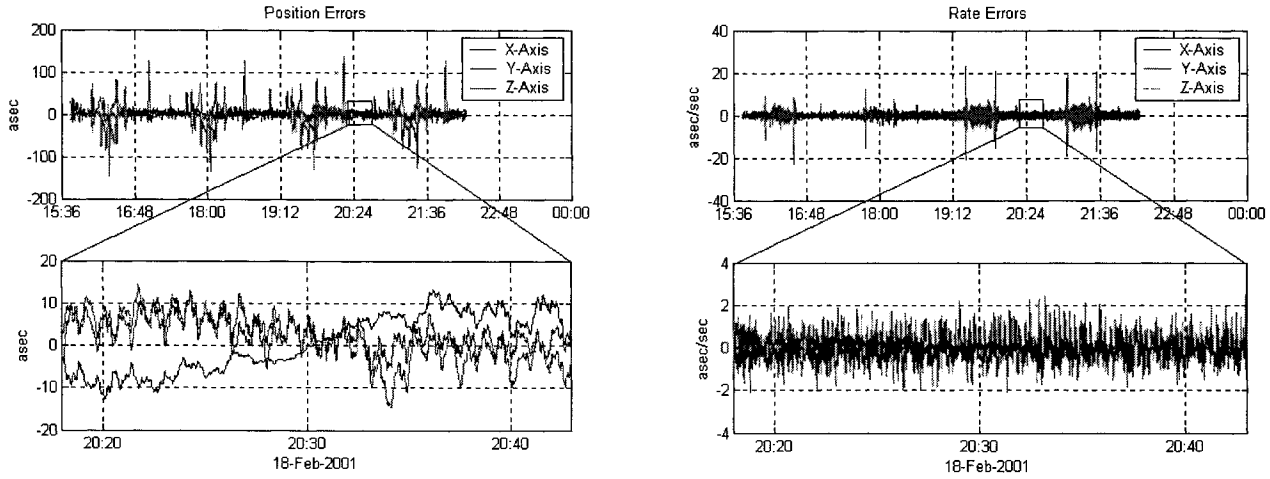


Figure 7 Position and Rate Control Errors During Image

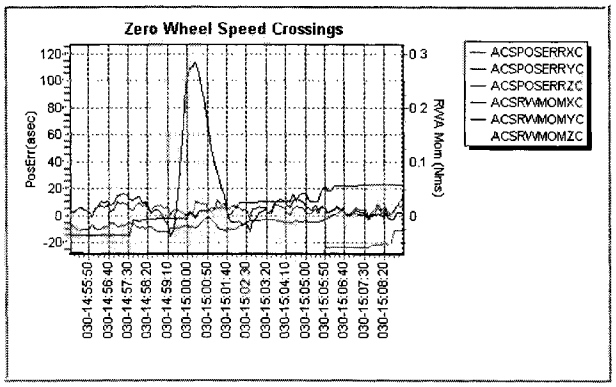


Figure 8 Position Errors During Reaction Wheel Y Zero Speed Crossing

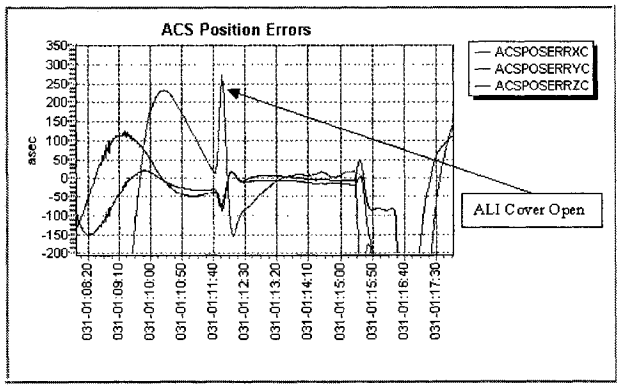


Figure 9 Position Errors During ALI Cover Open

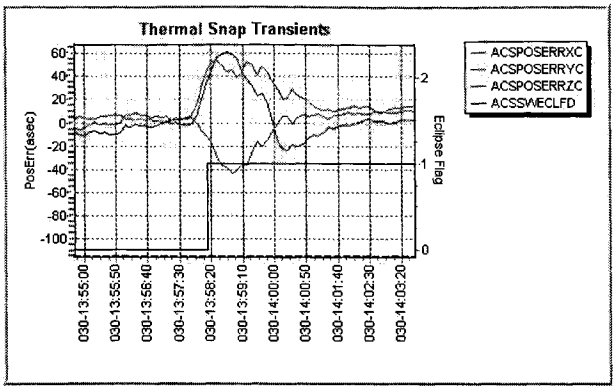


Figure 10 Thermal Snap Transients for Day to Night Transition

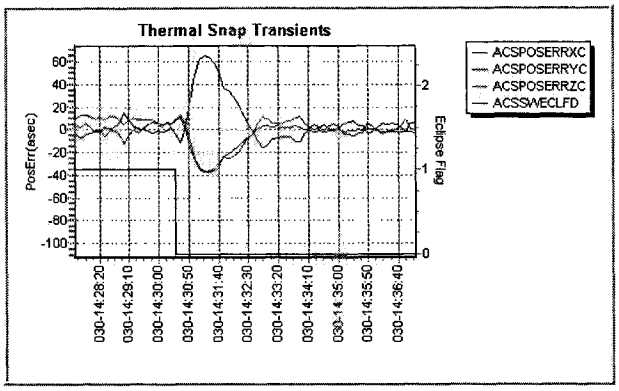


Figure 11 Thermal Snap Transients for Night to Day Transition

Delta-V Summary

Attitude control performance during thruster-based Delta-V mode has been excellent. The derived requirements for the Delta-V mode control were 5.0 degrees attitude error, and 0.2 deg/sec rate error, 3σ. The ACS phase plane plots indicate that the spacecraft is following the expected trajectories. A phase plane plot from the Day 00-346 17 minute Delta-V is shown in Figure 12. Attitude errors have predictable, steady-state values near [1.6 –1.0 0.0] deg in the X, Y, and Z-axes respectively; this is expected since there is not an integral term in the controller.

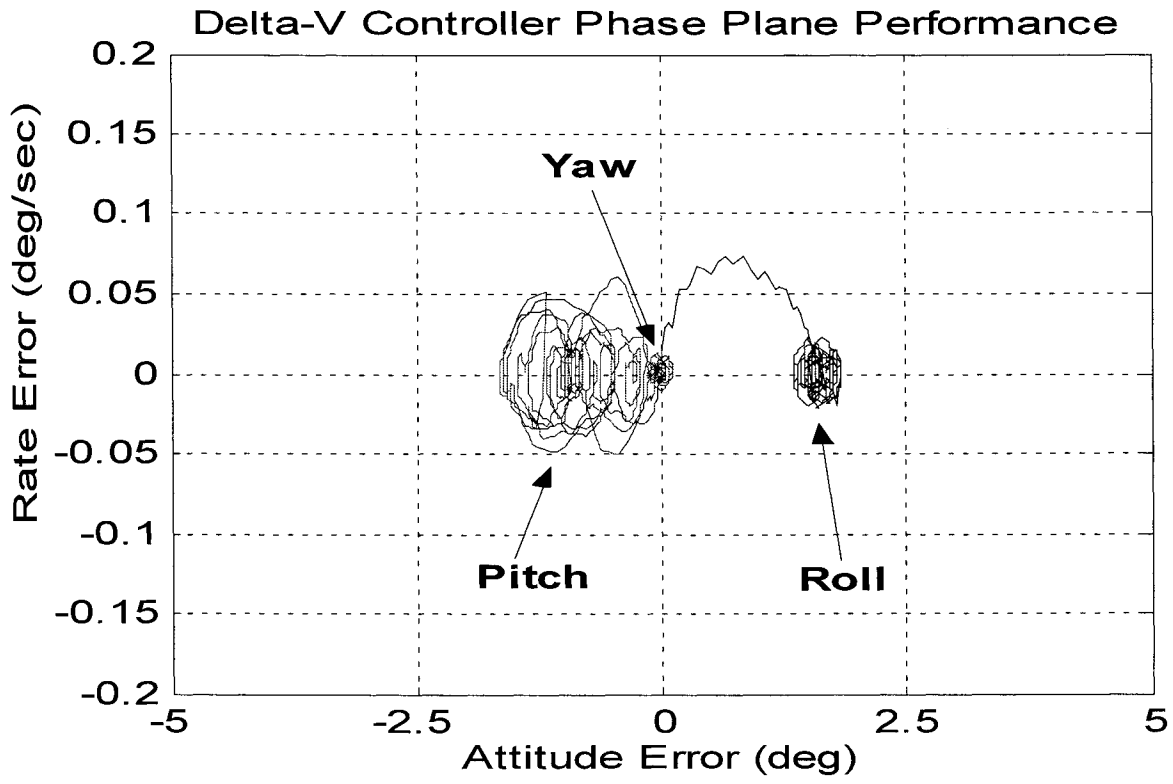


Figure 12 Delta-V Controller Phase Plane Performance

EO-1 SENSOR CALIBRATION

Calibration Slew Planning

The EO-1 calibration slews were performed on 00-328, the 3rd day of the mission. The slews consisted of three "out and back" 90 deg orbital referenced slew pairs about each of the spacecraft primary axes. Each slew pair included 20 minutes of constant rate data prior to returning to nadir pointing. The three slew pairs were separated by two orbits and centered on the orbital eclipses. The slews were designed such that the AST was not blocked by the sun, moon, or earth during any of the constant rate periods. This slew sequence not only provided the geometry required for the IRU calibration, but also provided the MTB activity required for the TAM calibration. Figure 13 shows the roll pitch and yaw angles, body rates and MTB dipole for the calibration slew sequence.

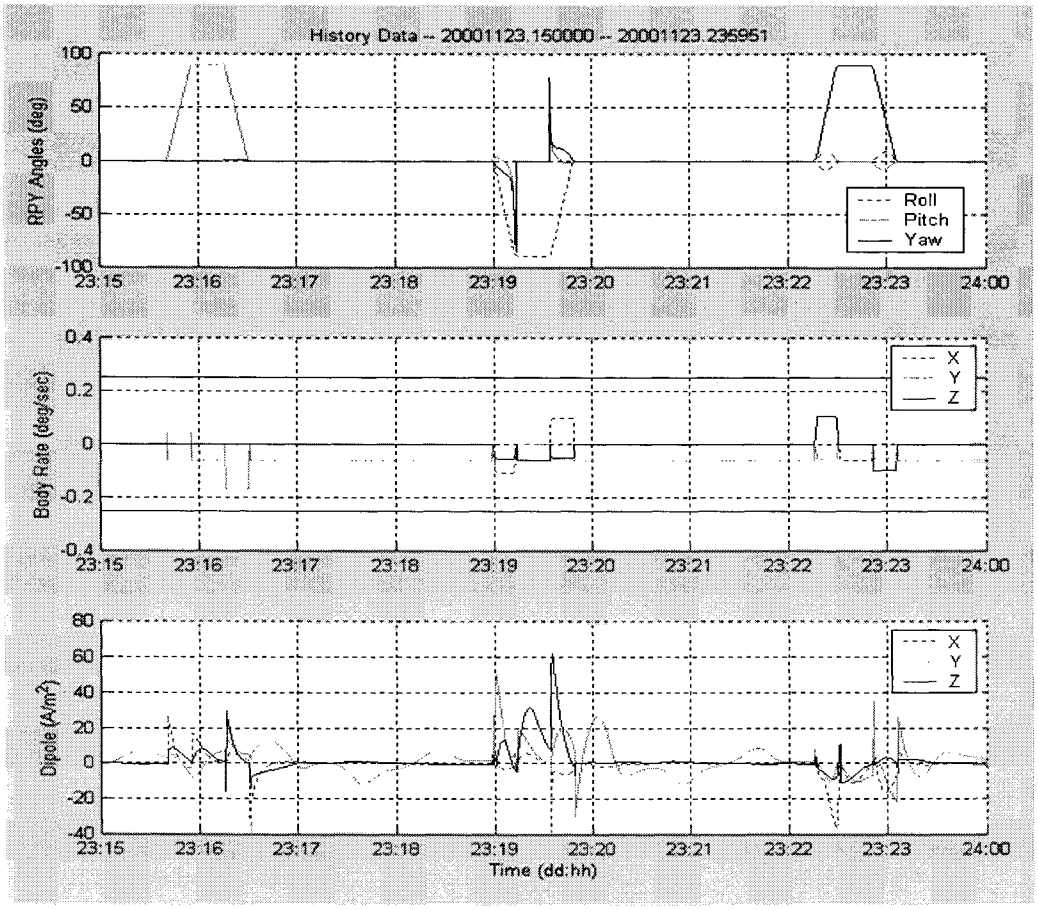


Figure 13 EO-1 Calibration Slews

IRU Calibration Results

IRU calibration parameters consisted of a 3x3 combined alignment/scale factor matrix and a three component bias vector. Input to the IRU calibration algorithm (Reference 3) consisted of ground attitude solutions during the constant rate sections and the IRU data. The resulting alignment/scale factor matrix showed very little scale factor error, but a rather large alignment shift, as shown in Table 3.

Table 3 IRU Calibration Results

	X	Y	Z
Scale Factor	1.000805	0.998989	1.000069
Rotation (arcsec)	301	753	183
Bias (deg/hour)	1.015376	0.080755	1.582949

Since the alignment is from the IRU to the spacecraft body (as defined by the AST), changes from the pre-launch values could be due to either a shift in the relative alignment between the IRU and the AST (caused by the launch forces) or pre-launch optical cube measurement errors for the AST or IRU. The IRU combined alignment/scale factor matrix was uplinked to the spacecraft on 00-339. Shortly after the table uplink, the ACS Kalman Filter IRU bias estimates shifted to the expected values in the Table 3.

TAM Calibration Results

TAM calibration parameters consisted of a 3x3 combined alignment/scale factor matrix, a 3x3 MTB coupling matrix, and a three component bias vector. Input to the TAM calibration algorithm (Reference 4) consisted of the ACS Kalman Filter attitude solution, TAM data and MTB data. Residuals and results of the TAM calibration can be seen in Table 4.

Table 4 TAM Calibration Results

	X	Y	Z
Scale Factor	0.963434	0.971647	1.002988
Rotation (arcsec)	3695	9414	3212
Bias (mG)	-15.141	10.267	-13.768
Coupling Error @ 62 Am^2 (mG)	-8.308	-11.222	-52.886

Kalman Filter Woes

The overall operation of the six-state Kalman filter in the ACS flight software for attitude determination and gyro drift bias estimation has exhibited features not anticipated before launch. However, the performance of the Kalman filter has been steadily improved as a result of on-board parameter updates and changes to the stored command sequences for attitude maneuvers.

Prior to the IRU alignment calibration and resultant upload on 00-339, the Kalman filter performance was limited by the relative misalignment between the AST and IRU. Many of the early attitude maneuvers also suffered from the complete loss of AST updates to the filter due to the AST residuals exceeding the 3-sigma expected magnitude. As these AST outages were induced by attitude slew maneuvers, a change was added to the science DCE sequence generation to disable filter updates during the image prep slews. Another early orbit issue with the Kalman filter was FDC test trips due to the attitude covariance exceeding a threshold. These Kalman filter operational issues led to closer evaluation of the design and a subsequent two-stage parameter modification.

The early Kalman filter operation exhibited IRU drift bias estimate signatures that were inconsistent with the expected IRU bias stability performance. During nominal, Earth pointed operation, the IRU drift bias estimates varied on the order of 1.0 deg/hr, and during many slew maneuvers exceeded 10.0 deg/hr. When the ACS team investigated the filter tuning parameters for the expected IRU noise/stability, it was found that the three flight table values were incorrect – apparently left over from Kalman filter development in the spacecraft CDR timeframe. An effort was immediately undertaken to analyze the EO-1 IRU acceptance test data to determine the actual noise/stability performance and re-design the filter tuning gains to properly account for the expected drift bias variation. These three new IRU noise parameters were uploaded on 01-018 and the Kalman filter was re-initialized. The IRU drift bias estimate signatures became more consistent with expected IRU drift stability performance. As a related benefit, the Kalman filter covariance divergence rate, which is directly dependent on these IRU noise parameters, also improved – reaching the FDC test trip threshold will now require about 10 orbits without AST

updates, compared to the 1 orbit of the original design. The commands to disable filter updates during the image prep slews was subsequently removed from the science DCE sequence generation.

As of 01-019, the Kalman filter IRU drift bias performance was improved significantly and allowed further details of the filter operation to be analyzed. Another symptom of a Kalman filter systematic error became apparent: during the image prep attitude slews, the AST residuals exhibited values which were proportional to the additional angular rate (i.e. with compensation for orbit rate in the pitch axis). The ACS team developed four possible causes for this type of behavior, which are listed below along with an evaluation of likelihood:

1. Uncalibrated IRU scale factor error *[the ground calibration had previously determined the scale factors very close to expected values, and this type of error does not match the residual profile]*
2. AST performance degradation during slews *[AST acceptance testing demonstrated performance much better than the residual profiles, and this type of error does not completely match the symptoms]*
3. AST propagation time calculation error *[the AST residuals are consistent with a +1.0 second propagation time error; since the propagation time does involve using time measurements from the AST, the ACE and the M5, it is a good candidate (more below)]*
4. Systematic time-tagging or M5 CDH/ACS error *[needs to be considered further]*

On 01-029, the ACS variable that holds the AST propagation time was inspected using the M5 memory dwell feature. The expected value was 0.626 sec, and the measured values were near 0.622 sec. This indicated that the additional 1.0 sec offset was not an error in the actual propagation delta time calculation. After an evaluation in the EO-1 software maintenance lab, the AST delta time on-board the spacecraft was adjusted by –1.0 second with a table load on 01-030. The Kalman filter residual performance during a subsequent instrument solar calibration slew demonstrated that the –1.0 sec time adjustment corrected the problem. As of this time, the ACS team is continuing to pursue the possible causes of the 1.0 sec error. Since the signage indicates that the AST attitude data is 'in the future', it is not a classic data latency problem.

INSTRUMENT CALIBRATION MANEUVERS

Solar Calibration

The instrument solar calibration requires a single slew maneuver to point the instruments toward the sun followed by an inertial hold and then an earth acquisition slew back to nadir pointing.

Lunar Raster Scan

The EO-1 spacecraft is required to perform a monthly raster scan of the moon for instrument calibration. There are five detectors that must be scanned across the moon with a spacecraft pitch rate of 0.06875 deg/sec [$1/8^{th}$ of the normal Earth scan rate of 0.55 deg/sec]. The raster scan is made up of a series of alternating roll and pitch slews. The roll slews are used to move from one detector to another, and the pitch slews are used to scan each detector across the moon. The slew maneuver command includes a string of parameters specifying target attitude, scan rate and slew duration. The pitch scans are designed to have a rate profile with an acceleration, coast and deceleration phase so that the scan of the detectors across the moon are during the coast phase with a constant rate.

The ACS spacecraft position errors, spacecraft body rates and control modes for a lunar calibration are shown in Figure 14. The position errors for the maneuver remained below 600 asec, or 0.167 degrees, during the lunar calibration slews.

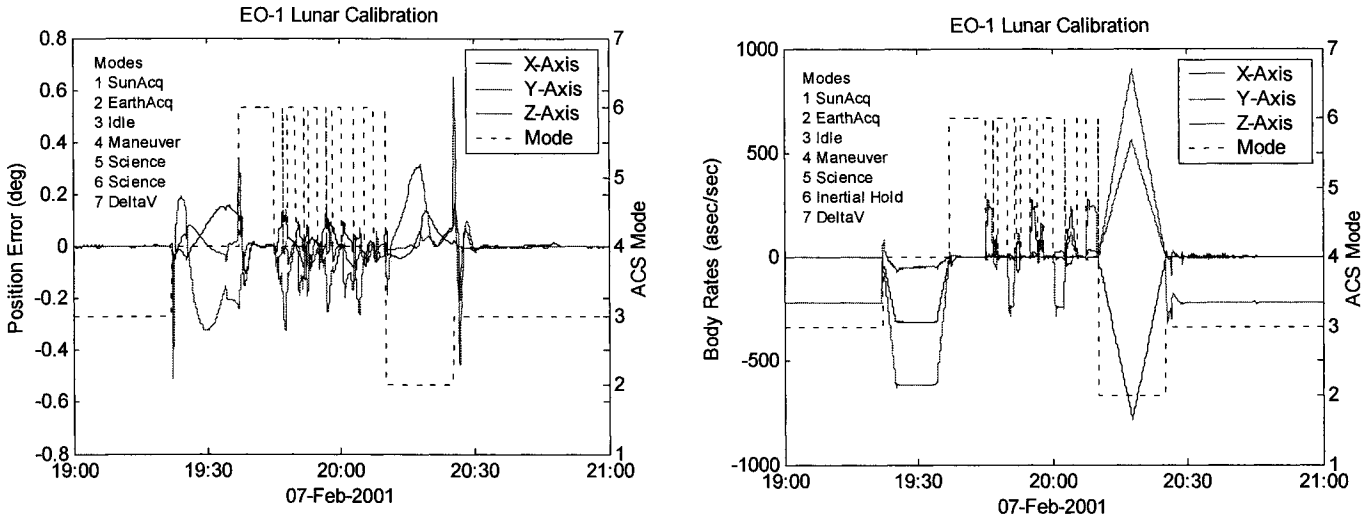


Figure 14 Lunar Calibration Position Errors and Rates

COMPONENT PERFORMANCE

Lockheed Martin AST-201 Star Tracker

Overall, the EO-1 Autonomous Star Tracker performance has been better than expected. All initial AST attitude acquisitions have been successful. The mean AST RMS error has been approximately 30 urad, as shown in Figure 15, which correlates closely with ground processing estimates. The AST RMS error has varied between 10 and 100 asec on average with spikes as high as 500 asec. The AST has been tracking between 10 to 45 stars, with a mean of 25 stars. The AST Effective Focal Length (EFL) error telemetry indicates a mean of less than 10 μm of error with no need for updates to the EFL. The AST has shown susceptibility in the South Atlantic Anomaly, with a range of 0 to 4 events per day – either coast frames, reacquisitions or reboots – all with autonomous recovery to track mode.

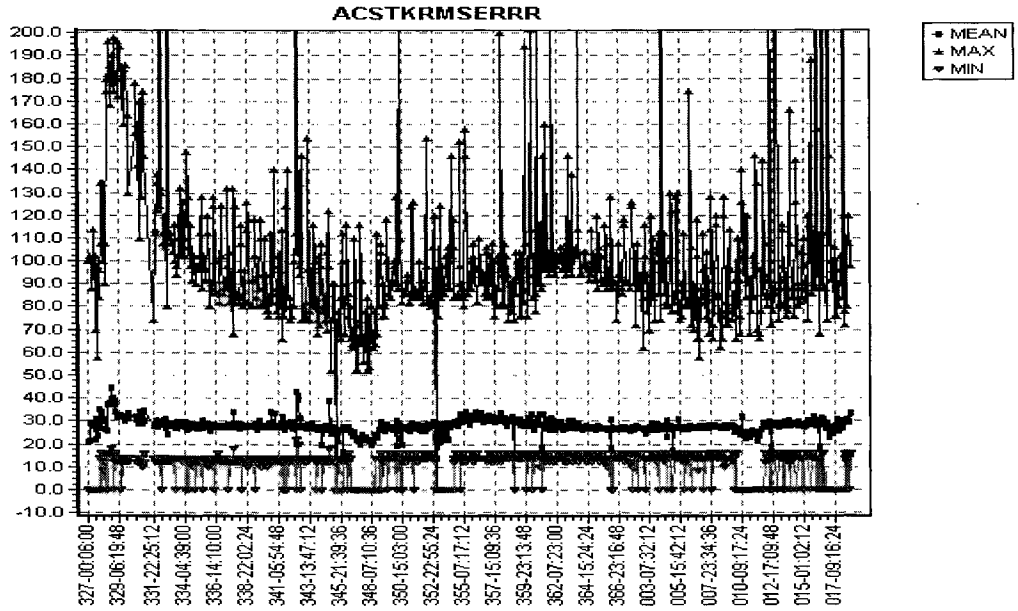


Figure 15 Star Tracker RMS Errors (urad)

Litton G&C Space Inertial Reference Unit

The IRU has performed nominally since launch. There have been no lost packets, packet checksum errors, or invalid packets. The IRU calibration identified negligible scale factor error and the IRU-to-AST on-orbit alignment calibration resulted in less than 0.2 degree (720 asec) change. The gyro drift biases were estimated at [1.02 0.08 1.58] deg/hr and have remained constant.

