A Modeling Strategy for Analysis of Optomechanical Systems

A. E. Hatheway

Alson E. Hatheway, Inc.
Consulting Engineers
1041 East Green Street
Pasadena, California 91106
(818) 795-0514

The interactions between optical and mechanical phenomena are becoming more important to the optics industry. In many cases the optical performance of systems is limited by the ability of the structural system to maintain the appropriate geometry. It, therefore, becomes important to quantify the interactions between the structural system and the electromagnetic waves that transit the optics. This paper compares optical and structural analysis procedures and describes a modeling strategy for including optical phenomena in structural finite element models.

Introduction

Optical instruments have grown in complexity, size and precision over the last decade. This process is continuing and shows no likelihood of abating in the next few years. The government is currently funding a spaced based astronomical telescope for deployment in the 1990's. This system, called Large Deployable Reflector (LDR), has a primary mirror that is twenty meters in diameter and is diffraction limited at thirty micrometers. It will be able to resolve planets in solar systems of nearby stars. Fundamental physics indicates this instrument will work if it can be supported by a suitable structural system. LDR's performance depends on our ability to design precision structures. Similarly, in terrestrial and airborne instruments the requirement for ten to one hundred nanoradian class stable platforms is increasing. One can anticipate that we will shortly see the requirement for subnanoradian pointing stabilities.

The performance of these instruments relies on the favorable behavior of their supporting structural subsystems. In order to analyze these systems, we need to link their optical performance to their structural behavior. In response to this challenge the industry has developed data base management systems to transfer data from one solution routine to another. This allows each of the disciplines to work in its own software codes with a minimum of manual effort in reformatting and interpreting data between routines. They have also extended the capability of some of the design codes to include the analysis of deformed optical shapes. The most advanced optical codes include interpolation routines that will allow the analyst to define a deformed optical surface at a number of points and the code will derive the position and slope of each of the ray intercept points based upon an interpolation algorithm and differentiation of the resulting curves. In order to assess the effectiveness of the industries' efforts, one must evaluate the accuracy and economy of these methods in simple terms.

Deformed Optical Surface Analysis

Analysis of systems with deformed optical surfaces relies upon interpolation procedures which first transform the finite element displacements into formats and coordinate systems that are compatible with the optics design code. The optics code then determines the displacements and slopes at each ray intercept point by further interpolation from data for the transformed grid points. Additionally, only displacements normal to the surface are used in the optics codes. The other five motions at each point are derived when needed from the normal displacements using the interpolation process.

Many interpolation systems are possible but we shall consider three which have found wide use in the industry:

Zernike polynomials,
linear four point interpolation and
cubic spline functions.

Zernike polynomials are an infinite set of polynomials that are capable of describing any continuous surface function inside a unit circle. The polynomials combine trigonometric and power series functions and they increase in complexity significantly as their order increases. The form of some of the polynomials is identical to the Seidel aberations of an optical wave front and all of the polynomial functions have a direct wave mechanics analog.

The Zernike polynomials interpret the deformed surface in terms of the optical aberations. This is convenient for the optical scientist or engineer, especially if he can adjust the position of some of the elements to accommodate the indicated aberations. Unfortunately, it tells the structural engineer very little about the interaction between the structural deformation and the optical performance. Additionally, the Zernike polynomials are reasonably long expressions and the analysis is usually limited to the first thirty-six. This procedure requires a curve fitting process (usually it is a least-squares type) because there are more data than there are polynomials. The result is an approximate evaluation of the optical aberations and a residual error vector that has not been included in the polynomials. The accuracy of a Zernike polynomial analysis will depend upon both the number of polynomials used and the regularity of the surface deformations. The polynomials are most frequently used as an interpretation scheme but they can also be used as interpolation functions, in either of which cases the analyst should be careful to review the residual error vector to determine the extent to which significant data may be missing from the analysis.

The linear four point method is, just as its name describes, a linear interpolation from the points nearest the ray intercept point. The errors incurred in the linear interpolation scheme can be assessed if one assumes a well behaved cylic function. Figure 1 is a plot of the displacement error in fitting a simple parabolic function with a straight line segment for various numbers of waves of cylic surface distortion. To achieve a .01 interpolation error in a simple structure with simple loads, the analyst needs to use about four hundred grid points to describe the surface deflections. For the same accuracy in a complex structure with complex loads, the analyst may need sixteen hundred or more grid points in the surface. The lower curve marked "SPH SHELL" is the upper bound of the error for an MSC/NASTRAN model using QUAD4 shell elements as documented in the MSC/NASTRAN Application Manual. The structural model achieves .01 error with only fifty-eight grid points.

The optical ray tracing procedures also use the slope of the optical surface. The slope rotations are derived by differentiating the interpolation function. Figure 2 is a plot of the slope error over the same range as Figure 1. Limiting these errors to .01 requires forty thousand to one hundred sixty thousand surface grid points be input to the interpolator. This procedure faces a severe problem in reducing the errors to less than about .10. The "SPH SHELL" curve is the same as in Figure 1 and the figure shows the loss in accuracy caused by the combined interpolation and differentiation procedure. Note that we have not yet addressed the uncertainties caused by boundry conditions and inhomogenieties in the structure.

One may question the use of a parabolic function in the above analysis. Several other functions were investigated:

> Circular arcs gave results nearly identical to those shown in Figures 1 and 2 with some degradation in the region where errors are greater than .10 and
>
> Cubic functions were found to offer some improvement in the displacement errors but at the price of significant degradation in the slope errors.

Since structural deformation curves may be a combination of many higher order terms, the actual interpolation error will be a function of the specific curve being fitted. The parabolic function has been presented only as a typical curve and an analyst should realize that the actual errors may be greater or smaller depending on the specifics of the problem.

To improve on the accuracy and hopefully to provide a viable tool for optics designers, the cubic spline interpolation system was developed. The cubic spline interpolation procedure requires some detailed attention because it has been designed especially to address the problems of structurally deformed optical surfaces. The procedure uses the surface displacements (normal to the surface) calculated at a number of places on the optical surface. The finite element method is ideal for calculating these displacements. A third degree polynomial is assumed to fit between each pair of adjacent points on the surface and the equations are set up for simultaneous solution using matrix algebra. However, additional data are required for the solution of these equations. In order to complete the calculations the routine assumes,

1) the slopes of the cubic curves are identical at each data point,
2) the curvatures are identical at each data point and
3) the curvatures are zero at the edges of the field of points.

The third degree polynomial was a reasonable choice since this form is used in describing deformations in some structures. However, the benefit of using a good polynomial may be lost if improper boundry conditions are applied. For instance, although assumption 1) is always valid for elastic systems, assumptions 2) and 3) are not. Curvatures do not vary smoothly and uniformly so as to be identical at each data point. In fact, curvatures may change abruptly at the data points since this is where the material and section properties of the structure are likely to change, reflecting structural inhomogenieties. Additionally, the very forces that give rise to the displacements normal to the surface also give rise to in-plane forces, torques and moments. These reactions will tend to twist the surface contour while it is constrained against normal translation at each of the data points. Also, assumption 3) eliminates thermal gradients as a potential source of deformation since these can easily give curvatures at the edges of a structure.

The cubic spline interpolation procedure improves on the accuracy of the linear procedure by about one or two orders of magnitude for a well behaved problem. However, if the boundry conditions and structural homogeneity are unfavorable, the increased accuracy may be completely lost, particularly in slow (long focal length) systems that are dominated by the slope error contributions. Table A shows the gains and losses that one might experience using the cubic spline method instead of the linear method (and accounting for boundry conditions and surface inhomogenieties). The table shows the assumed range of the effect in parenthesis. An average value is used in the calculation. The formatting procedure is the conversion of data from a polar coordinate system (typical of circular optical elements in the finite element model) to a rectangular coordinate system required of ACCOS V. This formatting is performed either by the analyst or a database manager/manipulator in the computer and is equivalent to an additional linear interpolation step. The analysis shows a thirty-eight percent improvement in accuracy by the cubic spline method over the linear method after accounting for boundry conditions and surface inhomogenieties. One should realize that all the factors used in the analysis are problem dependent and may have large variations in value from problem to problem. Also, there is no simple way to bound the magnitude of these variables so that one may not be able to assess the overall accuracy of the solution. One way to improve the accuracy of interpolation is to provide more data points but this requires an increase in the size of the structural model. With the above uncertainties about the accuracy of the analysis, it may be difficult to justify the expense of additional degrees of freedom in the finite element solution.