The IRU parametric drive voltages, which were a concern because of the possibility of helium venting on the launch pad, have continued to improve since launch. On orbit voltage statistics show an improvement of 0.18 volt for gyro A, 0.14 volt for gyro B and 0.24 volt for gyro C over the 3 month duration since launch.

Loral Space Systems GPS Tensor

The GPS receiver successfully performed two cold start acquisitions in the sun pointing orientation. The receiver remained in track during all of the IRU calibration slews, Delta-V slews, and imaging slews. It had been anticipated that the receiver may have difficulty during these off-nadir operations, but the orbit determination Kalman Filter within the GPS Tensor has demonstrated robustness through these events. The ACS was configured to utilize GPS ephemeris updates on 00-329 after two days of monitoring the GPS performance compared to the ACS propagated ephemeris, as well as the ephemeris generated using a ground solution. The GPS position and velocity navigation vector performance has been well within the 100m spec, as shown in Table 2. Reference 5 provides a more in-depth evaluation of the GPS testing and performance for EO-1.

CONCLUSION

The EO-1 Attitude Control Subsystem design and implementation had numerous challenges as part of a 'faster, better, and cheaper' mission. The on-orbit pointing accuracy, with attitude controller errors and knowledge errors combined via a Root-Sum-Square, is approximately 40 arcsec, 3-sigma during science observations. This performance has enabled the instrument team to perform alignment calibrations using nadir scenes. The lunar raster scan maneuver has exceeded expectations and provides a significant radiometric calibration source for instrument detectors. The performance and predictability of the Delta-V mode has conserved propellant usage and set the stage for extended mission operations with the Enhanced Formation Flying experiment.

REFERENCES

1. P. Sanneman, K. Blackman, M. Gonzalez, and D. Speer, "New Millennium Earth Orbiter-1 Mission: Attitude Control Requirements & Capabilities", 21st Annual AAS Guidance & Control Conference, AAS 98-002, February, 1998.

2. D. Speer and P. Sanneman, "Attitude Determination and Control for the New Millennium EO-1 Spacecraft", 1998 IEEE Aerospace Conference Proceedings, March, 1998.

3. J. Keat, "Gyro Calibration Analysis for the High Energy Astronomy Satellite-A (HEAO-A)", Computer Sciences Corporation, CSC/TM-77/6082 (1977).

4. G. M. Lerner and M. D. Shuster , "In-Flight Magnetometer Calibration and Attitude Determination of Near-Earth Spacecraft", Journal of Guidance and Control, Vol. 4, No. 5, 1981, pp 518-522.

5. D. Quinn, P. Sanneman, S. Shulman and J. Sanger, "GPS Systems on Orbit: The Integration, Testing, and Flight of the EO-1 GPS", 2001 GSFC Flight Mechanics Symposium, NASA CP-____, 2001 (Paper 36).

SURFACE TENSION LOCKUP IN THE IMAGE NUTATION DAMPER – ANOMALY AND RECOVERY

Carl Hubert
Hubert Astronautics, Inc.

Daniel Swanson
Lockheed Martin Space Systems

ABSTRACT

Early telemetry from the spin-stabilized IMAGE spacecraft indicated that the vehicle's initial nutation was not decaying. This behavior was especially puzzling because the spacecraft's passive nutation damper behaved as expected while IMAGE was attached to the spinning upper stage. The lack of damping was also puzzling because the damper was a tubular ring partially filled with liquid mercury; a simple, reliable device with a long flight history. In a partially-filled ring damper, the excess kinetic energy associated with nutation is dissipated by fluid viscosity when inertial forces cause the liquid to move through the tube. However, post-launch analysis indicated that the IMAGE damper liquid was immobilized by surface tension. This was an unanticipated consequence of the vehicle's low spin rate. When it became apparent that passive damping did not work, a ground-commanded open-loop damper was developed using the spacecraft's magnetic torquer and onboard logic that was intended for ground test of the torquer. This work-around successfully resolved the IMAGE nutation damping problem.

HUBBLE SPACE TELESCOPE SERVICING MISSION 3A RENDEZVOUS OPERATIONS

S. Lee, S. Anandakrishnan, C. Connor, E. Moy, D. Smith
Lockheed Martin Technical Operations/Greenbelt, MD

M. Myslinski
Honeywell Technology Solutions Inc./Greenbelt, MD

L. Markley, A. Vernacchio
NASA Goddard Space Flight Center/Greenbelt, MD

ABSTRACT

The Hubble Space Telescope (HST) hardware complement includes six gas bearing, pulse rebalanced rate integrating gyros, any three of which are sufficient to conduct the science mission. After the loss of three gyros between April 1997 and April 1999 due to a known corrosion mechanism, NASA decided to split the third HST servicing mission into SM3A, accelerated to October 1999, and SM3B, scheduled for November 2001. SM3A was developed as a quick turnaround "Launch on Need" mission to replace all six gyros. Loss of a fourth gyro in November 1999 caused HST to enter Zero Gyro Sunpoint (ZGSP) safemode, which uses sun sensors and magnetometers for attitude determination and momentum bias to maintain attitude stability during orbit night. Several instances of large attitude excursions during orbit night were observed, but ZGSP performance was adequate to provide power-positive sun pointing and to support low gain antenna communications. Body rates in ZGSP were estimated to exceed the nominal 0.1 deg/sec rendezvous limit, so rendezvous operations were restructured to utilize coarse, limited life, Retrieval Mode Gyros (RMGs) under Hardware Sunpoint (HWSP) safemode. Contingency procedures were developed to conduct the rendezvous in ZGSP in the event of RMGA or HWSP computer failure. Space Shuttle Mission STS-103 launched on December 19, 1999 after a series of weather and Shuttle-related delays. After successful rendezvous and grapple under HWSP/RMGA, the crew changed out all six gyros. Following deploy and systems checkout, HST returned to full science operations.

INTRODUCTION

During normal science operations, the Hubble Space Telescope (HST) employs a set of high-performance, low noise Rate Integrating Gyros (RIGs) in conjunction with Fine Guidance Sensors (FGSs) and reaction wheels to provide precision pointing with jitter of less than 7 mas (Refs. 1 – 3). The RIGs use a brominated fluorocarbon (BTFE) flotation fluid to achieve exceptional noise performance. HST lost three RIGs between April 1997 and April 1999 due to BTFE-caused corrosion of flexleads carrying electrical power and signals to the gyros, leaving the vehicle with the minimum complement of three RIGs. Failure of any of these gyros would cause HST to enter Zero Gyro Sunpoint (ZGSP) safemode, rendering it unable to conduct its science mission. In response to this situation, NASA divided the planned third HST servicing mission into two missions: SM3A and SM3B. The primary objective of SM3A, accelerated to October 1999, was to replace all RIGs on-board HST.

With HST under normal mode control with three functioning RIGs, a nominal R-bar rendezvous was planned for SM3A. However, on November 13, 1999, a fourth RIG failed due to flexlead failure and HST entered ZGSP. In this mode, HST utilized rates derived from Coarse Sun Sensors (CSSs) and the Magnetic Sensing System (MSS) along with CSS position to maintain a power positive sunpoint attitude. Momentum bias was established using the reaction wheels to provide gyroscopic stiffness during orbit night when sensor data was unavailable. Consideration was given to conducting the rendezvous in ZGSP, but expected body rates exceeded the 0.1 deg/sec rendezvous flight rule. However, Shuttle Engineering Simulation (SES) training runs demonstrated the crew's ability to execute a ZGSP rendezvous as a contingency.

With only two RIGs available, rendezvous in Hardware Sunpoint (HWSP) safemode using the Retrieval Mode Gyro Assembly (RMGA) was baselined. HWSP utilizes HST's secondary computer system with CSSs to provide closed-loop sunpoint control. Rate data is provided by the RMGA which houses three coarse, limited-life Retrieval Mode Gyros (RMGs). To verify the performance of this control mode, an on-orbit test was conducted on December 3, 1999. Rendezvous with HST under HWSP required the crew to execute an orbiter "flyaround" to place HST's grapple fixtures in position for capture by the Orbiter Remote Manipulator System (RMS).

To mitigate the uncontrolled vehicle spin-up which would result from a hardover RMGA failure, a ground-based monitoring and recovery procedure was developed. This procedure allowed ground controllers to quickly identify a hardover failure and command HST to drift mode. Strategies for manual reaction wheel commanding were developed to provide post-drift rate damping to allow recovery back to ZGSP.

STS-103/SM3A launched on December 19, 1999. On Flight Day 3, HST was commanded to HWSP using RMGAs. The crew successfully conducted the flyaround maneuver and captured HST on schedule. On Extravehicular Activity (EVA) Day 1, all six RIGs were replaced by the EVA crew. After two additional EVA days, HST was successfully deployed and shortly thereafter resumed science operations.

NORMAL POINTING CONTROL SYSTEM AND RATE GYROS

Description of Normal Pointing Control System

The HST Pointing Control System (PCS) is a multi-rate, multi-loop, discrete time system with multiple modes of operation that provides attitude reference, attitude control and stabilization, and maneuverability in support of HST science operations and vehicle momentum management. Functional control of the system resides within PCS software modules executed by the Flight Computer (see Figure 1). Ground support is required for uplink of data base updates and vehicle pointing requirements. The PCS operates in different mission modes to provide both coarse and fine pointing and a stable platform for scientific operations. In Normal mode operation, the primary actuators are four skew-mounted Reaction Wheel Assemblies (RWAs) and the primary sensors are single degree-of-freedom, fluid floated, gas bearing, Rate Integrating Gyros (RIGs). HST contains six RIGs housed in three Rate Gyro Assemblies (RGAs). Any combination of three or four RIGs can be used to provide three-axis PCS Normal mode control. Momentum management is accomplished using four magnetic torquer bar actuators with local B_{earth} field sensing provided by redundant, three-axis magnetometers. In the event of major failures, a Safing system provides Sunpoint attitude control to allow for HST survival until the system can be restored via ground commanding or by component replacement during servicing missions.

PCS software modules resident in the Flight Computer process the PCS sensor data and command the PCS actuators. These software modules provide for the calculation of the PCS control law, attitude updates, the momentum management control law, and the command generator. The command generator acceleration and incremental position commands are provided to the control system software together with the RGA incremental rate data. The control system uses position, rate, and integral compensation, as well as a rate filter to stabilize high frequency bending modes. An estimate of external torque is used to eliminate cross-coupling between axes, effectively decoupling the control torques in the vehicle axes, denoted V1, V2, and V3. A feedback loop provides an error calculation path to account for variations in parameters, such as the vehicle inertia estimate and RWA feed forward gain.

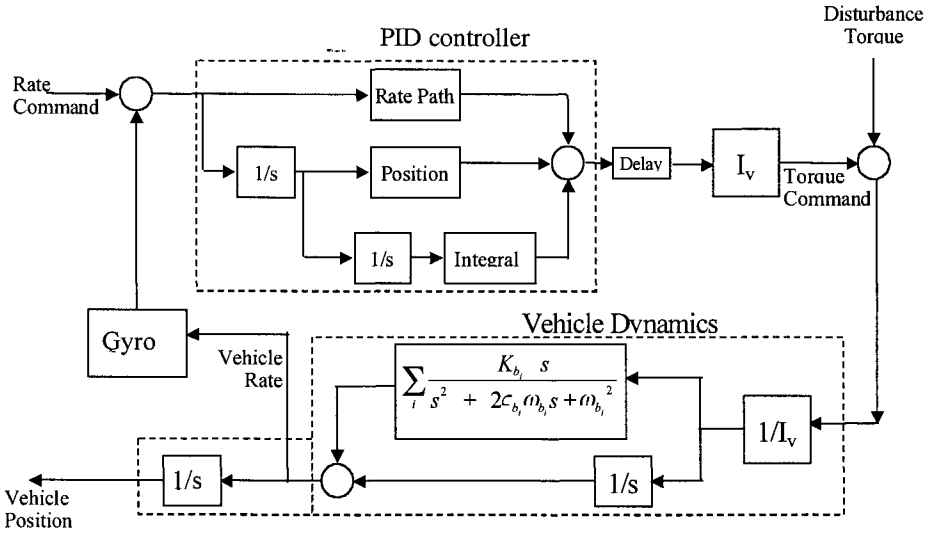


Figure 1: Normal Mode Block Diagram

Overview of Rate Gyro Assemblies

The RGA is a strapdown reference gyro package that senses vehicle motion using two ultra-low noise RIGs to provide two-channel digital attitude and analog rate information. The resulting short term attitude and attitude rate data are used during normal operations for fine pointing and spacecraft maneuvers. In addition, the RGA is used by the Safing system, for attitude and rate information during safe mode operations, provided that the RGAs are not the cause of safemode entry. A single RGA is composed of a Rate Sensor Unit (RSU) that houses the RIGs, and an Electronics Control Unit (ECU) that contains two independent sets of control electronics. Each RSU/ECU set provides two independent electronic channels and RIGs. The RSU and ECU can be replaced during extravehicular activity (EVA) on servicing missions.

The RGAs are configured in quantized, pulse rebalance loops. The HST data management system provides an interface between the gyros and the flight control computer. The gyros are driven by a two phase hysteresis motor, which is capable of maintaining rotor speed on a single phase under certain conditions. However, motor start up is not possible with a single phase, and loss of spin motor rotation (run down) will occur with loss of both phases.

Flexlead Corrosion Failure Mechanism

The gyroscope used in the RGAs is the 64 PM RIG built by L3 Space and Navigation (L3SN). This gyro is composed of a hermetically sealed housing and a float assembly, which contains the gyro wheel, drive motor and torquer coil. The float is attached electrically and mechanically to the housing via flexible leads of very small cross section. These flex leads, which transport power to the motor, magnetic torquer and signal generator, are composed of a highly conductive alloy. The flotation fluid is specially formulated for density match with the float and for viscosity consistent with the required gyro gain. This fluid, in the presence of oxygen (possibly introduced during the gyro fill process, or generated by interaction of the fluid with the flex leads), is very corrosive to the flex leads (Ref. 4). Over time, at operating current and temperature (which accelerate the corrosion process), the flex leads become porous, and structural integrity and current carrying capability are compromised. The effect of localized heating, due to the relatively high current in the porous flex leads, is believed responsible for the failures.

Chronology of Flex Lead Related Gyro Failures

Since HST was launched in April 1990, several on-orbit failures and anomalies have been experienced due to degraded flex leads (Table 1). Although five gyros have failed, all five had already achieved operating life expectancy at the time of failure. In two cases, the gyros continued to operate following loss of a single motor phase to flex lead failure. However, bias instability exceeded operational limits in one of these cases, making rate data unusable for normal science activity. In both of these cases, the second phase was lost within months of the first phase failure, due to increased current load in the surviving phase. In the other three cases, loss of a single phase resulted in rotor spin down and loss of the gyros.

Corrective Action

In response to these failures, L3SN undertook an aggressive gyro production program aimed at replacement of the failed units during SM3A. The gyro fill process was modified to preclude introduction of oxygen into the fluid, and the fluid was carefully screened to ensure that all parameters were well within specification, particularly those associated with flex lead corrosion. The program was completed as planned and the hardware was ready for the SM3A launch.

Additional reaction to the gyro failures led to development of enhanced flex leads, which are resistant to corrosion by the the flotation fluid while retaining the mechanical and electrical characteristics of the original configuration. This ambitious effort

Date	Gyro	Remarks
10/7/92	G6	Replaced during SM-1, not in control loop at time of failure
4/9/97	G4	Telemetered lead (Phase B) failed, resulting in rotor spin down. Second phase failed within hours.
7/28/98	G6	Phase A failure, motor continued to run with significantly increased current on single phase, gyro returned to control loop
10/22/98	G6	Phase B failed resulting in rotor spin down.
1/25/99	G3	Phase A failure, motor continued to run with significantly increased current on single phase, gyro unusable due to bias instability
4/10/99	G3	Phase B failed resulting in rotor spin down.
11/13/99	G1	Phase A failure resulted in rotor spin down

Table 1: RGA Gyro Flex lead Failure History

was made possible through active participation of LMMS and GSFC personnel, in conjunction with L3SN engineers. Enhanced flex leads were produced and subjected to accelerated life testing, which has yielded excellent results to date and is still in progress. Enhanced flex leads have suffered no failures thus far, while one of the original configuration has failed as expected. Enhanced flex leads have also been installed in a first flight gyro, which has yielded nominal data thus far. A second gyro is awaiting flex lead installation. These and subsequent units will be available to support SM4 and beyond.

SERVICING MISSION 3A DECISION

HST experienced the loss of three RIGs between April 1997 and April 1999. This left the vehicle with the minimum complement of three RIGs in a single string configuration. Failure of any remaining gyro would cause HST to enter Zero Gyro Sunpoint (ZGSP) safemode, rendering it unable to conduct its science mission.

Prior to the Gyro 3 phase A flexlead failure in January 1999, plans were underway to conduct the third servicing mission to HST in July 2000. The manifest for this servicing mission, named SM-3, included the following:

- Replacement of all three RSUs
- Installation of Voltage Improvement Kits (VIKs)
- Replacement of the existing DF-224 Flight Computer with an Intel 486-based Flight Computer
- Replacement of Fine Guidance Sensor 2 (FGS-2)
- Replacement of S-Band Single Access Transmitter 2 (SSAT-2)
- Replacement of Engineering and Science Tape Recorder 3 (ESTR-3) with Solid State Recorder 3 (SSR-3)
- Installation of New Outer Blanket Layers (NOBLs)
- Installation of a new instrument, the Advanced Camera for Surveys (ACS)
- Installation of new, rigid solar arrays
- Installation of two cooling systems, the NICMOS Cooling System (NCS) and the Aft Shroud Cooling System (ASCS)

As a result of the loss of the Gyro 3 phase A flexlead and the possibility of lost HST science, NASA decided to split SM-3 into two servicing missions. The first, SM3A, was accelerated to October 1999 as a "Launch-On-Need" mission. The primary objective of SM3A, designated STS-103 by the Space Shuttle Program, was to replace all six gyros on-board HST. In addition, the VIKs, 486, FGS-2, SSAT-2, SSR-3, and NOBLs were manifested. The remaining hardware from SM-3 was manifested for the second mission, SM3B, scheduled for November 2001.

NOMINAL RENDEZVOUS PLANS

Since three gyros were available, initial planning for SM3A baselined a nominal R-bar rendezvous with HST in normal mode. Figure 2 provides a graphical representation of the nominal R-bar rendezvous sequence. Throughout this sequence, the Orbiter maintains a local vertical attitude with the +Z axis aligned to the R-bar (i.e. Earth nadir). HST conducts a short series of two maneuvers to first establish an inertial rendezvous attitude followed by an inertial capture attitude as described below.

After launch into the HST orbit plane, the Space Shuttle Orbiter conducts a short series of burns to place it in a co-orbit with HST. In preparation for rendezvous, HST closes its aperture door, slews its solar arrays to 90 degrees (i.e. in the V1-V2 plane), and maneuvers to a rendezvous attitude. Figure 3 shows the HST reference frame and depicts the solar arrays at 0 deg. This attitude minimizes the impact of plumes from the Orbiter's upward firing, norm-Z Reaction Control System (RCS) jets while providing a power-positive sun angle on the HST solar arrays. The rendezvous attitude is an inertial attitude resulting in the +V1 axis (i.e. HST boresight) being aligned with R-bar (i.e. Earth nadir) with a roll angle established to place the sun in the V1/+V3 half-plane. This attitude effectively "feathers" the solar arrays edge-on to the Orbiter and is established at 59 degrees prior to orbit noon where the Orbiter reaches its closest norm-Z RCS braking gate. After this point, the orbit switches to low-Z RCS mode which avoids firing of the norm-Z jets, eliminating the threat of plume impingement on HST.

Prior to orbit noon, HST is maneuvered to a capture attitude in preparation for grapple by the Orbiter's Remote Manipulator System (RMS). This attitude is defined such that at grapple time, shortly after enter orbit night, the HST -V1 axis is aligned to the Orbiter +Z axis (i.e. HST's aft bulkhead into the payload bay). In addition, the HST grapple fixtures, which are aligned with the -V3 axis, point 52 degrees out-of-plane north, as measured from the anti-sun side of HST (see Fig. 4). Thus, relative to the rendezvous attitude, capture attitude is achieved by rolling HST by β_{solar} + 52 degrees. This series of HST maneuvers optimizes propellant usage by achieving capture without requiring Orbiter attitude maneuvers.

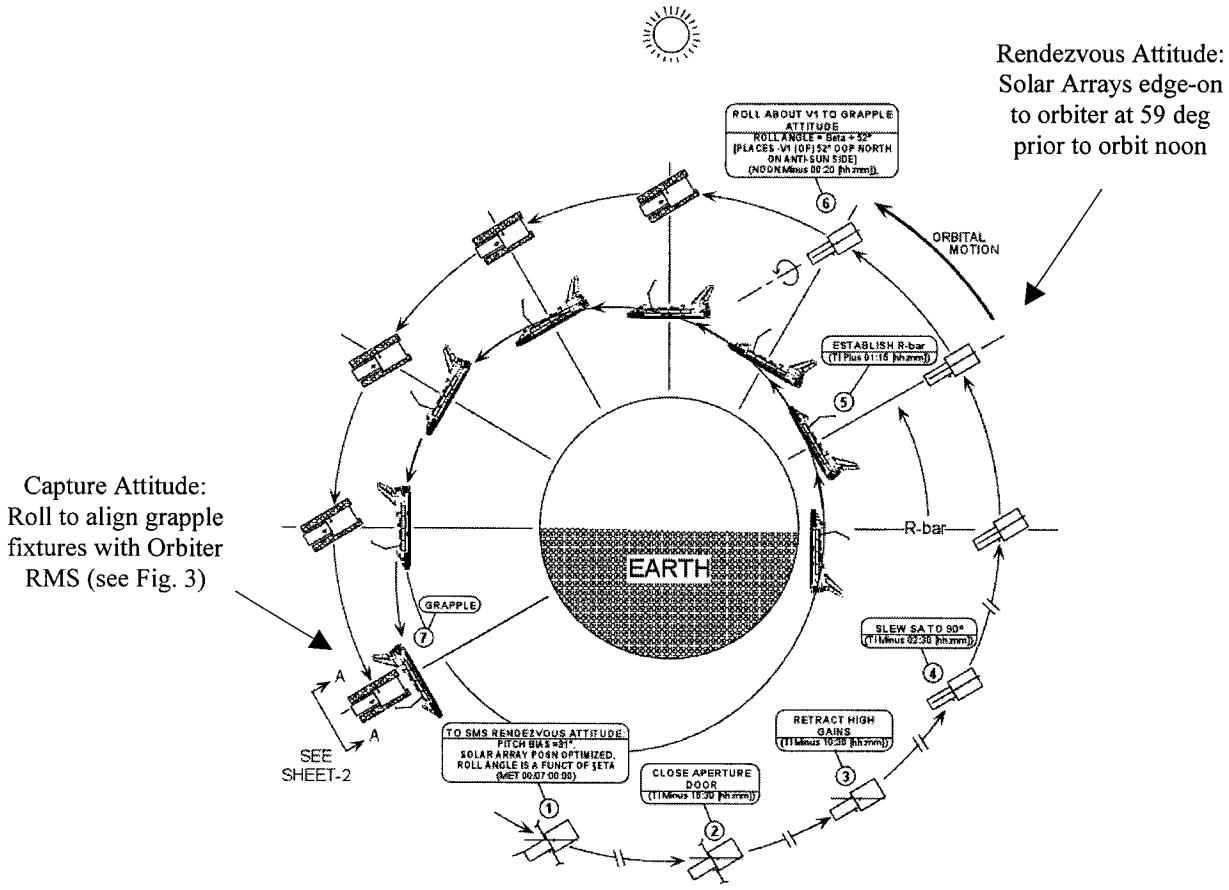


Figure 2: R-bar rendezvous sequence, zero beta case.
Inertial frame depiction, not to scale

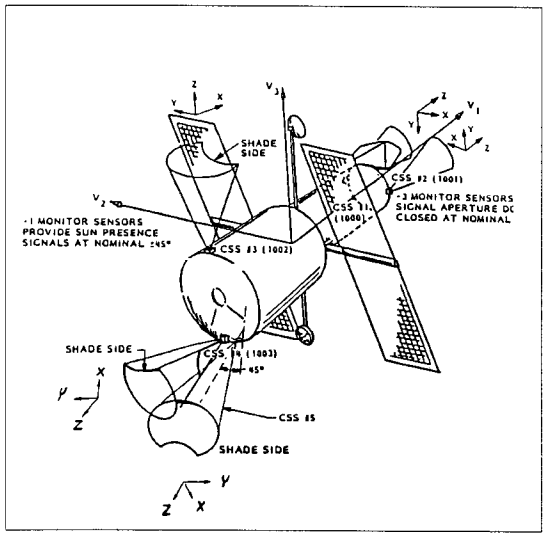


Figure 3: Vehicle frame definition and CSS fields of view

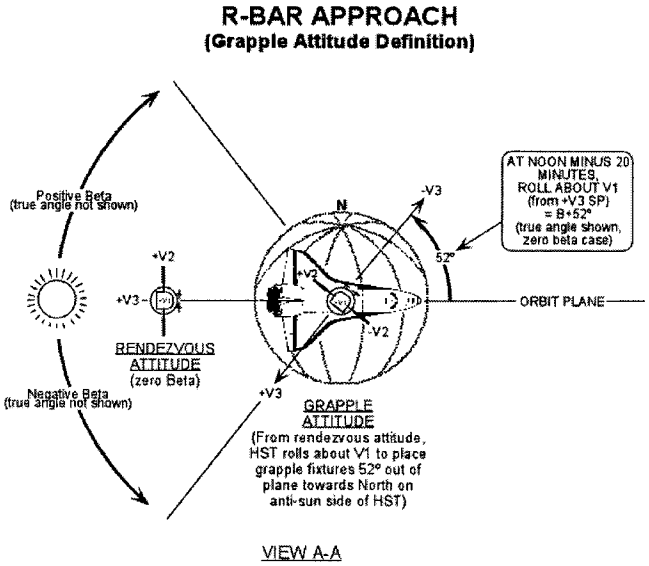


Figure 4: Grapple attitude definition

RGA GYRO 1 FAILURE AND ZERO GYRO SUNPOINT

On 11/13/99 a fourth gyro failed due to flexlead corrosion. With only two viable single-axis gyros available, three-axis attitude control with gyros alone was not possible. As a result, HST entered Zero Gyro Sunpoint (ZGSP) which ensures a power-positive and thermally safe state without the use of rate gyros. HST remained in ZGSP for approximately 570 orbits, until SM3A. ZGSP is a robust, long-term safemode requiring no ephemeris information. Since ephemeris and attitude knowledge are unavailable while in ZGSP, the high gain antennas are not used and communications are established through two hemispherical S-Band antennas. The aperture door is closed at ZGSP entry to preclude the possibility of sunlight down the boresight during the initial capture. ZGSP uses a set of five Coarse Sun Sensors (CSSs) which provide 4π steradian coverage as illustrated in Figure 3 along with the MSS as attitude sensors. The Reaction Wheel Assembly (RWA) is used for primary attitude control, and magnetic torque rods are used for momentum management.

While HST is in sunlight, ZGSP uses CSS data for position and rate information about the axes perpendicular to the Sun line, and the magnetic field vector for rate damping about the Sun line. Vehicle attitude about the Sun line is not actively controlled. Since the Earth's magnetic field is not inertially fixed, magnetometer-derived rates are coarse. However, they are adequate to maintain rates around the Sun line at an acceptably low level. During orbit night, CSS data is unavailable and ZGSP disables active control. As a result, all three axes drift, but the drift away from the Sun line is limited by a momentum bias in the direction of the sunpoint axis. This momentum bias is commanded after initial sun capture once sunpoint errors, as measured by the primary CSS, have been reduced below a database-specified value. Momentum bias is established in the RWAs by using the magnetic torque rods to load the wheels to 250 Nms. For a detailed description of ZGSP see Reference 5.

During the extended period under ZGSP control prior to SM3A, HST would occasionally Enter Orbit Day (EOD) with a large attitude error/excursion (Sun more than 41° from the +V3 axis at EOD) due to drift during orbit night - the uncontrolled part of the orbit. Except for one orbit on DOY 1999.350, ZGSP control captured the Sun (defined as the Sun being within 3 deg of the Sun point, i.e., +V3, axis) within 37 minutes following EOD. During the time that HST was in ZGSP, each orbit day was approximately 59 min. long. The DOY 1999.350/17:58 orbit entered orbit day with a large excursion and took approximately 57 min to capture the Sun. The large attitude excursions and capture times govern the time it takes the solar arrays to fully charge the batteries. While in ZGSP, on average, it took 28.5 min. from EOD for the batteries to reach trickle charge, compared to 15 – 20 min. in Normal mode. The lowest battery State Of Charge (SOC) was as low as 335 A-hr on one orbit, but still well above the 225 A-hr SOC safemode threshold.

While ZGSP met its prime objective of maintaining HST in a power positive and thermally safe state, it was not designed for and did not meet the performance requirements for nominal Orbiter rendezvous. The ZGSP HST-STS rendezvous option is discussed in detail in the next section. However, the parameter of primary importance for rendezvous is HST rates. Due to the large attitude errors at EOD, the transients can be quite large and the rates are quite often not favorable for rendezvous. While the vehicle rates, especially about the V1 and V2 axes, can be quite low after Sun capture, the time between Sun capture and Enter Orbit Night (EON) is not predictable.

Examination of the orbit night data from the on-orbit test of ZGSP showed that at EON, the largest V1 and V2 rate magnitudes were less than 0.03 deg/s. However, the V3 rate magnitude could be as high as 0.15 deg/s. This is not unexpected, given that the rate about V3 is computed from the measured magnetic field, using the simplifying assumption that the Earth's magnetic field and the V3 axis are inertially fixed. The resulting derived rates are inaccurate because the Earth's magnetic field describes a roughly conical motion at an average rotation rate of two revolutions per orbit due to spacecraft orbital motion. The 5787 sec HST orbit period gives an orbit rate of 0.06 deg/sec and can result in V3 rates as high as 0.12 deg/s. In addition, the control is low gain rate damping with a limited torque of 0.3 N-m.

The rotational motion around V3 does not have any effect on Sun pointing; but can affect drift during eclipse, since the total body momentum can be reduced by the body rates, lowering the stiffening effect of the momentum bias. ZGSP relies on the magnetic torque rods and the reaction wheels to establish and maintain a wheel momentum bias of 250 N-m-s about the V3 axis. However, the magnetic torque rods are ineffective near the South Atlantic Anomaly (SAA) and the wheel momentum bias vector at EON may be as much as 30 deg away from the vehicle momentum vector and 50% less than the target wheel momentum bias. This can cause the vehicle rate vector to precess about the net momentum vector, resulting in the possibility of rates as high as 0.15 deg/s about any axis during the uncontrolled orbit night pass.

HST-STS RENDEZVOUS OPTION WITH HST IN ZERO GYRO SUNPOINT

Following the failure of Gyro 1 and HST entry into safemode, consideration was given to conducting the SM3A rendezvous in ZGSP. Nominal Orbiter rendezvous flight rules restrict maximum HST body rate to 0.1 deg/sec during rendezvous. However, it became desirable to re-evaluate this restriction to provide options for HST retrieval during SM3A.

Rendezvous operations with HST in ZGSP require the Orbiter to conduct a "flyaround" to position HST's grapple fixtures for RMS capture. The initial Orbiter approach in this case is identical to the nominal R-bar approach until the Orbiter reaches 120 ft range at which point the flyaround is initiated. The flyaround is conducted at the mission commander's discretion based on the orientation of HST relative to the Orbiter at capture time. The commander flies the Orbiter into a position to place either the –V1 or +V1 axis into the payload bay. In addition, the –V3 axis is pointed towards the port forward region of the payload bay to expose the grapple fixtures for RMS capture.

To assess the capability of the Orbiter crew to support ZGSP rendezvous, a series of simulations were conducted at NASA/JSC using the Shuttle Engineering Simulator (SES). The SES is a six degree-of-freedom, real-time on-orbit simulator which includes a full scale aft flight deck and large visual graphical displays for crew training. On November 23, 1999, the SM3A crew commander, pilot, and flight engineer/RMS operator conducted a series of three simulated ZGSP rendezvous. HST rate profiles were simulated based on results of the 1993 ZGSP on-orbit test. Simulated rates were scaled up by 30% from the on-orbit data to provide a reasonable worst-case rate profile. As a result, maximum HST rates of 0.18 deg/sec were simulated.

Figure 5 provides an overview of the HST rate profile measured during the 1993 on-orbit test, scaled up by 30%. Four runs were conducted during the November 23 SES session. The first simulated a HWSP rendezvous. The second was selected as representative of a typical, low HST rate interval. In this run, the crew accelerated the R-bar approach to arrive at 120 ft 25 minutes prior to enter orbit night to allow capture shortly after orbit night when lighting conditions are optimal for the RMS operator. Flyaround and grapple was successfully completed 12 minutes before sunset. This run demonstrated the feasibility to attempt ZGSP grapple prior to enter orbit night although additional propellant was required during low-Z braking operations.

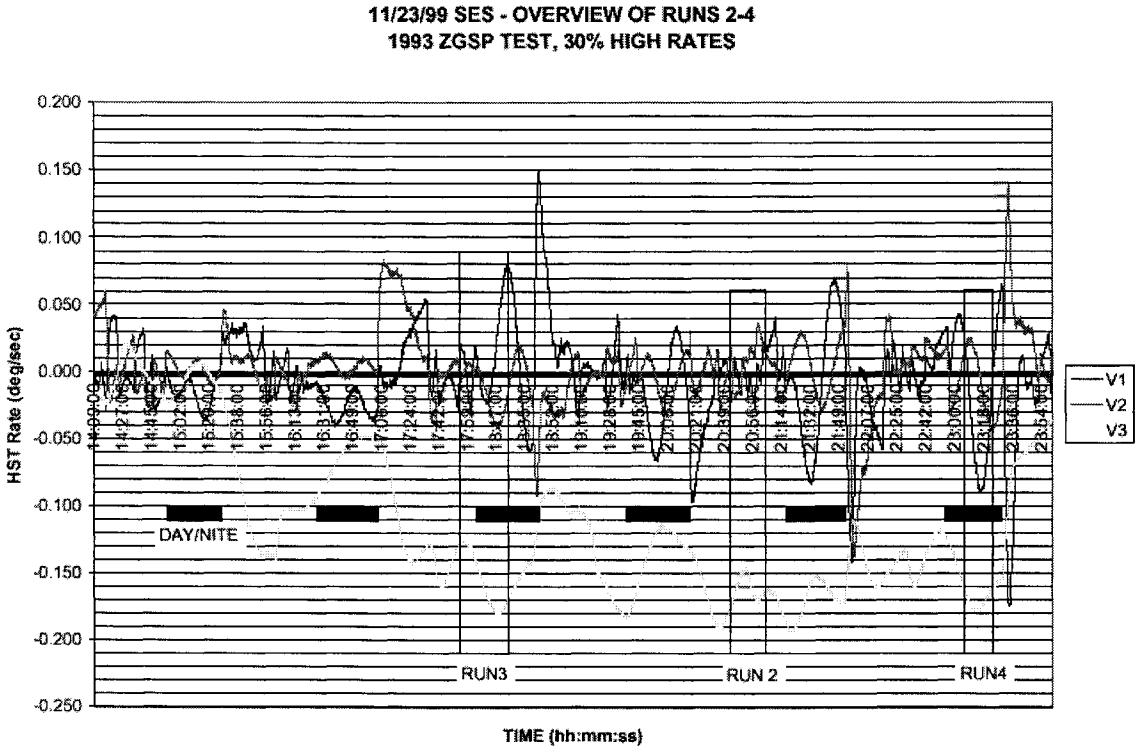


Figure 5: ZGSP SES Rendezvous Runs

The third run utilized a worst-case rate profile during transition from orbit day to orbit night. In this case, a nominally timed R-bar approach was executed followed by successful flyaround and capture shortly after enter orbit night. The final run simulated a combination of large and changing HST rates during orbit night. Maximum V1 rates of -0.085 deg/sec were experienced. In this case, due to the large HST roll rate, the grapple pin was rotating away from the RMS. As a result, the crew chose to maintain an Orbiter yaw rate of -0.25 deg/sec in free drift at grapple. Again flyaround and grapple operations were successful.

Additional ZGSP simulations were conducted in the weeks leading up to launch of SM3A. In each case the crew successfully executed Orbiter flyaround and capture. These simulations demonstrated the feasibility of ZGSP rendezvous. Ultimately, however, rendezvous under Hardware Sunpoint (HWSP) with Retrieval Mode Gyro Assemblies (RMGAs) was baselined for SM3A due to the lower and more predictable body rates provided. ZGSP rendezvous was retained as a contingency in the event of any anomalies which precluded HWSP/RMGA rendezvous.

HARDWARE SUNPOINT WITH RETRIEVAL MODE GYROS

Overview of HWSP and RMGAs

HWSP functionality is contained within the on-board secondary computer system known as the Pointing and Safemode Electronics Assembly (PSEA) which operates independently of the primary Flight Computer. The PSEA's primary purpose is to provide autonomous action in the event of a malfunction of the primary Flight Computer. HWSP provides attitude control to position the spacecraft at a power positive attitude with the sunline along the +V3 axis and the solar arrays at 90 deg. Alternatively, the PSEA can be configured for –V1 sunpoint with the solar arrays at 0 deg. In addition, power loads are shed by turning off hardware not critical for HWSP operations.

HWSP consists of a position and rate feedback system to provide immediate response to changes in vehicle attitude. The block diagram for HWSP is shown in Figure 6. Attitude sensing is provided by the CSSs. Rate data can be provided either by RGAs or by coarse gyros housed in the Retrieval Mode Gyro Assembly (RMGA). The RGAs are the primary gyros used for normal operations and measure rates as low as 3.91e-3 deg/sec. The RMGA is provided as a backup rate gyro in the event of significant failures of the RGA. The RMGA consists of three limited life, single-axis gyros configured along the vehicle axes which can measures coarse rates as low as 7.81e-3 deg/sec. The alignment of the gyros input axes are such that polarity of the measured rates are the opposite of vehicle rates. As with other modes, RWAs and MTS provide actuation for HWSP.

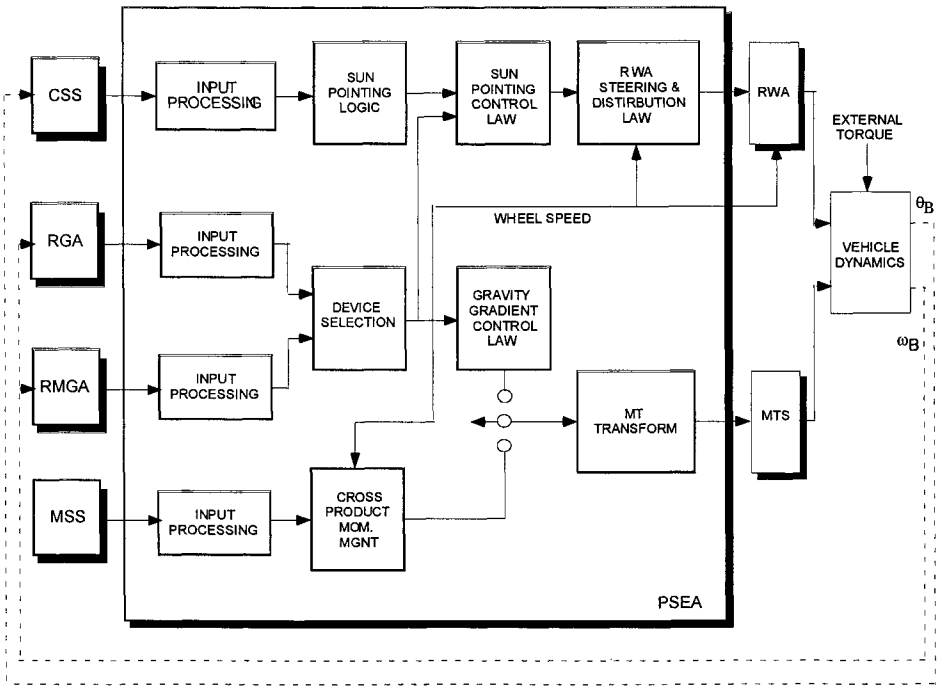


Figure 6: HWSP Block Diagram

Pre-SM3A On-Orbit Test of HWSP/RMGA

With their short life cycle, the RMGA is operated infrequently to measure their drift rate and update the associated bias, if necessary. Polarity checkout of the RMGA output had also been performed immediately after launch to ensure proper RMGA values. However, prior to SM3A, the RMGA had never been operated in the closed-loop HWSP control system. As a result, an orbit test was planned to ensure their performance as a viable configuration for SM3A. The HWSP/RMGA on-orbit test was performed on 12/3/99.

Prior to activating HWSP control, the RMGAs were powered on to assess their performance before committing to HWSP. PSEA Testmode was activated to assess the gyro's performance. Testmode is a pseudo operating mode within the PSEA which can be used to monitor hardware interfaces without HWSP entry. RMGA-measured rates were then compared with the rate output of the remaining RGA gyros (G2 and G5). This was accomplished by transforming body frame RMGA rates into G2 and G5 frame and evaluating the difference with G2 and G5 outputs. The largest error observed was 0.038 deg/sec with a mean/bias of 0.016 deg/sec. The largest standard deviation was 0.007 deg/sec. RMGA rates during this checkout period are shown in Figure 7. Note that body rates were relatively large (maximum of V1 = 0.072 deg/sec, V2 = 0.04 deg/sec, and V3 = 0.064 deg/sec) since the vehicle was still under ZGSP control. Based on this performance, the RMGAs were declared functional for the on-orbit test.

Following completion of testmode, HWSP was activated. Since the vehicle was fairly close to the desired sunpoint attitude when HWSP was initiated, capture occurred almost immediately. Figure 8 shows the body rates throughout the test. HWSP performance kept body rates relatively low with maximums of V1 = 0.048 deg/sec, V2 = 0.040 deg/sec, and V3 = 0.040 deg/sec.

Due to the relatively high noise introduced into the control law by the RMGA, the RWAs were torqued at a high duty cycle in HWSP, and resulted in wheel temperature increases. Initial trends indicated a potential for the RWA temperatures to eventually exceed their 49°C upper operating limit (see Figure 9). As a result, the on-orbit test was terminated and the vehicle commanded back to ZGSP after two hours as RWA temperatures reached 44 °C. Overall, the on-orbit test proved that HST body rates under HWSP/RMGA remained well below the 0.1 deg/sec rendezvous flight rule.

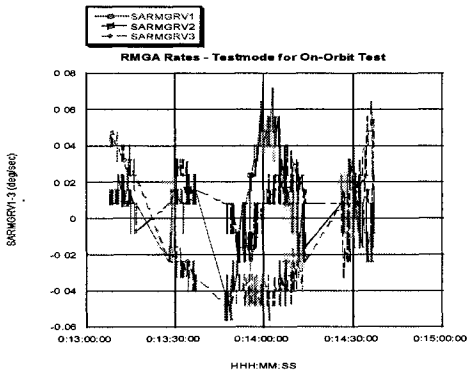


Figure 7: RMGA Rates during Checkout

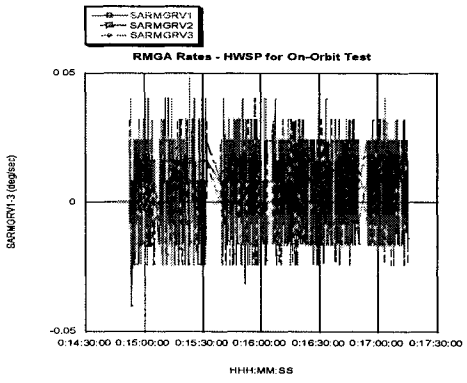


Figure 8: Rates during HWSP On-Orbit Test

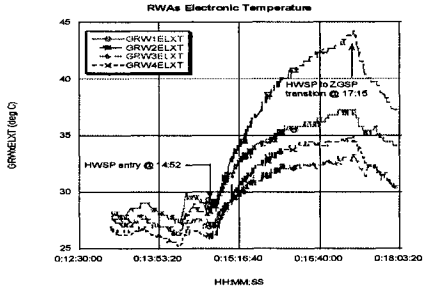


Figure 9: RWA Temps during On-Orbit Test

HWSP/RMGA Rendezvous

In order to support a HWSP/RMGA rendezvous, SM3A timeline changes were necessary based on knowledge of HWSP and lessons learned from the on-orbit test. Post-test analysis of RWA temperature trends by the vendor, Honeywell Inc., indicated that such high temperatures would not degrade performance.

The SM3A timeline was modified to reconfigure HST to the desired configuration for HWSP/RMGA rendezvous. This configuration included preparing the vehicle for rapid recovery to ZGSP as a contingency for rendezvous. In addition, as the Orbiter approaches HST, communication is switched from TDRSS relay to direct Orbiter communications. Vehicle loads are reduced and power relays are opened in preparation for receiving external power from the Orbiter after berthing in the payload bay. Immediately following grapple, the Orbiter crew commands HST to free drift via aft-flight deck commands to prevent the control law from fighting the RMS.

HWSP/RMGA FAILURE MONITORING AND GROUND-BASED RECOVERY PLANS

RMGA Hardover Failure Scenario

Consideration was given to contingency rendezvous strategies in the event of HWSP/RMGA failures or performance anomalies. Failure of the PSEA or RMGA zero-output failures would result in HST drift. Procedures were developed to rapidly recovery to ZGSP. Although body rates under ZGSP were higher than prescribed by the flight rules, the crew was trained to conduct a ZGSP rendezvous as a contingency described previously. However, it was recognized that a hardover failure of the RMGA could result in body rates too high to allow expeditious recovery to ZGSP.

An RMGA hardover failure would manifest itself as saturated positive or negative rate output. The response of the HWSP control law would be to command maximum RWA torque in the corresponding axis. The subsequent sequence of events is summarized in Figure 10 and described below.