Table A
Cubic Spline Gains and Losses

Gains:

50x for a better curve (10x to 100x)

Losses:

4x	for formatting the input (1x to 8x)
3x	for boundry conditions (1x to 6x)
3x	for inhomogenieties (1x to 6x)
1.38x	Net Gain.

The industry has been limited in its ability to analyze the behavior of optical systems with deformed optical surfaces. This has resulted in the continuation of the traditional methods of sizing the elements to have a maximum diameter to thickness ratio (based upon the material) and the use of a few proven mounting methods that both minimize and distribute the reaction loads between the elements and the supporting structure.

Time and Frequency Domain Analysis

The optical design codes are also used to assess the performance of optical systems in dynamic environments. In this case the displacements at a number of time steps or frequencies are derived by the finite element method. These are used for optical computation at each time step or frequency and the results are transformed by a post processor into transient, frequency or random response format. If the optical system can be conceived as a set of rigid optical elements, each with some rigid body motion that can be defined for the optical computation, this procedure is reasonably accurate. However, it can consume large amounts of resources. A frequency response problem that has two hundred frequency intervals in the solution may require sixteen hundred passes through the optical design code in order to determine the phase and magnitude of the maximum optical response for both the position and the path difference. Then this data must be transformed by a post processor to calculate the desired quantities. To aid in this process the industry has developed large data base manipulators that help to interface the various software codes, manipulate and format data from one input into another and transform the final output data into the required type of results. The procedure can be quiet accurate if interpolation procedures are avoided in the data base manager as well as the optics code itself.

Proprietary Codes

Additionally, every sizeable firm in the industry has a set of proprietary codes that perform many of the above tasks as well as others in an attempt to simulate end-to-end system optical performance. None of these proprietary codes improve upon the accuracy or cost of the above procedures. Although many of the codes offer special features, such as solving diffraction integrals or accommodating atmospheric effects, they all interpret the structure through some simplified and reduced set of displacement data. Some rely on interpolation and differentiation algorithms to evaluate the displacements and slopes at the required points while others use Zernike polynomials. None of them attempt time or frequency domain analysis.

Is there any way to improve on this situation? Let us look at both optical and structural computation methods, understand their differences and similarities and see if we can draw a few reasonable conclusions.

An Optomechanical Modeling Strategy

Optical calculations are in the realm of wave (Maxwellian) mechanics and determine static positions in a continuous medium. Structural calculations are in the realm of Newtonian mechanics and determine dynamic (statics representing a special case) displacements in a discontinuous medium. The seeming incompatibility of these two disciplines may be resolved if one considers some of the implications of the above statements, i.e.,

1) Structural calculations are in the first partial derivative domain with respect to optical calculations,

2) Optical displacements are very small so that a continuum is not necessary for calculating the effects of optical displacements,

3) Optical displacements are well within the limits of linear elastic system behavior, and

4) Optical waves and rays resemble rigid elements in that they are massless, rigid and their effects propogate very fast with respect to the effects of Newtonian mechanics.

One could presume from these observations that optical phenomena might be incorporated in a finite element model. In fact, optical phenomena are readily analyzed by the finite element method. To take advantage of this capability, the analyst must understand the optical phenomenon of interest, the finite element code to be used and the principles of structural mechanics.

To prepare a model of an optical phenomenon the analyst must first understand the equations that describe it. He must transform these equations by taking the first partial derivative of the quantity with respect to each of the displacements in the problem, these define the behavior of the phenomenon in the displacement vector. The transformed equations are then coded onto MPC cards to "drive" the phenomenon. An accurate geometric model of the optical surfaces and devices is prepared from the optical prescription developed by the optics designer. These surfaces are composed of rigid elements and the surfaces can be interconnected (with rigid elements) to duplicate rigid optical elements. This procedure will generate a finite element model of the optical system that can be "exercised" to determine its agreement with the prescription's optical behavior.

MSC/NASTRAN is an ideal code for performing optomechanical analysis. The free form of the Bulk Data Deck allows the optical model to be developed and debugged separately from the structural model. The variety of rigid elements and the Multipoint Constraint (MPC) features allow accurate representation of the Maxwellian waves and rays. The use of base excitation in dynamics duplicates the actual conditions and allows accurate calculation of relative displacements between stationary objects and moving ones. The code is optimized to solve the larger problems typical of optical systems. The plot module is flexible and has utilities that further enhance the interpretation of optical phenomena.