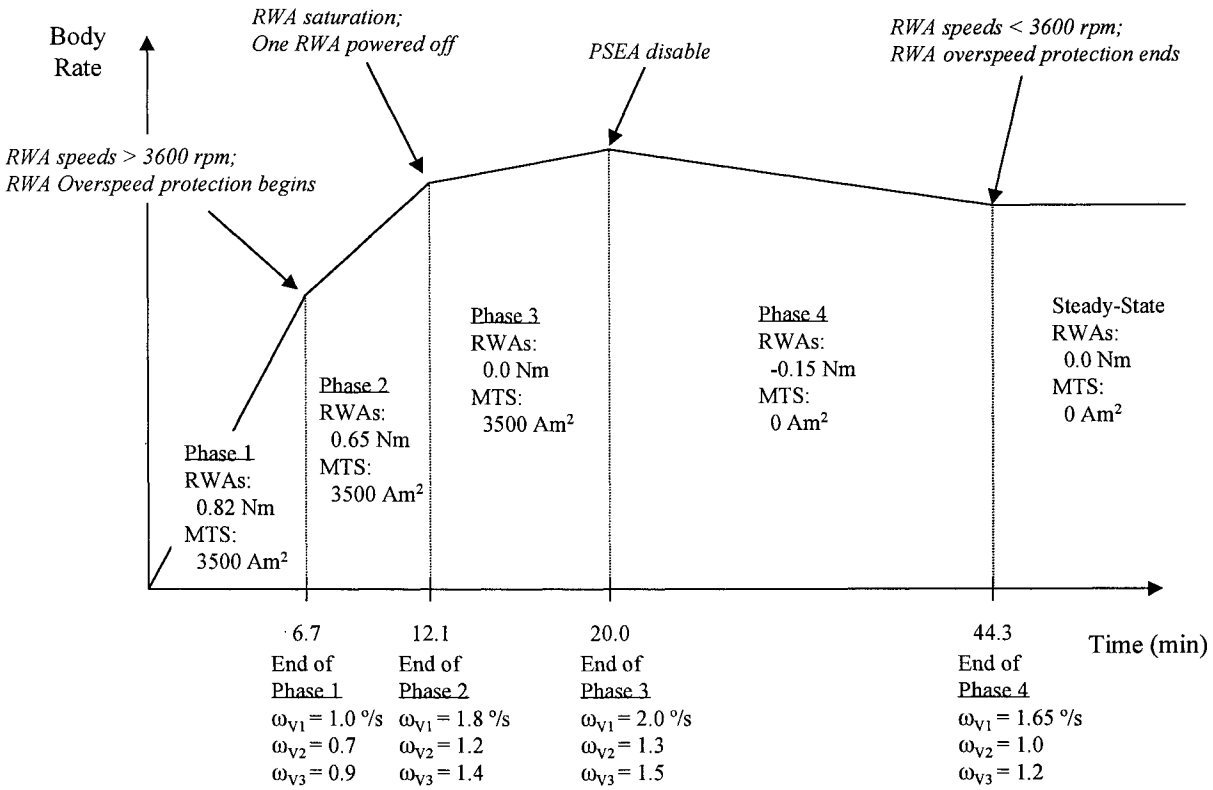


Figure 10: RMGA Hardover Rate Profile

During phase 1, immediately after the RMGA hardover failure, maximum torque commands are issued by the control law. Due to the symmetrical distribution of the RWAs, the result is maximum (0.82 Nm) torque to the RWAs. At the same time, momentum unloading attempts to reduce RWA wheel speeds by commanding maximum current to the MTS. This phase continues until the RWAs reach 3600 rpm approximately 6.7 minutes after the failure. Phase 2 begins when with the engagement of an RWA overspeed protection circuit, which subtracts 0.15 - 0.17 Nm from the torque command for speeds over 3600 rpm.. As wheel speeds approach saturation (6000 rpm) towards the end of this phase, a Safing system RWA momentum test declares one RWA failed. This is predicted to occur once wheel speeds reach approximately 5900 rpm. To simplify the analysis, this reconfiguration was not assumed. Thus, phase 2 ends 12.1 min after the failure when all RWAs saturate at 6000 rpm. During phase 3, all RWAs are saturated and provide no further torque to HST. However, momentum unloading continues to attempt to reduce RWA wheel speeds. As a result, HST body rates continue to build in response to MTS unloading commands. Left unchecked, this phase continues indefinitely. Thirty seconds are required to establish a lock on one of the two C-band antennas, which provide hemispherical coverage. Thus, once body rates exceed 6.0 deg/sec in either the V2 or V3 axis, communication with the spacecraft is no longer possible.

RMGA Hardover Failure Monitoring

To mitigate the RMGA hardover failure scenario described above, monitoring criteria were established to identify the failure and allow disabling of the PSEA. RMGA rates were not monitored directly because certain telemetry formats did not cover the full range of RMGA outputs (i.e. T-format limited to 0.25 deg/sec vs full scale range of 1.0 deg/sec). Since the response of the HWSP control law would be to command maximum RWA torque in the corresponding axis, RWA torque command became the primary monitoring parameter. This approach would also catch any anomaly in the control path which resulted in an erroneous, prolonged torque command.

A baseline criterion of 60 sec was established to preclude inadvertent declaration of an RMGA hardover failure for a momentary, nominal RWA torque command saturation in response to environmental disturbances or Enter Orbit Day (EOD) excursions. A second criterion was identified to account for any transients experienced during the initial capture under HWSP. This second criterion was based on simulations conducted with 50 run cases in HWSP. Each case provided different initial body rates and RWA wheel speeds. Results were evaluated to identify the case with the longest saturated RWA torque command time. This case is shown in Figure 11 with ω_{V1} = 0.10 deg/sec, ω_{V1} = -0.16 deg/sec, and ω_{V3} = -0.16 deg/sec, which

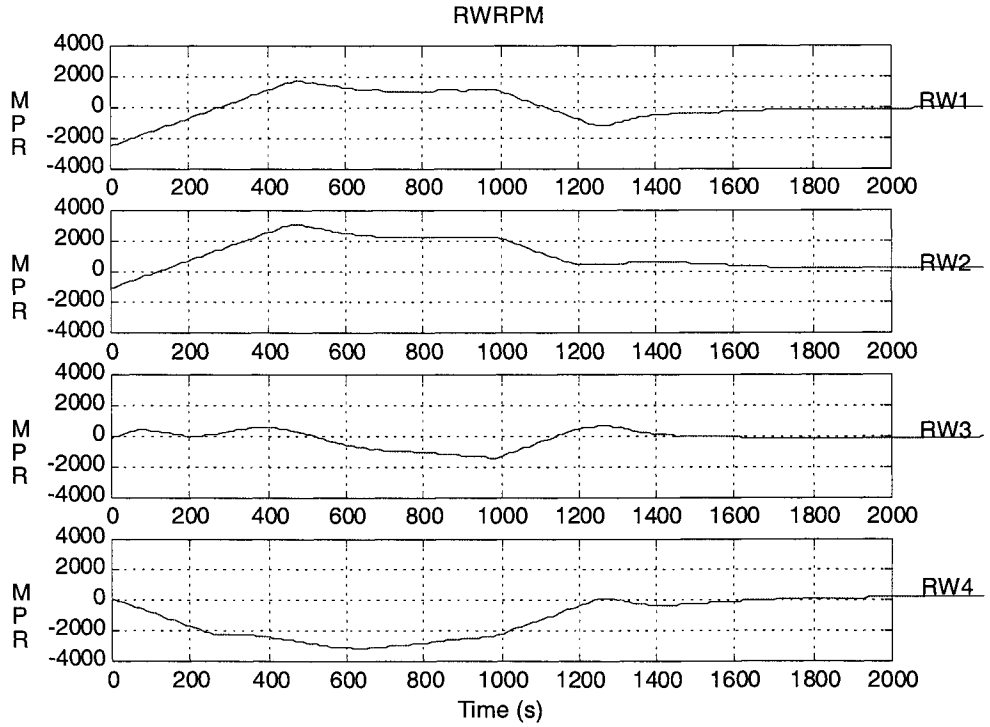


Figure 11: RWA Speeds HWSP Simulation with Longest RWA Saturation

results in saturated RWA torque commands for 460 sec. Therefore, 460 seconds was allowed for saturated RWA torques following HWSP initiation. In addition, a third criterion was added to preclude any unidentified failure resulting in slow divergence from sun capture.

The HWSP/RMGA monitoring criteria are summarized below:

1) Immediately after HWSP initiation, any RWA torque command > 0.70 Nm for 460 sec
2) Subsequent to initial HWSP capture, any RWA torque command > 0.70 Nm for 60 sec;
3) Loss of sun presence in CSS3 monitor cell during orbit day

The response to any of these criteria was to be the immediate commanding of PSEA disable, resultinhg in HST drift. Phase 4 of Figure 10 depicts this drift period. Note that, for purposes of this analysis, 20.0 minutes is assumed for PSEA disable (to account for potential communications outages and time to implement the PSEA disable command). At the end of phase 3, HST body rates are ω_{V1} = 2.0 deg/sec, ω_{V2} = 1.3 deg/sec, and ω_{V3} = 1.5 deg/sec. During phase 4, the PSEA is disabled and HST is free to drift. However, the RWA overspeed protection circuit remains active applying -0.15 Nm torque to each RWA. 44.3 minutes after the failure, the RWA wheel speeds fall back below 3600 rpm and the overspeed protection circuit is disabled. At this time, body rates are ω_{V1} = 1.7 deg/sec, ω_{V2} = 1.0 deg/sec, and ω_{V3} = 1.2 deg/sec.

RWA Speed Management

ZGSP was originally designed to support capture at HST body rates up to 0.2 deg/sec. Since body rates are expected to exceed 0.2 deg/sec following an RMGA hardover failure and subsequent PSEA disable, a method was developed to reduce body rates through direct commanding of individual RWA wheel speeds. The method employs the following three stages:

1) Power Management Stage: Reduce wheel speeds at low torque to prevent bus overvoltage due to detorque power return
2) RWA Momentum Management Stage: Return RWAs to pre-failure speeds
3) Body Rate Management Stage: Reduce body rates to < 0.2 deg/sec using rates measured by a combination of RMGA and RGA gyros

Power Management Stage: When an RWA is detorqued (i.e. magnitude of wheel speed is reduced), power is returned to the main bus as a function of detorque magnitude and wheel speed. Under circumstances of high wheel speeds, large detorque magnitude can result in bus voltages which exceed the operational limit of 32.7V. Exceeding 32.7V is not expected to damage any hardware (survival limit = 35V), but its effect on hardware operability is unknown. With a minimum HWSP load condition of 32A, maximum acceptable detorque power is calculated to be 1046W (32.7V x 32A). To provide system margin, 900 W is used as a practical upper limit. As a result, an individual RWA's detorque power is limited to 225W (900 W total ÷ 4 RWAs). By comparison, an RWA at 6000 rpm detorqued at -0.82 Nm would return 445W to the bus.

To satisfy the overvoltage restriction, a power management stage has been devised for conditions when one or more RWA wheel speed exceeds 3500 rpm. During this stage, RWA wheel speeds are reduced with a limit on the detorque command to restrict power returned to the EPS bus. RWA speeds should be reduced by the same amount to reduce the total RWA momentum vector magnitude while maintaining its direction.

The RWA vendor, Honeywell Inc., provided the calibrations for determining RWA power as a function of wheel speed and torque command. Based on this spreadsheet, the following table outlines the maximum allowable torque as a function of wheel speed:

ω_{RWA} (rpm)	τ_{max} (Nm)	P_{bus} (W)
4000	-0.68	224
4500	-0.52	224
5000	-0.42	225
5500	-0.34	223
6000	-0.29	225

Power Management Stage operations consider only the highest speed RWA and assume the same power return for all four RWAs. For future operations, it may be possible to refine these limits by calculating a total power return for all four RWAs given each individual wheel speed. This would allow higher detorque commands and subsequently reduce the time required to damp body rates and recover to ZGSP. Note that to support this stage operationally, the Flight Computer must be recovered and configured to allow visibility into the RWA wheel speeds. In addition, any RWAs pulled off-line should be powered back on.

RWA Momentum Management Stage: During this stage, all RWA wheel speeds are commanded to the state at the time just prior to the RMGA hardover failure. In other words, total RWA momentum is returned to its value prior to the failure. However, residual body rates will exist due to (1) torquer bar application during maximum RWA torque commanding and (2) environmental disturbances since the PSEA/RMGA failure. Worst-case residual body rates from torquer bar application were calculated to be V1 = 0.72 deg/sec, V2 = 0.29 deg/sec, and V3 = 0.29 deg/sec. Note that during this stage RWA wheel speeds are low enough (<3500 rpm) to preclude overvoltage conditions on the bus even for maximum detorque.

If the PSEA is disabled such that residual body rates after the RWA Momentum Management Stage are less than 0.2 deg/sec per axis, ZGSP can be commanded directly. However, if body rates exceed 0.2 deg/sec per axis, body rates must be reduced further prior to commanding ZGSP.

Body Rate Management Stage: In the event residual body rates after the RWA Momentum Management Stage exceed 0.2 deg/sec, it is necessary to further reduce body rate. In the Body Rate Management Stage, the remaining RGA and RMGA gyros are employed to measure body rate. This is done by formulating a non-orthogonal hybrid RGA-RMGA transformation matrix, $T_{RGA\text{-}RMGA/Body}$, whose rows correspond to the row of the appropriate RGA or RMGA transformation matrix. Then:

$$\boldsymbol{\omega}_{RGA-RMGA} = T_{RGA-RMGA/Body}\boldsymbol{\omega}_{Body}$$

and

$$\boldsymbol{\omega}_{Body} = T_{Body/RGA-RMGA}\boldsymbol{\omega}_{RGA-RMGA}$$

For example, if RGA 2 is to be used with RMGAs 1 and 2:

$$T_{RGA-RMGA/Body} = \begin{bmatrix} -0.526 & 0 & 0.851 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

so

$$\begin{bmatrix} \omega_{V1} \\ \omega_{V2} \\ \omega_{V3} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1.175 & -0.618 & 0 \end{bmatrix} \begin{bmatrix} \omega_{RGA2} \\ \omega_{RMGA1} \\ \omega_{RMGA2} \end{bmatrix}$$

A tool was developed to provide these hybrid rates in near real-time. The tool imports RGA and RMGA gyro rates and applies the appropriate transformation to determine three-axis body rates. Of course, prior to utilizing this method, any RMGA failures must be isolated.

Note that RGA gyros saturate at 0.5 deg/sec while RMGA gyros saturate at 1.0 deg/sec. In the event any gyro is saturated, this method can still be applied since the saturated signal will maintain its polarity. In this case, two or more iterations of this stage may be necessary to reduce body rates below 0.2 deg/sec per axis.

SM3A MISSION SUMMARY

SM3A/STS-103 launched successfully at 00:50 GMT on December 19, 1999. Rendezvous operations commenced at Mission Elapsed Time (MET) of 01/16:14 with power-on of the RMGA. As with the on-orbit test, PSEA Testmode was activated to assess the RMGA's performance relative to the remaining RGAs. The largest error observed was 0.041 deg/sec with a mean/bias of 0.017 deg/sec. The largest standard deviation was 0.009 deg/sec. Figure 12 shows the RMGA rates during this checkout period. Based on this performance, the RMGAs were declared functional for rendezvous operations.

Next, at MET 01/18:26, HWSP was initiated. Since this occurred during orbit day, immediate capture was observed with residual errors less than 0.36 deg, as measured by CSS #3. After an initial transient, body rates diminished to less than 0.02 deg/sec (see Fig. 13). The 250 Nms momentum bias that had been established in ZGSP was unloaded by the MTS within 45 minutes. Coarse attitude determined by ground controllers based on CSS and MTS data indicated that HST was in a stable position for capture with the +V1 axis into the payload bay.

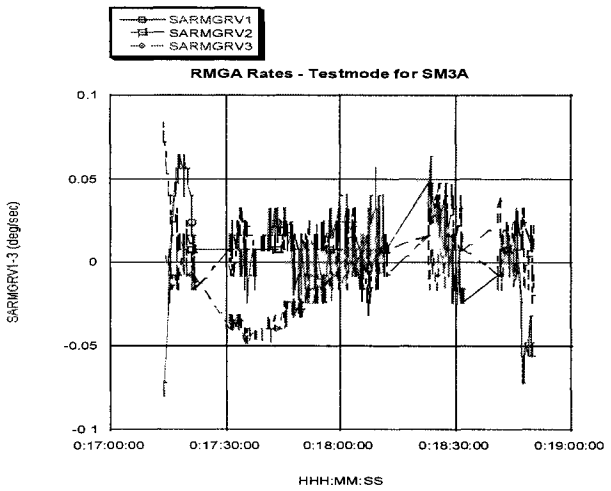


Figure 12: RMGA Checkout Rates during SM3A

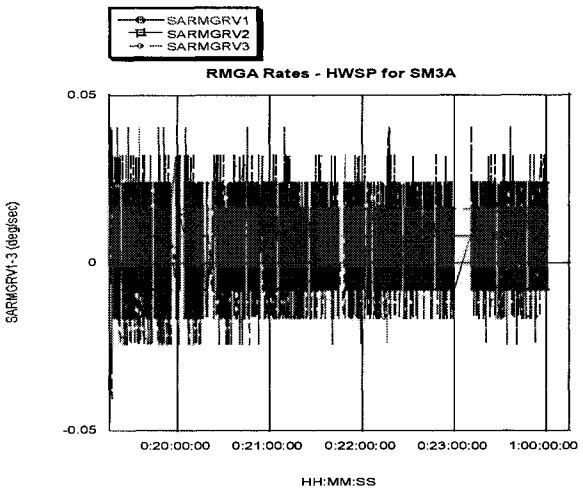


Figure 13: HWSP Initiation during SM3A

HWSP/RMGA performance was monitored continuously throughout the rendezvous. Performance remained nominal with no anomalies observed during this period. The Orbiter crew initiated the flyaround as planned on the rendezvous orbit leading to successful RMS capture at MET 01/23:44. Following aft flight deck commands to disable HWSP control, RWA speeds of less than 100 rpm were observed. HST was then berthed into the payload bay and external power applied from the Orbiter.

On the following day, the first EVA was conducted to replace all RSUs. All RSUs were powered off at 02/19:23. After the crew replaced each RSU the new RIGs within it were powered on for an aliveness test. All six gyros powered on nominally as indicated by motor current, temperature, and analog rate. Following the EVA, a functional test was conducted by maneuvering the orbiter to a local-vertical orientation which put rates on all RIGs. All gyros matched expected rates within the 72 deg/hr uncertainty band. The remaining EVAs resulted in the successful installation of VIKs, the 486 Flight Computer, FGS-2R, SSAT-2, SSR-3, and NOBLs.

On Flight Day 7, HST was prepared for release. As planned, the crew unberthed HST and positioned it for release into the Bright Earth Avoidance (BEA) sunpoint attitude. This attitude avoids allowing light reflected from the bright Earth to enter the boresight. This light contains ultraviolet radiation which can polymerize outgassed monomers from the newly installed hardware. Any such polymers in the vicinity of HST's optics can degrade optical performance. All outgassing is completed in 12 days after which bright earthlight can enter the boresight and the BEA attitude restriction no longer applies.

Following HST positioning for release, the aperture door was opened. As planned, a five minute drift timer was initiated at MET 05/22:11:30, 90 seconds before release. This timer kept HST in drift mode for 3.5 minutes after release to allow the Orbiter to perform a clearance maneuver prior to active HST control. Software Sunpoint (SWSP) safemode, which uses RIG and CSS sensor data under Flight Computer control to maintain a sunpoint attitude, was initiated autonomously as the timer expired. Only small transients were observed and all rates were completely damped by MET 05/22:27. Transition back to normal mode occurred at 05/22:29. HST resumed full science operations shortly after the 12 day BEA period. The first post-mission Early Release Observation (ERO) image revealed previously unresolved detail in NGC 2392, "The Eskimo Nebula."

SUMMARY

A description of the mechanism leading to the loss of three of HST's six RIG gyros has been presented. The failure of these gyros led NASA to plan a Launch-On-Need HST servicing mission, SM3A. Loss of a fourth gyro caused HST to enter ZGSP forcing a change in SM3A rendezvous strategy. Baseline plans for rendezvous in HWSP/RMGA were established with ZGSP available as a contingency. Procedures were developed to monitor HWSP/RMGA performance and recover to ZGSP even if hardover RMGA failures were encountered. Rendezvous for the STS-103/SM3A mission conducted in December 1999 was successful leading to the changeout of all RSUs. Following deployment, HST resumed its astronomical science mission.

REFERENCES

1. Dougherty, H. J., Tompetrini, K., Levinthal, J., and Nurre, G., "Space Telescope Pointing Control System," Journal of Guidance, Control, and Dynamics, Vol. 5, No. 4, July – August 1982, pp. 403 - 409
2. Beals, G. A., Crum, R. C., Dougherty, H. J., Hegel, D. K., Kelley, J. L., and Rodden, J.J., "Hubble Space Telescope Precision Pointing Control System," Journal of Guidance, Control, and Dynamics, Vol. 11, No. 2, March – April 1988, pp. 119 - 123
3. H. Dougherty, C. Rodoni, J. Rodden, and K. Tompetrini, "Space Telescope Pointing Control," paper 83-365, AAS/AIAA Astrodynamics Specialist Conference, Lake Placid, NY, August 1983
4. K. Kumar, "A Microstructure Study of Corrosion in Ag-Cu Flex Leads," Journal of Electrochemical Society, Vol. 127, No. 4, April 1980, pp. 906 - 910
5. Markley, F. L. and Nelson, J. D., Zero-Gyro Safemode Controller for the Hubble Space Telescope," AIAA Journal of Guidance, Control, and Dynamics, Vol. 17, No. 4, July – August, 1994

On-Orbit Performance of Autonomous Star Trackers*

V. Airapetian, J. Sedlak, and J. Hashmall
Computer Sciences Corporation
Lanham-Seabrook, Maryland, USA 20706

Abstract

This paper presents the results of a performance study of the autonomous star trackers (ASTs) on the IMAGE and the EO-1 spacecraft. IMAGE is a spinning spacecraft without gyros or redundant precision attitude sensors, so the statistical properties of the AST are estimated simply by comparing the output observed quaternions with a rigid rotator model with constant angular momentum. The initial conditions are determined by a least-squares fit to minimize the AST residuals. An additional fit is used to remove the remaining systematic error and to obtain the inherent sensor noise. Gyro rate data are available for the EO-1 mission, so the AST noise statistics are obtained from the residuals after solving for an epoch attitude and gyro bias also using a least-squares method.

INTRODUCTION

Current and future missions need continuous and accurate three-axis attitude knowledge onboard to achieve better pointing and stability. Recently autonomous star trackers (AST) have been used onboard a number of spin-axis and three-axis stabilized spacecraft such as the New Millennium Program Deep Space-1 (DS-1), the Imager for Magnetopause-to-Aurora Global Exploration (IMAGE), and the Earth Observing-1 (EO-1) missions. The Lockheed Martin ATC AST-201R Autonomous Star Tracker (AST) is used as a primary sensor for attitude determination for these three missions. IMAGE and DS-1 are spin-stabilized spacecraft, and EO-1 is a three-axis stabilized spacecraft. Because autonomous star trackers represent a relatively new type of sensor for attitude determination and may be used in future missions such as SIRTF and MAP, it is useful to examine the statistical parameters reflecting their performance.

The AST is a "starfield-in, attitude-out" Charge Coupled Device (CCD)-based sensor. It outputs an attitude without requiring any a priori attitude knowledge. An AST-based attitude determination system can be viable without use of a digital Sun sensor (as on the EO-1 spacecraft) and even is able to perform in case of a gyro failure.

This paper presents a study of AST performance. It focuses on sensor noise statistics from the IMAGE and EO-1 missions. Some reported characteristics for DS-1 are included for comparison, but no independent DS-1 analysis was done for this paper. The IMAGE spacecraft does not carry gyros for inertial reference or other attitude sensors with enough accuracy to provide redundant

* This work was supported by the National Aeronautics and Space Administration (NASA) / Goddard Space Flight Center (GSFC), Greenbelt, MD, Contract GS-35F-4381G, Task Order no. S-43411-G.

NASA/GSFC, Guidance, Navigation and Control Center, *Flight Mechanics Symposium*, Greenbelt, MD USA, June 2001.

information for determining the AST accuracy directly, while rate data are provided from gyros on the EO-1 mission. The purpose of this paper is to determine errors in the autonomously derived attitude from the AST sensor based on multiple star measurements from analysis of IMAGE and EO-1 attitude data. The DS-1 spacecraft also carries an AST for attitude determination and, according to a recent report, performed very well during the first year of operation. It achieved accuracy (sensor noise) of about 2 arcsec in pitch and yaw and 40 arcsec in roll about the AST's optical axis (Ref. 1).

AST PERFORMANCE ON IMAGE SPACECRAFT

The IMAGE mission was launched at 20:34:43.929 UTC on March 25, 2000 from Vandenberg Air Force Base, California, aboard a Boeing Delta II 7326 launch vehicle. The IMAGE spacecraft is spin-stabilized about its Z-axis, with closed loop spin-rate control. It is an octagon-shaped spacecraft with 2.25 m diameter and 1.52 m height. Four thin wire antennas positioned 90 deg apart define the spacecraft X-Y plane. In a fully deployed configuration, the radial antennas are 250 m long. That makes the spacecraft's inertia tensor nearly diagonal.

The IMAGE attitude hardware consists of one Lockheed Martin AST, one Adcole Sun sensor, one three-axis magnetometer, a magnetic torque rod, and a nutation damper. The AST serves as the main attitude sensor providing a quaternion representing the three-axis orientation relative to the J2000 geocentric inertial frame (GCI) and angular rotation rates about each of its axes. The AST has maximum accuracy at spin rates up to 0.6 revolutions per minute (rpm). If the spin rate exceeds 1 rpm, then the star tracker is expected to lose track of stars.

The AST-to-body transformation matrix (the nominal alignment matrix) is given as (Ref. 2):

$$N_{AST-to-body} = \begin{bmatrix} 0.386710 & -0.908103 & -0.160635 \\ -0.922202 & -0.380781 & -0.067459 \\ 0.000094 & 0.174225 & -0.984706 \end{bmatrix} \tag{1}$$

Flight data from July 21, 2000 were obtained from the IMAGE website (Ref. 3) for this study. This is a time after the deployment of the radial antennas. The data file provides time and the corresponding four components of the quaternion in the sensor frame. To evaluate the performance of the AST, the observed quaternions are compared with a model. The standard deviation of the quaternion residuals is a measure of the sensor noise.

The observed AST quaternions were provided in the sensor frame. They are rotated into the body frame using the nominal alignment and an additional misalignment needed to improve the fit. The quaternions can be converted into a direction cosine matrix, $A(q)$, and then rotated so that

$$A(q_{obs}^{body}) = N_{AST-to-body} M_{123}(\varphi,\theta,\psi) A(q_{obs}^{AST}) \tag{2}$$

where the misalignment matrix, M, is expressed as a function of φ,θ,ψ, the three Euler angles for a 1-2-3 sequence.

To calculate the modeled quaternions, we assume a simple model of an axial-symmetric spinning spacecraft as described, e.g., in Ref. (4). The use of this model can be justified by the fact that in the fully deployed phase of flight, IMAGE's X- and Y-components of the moment of inertia are nearly equal to each other (within ~ 7 percent). Then, assuming torque-free rotation about the body Z-axis, the body frame attitude can be expressed analytically from the kinematic equations of motion:

$$q_{\mathrm{mod}}(t) = \begin{bmatrix} q_4' & q_3' & -q_2' & q_1' \\ -q_3' & q_4' & q_1' & q_2' \\ q_2' & -q_1' & q_4' & q_3' \\ -q_1' & -q_2' & -q_3' & q_4' \end{bmatrix} q_0 = q_0 q', \tag{3}$$

where

$$q_1' = u_1 \cos\alpha \sin\beta + u_2 \sin\alpha \sin\beta$$
$$q_2' = u_2 \cos\alpha \sin\beta - u_1 \sin\alpha \sin\beta$$
$$q_3' = u_3 \cos\alpha \sin\beta + \sin\alpha \cos\beta$$
$$q_4' = \cos\alpha \cos\beta - u_3 \sin\alpha \sin\beta$$

$$\alpha \equiv 0.5\,\omega_p t$$
$$\beta \equiv 0.5\,\omega_l t$$
$$u \equiv L_0 / |L_0| = [u_1, \quad u_2, \quad u_3]^T$$
$$L_0 \equiv [L_{01}, \quad L_{02}, \quad L_{03}]^T$$

In this solution, all the constants of the motion are expressed in terms of the initial values of the quaternions and L_0, the angular momentum vector in body principal coordinates.

In these equations, ω_p is the body nutation rate and is given as

$$\omega_p \equiv (1 - I_3 / I_T)\omega_3,$$

where ω_3 is the axial angular velocity of the spacecraft, I_3 is the axial moment of inertia, and I_T is the transverse moment of inertia. For the case of an axially symmetric spacecraft, I_T is equal to the X- and Y-components of moment of inertia. Moments of inertia for the IMAGE spacecraft (Ref. 2) are

$$I_x = 14710.5\ \mathrm{kg\text{-}m^2},$$
$$I_y = 14688.8\ \mathrm{kg\text{-}m^2},$$
$$I_z = 29041.9\ \mathrm{kg\text{-}m^2}.$$

Because the values of I_x and I_y are close, it is valid to approximate $I_x = I_y = I_T$ and equal to 14700 kg-m^2.

Also, ω_l is the inertial nutation rate, which can be expressed in terms of the body nutation rate as

$$\omega_l = \frac{I_3}{(I_T - I_3)\cos\theta}\omega_p$$

where $\cos\theta = L_3 / L$, and L_{01}, L_{02}, L_{03} are the X-, Y- and Z-components of the angular momentum vector, and L is its absolute value.

The residuals, dq, between the observed, q_{obs}^{body}, and modeled, q_{mod}, quaternions are

$$dq = (q_{obs}^{body})^* q_{\mathrm{mod}}, \tag{4}$$

where $(q_{obs}^{body})^*$ is the conjugate of the observed quaternion rotated into the body frame. A loss function, F, is constructed from the squares of the vector parts of the quaternion residuals:

$$F = \frac{1}{n}\sum_{i=1}^{n} d\vec{q}(t_i)\cdot d\vec{q}(t_i) \tag{5}$$

where n is the total number of data points.

We then solve for 6 independent parameters: 3 components of the quaternion and the components of the angular momentum vector, so as to minimize the loss function specified in Eq. (5). The MATLAB subroutine ***fminsearch*** was used to find the set of parameters minimizing the loss function. Initial guesses for the x and y components of the angular velocity were taken to be zero, while the initial axial component of the angular velocity was derived directly from the periodicity of the observed quaternions.

The resulting quaternion residuals are very small for the z-axis but were found to have mean value of roughly 0.2 deg on the X- and Y-axes. Note that the negative Z-axis is only 10 deg from the AST boresight. Thus, the minimization finds the rotation phase angle but is less successful with the spin direction. This indicates there may be a misalignment between the AST-attitude (after rotation to the body frame using nominal alignment) and the principal rotation axis. This may be due to a combination of sensor misalignment and an offset of the principal axis from the body Z-axis causing some coning.

The variance in the attitude residuals is caused by the true sensor noise plus model errors. Besides the sensor misalignment, there may be a deviation from uniform rotation due to environmental perturbations or the non-symmetry of the transverse axes. Therefore, the standard deviation of the residuals represents only an upper bound for the AST error.

The mean residuals can be used as a starting point for a search for a misalignment matrix. The best misalignment found has the Euler angles $\varphi = 0.2032$ deg and $\theta = -0.1050$ deg. The ψ rotation is kept at zero since it corresponds to an unobservable phase (absorbed by the value of the epoch quaternion found by the minimization).

After correcting for the misalignment, the mean attitude residuals are greatly reduced. The X-, Y- and Z-axis residuals from the vector part of the *dq* in Eq. (4) are shown in Fig. 1. The standard deviation on the X-axis (the transverse axis with the least systematic error) is 61 arcsec. For the Y- and Z-axes, the standard deviations are 67 and 80 arcsec, respectively. Because the AST boresight lies mostly in the X-Z plane of the spacecraft body system, the body Y-axis will be the least sensitive to the rotation errors.

It is clear from Fig. 1 that some systematic error remains in the residuals and that various "oscillation" frequencies are present in X-, Y- and Z-components of the quaternion residuals. To account for them, we have selected ten of the local maxima and applied a parabolic fit to each. By subtracting the parabola, the remaining residual errors are "flattened" and the inherent sensor noise can be determined by calculating the standard deviation.

The time span around each peak used for the parabolic fit was varied to be sure of the robustness of the noise estimate. Time spans ranging from 5 sec to 160 sec were tried. The means of the standard deviations from all ten peaks are shown in Fig. 2 as a function of the time span of the fit.

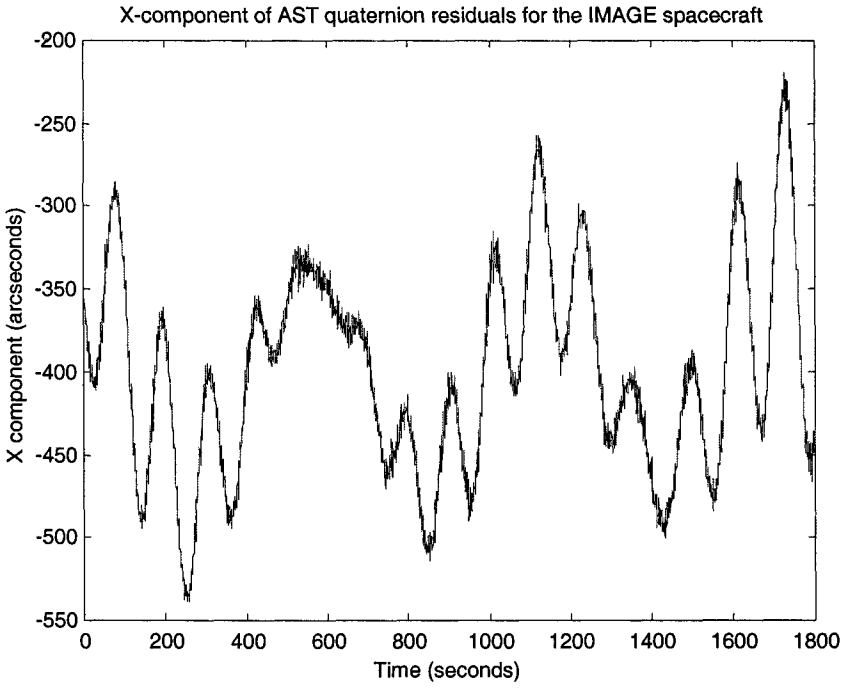


Figure 1a. Attitude residuals (X-component) comparing the AST and the torque-free rigid rotator model for the IMAGE spacecraft.

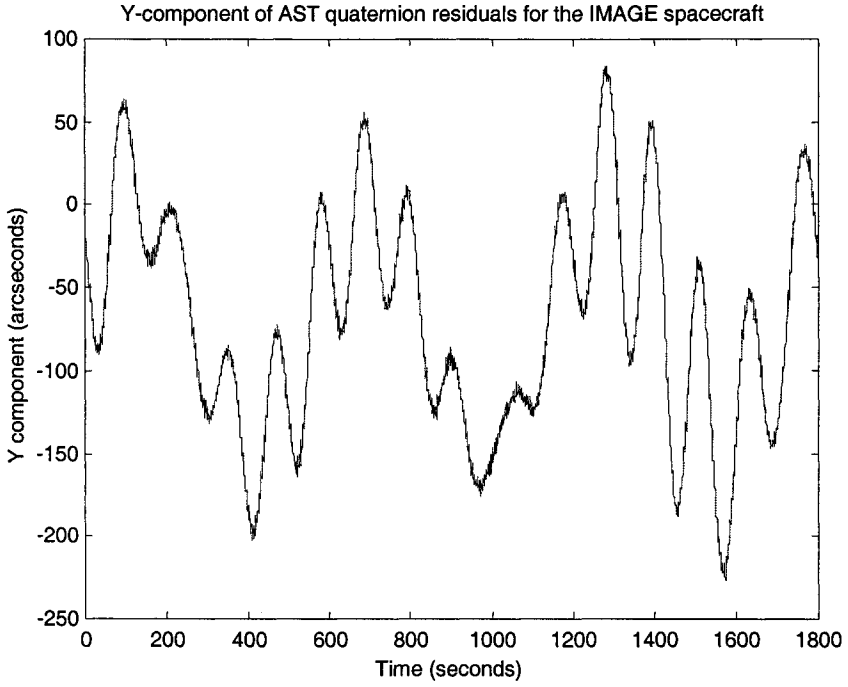


Figure 1b. Attitude residuals (Y-component) for the IMAGE spacecraft.

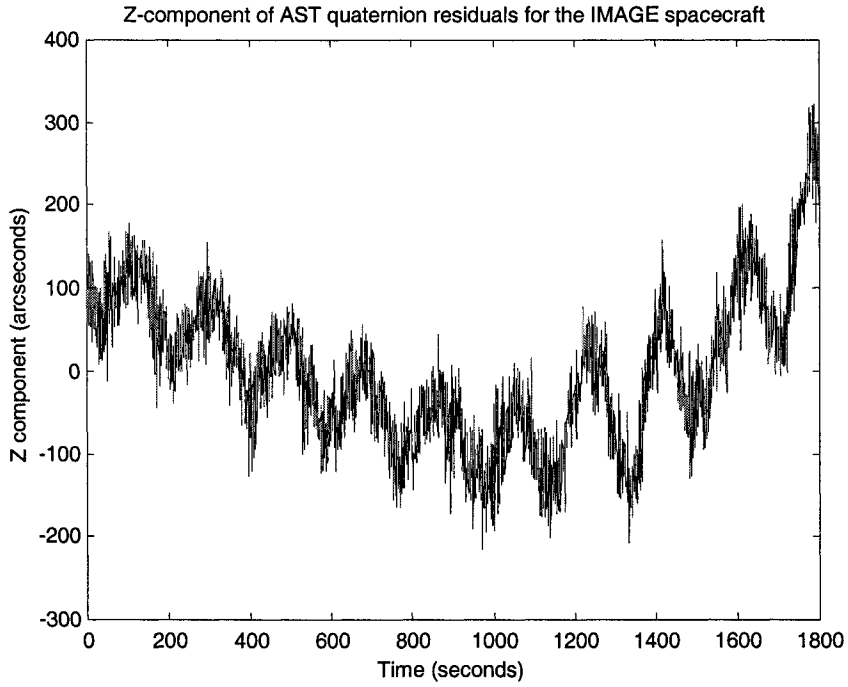


Figure 1c. Attitude residuals (Z-component) for the IMAGE spacecraft.

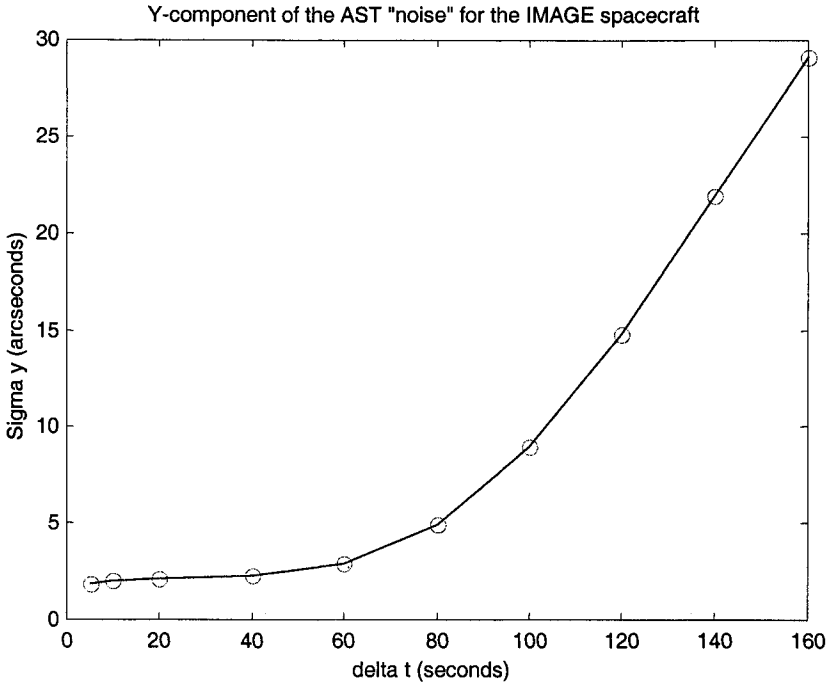


Figure 2. Y-component of the apparent AST noise averaged over ten peaks from Fig. 1 versus time span used for subtracting the systematic error.

Figure 2 reveals a characteristic plateau for parabolic fit time spans up to about one minute. The apparent sensor noise from this flat part of the graph is about 2.1 arcsec for the Y-axis. Similar analysis gives 4.7 arcsec and 22.9 arcsec for the sensor X- and Z-axes. These are our estimates of the inherent noise for the AST onboard IMAGE.

AST PERFORMANCE ON EO-1 SPACECRAFT

The EO-1 mission, managed by NASA's Goddard Space Flight Center, has three revolutionary land-imaging instruments collecting multispectral and hyperspectral scenes. EO-1 was launched on a Boeing Delta II rocket from Vandenberg Air Force Base on November 21, 2000.

The EO-1 is three-axis stabilized for nadir pointing. The pointing accuracy is 0.03 deg on all three axes and jitter is less than 5 arc-seconds. The EO-1 hardware consists of one Lockheed Martin Autonomous Star Tracker (AST 201R) which provides 3-axis attitude knowledge and gyros providing 3-axis body rates.

Analysis of the EO-1 AST performance was accomplished using attitude data (AST quaternions and gyro rates) for November 23, 2000 from 16:24:01 to 18:42:01 UT. The analysis was performed using the Matlab-based Attitude Determination System (ADS) (Ref. 5). The ADS estimates an epoch attitude and gyro bias vector from the AST and gyro input data using a batch least-squares method. This attitude is propagated using the bias-corrected gyro data to generate an attitude history.

Two intervals with nearly constant rates were selected. During such periods, the accuracy of the attitude solution is insensitive to errors in gyro calibration since gyro biases are included in the state vector. Such solutions were found and used as "truth" in determining the AST performance.

To determine the sensor performance, the attitude solutions were compared to the AST observations, and statistics on their differences (sensor residuals) were accumulated. The mean residuals are zero, as expected, because the AST is the only sensor used. The plots of the AST residuals expressed in the sensor frame are presented in the Fig. 3.

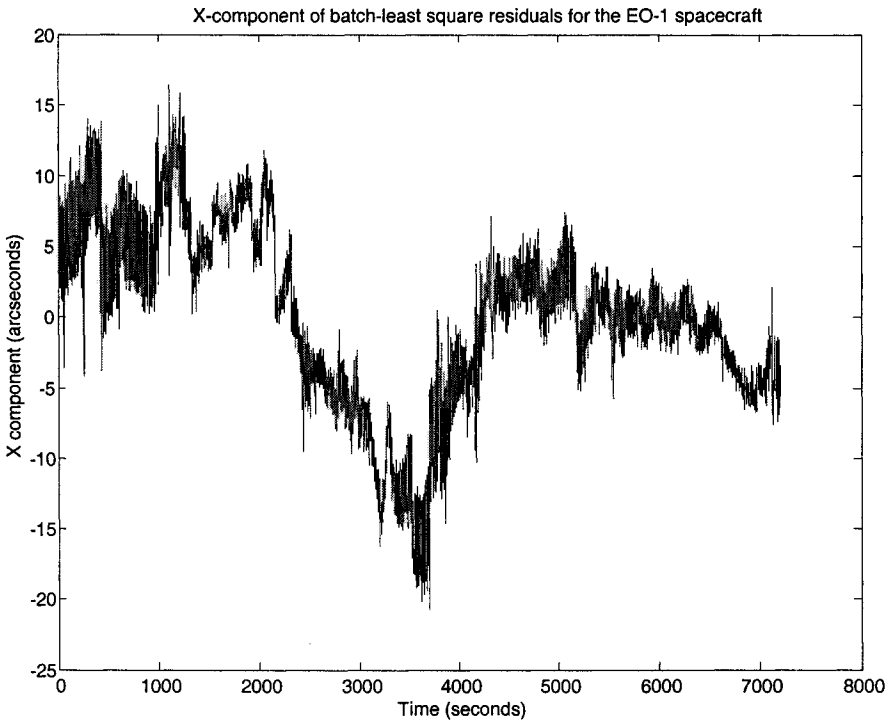


Figure 3a. Attitude residuals (X-component in the sensor frame) comparing AST and ground attitude solution for EO-1.

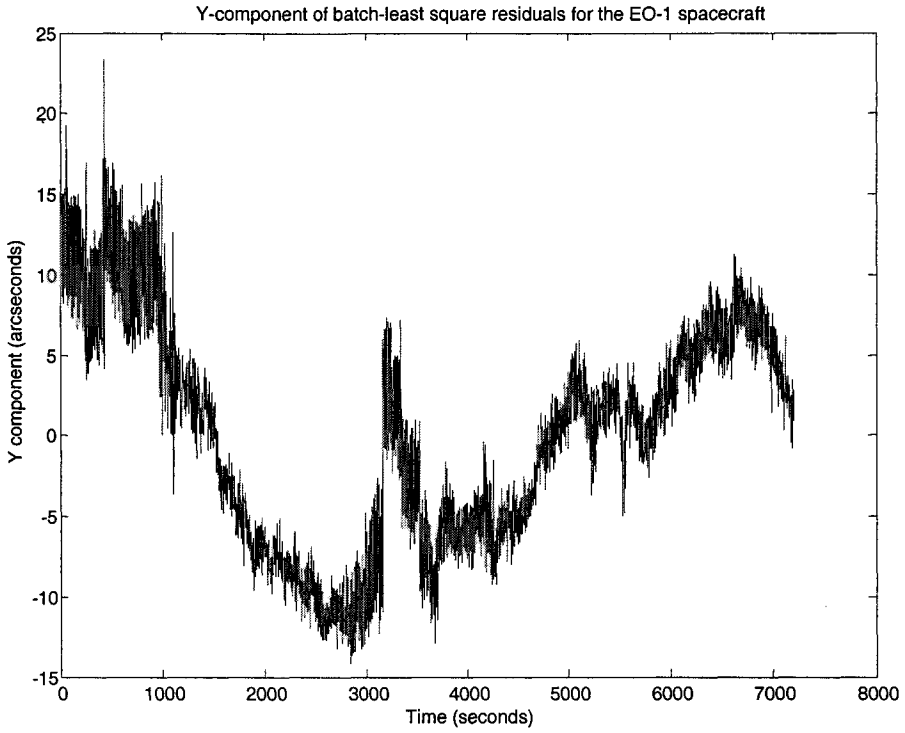


Figure 3b. Attitude residuals (Y-component in the sensor frame) for EO-1.