The analyst can attach the optical model to any suitable structural model (using rigid elements) and the combined model will calculate and display the resulting optomechanical behavior. A suitable structural model is one that is designed to calclulate accurate displacements at each of the optical principal points in the anticipated structural environment. The structural model may have many fewer degrees of freedom than one required to determine accurate stresses and the strategies for developing such structural models are well documented.

Applications of the Modeling Strategy

Figure 3 is an example of an optical model prepared according to the above guidelines. Reducing each optical element to a single point was made possible by only analyzing the behavior of a single axial chief ray. As simple as this optical model was, it proved to be very powerful in understanding the critical features of the stable platform that supports it (Figure 4). When the model was run in Solution 30, MSC/NASTRAN produced the response PSD plots of the optical ray that were required to evaluate the stable platform's ability to hold a steady line-of-sight. No data base manipulation, interpolation, formatting nor post processing was required.

The whole effort, including the optical model (which was very simple) used about four man-months. Earlier, a team of engineers using the conventional methods described previously needed about eighteen man-months. They abandon the effort after doing just one axis of vibration because they had no way to validate their methods or verify their results. If we remove the cost associated with just the structural model, about three man-months, one can derive the labor cost for extending the calculations to include the optical behavior. This results in a 45:1 cost advantage for MSC/NASTRAN optomechanical analysis methods over the conventional methods.

Figure 5 shows an optomechanical model of a large primary mirror. This model was used to determine the effects of gravitational and thermal loading of the mirror structure on the ability of the mirror to create a sharp image. The Figure shows the ability of MSC/NASTRAN to duplicate both the optical aberations and the structural/thermal aberations. The optical aberation was used as a test to validate the model. The magnitude of the sagital coma agreed to four decimal places with optical theory for off-axis illumination of a paraboloid. Comparison with optical theory provides assurance of the accuracy with which the aberations caused by the thermal and gravitational loads will be calculated.

As one can observe from Figure 7, the aberations caused by nonuniform temperature distributions and transverse gravitational loads are not necessarily regular or convenient optical aberations. This points out the uncertainty caused by interpreting these aberations with Zernike polynomials. For this design the gravitational aberation is a reasonably regular shape with some symmetry. It can probably be well represented by the Zernike polynomials. However, the thermal aberation is very irregular and much of it will be left in the residual displacements after the least squares fitting procedure is complete. On the other hand, the optical, thermal and gravitational aberations are handled with equal ease and accuracy by including the optical imaging phenomenon in the MSC/NASTRAN model.

Conclusion

Current methods for analyzing optomechanical systems suffer an unfavorable cost versus accuracy trade off. Some of the methods also require considerable calendar time for execution. Both of these shortcomings can be overcome by performing the analysis in the finite element medium instead of the optical design medium. Optical phenomena are readily accepted by some finite element codes. This paper presents a modeling strategy that allows the analyst to model optical phenomena in MSC/NASTRAN. Two examples of applying the strategy have been used to illustrate the accuracy and economy of the method. Although the examples have been limited to ray tracing, virtually any optical phenomenon may be modeled.

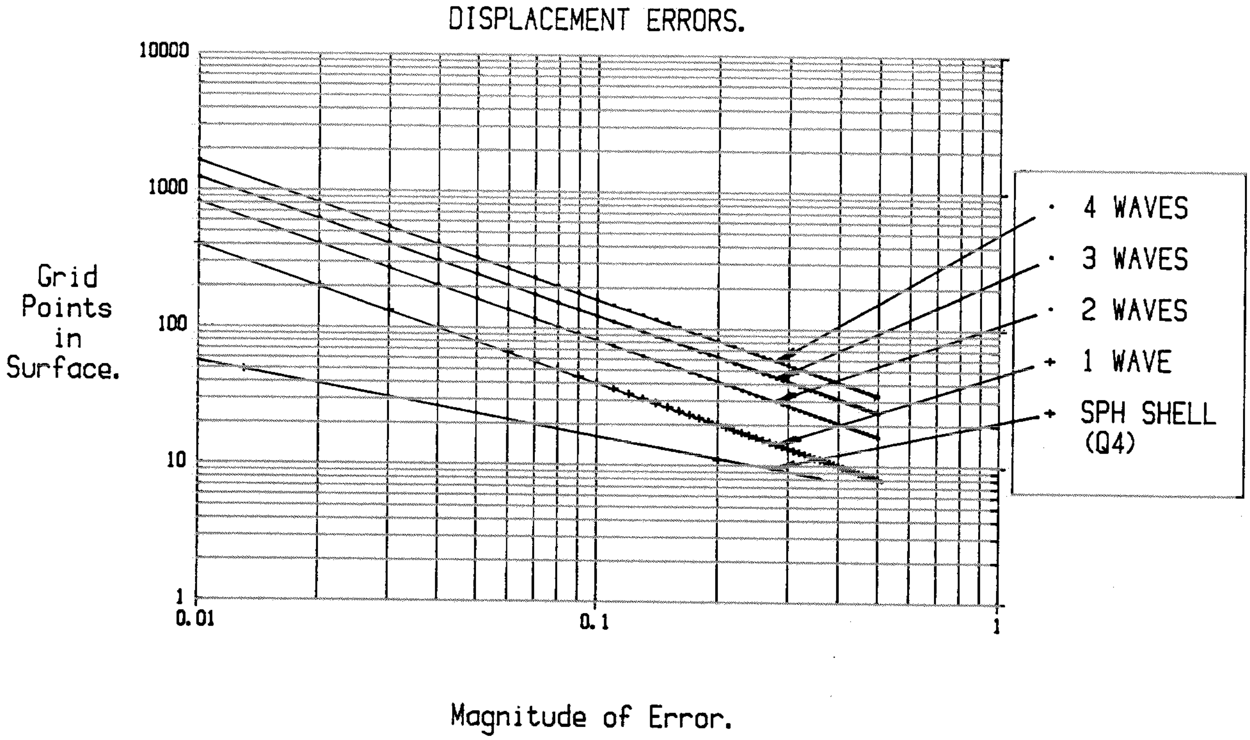


Figure 1 Displacement Interpolation Errors

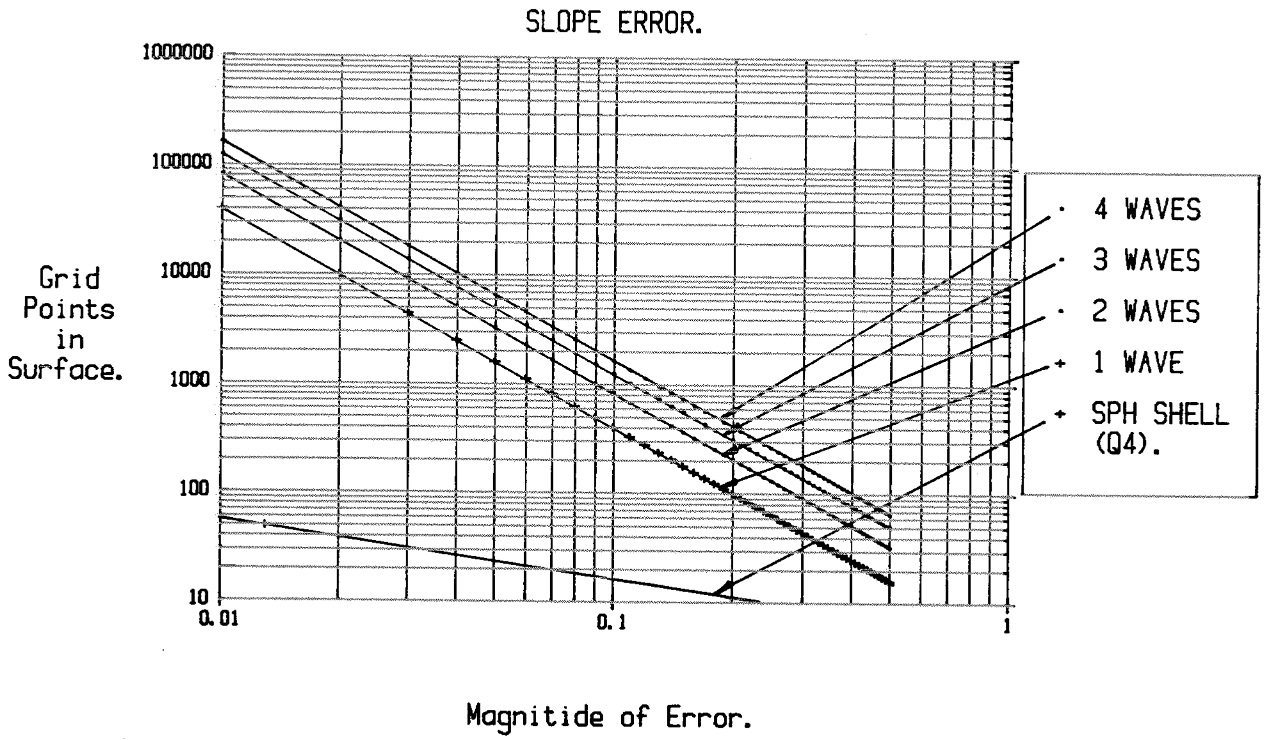