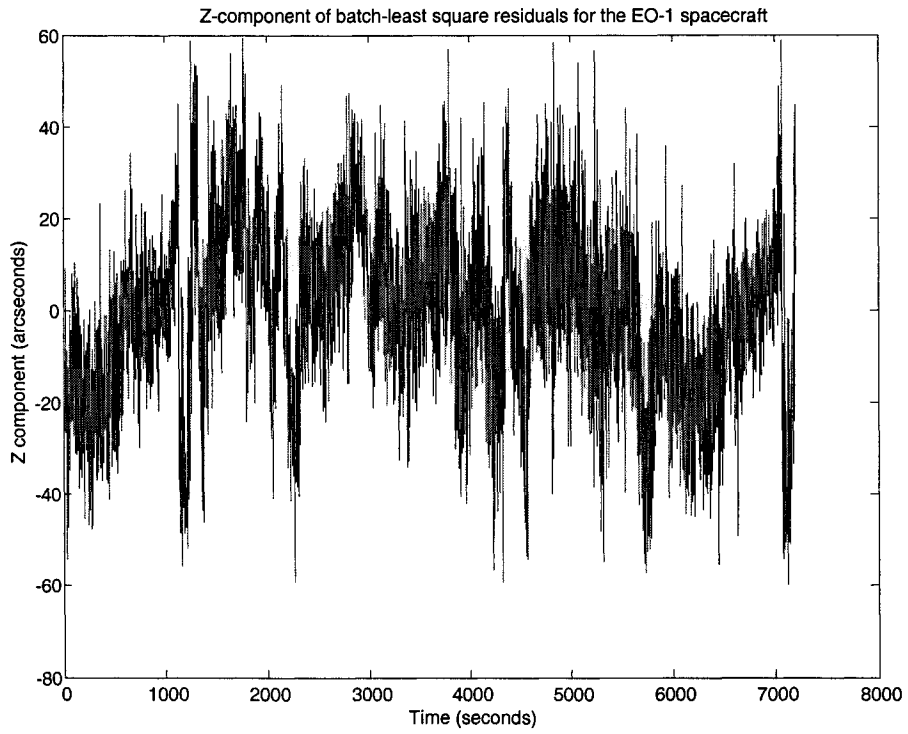


Figure 3c. Attitude residuals (Z-component in the sensor frame) for EO-1.

Low-frequency variations present in the residuals indicate systematic errors on time-scales of hundreds of seconds. These possibly are accounted for by variations in the stars crossing the sensor field of view at any given time. The number of stars, their magnitudes, and the distribution across the field of view all might affect the attitude estimate. The actual inherent AST sensor noise is probably better represented by the standard deviations of the higher frequency variations seen in Fig. 3.

The standard deviations of the AST residuals for the entire 2-hour EO-1 data set are 6.2, 6.9, and 18.1 arcsec for the X-, Y-, and Z-components, respectively. The root-sum-square of the transverse X- and Y-components is 2–3 times larger than found for the ASTs on IMAGE and DS-1, as discussed above. An improved estimate of the inherent AST noise can be found by first removing the effects of the systematic errors; however, it should be emphasized that actual spacecraft control is likely to depend on the total error and not just on this somewhat subjectively determined inherent error.

To estimate the inherent AST noise, a number, N, of relatively flat intervals were selected for each AST component (7 for X, 5 for Y, and 6 for Z). The mean residual was subtracted from each interval and the variances were computed. With the mean values removed, the total variance is

$$\text{var} = \frac{\sum_{i=1}^{N} N_i \, \text{var}(i)}{\sum_{i=1}^{n} N_i}, \tag{6}$$

where N_i is the number of observations and var(i) is the variance for the i-th interval. The square root of the total variance is our estimate of the inherent AST noise. Figure 4 shows a sample of typical "flat" intervals for each of the components. The selected intervals cover roughly 50 percent of the entire 2-hour data set for the X- and Y-axes and 30 percent for the Z-axis. The actual time spans range from 500 to 1000 sec. The noise standard deviations range from 1.0 to 2.6 arcsec for the transverse components and from 8.5 to 13.8 arcsec for the Z-axis. The total errors from Eq. (6) are found to be 1.7, 1.7, and 11.6 arcsec for the X-, Y-, and Z-components, respectively. These errors are consistent with the AST performance found for IMAGE and reported for DS-1.

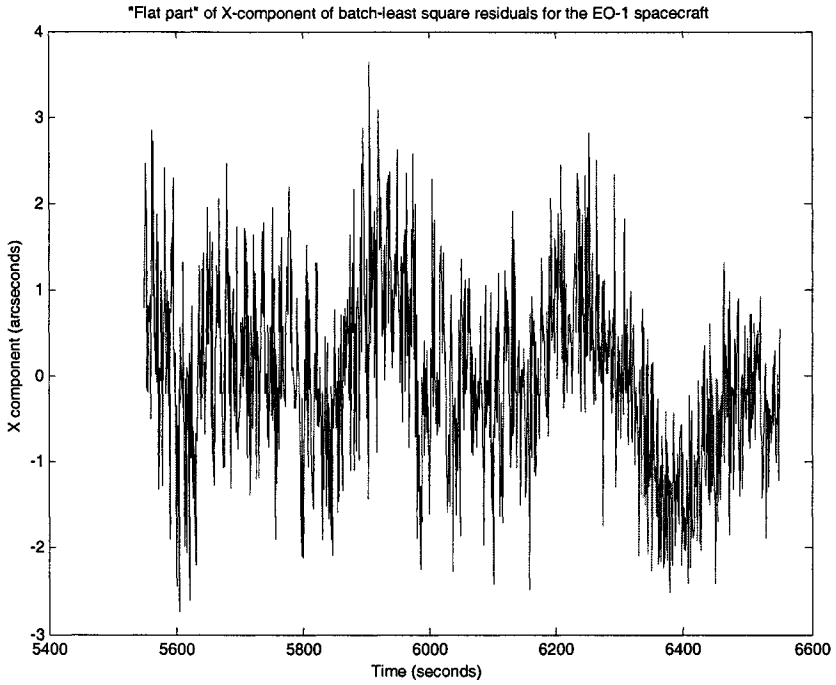


Figure 4a. Attitude residuals (X-component in the sensor frame) comparing AST and ground attitude solution for EO-1 showing a selected interval where the AST residuals are relatively flat.

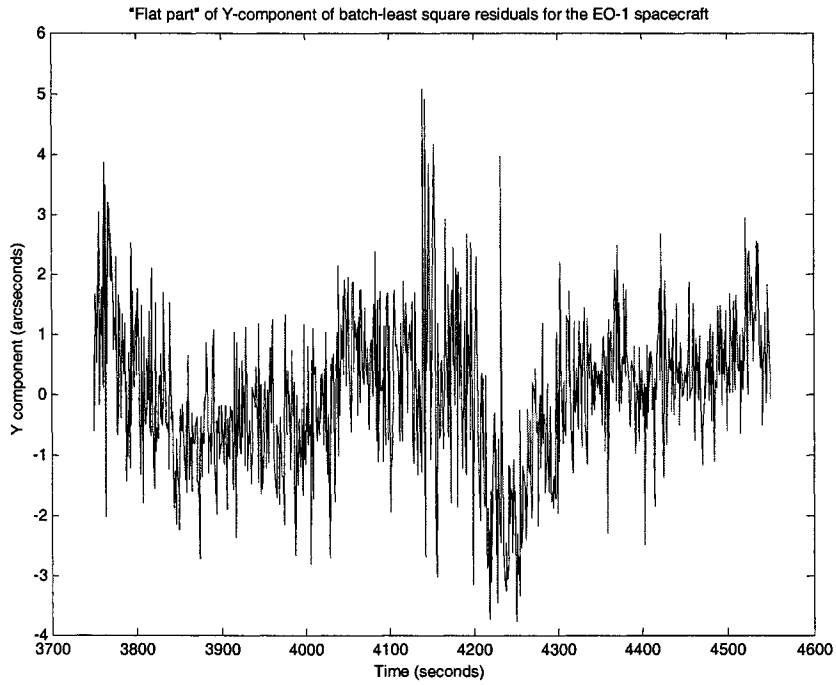


Figure 4b. Attitude residuals (Y-component in the sensor frame) for EO-1 showing a selected interval where the AST residuals are relatively flat.

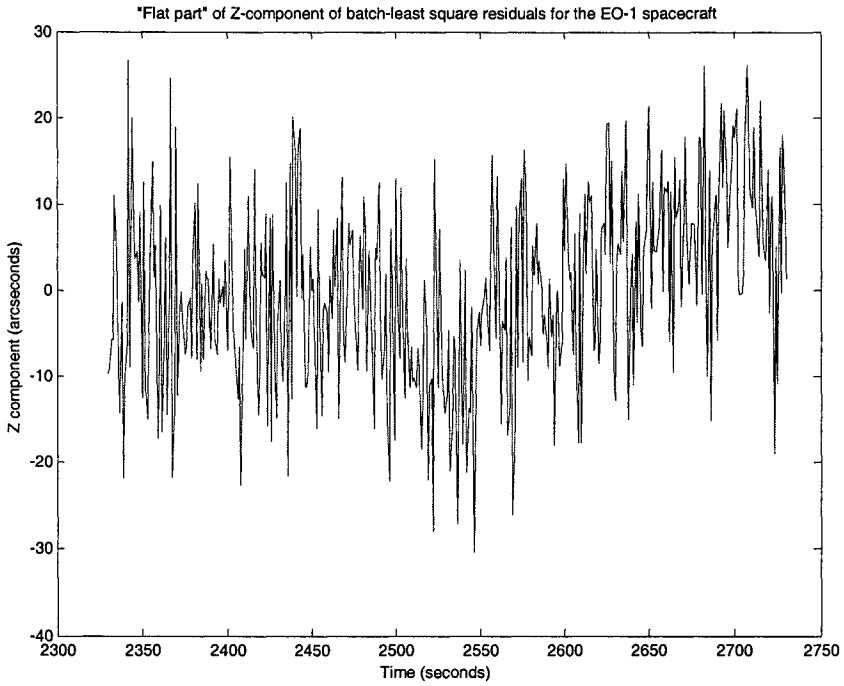


Figure 4c. Attitude residuals (Z-component in the sensor frame) for EO-1 showing a selected interval where the AST residuals are relatively flat.

CONCLUSIONS

In this study we have used in-flight data from the IMAGE and EO-1 missions to estimate the uncertainty of the AST. The noise of the AST onboard IMAGE has been estimated by comparing quaternions from the AST with reference quaternions calculated with a torque-free model for an axial-symmetric spacecraft. The initial quaternion and angular momentum are determined by minimizing a loss function as a function of these six solve-for parameters. A misalignment matrix was estimated to improve the match between observed and modeled attitudes. The remaining systematic errors were removed by fitting and subtracting parabolas from selected subsets of the data. For EO-1, both AST quaternions and gyro rate data are available. The AST noise statistics were obtained from the AST residuals after solving for an epoch attitude and gyro bias. Again, it was necessary to remove systematic errors by processing selected subsets of the data.

A summary of estimated noise parameters for ASTs onboard the IMAGE, EO-1, and DS-1 spacecraft is given in Tables 1 and 2. Table 1 shows the AST noise as determined from residual statistics over moderately long time periods (30 min for IMAGE and 140 min for EO-1). Table 2 shows noise for selected time intervals with systematic variations removed. However, these systematics may be a real source of onboard error, as seems to be indicated by the gyro-propagated attitude solution for EO-1, or they may be due to our estimation method (in particular, the symmetric rigid rotator model for IMAGE). In either case, Table 1 shows upper bounds for the AST noise. Table 2 indicates the apparent inherent noise after the somewhat

subjective subtraction of the systematic errors. The inherent noise parameters estimated this way (and the reported noise for DS-1) are all consistent with each other.

Table 1. AST Noise Including Systematic Errors.

	IMAGE	EO-1
1-σ AST X residuals (arcsec)	61	6.2
1-σ AST Y residuals (arcsec)	67	6.9
1-σ AST Z residuals (arcsec)	88	18.1
Attitude accuracy requirement (arcsec)	360	108

Table 2. AST Noise With Systematic Errors Excluded. AST noise as reported for DS-1 [Ref. 1] is presented for comparison.

	IMAGE	EO-1	DS-1
1-σ AST X residuals (arcsec)	2.1	1.7	2
1-σ AST Y residuals (arcsec)	4.7	1.7	2
1-σ AST Z residuals (arcsec)	22.9	11.6	40

REFERENCES

1. S. S. Lisman, NASA/Jet Propulsion Laboratory, private communication, January 20, 2001

2. M. Challa, et al., *IMAGE Attitude Support Post-Launch Report*, prepared for GSFC by CSC, June 2000

3. The IMAGE homepage is at http://image.gsfc.nasa.gov/

4. J. R. Wertz (ed.), *Spacecraft Attitude Determination and Control*, D. Reidel Publishing Company, Dordrecht, The Netherlands, 1978, p. 529

5. J. Landis, et al., *Multimission Three-Axis Stabilized Spacecraft (MTASS) Flight Dynamics Support System Functional Specifications, Rev. 1*, Computer Sciences Corporation, 554-FDD-91/070R1UD0, CSC/TR-91/6071R1UD0, September 1995

HORIZON SCANNER TRIGGERING HEIGHT ANALYSIS FOR ORBVIEW-2

Frederick S. Patt
Stephen Bilanow
SAIC General Sciences Corporation

ABSTRACT

The OrbView-2 spacecraft (OV-2) was launched in August 1997 from a Pegasus vehicle and maneuvered to a 705-kilometer altitude, Sun-synchronous orbit with a local noon descending node. The spacecraft carries the Sea-viewing Wide Field-of-view Sensor (SeaWiFS). The science data processing is performed by the SeaWiFS Project Office at NASA/GSFC. The definitive attitude determination for image geolocation is performed as part of Level 1 product generation by post-processing the spacecraft attitude sensor data. The geolocation accuracy requirement is 1 pixel (2 sigma), which corresponds to 1.12 km at nadir.

OV-2 is equipped with two horizon scanners for redundancy. In November 1998 one string of the attitude control system, including one of the scanners, was powered off to save electrical power and sensor lifetime. In order to meet the 1-pixel requirement with the single scanner, an accurate model of the horizon scanner triggering height was needed. Analysis of the triggering height was performed using data from both scanners prior to November 1998. This analysis showed significant variations in triggering height from the basic oblate Earth model. The height had a different oblateness than the solid Earth, and also showed latitudinal variations which had a seasonal dependence.

A model was developed to incorporate the observed variations in scanner triggering height. This model has been implemented in the SeaWiFS Project processing software, and has been instrumental in meeting the geolocation accuracy requirement for all data processed using the single scanner.

1. INTRODUCTION

The Sea-Viewing Wide Field-of-view Sensor (SeaWiFS) is an ocean color sensor which was launched on the OrbView-2 spacecraft (OV-2) in August 1997. The spacecraft (called SeaStar prior to launch) was built and launched by Orbital Sciences Corporation (OSC), who were also responsible for the orbit raising, initial activation and check-out of the spacecraft. After OV-2 was declared operational and routine SeaWiFS data collection was started, spacecraft operations responsibility was assumed by OrbImage, a subsidiary of OSC. The SeaWiFS Instrument collects science data both via an onboard recorder, which is dumped twice daily, and through direct broadcast of the data to ground stations.

The SeaWiFS Project Office at NASA Goddard Space Flight Center (GSFC) obtains the SeaWiFS science data in a unique arrangement called a "data buy". In this arrangement, the SeaWiFS Project schedules the SeaWiFS instrument operation and onboard data recording, and also collects direct broadcast data from authorized research stations. NASA is responsible for all research and educational use of the data, while OrbImage retains all commercial data rights.

The SeaWiFS Project processes all of the data and generates data products for use by the Ocean Science community. The data are processed from Level 0 to Levels 1A, 2 and 3, using software developed and operated by the SeaWiFS Project. The data products are transmitted to the GSFC Distributed Active Archive Center (DAAC) for distributed to authorized users. Details of the SeaWiFS data processing, product definitions and project activities can be found on the SeaWiFS home page (http://seawifs.gsfc.nasa.gov).

An integral part of the Level 0-to-1A data processing is the determination of the navigation information, e.g., the spacecraft definitive orbit and attitude, to support the instrument data geolocation. The spacecraft attitude is determined by post-processing of the telemetry from the onboard attitude sensors. The accuracy requirement for geolocation is 1 pixel (~1.6 mrad, or 1.12 km at nadir), 2 sigma.

The assessment of navigation accuracy is performed using the method of island target matching which was developed prior to launch (reference 1) and refined using flight data (reference 2). In order to meet the accuracy requirements, it has been necessary to utilize the navigation assessment results to develop numerous refinements to the navigation algorithms, including the horizon height model described in this paper.

The remainder of this section describes the SeaWiFS data collection, the OV-2 attitude control system (ACS) and the SeaWiFS Project attitude determination algorithms. Section 2 describes the initial development of the horizon angle model, and Section 3 discusses the use of the island target matching to assess and refine the model. The spacecraft orbit and sensor geometric characteristics are summarized in Table 1.

Table 1 - Spacecraft Orbit and SeaWiFS Viewing Geometry

Orbit Characteristics:	
Type	Polar, near-circular, Sun-synchronous
Altitude (nominal)	705 km
Inclination	98.2 degrees
Period	99 minutes
Equator crossing	Local noon (descending)
SeaWiFS Geometry:	
Instantaneous FOV	1.5897 mrad (0.091 degree)
Ground IFOV	1.12 km (at nadir)
Scan half-width	58.3 degrees
Pixels along scan	1285
Scan period	0.167 second
Sensor tilt	+/- 20 degrees
Sensor coverage	~40 minutes per orbit during the orbit day

1.1 SeaWiFS Data Collection

The SeaWiFS instrument is commanded to collect data for approximately 40 minutes each orbit during the sunlit part of the orbit, in order to observe locations where the solar zenith angle is 75 degrees or less. The instrument is also tilted to avoid Sun glint from the ocean surface, 20 degrees aft prior to the mid-point of orbit day and 20 degrees forward afterward.

The instrument can produce data in two formats: full-resolution, local-area-coverage (LAC data, and 4x4 subsampled, global area coverage (GAC data). The data are collected by two paths. An onboard recorder records data during scheduled instrument operation, and is dumped to a ground station twice daily. Due to storage limitations, the majority of recorded data are in GAC format, with a limited amount of LAC storage available for calibration. Full-resolution data are also direct-broadcast continuously in high-resolution picture transmission (HRPT) format during scheduled instrument operations, and can be received during spacecraft overpasses by stations with SeaWiFS direct broadcast licenses.

The SeaWiFS Project collects and processes both the onboard recorded data and direct-broadcast data from research HRPT stations. A SeaWiFS scene is a scheduled orbit of GAC data (40 minutes), a scheduled LAC segment or a single HRPT station overpass.

1.2 OrbView-2 Attitude Control System

The OV-2 ACS was designed by OSC (reference 3). The attitude sensor complement consists of three two-axis digital Sun sensors (DSSs), two infrared horizon scanners (HSs) and two three-axis magnetometers. For purposes of definitive attitude determination, the sensors of interest are the DSSs and HSs.

The attitude control hardware consists of redundant three-axis magnetic torque rod assemblies and pitch-axis momentum wheels. The torque rods are the sole source of roll and yaw control. The momentum wheels are used in concert with the torque rods for pitch control, and also to maintain a pitch momentum bias for attitude stabilization.

The spacecraft coordinate system is defined with the X axis downward, the Y axis opposite the orbit velocity direction and the Z axis toward the orbit normal. Thus, at the nominal spacecraft attitude, the spacecraft X-Y plane corresponds to the orbit plane and the spacecraft Y-Z plane corresponds to the local tangent plane.

The geometry of the three DSSs is illustrated in Figure 1. Each DSS has a field-of-view (FOV) of +/-64 degrees in each axis. All three DSSs are mounted with their boresights in the spacecraft X-Y plane and with one sensor axis parallel to the spacecraft Z axis. DSS-A is canted approximately 64 degrees in the direction of the spacecraft velocity, DSS-B is canted in the opposite direction and DSS-C is oriented toward the zenith. Each sensor provides two digital outputs (one for each axis) with a resolution of 0.0002 (about .011 degree at the boresight).

The HS geometry is illustrated in Figure 2. The HS design uses a conical scan with a half-cone angle of 45 degrees. HS-A is oriented with its rotation axis opposite the orbit normal, and HS-B is toward the orbit normal; both HSs have their rotation axes canted 5 degrees downward from the spacecraft plane corresponding to the local tangent plane. Each sensor provides time-tagged Earth width and phase measurements with a resolution of 0.0055 degree.

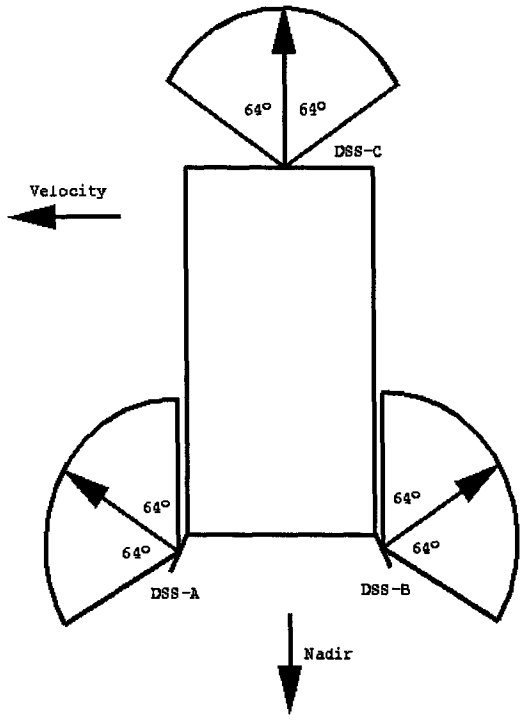


Figure 1. Digital Sun Sensor Geometry

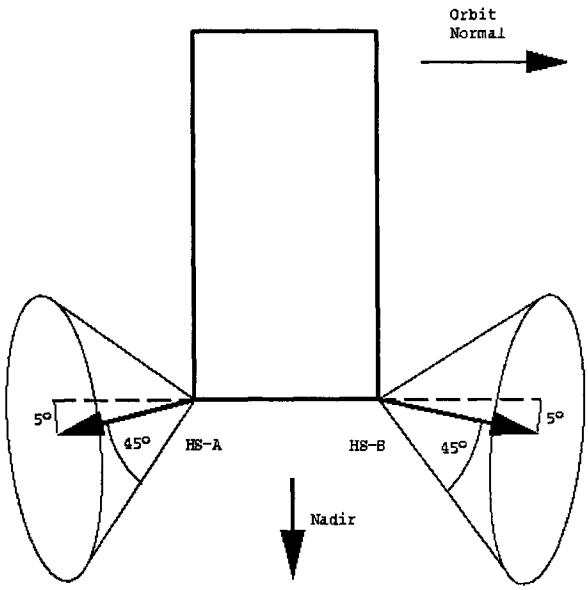


Figure 2. Horizon Scanner Geometry

The choice of orbit (local noon equator crossing) causes the Sun and nadir vectors to be nearly anti-parallel at the mid-point of the orbit day (the sub-solar point). The maximum angle varies seasonally, but remains close to 180 degrees, and twice yearly the two vectors become exactly anti-parallel. Thus the attitude geometry is always poor at the sub-solar point, specifically for the yaw angle.

The ACS is separated into two strings, with the primary string including DSSs A and B, HS-A, one magnetometer and one set of torquers, and the remaining components in the secondary string. The spacecraft was operated from launch

until November 1998 with both strings powered on, and therefore all sensors active. The secondary string, including DSS-C and HS-B, was powered off in mid-November 1998 to conserve lifetime and electrical power.

1.3 SeaWiFS Project Attitude Determination Algorithms

The following is a summary of the attitude determination algorithms used in the SeaWiFS Project science data processing software. A detailed description is being published as a NASA technical report. As previously stated, this processing is performed as part of the Level 0-to-1A conversion.

The determination of the spacecraft orbit is a precursor to the attitude determination. The orbit determination is based on the output of a GPS receiver carried on the spacecraft. The GPS orbit vectors are fitted to an orbit model which includes a high-order geopotential field model, third-body (Sun and Moon) attraction and atmospheric drag.