Figure 2 Slope Interpolation Errors

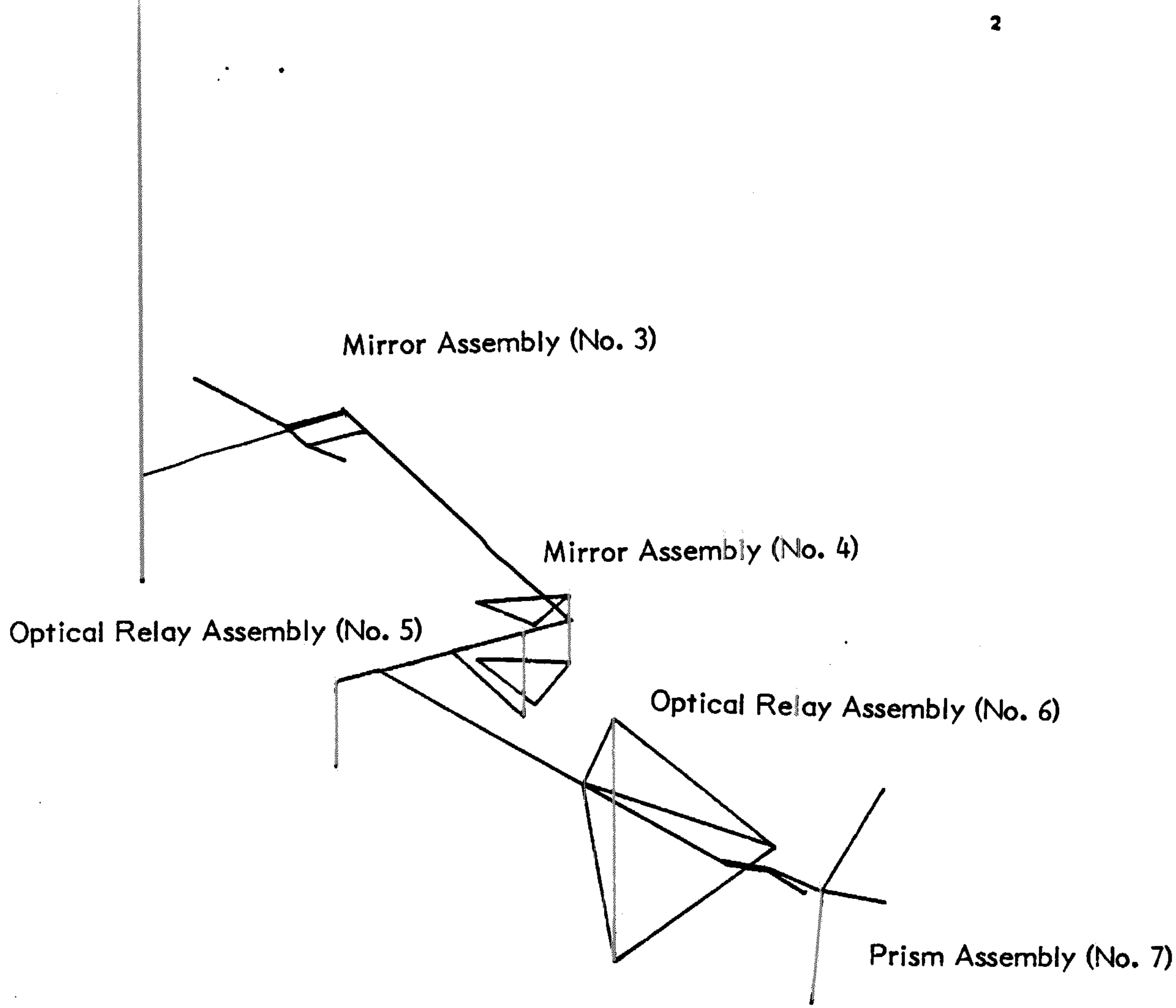
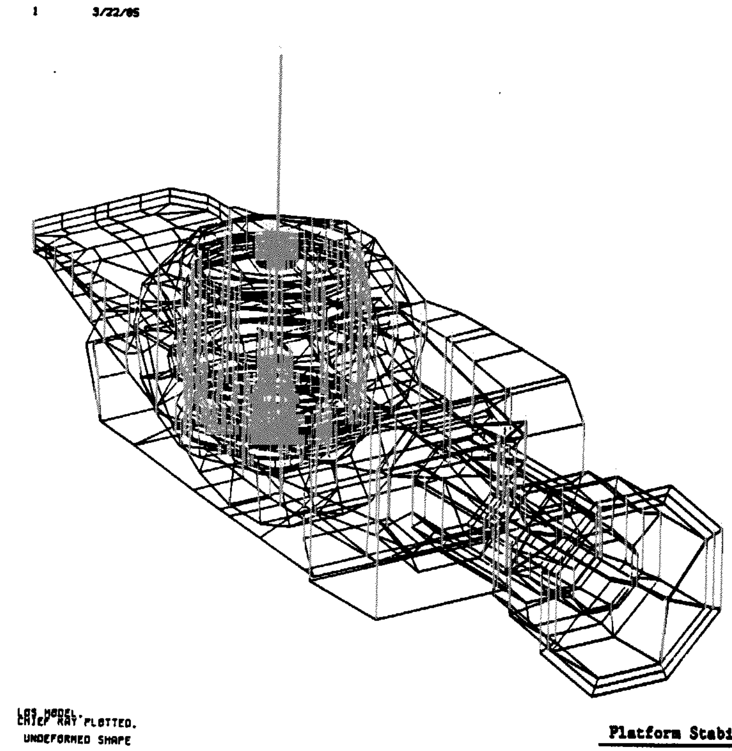
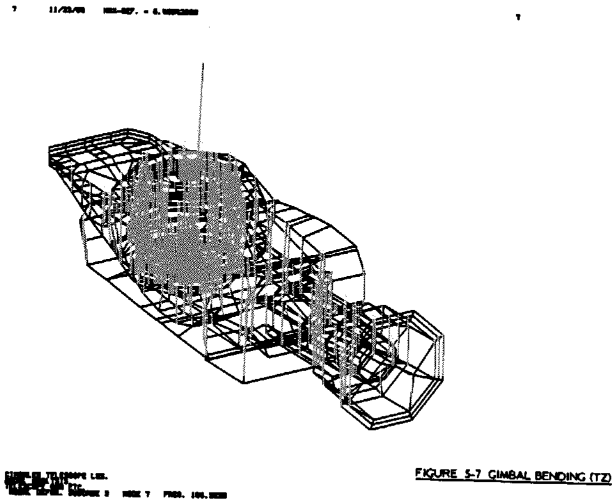


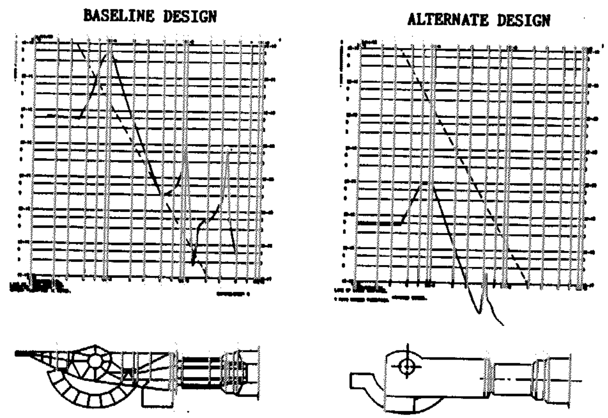
Figure 3 Optical System Model

UNDEFORMED SHAPE



Platform Stability Analysis.





Line-of-sight Stability PSD (Radians Squared per Hertz).

Figure 4 Stable Platform Optomechanical Model and Analysis

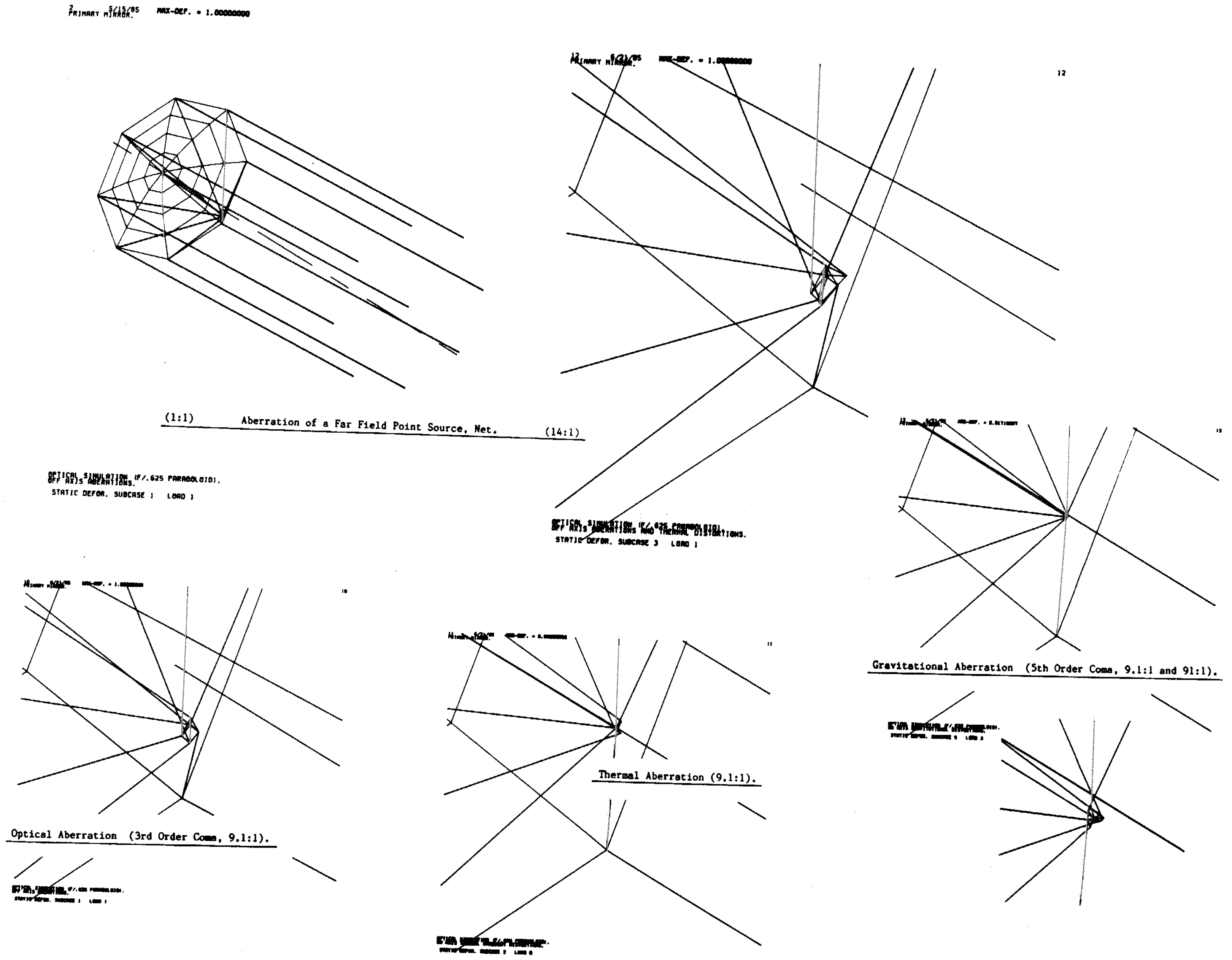


Figure 5 Thin Primary Mirror Optomechanical Model and Analysis