The attitude determination processing is performed individually for each SeaWiFS scene. The first step is to unpack the attitude sensor telemetry from the spacecraft data stream and convert the samples to engineering units. The data are validated using absolute limits and sample-to-sample consistency checks. They are then smoothed and interpolated to the SeaWiFS scan times using overlapping cubic polynomials. The measurements are combined to produce a measured Sun and nadir vector in the spacecraft frame for each scan line. For the HS data, the Earth width and phase angles are first converted to in-crossing and out-crossing angles, which are calibrated using vendor-provided tables and then used to compute horizon crossing vectors in the spacecraft frame. With data from both HSs a least-squares determination of the nadir is performed, while for a single scanner the nadir algorithm is necessarily deterministic.

The attitude determination is performed using a Kalman smoother, i.e., a two-pass (forward and reverse) Kalman filter. This algorithm reduces the discontinuities at transition points (i.e., start/end of DSS coverage), stabilizes the yaw angle through the sub-solar point and provides continuity of the attitude determination during periods of missing or invalid attitude sensor data (in particular, the Sun coverage gap at the sub-solar point following the DSS-C power-off). The state consists of the attitude (roll, pitch and yaw) angles, and the state propagation uses a simple dynamics model which assumes an inertial pitch rotation axis (due to the due to the stability provided by the momentum wheels) and a fixed pitch angle. The attitude is propagated to each scan line time and updated using the preprocessed DSS and HS data, with the measurement noise covariances selected to place more weight on the DSSs. The final attitude angles are computed as a weighted average of the two passes, with the weighting determined by the covariance from each pass.

The performance of the attitude determination depends upon the overall characterization of the sensors, including the biases and alignments (references 4 and 5). For the ACS single-string mode, the accuracy of the attitude depends upon individual sensors. In particular, the roll angle is based almost entirely on the HS data, and is directly affected by the model used for the scanner triggering height. The yaw angle is determined from a combination of HS and DSS data, and is also quite sensitive to the horizon model.

2. DEVELOPMENT OF THE HORIZON ANGLE MODEL

The nadir vector is computed from the horizon scanner crossing vectors using a model of the nadir-to-horizon angle. A model is given in reference 6, eqns. 4-22 and 4-23, which are used to determine the subsatellite radius, R, and the horizon angle, ρ, for a given latitude, geocentric distance and azimuth (psi). This model is based on an ellipsoidal Earth and is parameterized by the Earth equatorial radius, R_e, and the flattening factor, f, defined as

$$f = (R_e - R_p)/R_e \tag{1}$$

where R_p is the Earth polar radius (reference 6, p. 99). The values for R_e and f are 6378.137 km and 1/298.257, respectively, based on the WGS84 standard. The value of ρ given by this model is for the Earth surface ellipsoid. Since the HS actually triggers on the CO_2 height, not the Earth's surface, an additional bias is added to the computed value of ρ to get the final horizon angle. The value of this bias was set at 0.74 degree, based on a typical CO_2 height of 40 km and the OV-2 orbit radius of 7083 km. This model was implemented in the navigation software for OV-2 at launch.

In ACS dual-string model the use of two HSs allows systematic errors in triggering height to be averaged out; however for single-scanner navigation any height errors are directly manifested as corresponding attitude (roll) errors. Ideally we would want to characterize the horizon triggering height at the individual in-crossings and out-crossings; however, the spacecraft attitude motion is typically 1 degree, which tends to obscure the systematic HS errors. Instead, we analyzed the average triggering height for all four horizon crossings from the two scanners. The symmetric mounting of the HSs causes the two in-crossings to be almost exactly 180 degrees apart, and likewise for the out-crossings. Thus the effects of attitude motion are cancelled by averaging the horizon angles determined from the four crossings. This average was then compared with the average of the model angles.

We performed this comparison for HS data collected during the dual-string period (September 1997 to mid-November 1998). A typical comparison was performed using a single spacecraft recorder dump, consisting of about 12 hours or 7 orbits of data. An example of these results with the at-launch horizon angle model, for data from March 21, is given in figure 3a, which shows the average of the measured (scattered points) and model horizon angles vs. latitude for all of the HS samples over several orbits. The plot clearly shows that the measured angles decrease more rapidly than the model at the extremes of latitude. The differences for this same data set are shown in Figure 3b.

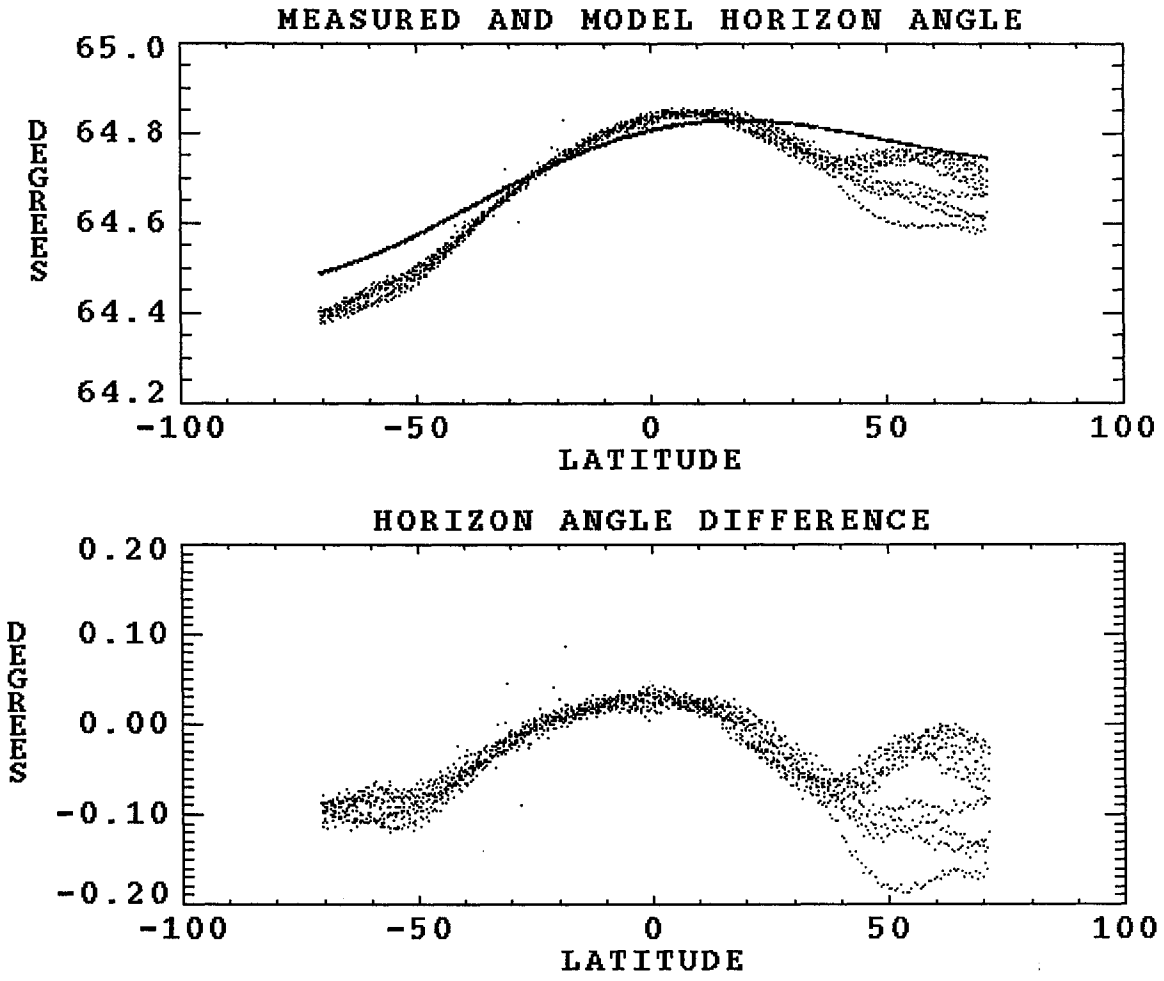


Figure 3. Horizon Angle Comparison with at-launch model for March 21, 1998; (a) Measured and model angles; (b) Difference between measured and model angles.

The systematic decrease in measured horizon angles towards the poles indicates that the shape of the CO_2 layer is more oblate than the Earth's surface, i.e., the difference between the equatorial and polar radii is larger. This type of result, which results from the CO_2 height being lower at the poles, has been documented previously (e.g., reference 6, pp. 90-98). It can be represented by an increase in the flattening factor, f. We estimated empirically that the apparent value of f for the measured horizon angles is about 1/200, which corresponds to about 10 km of additional flattening at the poles, over the 21 km in the standard Earth surface ellipsoid. The March 1998 angle comparison with the revised flattening is given in Figures 4a and 4b, which shows improved overall agreement. A small bias is present, indicating some combination of higher average triggering height and HS misalignment. The larger scatter to the north represents orbit-to-orbit variations in triggering height, indicating longitudinal dependence with polar asymmetric variability. This variability has been shown in previous studies (reference 6) to be large in the winter hemisphere, when the stratospheric temperatures show the greatest variability.

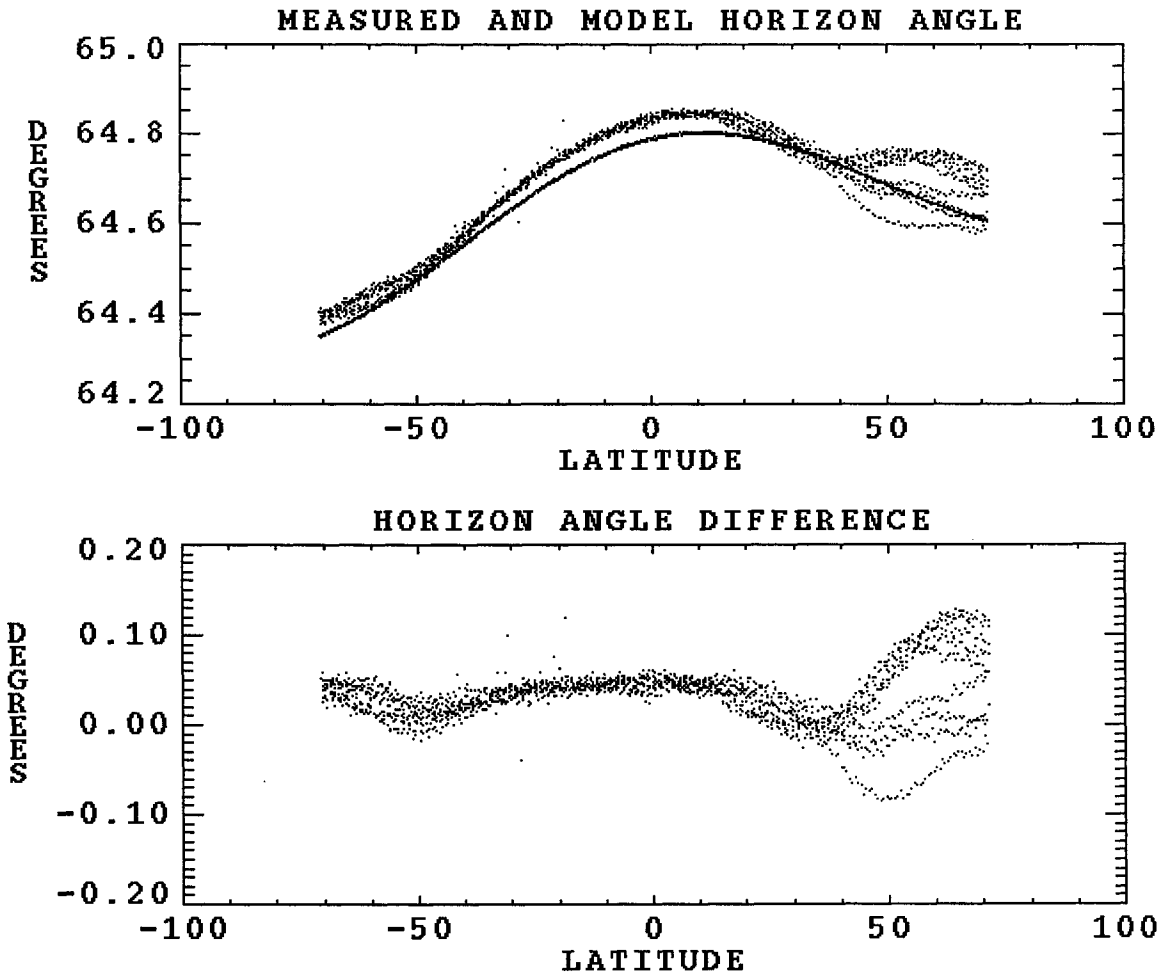


Figure 4. Horizon Angle Comparison with revised flattening factor for March 21, 1998; (a) Measured and model angles; (b) Difference between measured and model angles.

During other seasons the angle differences exhibited systematic variations. This can be seen in the differences for December 9, 1997 and June 12, 1998 (figures 5a and 5b). The December differences show a negative trend with latitude, while for June the trend is positive. The trend is not linear, but rather suggests a sinusoidal dependence. Again, this type of behavior has been observed previously (e.g., reference 6, p. 99).

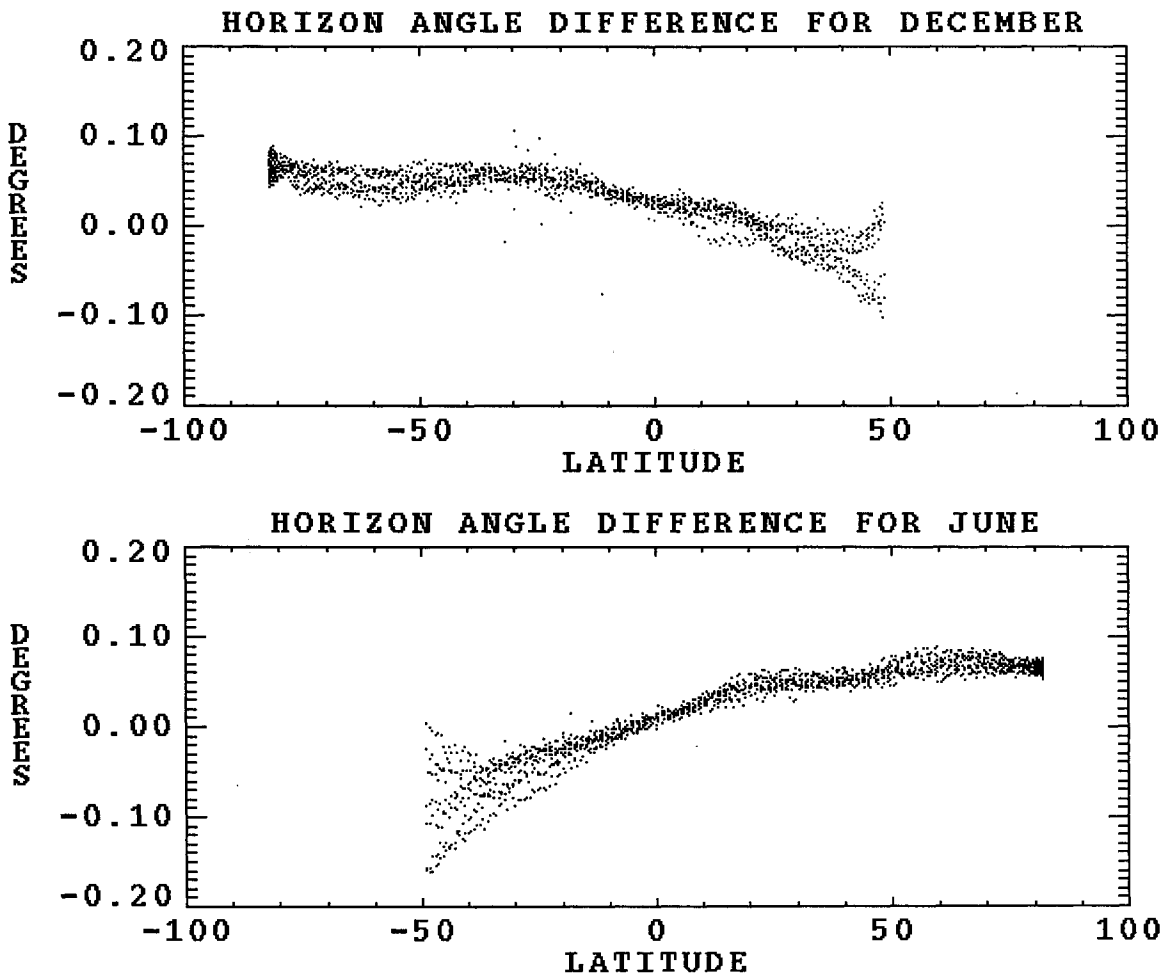


Figure 5. Horizon angle differences showing seasonal variations; (a) December 9, 1997; (b) June 12, 1998.

The observed variations can be modeled by shifting the ellipsoid north or south along the Earth's axis. However, this type of shift is not accommodated by the standard model. We re-derived the horizon angle computation (see appendix) to add the ellipsoid shift, based on previous work by one of the authors (reference 8). The resulting quadratic equation relating the horizon angle, ρ, and the horizon crossing azimuth, ψ, is:

$$(H^2\cos^2\psi - 4J\sin^2\psi - 4BJ\cos^2\psi) + (4JF\cos\psi - 2HI\cos\psi)\cot\rho + (I^2 - 4JC)\cot^2\rho = 0 \qquad (2)$$

where

$$B = \sin^2\lambda + \cos^2\lambda/(1-f)^2 \qquad (3a)$$

$$C = \cos^2\lambda + \sin^2\lambda/(1-f)^2 \qquad (3b)$$

$$F = -2\sin\lambda\cos\lambda + 2\sin\lambda\cos\lambda/(1-f)^2 \quad (3c)$$

$$H = -2P\sin\lambda\cos\lambda + 2P\sin\lambda\cos\lambda/(1-f)^2 - 2\delta z\cos\lambda/(1-f)^2 \quad (3d)$$

$$I = 2P\cos^2\lambda + 2P\sin^2\lambda/(1-f)^2 - 2\delta z\sin\lambda/(1-f)^2 \quad (3e)$$

$$J = P^2\cos^2\lambda + (P\sin\lambda - \delta z)^2/(1-f)^2 - (R_e + \delta r)^2 \quad (3f)$$

where λ is the geocentric orbit latitude, P is the orbit vector magnitude, δz is the ellipsoid shift along the polar axis and δr is a bias in the Earth radius.

We analyzed eleven HS data sets from October 1997 through October 1998 to estimate the value of δz which would minimize the systematic horizon angle differences, along with the value of δr to remove any residual bias. The sensitivity to δz and δr was determined from numerical derivatives using the above equations. The value of δz was found to correlate strongly with the sine of the solar declination (figure 6a), while δr also showed a small correlation (figure 6b).

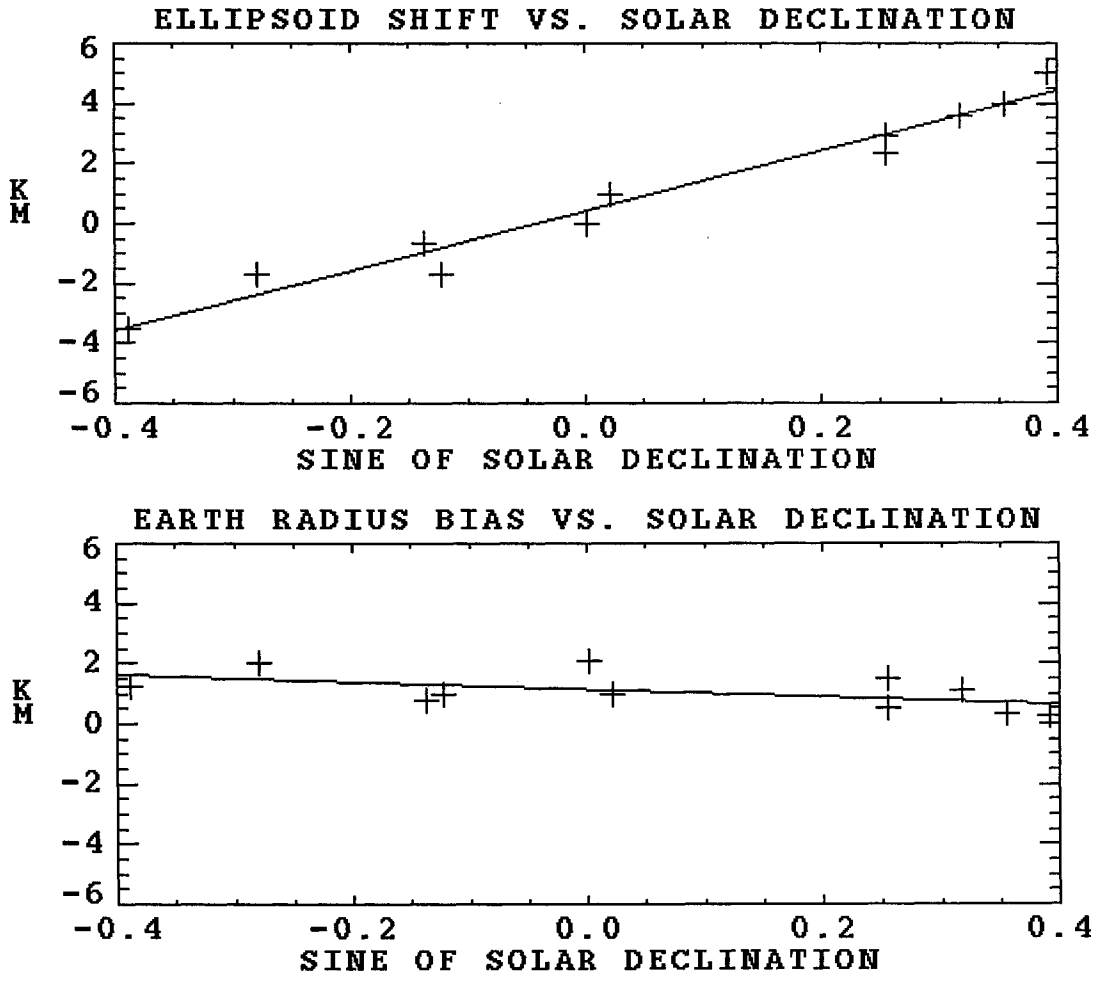


Figure 6. Estimated ellipsoid adjustments vs. solar declination; (a) Ellipsoid shift; (b) Earth radius bias

We computed linear fits of the ellipsoid shift and Earth radius bias to the sine of the solar declination; the results were

$$\delta z = 10 \sin(d) + .4 \quad (4)$$

$$\delta r = -1.2 \sin(d) \quad (5)$$

where d is the solar declination, and the average Earth radius bias is removed by adjustments to the HS alignments.

This model succeeded in removing most of the seasonal horizon angle variations, as shown in the revised analysis of horizon angle differences for December and June (figures 7a and 7b). The model was implemented in the operational navigation code prior to the start of single-string ACS operation and used for all subsequent SeaWiFS data processing.

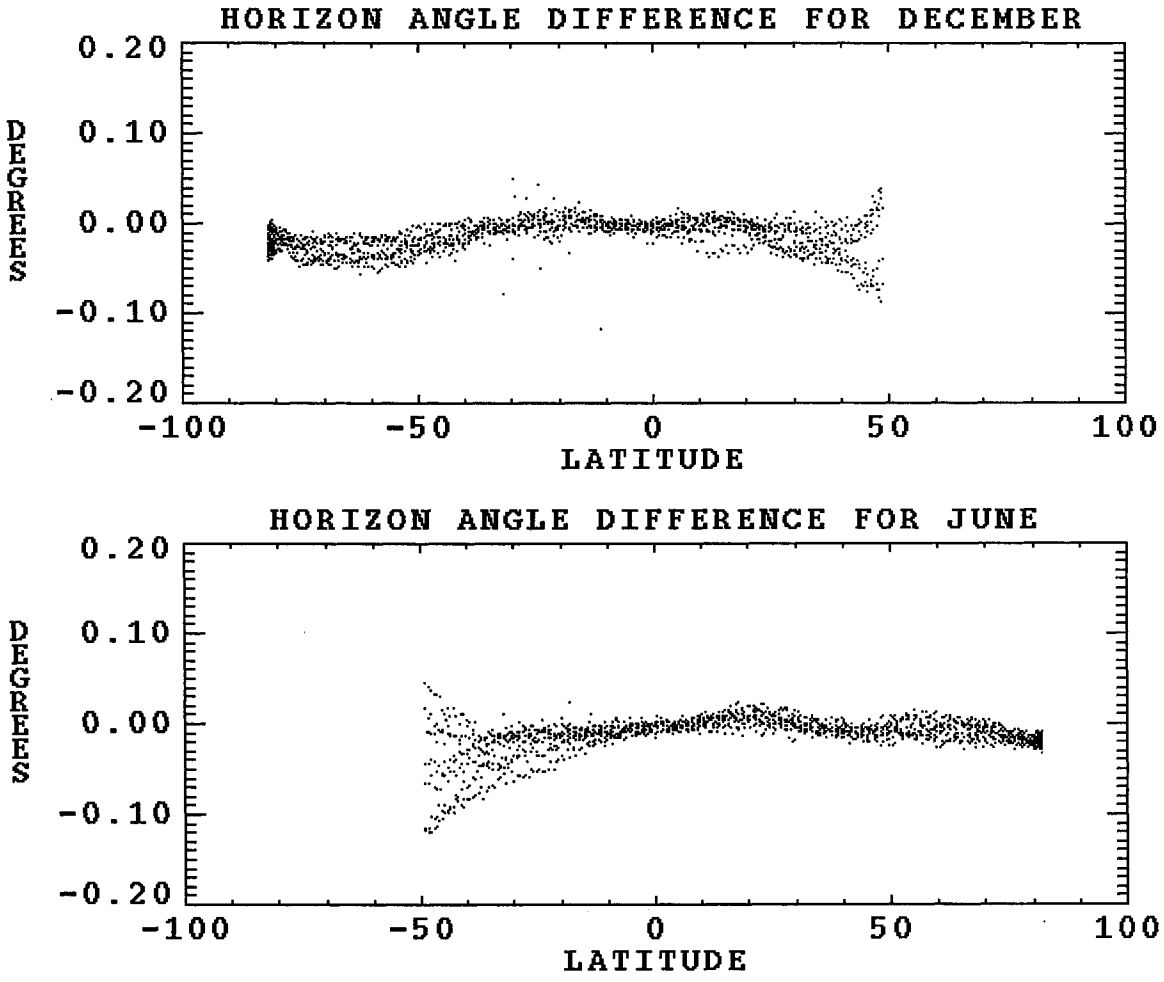


Figure 7. Horizon angle differences with seasonal model adjustments; (a) December 9, 1997; (b) June 12, 1998.

3. REFINEMENT OF THE MODEL USING ISLAND MATCHING RESULTS

As stated in the introduction, the accuracy of SeaWiFS navigation processing is assessed using the method of island target matching. The results of island target matching consist of the island location errors in latitude and longitude, along with the island coordinates, scan angle (i.e., pixel number), and other information (number of pixels, catalog ID number) for each matched island. Given a set of island matches spanning a range of scan angles, the average spacecraft attitude error can be estimated.

We have been using the island matching results to monitor the average attitude errors. Typically, the errors are estimated on a monthly basis as the data are processed. The average roll errors using a horizon scanner are directly influenced by the horizon angle model. Typical results for the roll errors vs. latitude are shown for June and December, 2000 (figures 8a and 8b).

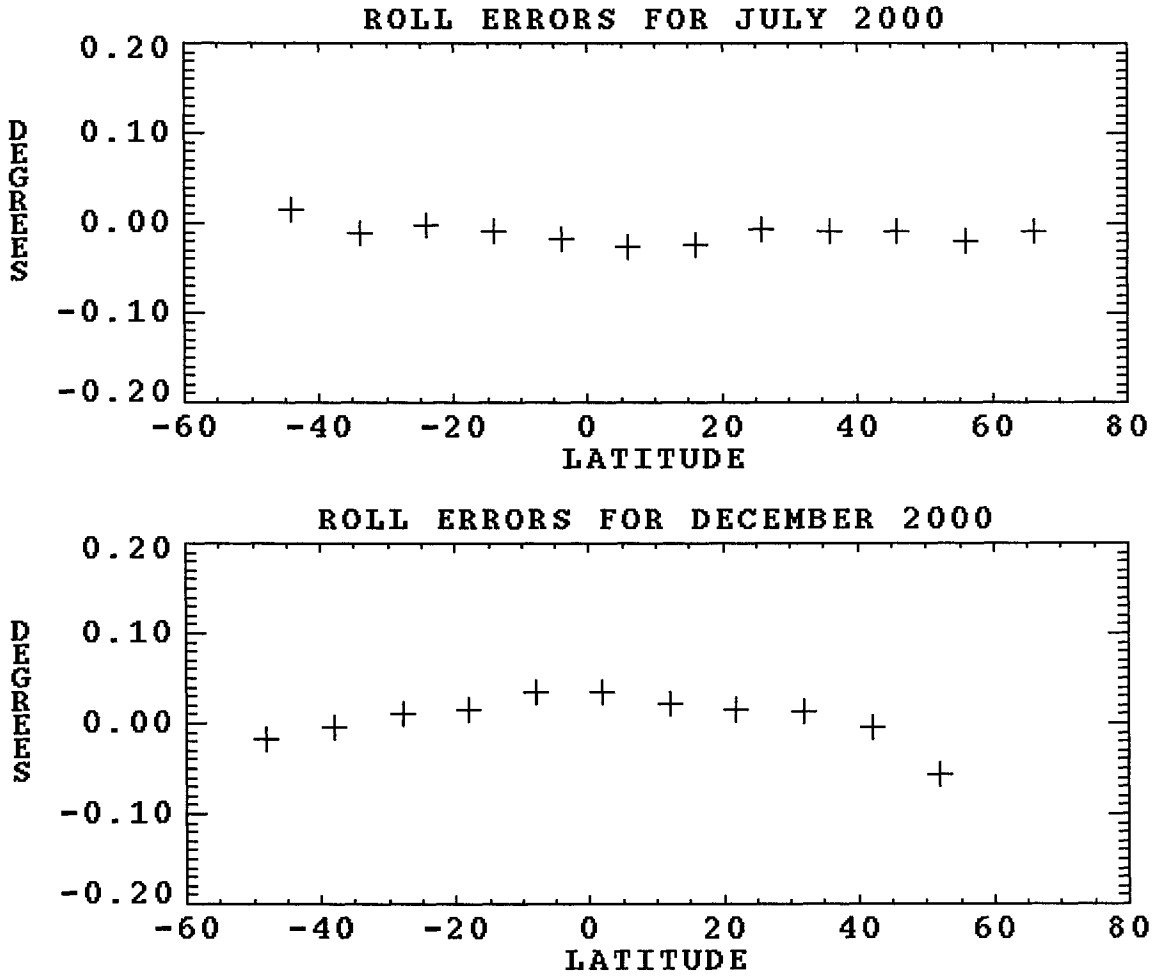


Figure 8. Average monthly roll errors determined from island matching results; (a) July 2000; (b) December 2000.

These results show that the average of roll errors for both periods is very close to zero. However, the July results show a slight trend, with errors being negative near the equator and increasing toward the poles. In December, the trend is opposite, with positive errors near the equator decreasing toward the poles. The results for summer months all appear similar to those for July 2000, and the winter months show a similar trend to that for December. The results in spring and autumn (not shown) indicated no appreciable trend.

These trends suggest an additional, seasonal variation in the horizon angle model, specifically in the flattening factor, f. The summer trend can be corrected by a small reduction in the value of f, while the winter trend indicates that f should be increased. We estimated that the value of f should be about 1/210 for the summer and 1/165 for the winter. Using these values we arrived at the following empirical expression for f as a function of solar declination, d:

$$f = 1/185 + \sin(d)/640 \quad (6)$$

We then repeated the analysis of the HS data from the dual-string period, to determine whether the seasonal adjustment of f would affect the ellipsoid shift and the Earth radius bias. The results (figures 9a and 9b) were a slight change in the ellipsoid shift, and a stronger correlation between the Earth radius and the solar declination. The revised expressions are:

$$\delta z = 11 \sin(d) \quad (7)$$

$$\delta r = -3.4 \sin(d) \quad (8)$$

These changes to the horizon angle model have just recently been implemented in the SeaWiFS navigation processing code. As of this writing we do not have any navigation assessment results with the new model.

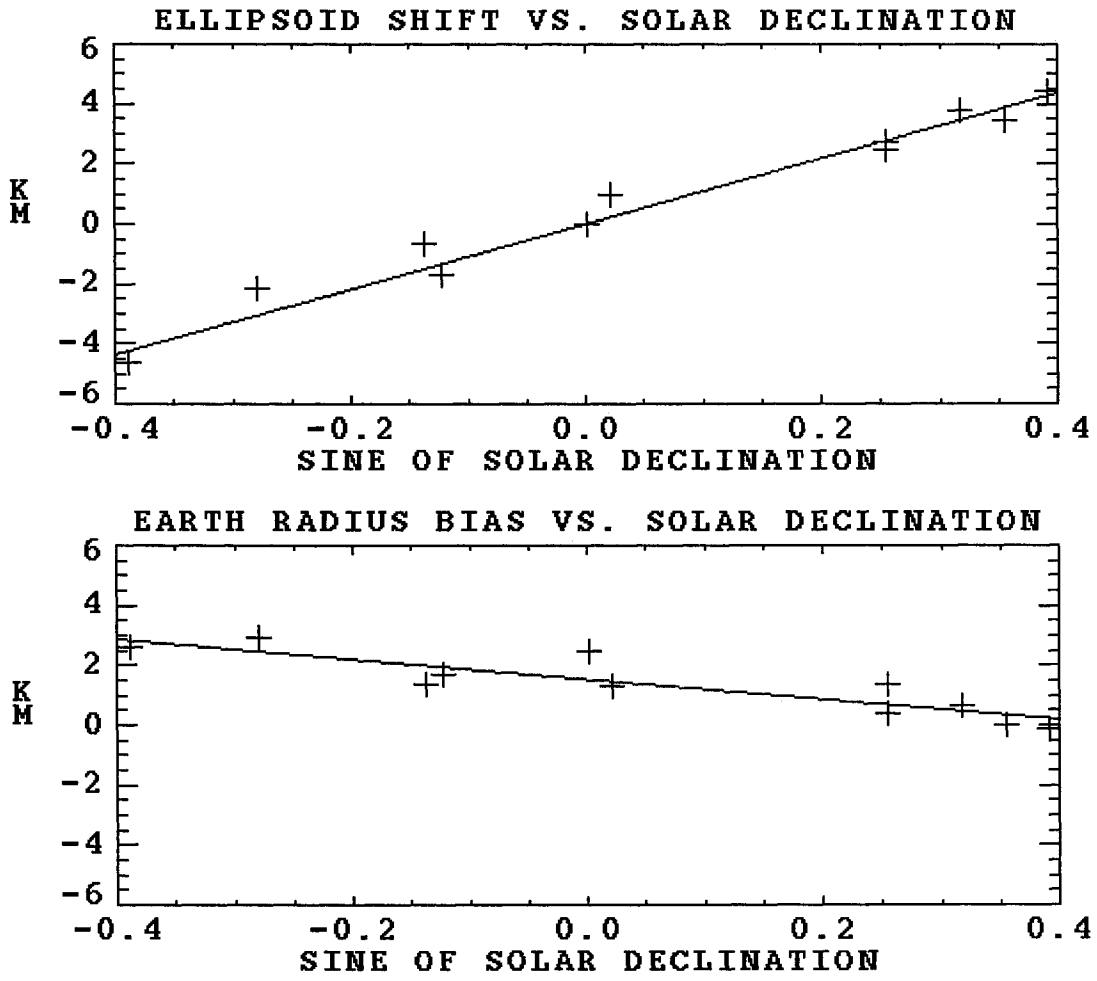


Figure 9. Estimated ellipsoid adjustments with seasonal flattening adjustment; (a) Ellipsoid shift; (b) Earth radius bias

4. CONCLUSION

An empirical, seasonal model has been developed for the Earth horizon angle, to account for apparent systematic variations in the OV-2 horizon scanner triggering height.. This model is required in order to meet navigation processing accuracy requirements using a single horizon scanner. It was initially developed using dual-scanner data from the start of the mission. Recently the model was refined based on analysis of navigation results with a single scanner. Navigation assessment processing and analysis will be continued to verify the accuracy of future processing

REFERENCES

1. Patt, F., R. Woodward, and W. Gregg, "An automated method for navigation assessment for Earth survey sensors using island targets", *International Journal of Remote Sensing*, vol. 18, no. 16, 1997

2. Patt, F., "Assessment of Geolocation for SeaWiFS and OCTS Using Island Targets", CNES Seminar on In-orbit Geometric Characterization of Optical Imaging Systems, Bordeaux, November 1999.

3. Anthony, T. and S. Sivapiragasam, "Design and Performance of the OrbView-2 Attitude Control System", 12th AIAA/USU Conference on Small Satellites, September 1998.

4. Bilanow, S., F. Patt and S. Kennison, "Attitude Sensor Bias Evaluation for OrbView-2", Proceedings of the 1999 Flight Mechanics Symposium, NASA/GSFC, NASA/CP--1999-209235.

5. Patt, F. and S. Bilanow "Attitude Sensor Bias Evaluation for OrbView-2", Proceedings of the 1999 Flight Mechanics Symposium, NASA/GSFC, NASA/CP--1999-209235.

6. Wertz, J. R., *Spacecraft Attitude Determination and Control,* D. Reidel, Dordrecht, Holland, 1978.

7. Bilanow, S. and L. Chen, "Landsat-4 Horizon Scanner Flight Performance", AAS/AIAA Astrodynamics Specialists Conference, Lake Placid, NY, August 1983.

8. Patt, F. and Gregg, W., "Exact closed-form geolocation algorithm for Earth survey sensors", *International Journal of Remote Sensing*, vol. 15, no. 18, 1994

APPENDIX -- COMPUTING THE HORIZON ANGLE

This section presents the derivation of a method for computing the horizon angle, i.e., angle from the geocentric nadir to the Earth horizon as a function of satellite orbit position and viewing azimuth. This approach allows the horizon ellipsoid to be shifted from the geocentric origin, thus providing a more general model.

Following the approach used for SeaWiFS geolocation (reference 8), a vector from an orbit position to the Earth surface ellipsoid can be solved for using the equation

$$AQ_1^2 + BQ_2^2 + CQ_3^2 + DQ_1Q_2 + EQ_1Q_3 + FQ_2Q_3 + GQ_1 + HQ_2 + IQ_3 + J = 0 \quad \text{(A1)}$$

Here **Q** is the view vector in the orbit position frame and A through J are defined as follows:

$$A = M_{11}^2 + M_{12}^2 + M_{13}^2/(1-f)^2 \quad \text{(A2a)}$$

$$B = M_{21}^2 + M_{22}^2 + M_{23}^2/(1-f)^2 \quad \text{(A2b)}$$

$$C = M_{31}^2 + M_{32}^2 + M_{33}^2/(1-f)^2 \quad \text{(A2c)}$$

$$D = 2M_{11}M_{21} + 2M_{12}M_{22} + 2M_{13}M_{23}/(1-f)^2 \quad \text{(A2d)}$$

$$E = 2M_{11}M_{31} + 2M_{12}M_{32} + 2M_{13}M_{33}/(1-f)^2 \quad \text{(A2e)}$$

$$F = 2M_{21}M_{31} + 2M_{22}M_{32} + 2M_{23}M_{33}/(1-f)^2 \quad \text{(A2f)}$$

$$G = 2M_{11}P_X + 2M_{12}P_Y + 2M_{13}P_Z/(1-f)^2 \quad \text{(A2g)}$$

$$H = 2M_{21}P_X + 2M_{22}P_Y + 2M_{23}P_Z/(1-f)^2 \quad \text{(A2h)}$$

$$I = 2M_{31}P_X + 2M_{32}P_Y + 2M_{33}P_Z/(1-f)^2 \quad \text{(A2i)}$$

$$J = P_X^2 + P_Y^2 + P_Z^2/(1-f)^2 - R_e^2 \quad \text{(A2j)}$$

where **M** is the geocentric-to-orbit frame transformation matrix, **P** is the orbit position vector, f is the ellipsoid flattening factor and R_e is the Earth equatorial radius.

Since the ellipsoid is longitudinally symmetric, this problem can be solved with **P** restricted to a plane without loss of generality; therefore

$$\mathbf{P} = \begin{vmatrix} P\cos\lambda \\ 0 \\ P\sin\lambda \end{vmatrix} \quad \text{(A3)}$$

where λ is the orbit geocentric latitude.

The matrix **M** is specified using the orbit latitude, as follows:

$$\mathbf{M} = \begin{vmatrix} 0 & 1 & 0 \\ -\sin\lambda & 0 & \cos\lambda \\ \cos\lambda & 0 & \sin\lambda \end{vmatrix} \quad \text{(A4)}$$

which defines the orbit frame basis vectors as the local East, North and zenith directions.

By comparing equations 2, 3 and 4, we can see that D, E and G are all 0. The other coefficients become:

$$A = 1 \quad \text{(A5a)}$$

$$B = \sin^2\lambda + \cos^2\lambda/(1-f)^2 \quad \text{(A5b)}$$

$$C = \cos^2\lambda + \sin^2\lambda/(1-f)^2 \quad \text{(A5c)}$$

$$F = -2\sin\lambda\cos\lambda + 2\sin\lambda\cos\lambda/(1-f)^2 \quad \text{(A5d)}$$

$$H = -2P\sin\lambda\cos\lambda + 2P\sin\lambda\cos\lambda/(1-f)^2 \quad \text{(A5e)}$$

$$I = 2P\cos^2\lambda + 2P\sin^2\lambda/(1-f)^2 \quad \text{(A5f)}$$

$$J = P^2\cos^2\lambda + P^2\sin^2\lambda/(1-f)^2 - R_e^2 \quad \text{(A5g)}$$

Clearly further simplifications are possible, e.g., H = PF.

The orbit-to-surface vector, **Q**, is specified in the orbit frame as

$$Q = \begin{vmatrix} Q\sin\psi\sin\rho \\ Q\cos\psi\sin\rho \\ -Q\cos\rho \end{vmatrix} \quad \text{(A6)}$$

where ψ is the azimuth of the vector relative to local North, and ρ is the angle between **Q** and the geocentric nadir vector. Substituting this into equation 1, with the zero terms omitted, gives:

$$(Q\sin\psi\sin\rho)^2 + B(Q\cos\psi\sin\rho)^2 + C(Q\cos\rho)^2 + F(Q\cos\psi\sin\rho)(-Q\cos\rho) + H(Q\cos\psi\sin\rho) + I(-Q\cos\rho) + J = 0 \quad \text{(A7)}$$

This can be rearranged as:

$$(\sin^2\psi\sin^2\rho + B\cos^2\psi\sin^2\rho + C\cos^2\rho - F\cos\psi\sin\rho\cos\rho)Q^2 + (H\cos\psi\sin\rho - I\cos\rho)Q + J = 0 \quad \text{(A8)}$$

If ψ and ρ were known, this quadratic could be solved for Q. In this case, we wish to determine ρ, given ψ, for the case where Q represents the horizon vector. To do this, we need to eliminate Q from equation 8.

Using the standard quadratic formula

$$Q = \frac{-b +/- (b^2 - 4ac)^{1/2}}{2a} \quad \text{(A9)}$$

where

$$a = \sin^2\psi\sin^2\rho + B\cos^2\psi\sin^2\rho + C\cos^2\rho - F\cos\psi\sin\rho\cos\rho \quad \text{(A10a)}$$

$$b = H\cos\psi\sin\rho - I\cos\rho \quad \text{(A10b)}$$

$$c = J \quad \text{(A10c)}$$

we note that Q has two solutions if $(b^2 - 4ac)$ is positive (i.e., view vector intersects the ellipsoid in two places), no solutions if it is negative (vector misses the ellipsoid) and one solution if this quantity is zero. This last case is the desired solution, since it represents the view vector exactly at the Earth horizon. Substituting equations 10 into the quantity gives

$$(H\cos\psi\sin\rho - I\cos\rho)^2 - 4J(\sin^2\psi\sin^2\rho + B\cos^2\psi\sin^2\rho + C\cos^2\rho - F\cos\psi\sin\rho\cos) = 0 \quad \text{(A11)}$$

Expanding the first term and rearranging gives

$$(H^2\cos^2\psi - 4J\sin^2\psi - 4BJ\cos^2\psi) + (4JF\cos\psi - 2HI\cos\psi)\cos\rho/\sin\rho + (I^2 - 4JC)\cos^2\rho/\sin^2\rho = 0 \quad \text{(A12)}$$

which can be solved as a quadratic in $\cos\rho/\sin\rho$, or $\cot\rho$.

This equation has been shown to produce identical results to equation 4-24 in Wertz and, with sufficient algebraic simplification, should reduce to that form. However, equation 12 is more suitable to the purpose of generalizing the horizon angle model.

One useful generalization is to allow the center of the ellipsoid to not be the coordinate system origin, i.e., to be shifted from the actual Earth center. This is useful in developing a seasonal model in which the effective triggering height, the CO_2 layer, is shifted North or South according to the time of year.

The ellipsoid shift can easily be introduced by subtracting it from the orbit position vector. For a simple North-South shift, this gives

$$\mathbf{P} = \begin{vmatrix} P\cos\lambda \\ 0 \\ P\sin\lambda - \delta z \end{vmatrix} \qquad \text{(A13)}$$

Using this expression in place of equation 3 in deriving coefficients H, I and J gives

$$H = -2P\sin\lambda\cos\lambda + 2P\sin\lambda\cos\lambda/(1-f)^2 - 2\delta z\cos\lambda/(1-f)^2 \qquad \text{(A14h)}$$

$$I = 2P\cos^2\lambda + 2P\sin^2\lambda/(1-f)^2 - 2\delta z\sin\lambda/(1-f)^2 \qquad \text{(A14i)}$$

$$J = P^2\cos^2\lambda + (P\sin\lambda - \delta z)^2/(1-f)^2 - R_e^2 \qquad \text{(A14j)}$$

The calculation of the horizon angle, ρ, as a function of azimuth is then performed as before using equation 12.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2001	Conference Publication

4. TITLE AND SUBTITLE
2001 Flight Mechanics Symposium

5. FUNDING NUMBERS
Code 572

6. AUTHOR(S)
Edited by John P. Lynch

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS (ES)
Goddard Space Flight Center
Greenbelt, Maryland 20771

8. PEFORMING ORGANIZATION REPORT NUMBER
2001-02668-0

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS (ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA/CP-2001-209986

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified–Unlimited
Subject Category: 13
Report available from the NASA Center for AeroSpace Information,
7121 Standard Drive, Hanover, MD 21076-1320. (301) 621-0390.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This conference publication includes papers and abstracts presented at the Flight Mechanics Symposium held on June 19–21, 2001. Sponsored by the Guidance, Navigation and Control Center of Goddard Space Flight Center, this symposium featured technical papers on a wide range of issues related to attitude/orbit determination, prediction and control; attitude simulation; attitude sensor calibration; theoretical foundation of attitude computation; dynamics model improvements; autonomous navigation; constellation design and formation flying; estimation theory and computational techniques; Earth environment mission analysis and design; and, spacecraft re-entry mission design and operations.

Government, industry, and the academic community participated in the preparation and presentation of these papers.

14. SUBJECT TERMS
Flight Mechanics, Attitude Determination, Mission Analysis, Spacecraft Dynamics, Orbit Determination

15. NUMBER OF PAGES
573

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL