REPORT OF THE HST STRATEGY PANEL:
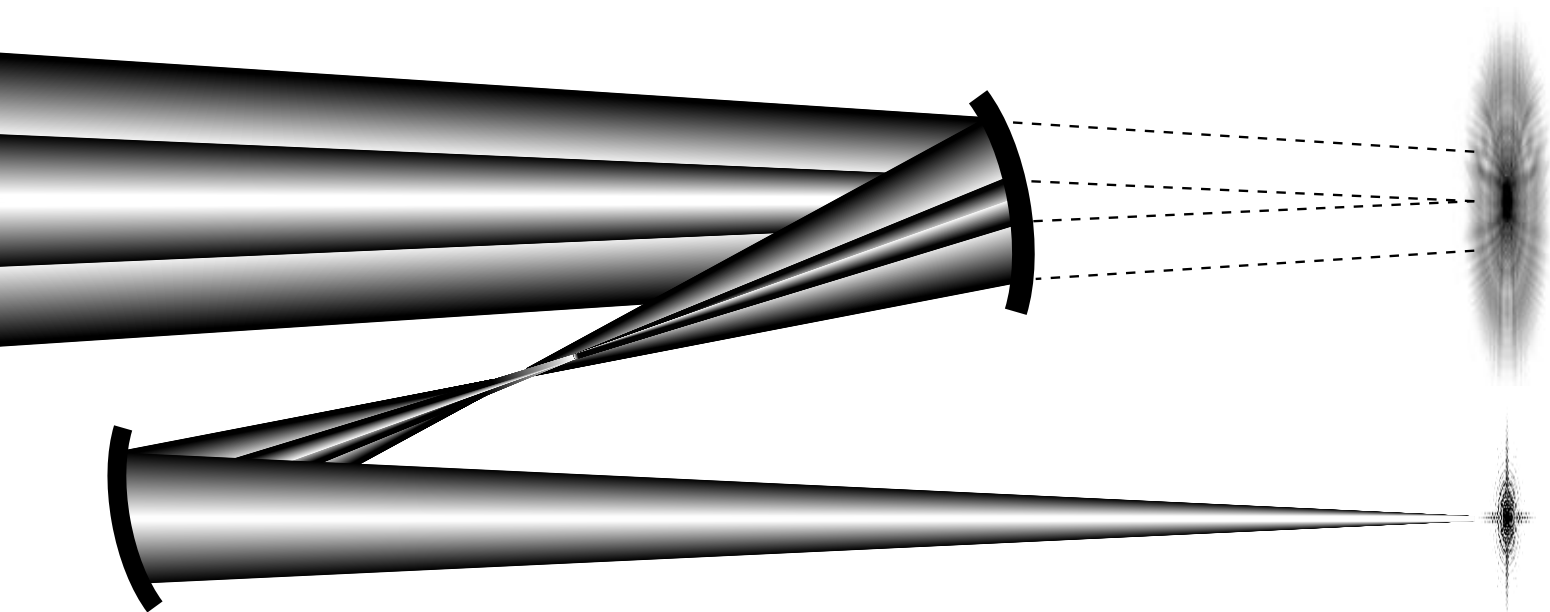
A STRATEGY FOR RECOVERY

The Results of a Special
Study
August–October 1990

EDITED BY R. A. BROWN AND H. C. FORD

REPORT OF THE
HST STRATEGY PANEL:
A STRATEGY FOR RECOVERY

THE RESULTS OF A SPECIAL STUDY
AUGUST–OCTOBER 1990

EDITED BY R. A. BROWN AND H. C. FORD

PANEL MEMBERS

Robert Brown, *Co-Chairman*   Space Telescope Science Institute
Holland Ford, *Co-Chairman*   Space Telescope Science Institute
Roger Angel   University of Arizona
Jacques Beckers   European Southern Observatory
Pierre Bely   Space Telescope Science Institute
Piero Benvenuti   ST European Coordinating Facility
Murk Bottema   Ball Aerospace
Christopher Burrows   Space Telescope Science Institute
James Crocker   Space Telescope Science Institute
Rodger Doxsey   Space Telescope Science Institute
Sandra Faber   Lick Observatory
Edward Groth   Princeton University
Shrinivas Kulkarni   California Institute of Technology
Bruce McCandless   Johnson Space Center/NASA (Ret.)
Francesco Paresce   Space Telescope Science Institute
Lyman Spitzer, Jr.   Princeton University
Raymond Wilson   European Southern Observatory

PANEL CHARTER

The Panel will identify and assess strategies for recovering HST capabilities degraded by spherical aberration. It will review the current state of the observatory, the Allen Board findings, the scientific potential of the ideal HST, and the tentative science program of the unimproved HST. It will develop a comprehensive framework for identifying possible improvements, including OTA-, instrument-, spacecraft-, and operations-level changes, and including hybrid combinations. Within this framework, the Panel will develop and debate the technical and scientific merits of particular improvements. On the basis of their findings, the Panel will formulate a set of recommendations and conclusions.

The Panel will cast its net widely, especially seeking the ideas and appropriate involvement from ST ScI staff. It is expected that the whole Institute will be informed regularly of the Panel's thinking and progress. As necessary in the course of its work, the Panel can request ST ScI support for short studies of specific technical or scientific issues that may arise. They may also request the support of outside experts including but not limited to NASA and NASA contractor personnel.

The Panel is appointed by and reports to the ST ScI Director, who will take the panel's findings to NASA.

CALENDAR OF MEETINGS

August 17 & 18, 1990   Space Telescope Science Institute, Baltimore
September 3 & 4, 1990   European Coordinating Facility, Garching
October 1 & 2, 1990   Space Telescope Science Institute, Baltimore
October 17 & 18, 1990   Space Telescope Science Institute, Baltimore

ACKNOWLEDGEMENTS

CONTENTS

PREFACE

ASTRONOMERS AND ENGINEERS REALIZED that there was a problem with the images of Hubble Space Telescope (HST) shortly after it was launched in April 1990. The quality of the images failed to improve despite attempts to adjust the alignment of the optics. NASA concluded in June 1990 that the HST primary mirror had been manufactured with the wrong shape. Compared with the desired profile, the mirror surface is too low by an amount that from the center to the edge grows from zero to 0.002 mm or four wavelengths of optical light. NASA convened an investigatory board in July 1990 under Dr. Lew Allen, which reported in November 1990 how the error probably occurred. In late 1980 or early 1981, a technician had improperly assembled a measuring device used to figure the primary mirror. Though tests at the time indicated a problem, the warning was not heeded, and the HST was assembled and launched with the flawed mirror.

The deformity of the HST mirror causes spherical aberration in the images. This means light rays come to a focus at different distances depending on the radius at which the rays strike the mirror, as shown in Figure 1. Light from the edge of the primary mirror comes to a focus about 38 mm beyond where the innermost rays converge.

Figure 1. Spherical aberration means that light rays from different radii on the primary mirror come to focus at different distances. The marginal focus is 38 mm below the focus of the innermost rays, which graze the secondary mirror. The paraxial focus is obscured. Currently, the adopted focus (not shown) is 12 mm below the paraxial focus.

No positions, orientations, or other adjustments of the primary and secondary mirror can produce the diffraction-limited images required by much of the HST science program. The center of a star image in visible light has a core of radius 0.1 arcsec containing about 15% of the light; 70% was expected. The rest is spread about in a complex halo of radius 3 arcsec. Since aperture diffraction sets the size of the image core, the size is smaller at shorter wavelengths. The size of the halo, on the other hand, is set by geometrical optics and is constant. (The *pattern* of the halo varies with wavelength because it is an interference pattern.)

Spherical aberration degrades the science capacity of HST. Good science *is* being accomplished with HST as it is, but many crucial investigations—including many of the original justifications for HST—are on hold until the problem is solved.
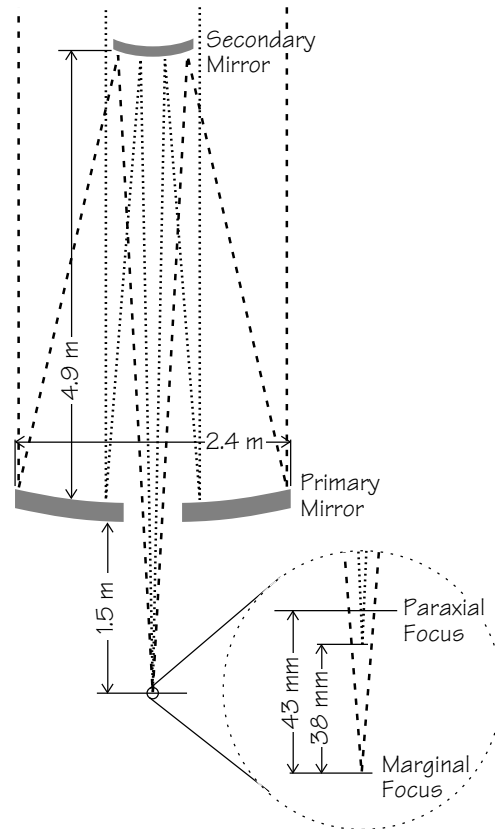
When the optical problem was announced, NASA began to seek solutions and develop a recovery plan. In the first phase, NASA focused on how to modify the scientific instruments already under development. These instruments are the Space Telescope Imaging Spectrograph (STIS), the Near-Infrared Camera and Multi-Object Spectrometer (NICMOS), and the second Wide Field and Planetary Camera (WFPC II), which NASA began to build in 1985 as a "clone" of the WFPC now in HST. NASA found it is feasible to correct these future instruments to compensate for spherical aberration. Based on this finding, NASA adopted an initial baseline plan to install the corrected WFPC II in place of WFPC on the first servicing mission in 1993, and later, on a second mission in 1996, to install STIS or NICMOS either to recover spectroscopic capabilities (in the case of STIS) or to add new infrared capabilities (with NICMOS).

This initial recovery plan of NASA restored faint source detection, one of the most critical capabilities crippled by spherical aberration. However, the plan delayed improving spectroscopy until the second half of the HST mission, and did not address full-resolution imaging at all. For these reasons, the HST Strategy Panel was formed in mid-August 1990 with a charter to search briskly for additional or alternative solutions.

In this second phase of NASA study, the HST Strategy Panel sought the best overall strategy to recover all primary HST science capabilities at an early time. The Panel did not adopt the WFPC II fix as a groundrule, but started "with a clean sheet of paper," and tried to identify and review *all* potential options to alleviate the negative effects of spherical aberration on the HST science program. However, the Panel's recommendations and deliberations were firmly rooted in the assumption that the schedule for the two second generation instruments, STIS and NICMOS, would be adhered to by NASA.

The HST Strategy Panel's findings and recommendations were presented to Dr. Riccardo Giacconi, Director of the Space Telescope Science Institute, on October 18, 1990. The Panel proposed a new program component as part of an augmented recovery strategy. The new component is the Corrective Optics Space Telescope Axial Replacement (COSTAR), a device to deploy corrective optics in front of the Faint Object Camera (FOC), High Resolution Spectrograph (HRS), and Faint Object Spectrograph (FOS). The strategy is to install both COSTAR and WFPC II into HST on the first servicing mission in 1993, and to fix the HST pointing problems. This strategy recovers essentially all the science capabilities expected at launch.

Dr. Giacconi endorsed the oral recommendations of the Panel and took the findings to NASA management. The Panel made a presentation at NASA Headquarters on October 26, 1990. In the following weeks, NASA conducted an intensive study of the feasibility and costs of COSTAR. In December 1990, NASA Headquarters authorized the implementation of the COSTAR program to proceed.

SYNOPSIS

THE HST STRATEGY PANEL HELD FOUR MEETINGS between mid-August and mid-October 1990. At these meetings, a wide variety of options for correcting spherical aberration were identified and debated. This report, as outlined below, presents the Panel's findings and recommendations.

The **OPTICAL PROBLEM** is now understood well enough to design and install a highly effective optical correction.

The **OPTICAL SOLUTION** is a pair of mirrors for each Science Instrument (SI) field of view. The first corrective mirror forms an image of the HST primary mirror on the second corrective mirror; the second corrective mirror has spherical aberration in precisely the same amount as the primary mirror—but with the opposite mathematical sign, thus cancelling the effect.

The **COSTAR** is the proposed device to carry and deploy the corrective optics for three scientific instruments, the FOC, HRS, and FOS. COSTAR would replace the High Speed Photometer (HSP).

The **POINTING** of HST must be improved to gain full value from the restored HST optical performance. The solar array "snap" that causes HST to lose pointing lock at day/night transitions must be fixed. The Panel further recommends that the operational parameters of the guidance system be adjusted to reduce jitter in the coarse tracking mode.

The **WFPC II** is being corrected with the same optical solution used in COSTAR. The Panel found that the alignment of the corrective optics is critical, which COSTAR can assure by special mechanisms. No comparable mechanisms exist in the original design for WFPC, and because WFPC II is a close copy of the original, the Panel recommends that the issue of WFPC II alignment be addressed with critical attention.

The **1993 SERVICING MISSION** can install the WFPC II and COSTAR. This currently planned mission can solve the spherical aberration problem for the SIs, fix the solar array disturbances, and replace other subsystems, as necessary.

The **RECOMMENDED STRATEGY** is to develop COSTAR on an urgent basis, continue WFPC II development with special attention to the alignment concerns, and improve the coarse track pointing performance by operational measures. Then, the 1993 HST servicing mission restores the scientific functionality expected at launch.

The **FRESH REASONS** to commit new resources to fix HST are abundant in the science program awaiting sharp images and precise pointing. This science program is the culmination of decades, even centuries, of maturing questions about the universe. It is also a program proposed largely by young astronomers, who need a restored HST to make the discoveries that will propel astronomical exploration into the twenty-first century.

The **APPENDICES** document the approach, options, background findings, and analyses of the HST Strategy Panel.

OPTICAL PROBLEM

THE OPTICAL TELESCOPE ASSEMBLY (OTA) consists of a 2.4 m diameter $f/2.3$ hyperbolic primary mirror and a 0.34 m hyperbolic secondary mirror separated by 4.9 m. In the design, the conic constants (negative of the squared eccentricities) of the primary and secondary mirrors were chosen to yield zero third-order spherical aberration and field coma on the focal surface. (This is the Ritchey-Chrétien criterion.)

The image quality of a precision optical system like HST is critically dependent on the correct alignment, or collimation, of the primary and secondary mirrors. For this purpose, the HST secondary mirror is designed so that it can be precisely positioned with all six degrees of rigid-body freedom.

To assist with the alignment of the OTA, three radial shearing interferometers, called Wave-Front Sensors (WFS), are located at the inner edges of the FGS fields of view. Their purpose is to provide measurements of the wavefront of light from a star. However, the WFS performance is badly degraded by spherical aberration. As a result, most of the diagnostic, focusing, and collimation efforts have relied on star images taken with the WFPC and FOC.

The on-axis image was expected to have wavefront aberrations totalling about 1/20 wave rms at 633 nm wavelength. The conventional definition of "diffraction limited"—Maréchal's criterion—is that the Strehl ratio of the peak intensity of the point spread function (PSF) to the theoretical limit in the absence of aberrations be greater than 0.8. This definition is equivalent to requiring that the wavefront error be less than 1/14 wave rms. According to optical metrology during the manufacturing of the OTA mirrors, HST was expected to be diffraction limited at wavelengths throughout the visible spectrum.

Figure 2a. The expected appearance of an HST visible-light star image.

Star images taken with the WFPC have a tight core containing about 15 percent of the light, a surrounding plateau containing most of the energy, and "tendrils" extending in apparently random directions from the core. The images are not consistent with those expected from a defocused and uncollimated telescope. However, the observations can be adequately reproduced by a computer optical model including spherical aberration.

The best fits to WFPC star images indicate the HST OTA has about 0.43 waves rms of spherical aberration at 633 nm wavelength. This error corresponds to a 4.6 μm optical path length error for marginal rays at the paraxial focus, a 1.6 arcsec diameter circle of least confusion, a 2.3 μm surface error at the edge of the primary, and a change of the primary mirror conic constant from the specified -1.0022985 to about -1.014.

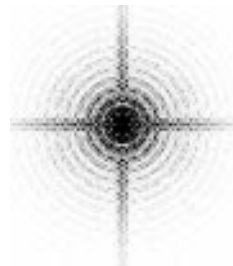The marginal focus is about 43 mm from the paraxial focus in the $f/24$ focal plane. The adopted fo-
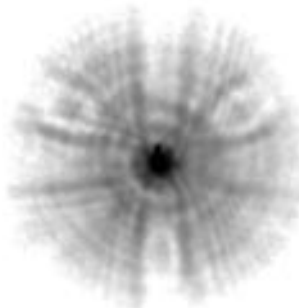

Figure 2b. An actual HST star image.

cus is in between, about 12 mm from the paraxial focus. Moving closer to the diffraction focus or to the circle of least confusion causes the outer halo to decrease in size, but the movement also causes the core energy to fall by about 40%.

The Strehl ratio of HST was predicted to be close to 0.9 at 633 nm. Due to spherical aberration, it is an order of magnitude lower, 0.10. Nevertheless, it is still an order of magnitude better than the value 0.012 for 0.5 arcsec ground-based seeing.

Because spherical aberration affects FOC images as well as WFPC images, the problem must be in the OTA, i.e., on either the primary or the secondary mirror. If the error were on the secondary mirror, it would cause coma with a linear dependence on field angle, which is not observed. For example, the expected 0.5 waves rms of coma at the FGS field position, 10-14 arcmin off-axis, would destroy the interferometer fringe visibility (S-curves). Also, the FOC, which has the largest off-axis distance (6.6 arcmin) of any SI, does not exhibit significant coma. Hence, the major—and perhaps the entire—error is on the primary mirror.

The NASA-appointed board chaired by Dr. Lew Allen, which has investigated the cause of the spherical aberration problem, has uncovered an error of 1.3 mm in the placement of a field lens in the reflective null corrector used in the manufacture of the HST primary mirror. This error alone would lead to the primary mirror having a conic constant estimated at -1.0132, or an on-axis wavefront error in the OTA of 0.40 waves rms at 633 nm.

Thus, the star images and fossil evidence agree on the value and location of the HST optical problem. The primary mirror has an incorrect conic constant, -1.013 or -1.014, instead of -1.00230. The estimated uncertainty in the "prescription" for the optical corrections envisioned by this report is the difference between the two estimates of the actual conic constant. Today, the error is about 10%. This is adequate to assure at least a 90%

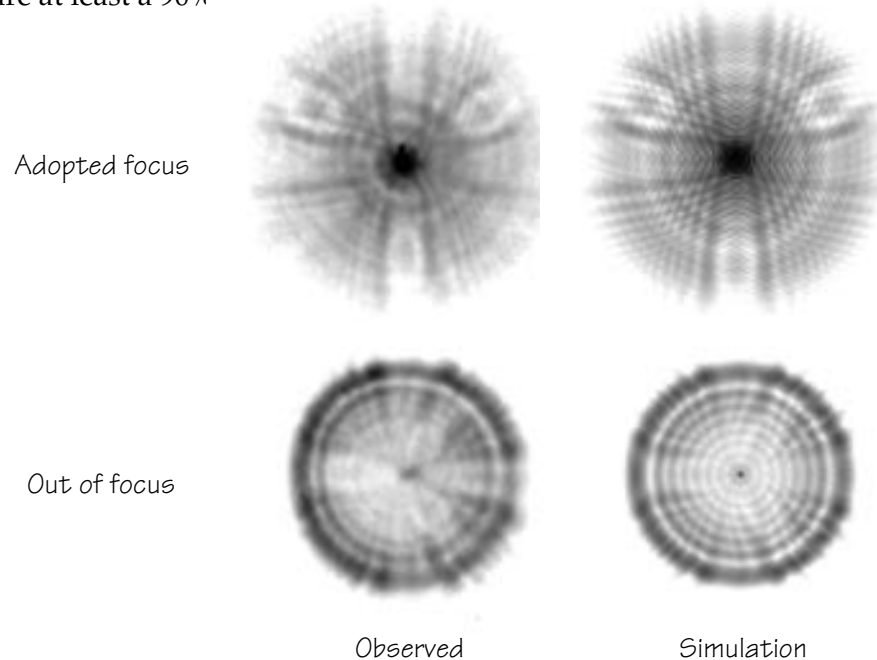Figure 3. Star images taken with the Planetary Camera near 5000 Å compared to simulations with spherical aberration. The close correspondence between computer model results and observations shows that the HST optical error is simply characterized.

OPTICAL SOLUTION

THE CORRECTOR SYSTEM proposed for the FOC, HRS and FOS consists of two mirrors, M_1 and M_2 (Figure 4). M_1 forms an image of the OTA pupil at M_2 and an image of the OTA field between the mirrors. The latter image is relayed by M_2 to the SI aperture. M_1 has the function of a field mirror and is a simple sphere. The correction of spherical aberration is done by M_2 and is fully equivalent to correction at the OTA primary mirror itself. This feature is unique among the SI-external optical corrector systems considered in this report. It has the advantage that the corrected
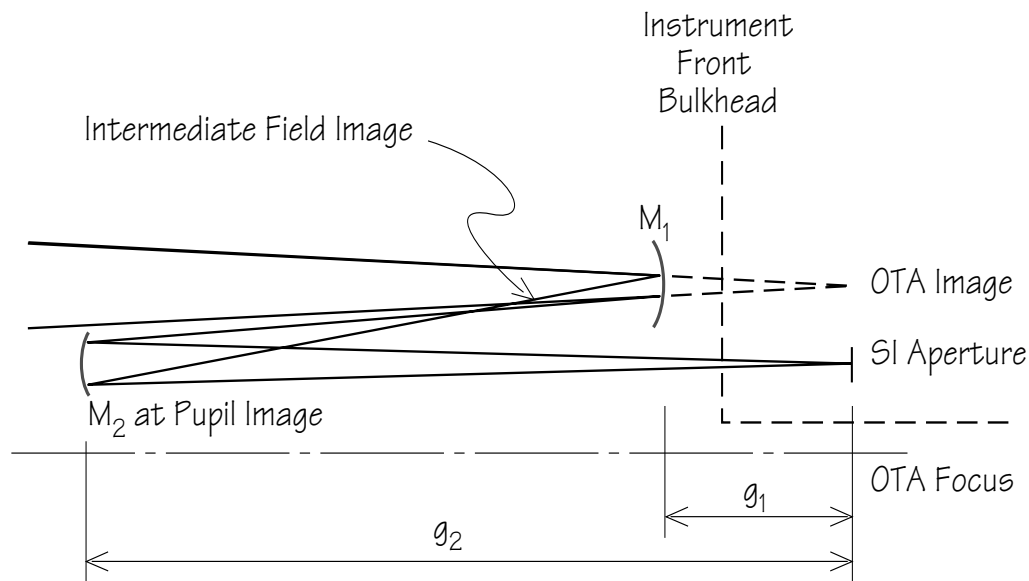
Figure 4. Schematic optical design for the COSTAR corrective optics, M_1 and M_2.

field is free of coma.

The principle of correcting the OTA spherical aberration at a pupil image is also planned in the new instruments, WFPC II, STIS and NICMOS. In STIS, this correction requires the addition of an internal corrector system before the slit. In NICMOS, an initial reimaging system is already present for the purpose of internal beam-steering, for which a beam-steering mirror is placed at a pupil image to direct the OTA beam to one of three cameras or one of three spectrographs. Beyond aspherizing this mirror, no changes in NICMOS are necessary.

In general, the output $f/$ ratio, F^*, of the corrector system is not the same as that of the OTA ($F_0^* = 24$). The ratio depends on the axial location of M_1 and is quantitatively given by

$$F^*/F_0^* = (1 - g_1/u) \Big/ (1 - g_1/g_2)$$

where u, g_1, and g_2 are the distances from the OTA focal plane of, respectively, the OTA exit pupil, M_2, and M_1. Evidently, F^* equals F_0^* only if $g_1 = 0$, i.e., when M_1 is placed in the OTA focal plane. This leaves two options:

(a) Place M_1 inside the COSTAR, which is the device for carrying and deploying the corrective optics. (COSTAR is described in the next section.)

(b) Place M_1 in front of the SI and accept an increase in $f/$ ratio.

The first solution is certainly the most attractive with regard to restoration of the HST potential. As will be shown below, it is quite suitable for the HRS and the FOS, but less so for the FOC. For the latter, option (b) seems more promising at the present time.
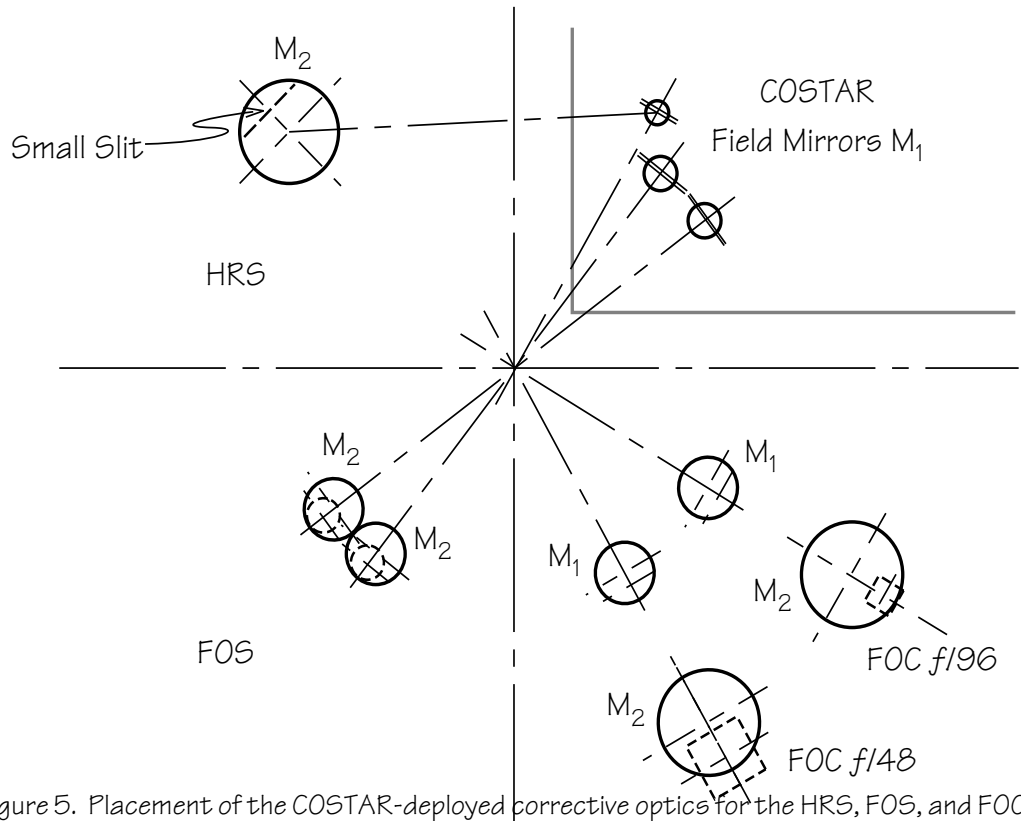
Figure 5. Placement of the COSTAR-deployed corrective optics for the HRS, FOS, and FOC.

A tentative arrangement of the corrector systems is shown in Figure 5. For the HRS, M_1 is placed a few centimeters inside the COSTAR on a line normal to the dividing line between the HRS and the COSTAR. In this manner the deviation angle of the chief ray at M_2 is kept to a minimum. For the FOS, two corrector systems are required. The two M_1 mirrors are placed directly opposite the FOS slits and at the same distance from the OTA axis. The FOC also requires two corrector systems. Here, each of the field mirrors is placed in front of the FOC enclosure on a line through the OTA axis and M_2.

For the present, the axial locations of the corrector mirrors have been chosen as shown in Table 1. In the HRS and FOS, g_1 has been selected to create an output $f/$ ratio, larger than $f/24$, in order to allow for possible alignment errors. In the HRS,

SI	g_1 (mm)	g_2 (mm)	F*
HRS	50	500	$f/27$
FOS	50	300	$f/28$
FOC	185	520	$f/36$

Table 1. Current values for the axial placement of the HRS, FOS, and FOC corrector mirrors, and the resulting output beam $f/$ ratios.

g_2 has been selected to place M_2 well below the WFPC pick-off mirror in order to prevent scattering by M_2 into the WFPC field. It is then necessary to prevent scattering from the WFPC mirror into the correctors, for which baffles should be provided to the extent possible. In the FOS, g_2 was selected to keep the two beams just separated. In the FOC, g_1 was selected to allow about 15 mm for deployment of the field mirrors in front of the enclosure. The choice of g_2 was about the same as for the HRS, and the result is an $f/36$ corrector output beam.

The principle used in the optical design of the corrector systems is to null spherical aberration at the center of the corrected field and to restore astigmatism at this point to the value for which the SI was designed. For correction of spherical aberration, a fourth-order asphere suffices. Restoration of astigmatism requires either a toroidal blank for M_2 or elliptical contours in the asphere. This conclusion is subject to further evaluation, including feasibility of fabrication.

Away from the center, the images become astigmatic. The tangential and sagittal image planes produced by the corrector are tilted with respect to those in the OTA. The largest contribution comes from M_2, which introduces a tilt difference between the tangential and sagittal image planes equal to twice the deviation angle of the chief ray. In addition, the tilt difference at the initial OTA image is not the same as is required at the corrector image.

In the HRS and the FOS, the fields are very small. The tilt effects are of little consequence and allow large margins for the deviation angle at M_2. Hence, positioning of M_1 within the COSTAR envelope is not critical. The aberration diameters corresponding to the arrangement in Figure 5 are given in Table 2 at the corners of the target acquisition fields. Evidently, the aberrations are small compared to the

SI	Corrector Field Size	Aberration Diameter in Corners
HRS	1.8 x 1.8 arcsec^2	0.04 arcsec
FOS	3.7 x 3.7 arcsec^2	0.08 arcsec

Table 2. The aberration diameters in the corners of the target acquisition fields of the spectrographs.

apertures, which leaves some margin for uncertainty in the actual value of the conic constant of the OTA primary mirror as well as other image imperfections. The optical solution fully restores the intended performance of the HRS and FOS.

In the FOC, the fields are much larger. It is essential to keep the tilt effects at a minimum; therefore, M_1 is placed in front of the FOC enclosure. The deviation angle at M_2 can be made as small as baffling allows. Placement of M_1 in the COSTAR would more than double the aberrations.

Table 3 shows the aberration diameters in the full FOC fields. Also shown are the aberrations in the "prime" area, i.e., the central quarter of the full field. Although significantly improved imaging is achieved with this optical solution, full restoration of performance is possible only in the central part of the field. The size of the well-corrected area will depend on how well the OTA spherical aberration can be characterized at the time the corrector systems are built.

	FOC Camera	
	f/48	f/96
Corrector f/ ratio	36.3	36.3
New camera f/ ratio	72.5	145
Corrector field size (arcsec^2)	29x29	15x15
Aberration diameter in corners (arcsec)		
- Full Field	0.25	0.13
- Central Quarter	0.08	0.04

Table 3. The FOC aberration diameters in the corners of the full field and central quarter

The optical alignment of the corrector systems is critical. The positions of the mirrors with respect to the SI fields can probably not be predicted to better than about 1 mm. On the other hand, the pupil image must be centered on M_2 within 0.1 mm or less to keep coma acceptably small. Hence, on-orbit alignment is necessary. The simplest method of alignment is to provide biaxial tilt or decentering for M_1 alone and to leave M_2 fixed. In this manner the pupil image can be centered on M_2. An additional condition to be met is that the chief ray must pass through the center of the SI slit. This is done simply by repointing the telescope.

In the above scheme the simplest alignment criterion is to check the image for the absence of coma after each iteration step. In practice, this may be a time-consuming procedure, especially with the HRS and FOS, which have limited imaging capability. However, if all is well, alignment has to be done only once after the corrector systems have been deployed. Hence, iteration efficiency may not be a serious concern. A different method is to use an array of three or four photodiodes at the periphery of M_2 to monitor centering of the beam directly. This would not be difficult to implement and could be the prime alignment device, with absence of coma serving for final verification.

The corrector mirrors M_2 are not moved in orbit. As a consequence, the chief ray into the SI may deviate a little from its intended course. To allow for this devia-

tion, the $f/$ ratio in the HRS corrector is set at $f/27$ and in the FOS corrector at $f/28$. In each case, this setting allows a deviation radius of 1.0 mm at M_2. Higher $f/$ ratios and larger positioning margins may be considered as the corrector systems are developed further. In stellar spectra, not much light is lost at higher $f/$ ratios as long as the core of the image remains smaller than the slit. For extended objects, the energy through the slit decreases inversely proportional to the square of the $f/$ ratio.

With regard to stability in orbit, it may be best to couple M_1 and M_2 rigidly and to execute on-orbit alignment by adjusting this assembly as a whole. In principle, the alignment procedure remains the same. Higher $f/$ ratios may be needed to allow for larger displacements of M_2. The trade-offs associated with $f/$ ratios will be the subject of future trade studies.

In principle, it would be possible to center M_2 on the original SI chief ray by first acquiring a bright, isolated object without the corrector system. Then M_2 is inserted and centered on the beam to the SI slit, guided by additional photo sensors on the OTA side of M_2. After M_2 is centered, alignment would proceed as before. For each M_2, an additional biaxial control mechanism would be needed. This extra complication does not seem warranted if vignetting of the beams inside the SI can be avoided by simply increasing the corrector $f/$ ratio. (Vignetting by structures inside the SI may give rise to strong ultra-violet (UV) scattering and should be avoided.)

To assure cofocality of the various SIs, provisions for corrective focus adjustments in the HRS and the FOS corrector systems are necessary. The FOC already has internal focus control, as will the STIS and NICMOS.

A key issue in the practicality of the corrector systems is the feasibility of fabricating the asymmetrical, aspherical corrector mirrors. This issue has already been investigated for STIS and a promising approach has been identified. The proposed fabrication method is to use a stress-polishing technique in which the blank is bent while a circularly symmetrical fourth-order asphere is generated by small figuring tools. A final large-tool lap assures surface smoothness. The STIS investigation is applicable because all corrector mirrors need an edge deviation which is equal and opposite to that in the OTA primary mirror. The deviation from the vertex radius is about 2.2 µm. The associated deviation from a best-fit sphere is about 0.6 µm. Generating such an asphere appears to be well within the state of the art. An alternative technique would be ion polishing, which might have the advantage of generating a smoother surface.

COSTAR

THE PROPOSED OPTICAL SOLUTIONS for the FOS, HRS, and FOC are pairs of small mirrors that must be carried to orbit, installed in HST, deployed above the SI entrance apertures, and aligned optically. The HST Strategy Panel discussed several possible ways of accomplishing this, and the best way by every measure is a device now called COSTAR. The concept is to use the existing STAR (Space Telescope Axial Replacement), which was built as a dummy replacement for any axial SI not ready for launch. COSTAR (Corrective Optics STAR) is this existing dummy—or a new, equivalent box— modified to implement the optical corrections for three of the four axial SIs.

The open volume in HST where the corrective mirror pairs must be placed is a cylinder about 25 cm along and 16 cm out from the OTA optical axis, located above the axial SIs and below the WFPC. Due to tight clearances, it would not be possible to attach the corrective mirrors on the FOS, HRS, or FOC as rigid extensions and then to reinstall the SI in the HST. Similarly, it would not be possible to attach the mirror pairs to the WFPC II pickoff arm and then to install it. Furthermore, those deployment schemes would be static, since neither the WFPC II nor the existing axial SIs have any provision for powering and operating external mechanisms to adjust the corrective mirrors. However, the positions of the corrective mirrors must be commandable by ground control to achieve their proper alignment and to allow them to be withdrawn if necessary. The COSTAR is the sure way to emplace the corrective optics in HST and to control their locations and orientations.

A schematic concept for COSTAR is shown in Figure 6a. An optical bench is located in a retracted position inside of the COSTAR during the Shuttle ascent to orbit. After astronauts install COSTAR in HST, a ground command would raise
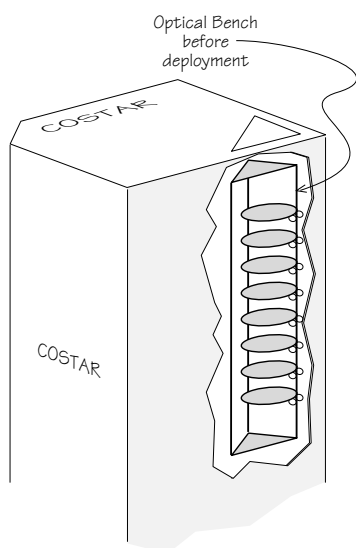

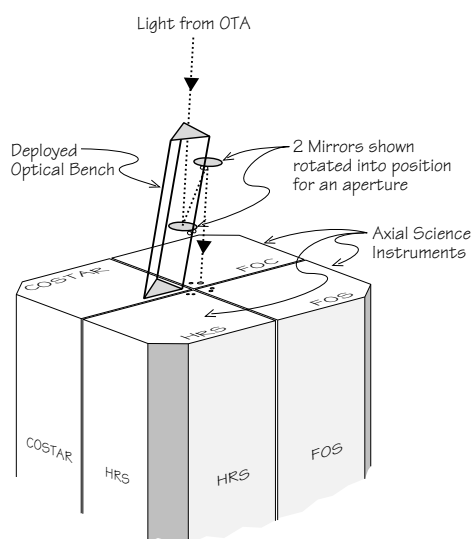Figure 6a. COSTAR carries the corrective optics in a retracted position.

Figure 6b. When COSTAR is installed in HST, ground commands activate mechanisms to deploy the corrective optics over the SI apertures of the FOS, HRS, and FOC.

the optical carrier, and the individual corrective mirror pairs would be deployed, as shown in Figure 6b. Based on subsequent tests, the proper placement of the corrective optics would be achieved and verified by commanding mechanisms to move the corrective optics.

Only a minimum of systems would be required in COSTAR, which, in addition to the corrective optics and a structure to hold them, would include (1) a system to raise the optics carriers into position, (2) a system to adjust and align the optics, (3) thermal control to maintain temperature, (4) a standard HST interface for command, control and power, and (5) electronics to drive actuators and determine their positions.

The existing STAR is a copy of the HSP without the detectors, optics, or electronics. It has standard HST interface wiring and an active thermal control system. As shown in Figure 7, an interior volume is clear and available for the optics carrier and requisite mechanisms. STAR is fitted with standard latches for precise installation into the HST. Whether STAR is actually used, or whether a new box is built for COSTAR, depends on the results of further NASA technical and cost studies.

Once the optical bench is deployed on orbit, it will be necessary to focus the new field image on the SI aperture and to align the primary mirror image formed by M_1 on M_2 precisely. This focus and alignment could be achieved with mechanisms by keeping M_2 fixed and adjusting the position of M_1 in tilt and translation along its optical axis. The information for these adjustments could come from SI data or from sensors mounted around M_2 to detect the edges of the primary mirror image, as discussed in the previous section.

The great beauty of COSTAR is that it fits seamlessly into both the design of the HST spacecraft and the philosophy of the HST program. Physically, COSTAR is like all the axial instruments, which are designed to be switched in and out of the telescope. The Shuttle interface for carrying replacement SIs to orbit is well defined.

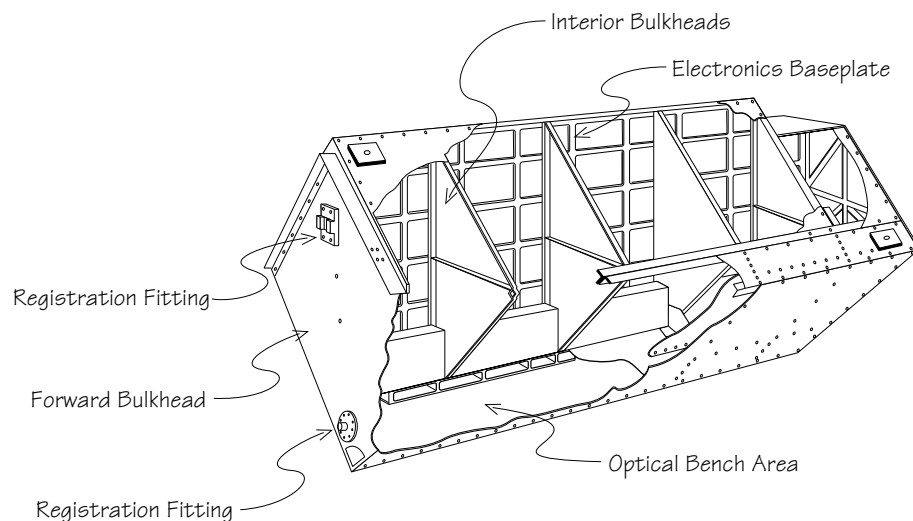

Figure 7. Cutaway of the existing STAR, which could be converted into COSTAR by adding optics, controls, and mechanisms.

Astronauts have practiced the SI installation procedures in watertank simulations for many years. Thus, although the need for COSTAR is a surprise, the basic concept of COSTAR is mature and comfortable from two critical standpoints: the space-

POINTING

THE UTILITY OF THE HST optics is no better than the pointing ability of the telescope. That is, jitter in the pointing will blur images as surely as optical aberrations. And, since the telescope points by locking onto guide stars, an inability to use faint stars for pointing would mean that some objects located in star-poor fields could not be studied. The vulnerable cases are among the most interesting HST targets: extra-galactic objects such as quasars and distant galaxies, which are visible precisely because there are few intervening stars that might serve for guiding. For these reasons, pointing-related recommendations are a critical part of the strategy for HST's recovery from spherical aberration.

COSTAR does not help the Fine Guidance Sensors (FGSs) in any way. In fact, COSTAR *depends* on improvements to FGS performance that must be achieved by other means.

The current performance of the guidance system during quiescent periods is summarized in Figure 8, which shows the rms jitter of the line of sight as a function of guide star magnitude in both fine lock and coarse track modes. The periods of day/night transitions during which the vehicle is excited by thermal shocks in the solar array are excluded from the data, since they are not representative of the performance of the guiding system under normal operational conditions. These day/night excitations are presently beyond the corrective capability of the pointing system, but it is anticipated that this situation will be improved by new flight software and, ultimately, fixed by replacement or modification of the solar array hardware.

Figure 8. Summary of current HST pointing performance during quiescent periods.

Fine-lock tracking performance is within specification (jitter < 7 mas rms), but the visibility of the fringe producing the error signal is poor, especially in two of the three FGSs. The result is a high rate of acquisition failures and frequent losses of lock unless the 2/3 aperture stop is in place in the FGS. This pupil stop reduces the influence of aberrations, especially the spherical aberration, but also reduces the guide star flux by almost one magnitude. In theory, the interferometric system used in fine lock should be insensitive to any axi-symmetric aberration such as focus and spherical aberration, but it is likely that the spherical aberration magnifies the effect of internal misalignments. In addition, residual aberrations in the telescope optics due to a still imperfect alignment of the secondary mirror can also contribute to the fringe visibility degradation.
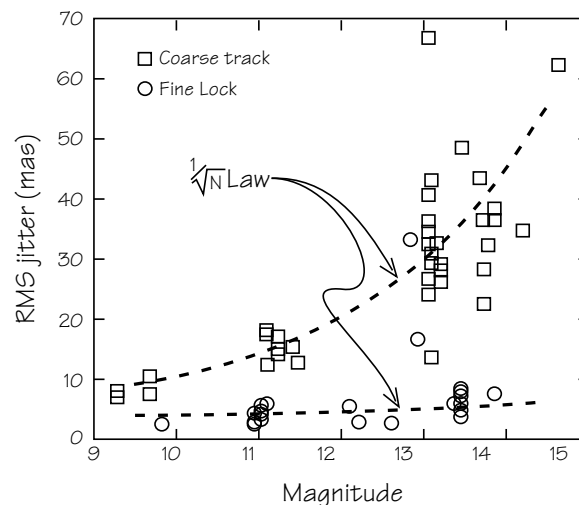
Stopping down the pupil, combined with the residual degradation in the fringe visibility, limits fine lock to guide stars brighter than 13th magnitude. The FGSs were designed to operate down to 14.5 magnitude in order to ensure at least an 85% probability of finding guide stars anywhere in the sky. With a reduction to 13th magnitude, adequate sky coverage can only be obtained for the axial SIs. In the case of the axial SIs (the lines of sight of which are off axis), rolling the spacecraft around the desired target significantly enlarges the area of the sky available to the FGS. For the WFPC, which is essentially on axis, the 13th guide star magnitude limit restricts the use of fine lock to the densest regions of the sky (in practice, galactic latitudes below 30°, or only 50% of the sky).

On the other hand, the coarse track mode is very robust. Acquisition in coarse track is, by and large, successful and does not require backup guide star pairs because of its insensitivity to binaries. Loss of guide star is rare, even in the presence of large vehicle oscillations such as those encountered during day/night transition. Unfortunately, the system is inherently sensitive to the spherical aberration and, as a result, coarse tracking accuracy is poor, about three times worse than originally predicted. Tracking is typically worse than 20 mas rms for stars in the 13 to 14.5 magnitude range, and often worse than 50 mas rms, which is not acceptable for many observations.

However, the coarse track system is still using pre-launch settings, which were established for diffraction-limited images and are not optimal for images with spherical aberration. Simulations of the coarse-track performance in presence of the spherical aberration indicate that two of the adjustable parameters (radius of nutation and gain) could be tuned to improve tracking performance. A test has been scheduled on the spacecraft to measure tracking performance as a function of these adjustable parameters. It is expected that a 20 mas tracking performance for the faintest guide stars (14.5 magnitude) can be achieved after optimization of the coarse track parameters. A jitter of 20 mas rms would degrade the light concentrated in the restored image core by 10%, which is scientifically acceptable.

The HST Strategy Panel makes the following recommendations with respect to pointing:

1. Coarse track should be optimized by adjusting the gain and the radius of nutation. The goal should be to reduce coarse tracking jitter to less than 20 mas rms on 14.5 magnitude stars.

2. Fine lock should be optimized in order to allow reliable operation on 13.5 magnitude stars. This might be achieved by a variety of steps, including: adjusting the various FGS internal parameters (in particular, the averaging time), refining the OTA optical alignment to minimize residual coma and astigmatism, or even adjusting the OTA optics to favor the FGS as a scientific compromise between guiding performance and image quality in the SIs.

3. The flight software should be enhanced to include mixed-mode guiding. Guiding with the dominant star in fine lock (pitch and yaw control) and the roll control star in coarse track would essentially be as effective as using the regular fine lock, and would greatly increase guide star availability.

4. Determine the reasons behind the poor fine-lock performance and modify the spare FGS as required. The fine-lock behavior should be analyzed theoretically to study the influence of external aberration and internal misalignment. These findings should be confirmed by experimenting on the spare FGS, which could then be modified appropriately to ready it for a future servicing mission.

5. Fix the solar arrays on the 1993 servicing mission so they no longer cause pointing disturbances.

WFPC II

THE WFPC is the main imaging instrument of HST. It reimages the central region of the OTA focal plane onto one of two possible camera modules, the $f/12.9$ Wide Field Camera (WFC) and the $f/30$ Planetary Camera (PC). A four-faceted pyramid is rotated by 45 degrees to switch between cameras. In each mode, the beam is divided by the four faces of the pyramid, then folded and reimaged by four small Cassegrain repeater cameras onto four thinned, backside-illuminated, 800 x 800 element CCD chips.

A backup instrument to WFPC is currently under construction. Intended originally as a carbon-copy emergency replacement for WFPC, WFPC II now offers a chance to correct the WFPC science capability for OTA spherical aberration.

Both WFPC and WFPC II contain optics that are designed to reimage the primary mirror pupil of the telescope onto or close to the secondary mirrors of the repeater cameras. On the WFC side, this reimaged pupil is already very close to the secondary mirror in the repeaters. An exactly compensating amount of spherical aberration can be added to each of the four WFC secondaries to give complete achromatic correction over the full Wide Field field of view with no extra reflections. In the PC, the reimaged pupil is not located directly on the repeaters' secondary mirror, with the consequence that unacceptable coma is introduced for off-axis field angles. However, by putting power on the folding flat mirrors that feed the repeater cameras, the pupil location can be moved onto the secondaries, where the correction would then be the same as in the WFC.

There are several technical issues that must be resolved in order to proceed successfully with this program of correcting WFPC II for spherical aberration. The main challenge is to align the reimaged pupil directly on the camera secondaries. This alignment must be achieved with a lateral tolerance of about 1% of the pupil diameter; otherwise, coma will be introduced. There are indications that the beams in the present WFPC repeaters are misaligned by typically 5-10%, which is clearly unacceptable when the corrective prescription is only the repeater secondary mirrors in WFPC II. To achieve the proper alignment, the cameras must not only be properly aligned internally, but the WFPC as a whole must be positioned accurately with respect to the OTA optics. A misalignment of 1% corresponds to a decenter of about 60 microns at the repeater secondary mirror, or to a tilt of the WFPC about the HST focal plane of about 1.5 arcmin, which is about 0.5 mm over a 1 m baseline. It will be extremely demanding, if possible at all, to position the camera passively to these tolerances on orbit. It is not clear that the instrument latch positions are known to this accuracy.

If passive alignment is not possible, it will be necessary to incorporate one or more movable optical elements into WFPC II to steer the beams onto the secondary mirrors. One possibility is to motorize the 45°-diagonal mirror that feeds the whole camera. Even with this, the internal alignments in the rest of the optics remain extremely stringent. On the bright side, it will be easy to diagnose any alignment error from the coma it produces, and the necessary correction to the tilts would be unambiguous. In-orbit alignment should thus be straightforward if the necessary mechanisms are provided.

The proposed change will also alter the f/ratio of the PC from f/30 to f/28. As the PC folding flat will now have power, the PC secondary mirrors will need to be changed in radius to compensate, and the camera despaced accordingly to achieve the correct focus. We do not regard this as a technical obstacle, but some mechanical redesign is required.

Although spherical aberration can be corrected, other imaging properties of the original WFPC design will not be recovered. Since the field light will now be out of focus as it hits the pyramid in the focal plane, light from a given object near the edge of a pyramid facet can strike two or more facets simultaneously and be directed to more than one camera. Multiple images may thereby be produced, each containing a fraction of the light of the star. The new images will be in the dead areas of the CCDs near the current pyramid boundaries, and images within ~3 arcsec of the edge of all chips will be vignetted. These areas will not be dark, as in the present design. Even if the extra image falls off a chip, stray light from it may be objectionable.

Some science programs that were planned with the WFPC may not be feasible with the WFPC II. For example, any program that uses the pyramid to occult a bright object to examine its environs may have trouble because of the aberrated light from the bright source. The function of the Baum occulting spot as an occulting spot will not be recovered, for similar reasons.

Preliminary evidence suggests small focus changes at least in the WFPC images on the timescale of a few orbits. They are equivalent to 5 μm OTA secondary mirror despace changes, and increase the outer rings in the image by about 5% in diameter. If these effects are internal to the cameras, then they may indicate a time dependence in the internal alignments. This possibility should be checked before a final design is adopted.

A final, major concern involves the pointing of the HST as relates to WFPC observations. The number of guide stars available for WFPC is reduced relative to the other instruments because of its location in the middle of the focal plane. Rolling the telescope thus does not increase the area of sky of the FGS pickles as it does for the off-axis instruments. Optimistic assumptions are (1) that reliable fine lock on 13.5 magnitudes stars will be achieved (presently, there is an operational constraint at 13th), (2) that the proportion of binaries is 10% (prelaunch, the proportion of unflagged binaries in the Guide Star Catalog (GSC) that were expected to fail to lock due to duplicity was 20% ; on-orbit data indicates that 10% may be a better estimate, but the guide stars were preselected by other means in certain cases), (3) that the spacecraft roll can be chosen freely (this seems unlikely because of terminator constraints, possible target position angle requirements, and the need to avoid the sun/moon/earth for faint targets), and (4) that the star density does not vary significantly at galactic latitudes exceeding 30 degrees. These assumptions imply a 30% loss of sky away from the galactic plane to fine lock with the WFPC. In reality the proportion is likely to be significantly higher. It is therefore a crucial element in fixing WFPC with WFPC II that coarse track be improved, and that fine lock on only one guide star be implemented (with coarse track to control the spacecraft roll). Without proper attention to these pointing concerns, the full potential of WFPC II will not be realized.

1993 SERVICING MISSION

THE FOLLOWING SCENARIO of extra-vehicular activity (EVA) shows how the recommendations of the HST Strategy Panel might be accomplished. The scenario includes a number of tasks not discussed elsewhere in the report, but which are also candidates for the first servicing mission. These are installation of a Rate Sensor Unit (RSU) to replace a gyro unit that failed in November 1990, replacement of the Engineering and Science Tape Recorders (ESTRs) and solar arrays (SAs), and possible internal work on FOC. (At the time this report is going to press, ESA is studying possible internal modifications to the FOC to be performed in orbit by astronauts.)

Assumed priorities for on-orbit replacement:

1: RSU #3. *To maintain redundancy in system mandatory for safe operations.*
2: WFPC II and COSTAR. *Required for realization of HST design optical performance.*
3: SAs (2). *Elimination of array-induced oscillatory attitude excursions.*
4: ESTRs (2). *Limited life; failure would degrade HST data handling capabilities.*
5: Internal FOC fix. *Only if formally requested by ESA.*

Efforts are currently in progress to develop a means for carrying WFPC II in its Science Instrument Protective Enclosure (SIPE), an axial SI in its SIPE and two Solar Arrays (SA) on the same Shuttle flight. Currently, however, this configuration presents insufficient clearance for the EVA crew members to remove and replace the SAs from the Solar Array Carrier (SAC), moves the Orbiter center of mass too far forward, and is at the high end of the permissible payload weight for the mission. In light of this, it is possible that an SA fix not requiring total replacement of both SAs might be developed since the electrical performance of the arrays is excellent and projected to remain so through the second M&R mission.

Any critical failure on the HST spacecraft has the potential for preempting the priority levels, probably as late as L - 3 months, although not without very strong justification.

This scenario assumes the availability of three 6-hour EVAs with two crew members (EV-1 and EV-2) in vacuum on each occasion, as well as a Remote Manipulator System (RMS) operator, and an IV-1 "coordinator." It does not presume to specify the means for achieving three EVAs, leaving that to the shuttle program office. Such capability will be required in support of space station assembly, however, and this would be an excellent opportunity to hone the technique. The following scenario should be taken as representative only; a myriad of other factors will influence the final detailed mission timeline.

EVA #1

Following egress from the airlock, the crew members would configure their tools to perform the RSU changeout task. To facilitate ground commanding after COSTAR installation, the HST remains powered up throughout this EVA, although the individual Orbit Replaceable Units (ORUs) will be powered down prior to breaking or making connections. While EV-2 is rigging the MFR in the RMS, EV-1 would retract the Fixed Head Star Tracker (FHST) light shields (3) from the cones and unlatch, open, and tether the -V3 Bay Aft Shroud doors. In order to provide access to RSU #3, EV-1 would then remove the -V2 cone from its FHST and stow it

temporarily out of the way. On the MFR, EV-2 would rig the Portable Work Light Assembly (PWLA) and proceed to remove RSU #3 (2 wing tab connectors and three bolts). With the assistance of EV-1, the new unit would be fetched from the ORU carrier and the old one secured. EV-2 would then install the new RSU, torque the bolts to specified values, remove the PWLA, reinstall the FHST cone, and collaborate with EV-1 in closing the -V3 Bay Doors.

While EV-2 extends the FHST light shields, EV-1 would start preparations for the WFPC exchange and COSTAR installation by unlatching and opening the +V2 Bay Doors, preparing the HSP for removal, and starting on WFPC preparation for removal. As soon as EV-2 is free, he or she would use the MFR to extract the old WFPC with EV-1 assisting. While they install it in the temporary stowage location on the back side of the ORU Keel Latch Support Structure, the FSS would be rotated ninety degrees clockwise (viewed from above) to line up the HSP for extraction. When in position, EV-1 and EV-2 would together extract the HSP and put it in a TBD temporary stowage location. Immediately thereafter, the COSTAR would be removed from the axial SIPE and completely installed. The +V2 Bay doors are left slightly ajar to facilitate checking the WFPC status lights later.

While both crew members are involved in putting the HSP into the axial SIPE, the FSS would be rotated ninety degrees counter-clockwise so as to again face the WFPC aperture forward in the Payload Bay. Ground command and telemetry links would be established. Following setup of TBD film or video cameras, extension of the COSTAR optical bench would be commanded and documented. In the event of a malfunction, the EV crew members might intervene in TBD fashion to extend the periscope manually.

Following successful verification of COSTAR functionality, WFPC II would be removed from its SIPE and immediately installed. The "old" WFPC would then be placed in the radial SIPE, the +V2 Bay doors would be secured, and the Payload Bay would be stowed for EVA termination. This is a full six-hour EVA, but it should be possible to do it all within one EVA with existing tools and the usual high caliber of crew performance.

EVA #2

This EVA is dedicated to replacement of both SAs with redesigned items that will not oscillate when encountering the rapid change in solar flux associated with the night-day terminator crossing. As this operation is not directly connected with the charter of the HST Strategy Panel and because the SAC) has not been fully specified yet, only a brief overview will be given.

Each SA is installed on the HST by means of a Marman (or manacle) clamp and three sets of ganged electrical connectors. Additionally, two sets of motor driven latch assemblies (with EVA manual override) restrain each SA during launch and reentry, both on the HST and on the SAC. Prior to removal, a portable grapple fixture will be installed on the array to adapt it for handling by the RMS. Subsequently the electrical connectors are broken and the loose cables secured. Then the Marman clamp is opened and both EVA crew members guide it into the RMS end effector, pre-positioned a few inches away. The RMS is used to move the SA to a temporary stowage location, and the basic task cycle is complete. The appropriate new SA is then removed from the SAC using the same operations, with the RMS positioning

it a few inches away from the open Marman clamp and latches on the HST. The reverse actions are used to install it, carefully torquing all fasteners to specified values and checking for precise angular alignment within the Marman clamp.

Subsequently the HST would be rotated one hundred and eighty degrees, and the process completed on the second SA. This process is greatly facilitated by use of the RMS for moving the awkwardly sized SA's around, but it can be done without one, albeit more slowly. There should be ample time in one six-hour EVA, especially if the RMS is fully functional, to complete the foregoing. If sufficient time remains, replacement of one or both of the two ESTR's could be completed. These units are located in System Support Module (SSM) equipment bays, and are installed using four bolts engaging keyhole slots and three low- torque electrical connectors. Payload Bay closeout is required at the end of every EVA.

EVA #3

If required, this EVA would first clean up any "left-overs" from the first two EVA's. Subsequently, the FOC would be removed to a temporary work location using the standard axial SI procedures. Once there, using a special Allen bit in an existing space-qualified power screwdriver, four screws and the entrance baffle would be removed. Next, an additional forty screws would be removed to allow the "upper channel cover" (cover on the edge of the FOC closest to and paralleling the V1-axis) to be pried off. This would be immediately replaced with a new one having apertures for installation of the corrective elements, and secured by a much smaller (e.g., ten) number of captive fasteners. Using these apertures as guides and access ports, the corrective optical assemblies would then be installed, engaging hard points in the upper channel proper and being secured to the new upper channel cover. Reinstallation of the FOC would follow standard axial SI procedures. The only difficulty foreseen is the possibility that some of the screw threads may already be stripped out, and the screws epoxied in place. These offenders would be cut off with a special tool carried for this purpose. Task complexity is assessed to be comparable to that of replacing the Coronagraph-Polarimeter Main Electronics Box on the Solar Maximum Observatory, successfully accomplished on the STS-41C mission.

On completion of the foregoing tasks EVA closeout of the Payload Bay for reentry would be commenced while the in-cabin crew grappled the HST for deployment. Following Maintenance Umbilical retraction and HST-FSS berthing latch release, the HST would be maneuvered to the appendage deploy position and the SAs and HGAs deployed. At the successful completion of these operations, the EVA crew members would be released from standby duty and allowed to return to and to repressurize the airlock. Barring a contingency EVA, this would conclude the EVA operations on the first M&R mission.

RECOMMENDED STRATEGY

THE HST STRATEGY PANEL *unanimously* makes the following recommendations:

1. *Solve the HST pointing problems.* That is, fix the solar array-induced loss of lock either with flight software or hardware changes, improve the fine lock performance to assure lock on 13.5 magnitude stars, and promote operation of FGS in coarse track with a jitter performance goal of 20 mas rms on 14.5 magnitude guide stars.

2. *Continue WFPC II development for the first servicing mission (in 1993).* Assure the proper alignment of the spherical aberration correction.

3. *Develop COSTAR to repair FOC, HRS, and FOS on the first servicing mission (in 1993).* Taken with the recommendations on pointing and WFPC II, COSTAR fills out a strategy to recover the initial science capabilities of HST fully at the earliest possible time.

The Panel's unanimous support for this strategy is based on these assumptions:

1. *WFPC II can be completed on schedule, even if design changes are necessary to assure proper optical alignment.*

2. *The pointing problems can be fixed.*

3. *COSTAR can be developed in time for the first HST servicing mission in 1993.*

FRESH REASONS

WHEN LYMAN SPITZER FIRST PROPOSED a great, earth-orbiting telescope in 1946, the nuclear energy source of stars had been known for just six years. External galaxies and the expanding universe were about twenty years of age in the human consciousness. Pluto was seventeen and Seyfert galaxies were three. Quasars, black holes, gravitational lenses, and detection of the Big Bang were still in the future—together with much of what constitutes our current understanding of the solar system and the cosmos beyond it. In 1990, forty-four years after its conception in a forgotten milieu of thought, Hubble Space Telescope is a reality. Is it still relevant? That the answer is a resounding "Yes!" is wonderfully instructive of the dynamic nature of learning, and of the "revolutions" in science caused by new instruments.

Revolutionary advances in science occur whenever we improve capabilities by an order of magnitude or open entirely new physical regimes for exploration. The 60-inch and 100-inch telescopes built at Mount Wilson in the early part of this century led to the realization that our sun is but one of billions of stars in a galaxy, which is itself imbedded in an expanding universe containing billions of galaxies like our own. When we first built radio telescopes, we found that the disk of our galaxy is filled with otherwise invisible clouds of cold hydrogen gas from which stars are born. Because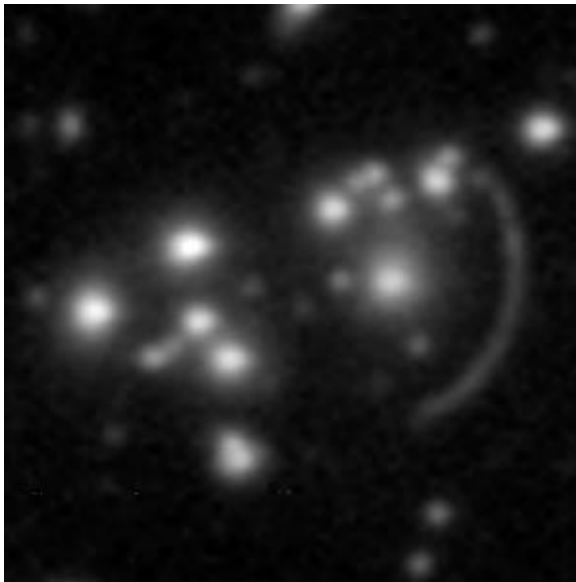 of greater transparency at long wavelengths, these radio telescopes showed us the center of our galaxy for the first time, and revealed in the centers of many other galaxies the release of enormous amounts of energy, often transported and deposited hundreds of thousands of light years away. The combination of the new radio telescopes built in the 1950s and the largest U. S. optical telescope, the Palomar 200-inch, led in the 1960s to the startling discovery of quasars or quasi-stellar objects (QSOs). QSOs produce up to 10,000 times the light of our galaxy from a volume only 50 to 100 times larger than the solar system. The extraordinary amount of energy released in the centers of galaxies and in QSOs led to the possibility that enormous black holes with masses up to a billion times the mass of the sun reside in the centers of some galaxies. Radio telescopes also led to one of the most fundamental discoveries of

Figure 9. A luminous arc in the cluster of galaxies Cl 2244-02. The rich cluster of galaxies is at a redshift z =0.328. The arc is most likely a high redshift quasar or galaxy. Its image is formed by the gravitational field of the luminous and dark matter in the cluster. A cluster mass equivalent to 100,000 billion suns is needed to bend and amplify the light of the background object. Photo courtesy of R. Lynds.

this century, the observation that the universe is filled with fossil light released by the cosmic fireball from which the universe was born. The nature of elementary particles established by the latest generation of ultra-high energy accelerators and advances in theoretical and observational cosmology are showing that the nature of the universe—and of matter itself—were determined in the first nanosecond of the cosmic fireball. Advances in technique have similarly propelled our understanding of the solar system. The planetary probes and flybys launched by NASA in the last thirty years have revealed worlds unimagined by scientists and layman alike. We found continent-spanning canyons and shield volcanoes higher than Everest and as large as Arizona on Mars, vulcanism on Io, Galilean satellites with icy mantles and possibly oceans beneath the ice, rings around Jupiter, and "spokes," braided rings, and shepherding satellites in Saturn's rings. The flybys showed hurricane force winds girdling the outer planets, and storm systems large enough to swallow the earth many times over. We found each planet and moon with an astonishingly different face.

These advances illustrate the power of qualitatively superior scientific instruments. In responding to the question of HST's continued relevance, perhaps it is enough to say that no other telescope has achieved the power expected of HST, nor will any until the successor to HST is built. HST's four "crown jewels"—grasp of faint objects, acuity of vision, specificity of spatial address in spectroscopy, and UV sensitivity—will be unsurpassed if they are relieved of spherical aberration. Great discoveries are truly to be expected of HST! In 1946, Dr. Spitzer wrote, "the chief contribution of such a radically new and more powerful instrument would be, not to supplement our present ideas of the universe we live in, but rather to uncover new phenomena not yet imagined, and perhaps to modify profoundly our basic concepts of space and time." More than four decades later this promise is one of the most compelling arguments for restoring the HST to its full potential.

If fresh reasons were needed to justify the costly but necessary restoration of HST, they are to be found in the 200 selected observing programs that today wait their turn for HST time. In the majority of cases, they await the correction of spherical aberration. Whereas the future potential for unknown discoveries may seem insubstantial or hard
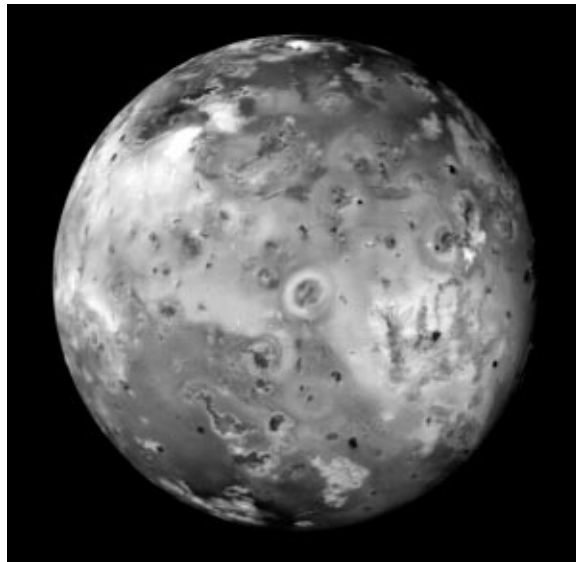

Figure 10. Io, a moon of Jupiter discovered by Galileo in 1610, This image by the Voyager space spacecraft in 1979 shows current geologic activity, including active volcanoes. The magnetosphere of Jupiter, filled with atoms and ions escaping from Io, is the largest structure in the solar system. Several HST programs are directed at Io and its influence on the Jovian magnetosphere.

to evaluate, the definite and specific questions we know to ask now with HST illustrate the general dynamism of science and, specifically, the scientific benefits offered by a corrected spherical aberration problem.

Lyman Spitzer began building HST's agenda when, in 1946, with extraordinary prescience, he discussed what might be done with a "large reflecting satellite telescope." He began by noting that the "powerful instrument envisaged here would help answer the questions whether space is curved, whether the Universe is finite or infinite." These thoughts were a direct antecedent to a Key Project on the HST today. Its intent is to measure how fast the Universe is expanding, and how fast the expansion is slowing down.

NASA's Announcement of Opportunity (AO) for Space Telescope in March 1977, which called for proposals to build the scientific instruments to conduct scientific investigations, listed by topic the outstanding problems in astronomy at that time. The list began with the "precise determination of distances to galaxies out to expansion velocities of order 10^4 km/s and calibration of distance criteria applicable at cosmologically significant distances." In simpler words, investigators should propose instruments and research to measure the Hubble Constant, H_0, to learn how fast the universe is expanding. Next, the AO called for determining how the expansion rate changes in time, as measured by the cosmological constant, q_0. The two numbers called for in these investigations state the age and fate of the universe; they will tell us when the universe began and whether or not the expansion eventually will be slowed, stopped, and then reversed by the mutual gravitational attraction of the mass contained within the universe. The AO called for the observation of galaxies in the distant past to establish how galaxies and their constituents of stars and gas evolve with time. Within our galaxy, the AO called for research on the dense cores of globular clusters to learn if they harbor massive black holes. The AO indirectly addressed the uniqueness of life on our planet by a call for "direct imaging and astrometric search for planetary companions of nearby stars."

The AO's ambitious science goals and many more are addressed in the observational programs of the Guaranteed Time Observers (GTOs). The GTOs are the Investigation Definition Teams (IDTs) who built HST's scientific instruments, the Observatory Scientists (OSs), who monitored the OTA/FGS/OCS development, and Interdisciplinary Scientists, who brought a wide range of science perspectives to the complex telescope project. The GTO programs today are founded in the vision of the 1977 AO. The WFPC team plans to observe clusters containing hundreds to thousands of galaxies at distances so large that the light we receive left the clusters 5 to 7 billion years ago. The WFPC pictures will show how these enormous aggregates of galaxies began to grow and emerge from their surroundings in the early universe and how the galaxies within these clusters have aged from then to now. Another WFPC GTO program will study the density of stars in the deep gravitational wells at the very centers of galaxies, looking for evidence of quiescent, massive black holes in normal galaxies. An OS plans to take WFPC pictures of QSOs in order to see the underlying galaxy in which the brilliant source of light is imbedded. Such pictures, studied together with spectra obtained with the FOS, will show the circumstances and types of galaxies that produce beacons bright enough to be seen across the universe. The FOS IDT plans to use a combination of WFPC images and FOS spectra to look for the gravitational signature of massive black holes

in the centers of galaxies. The FOC IDT plans to use a narrow occulting bar within their ultra-high resolution camera to block the light of nearby stars in order to suppress the glare of the star so they can search for disks from which planets could be forming. The HRS IDT will observe QSOs at large lookback times in order to detect intervening gas and galaxies in the early universe. The HSP IDT will use their instrument to investigate the structure of the ultra-dense neutron stars that form pulsars.

After the GTO science programs were finalized, the Space Telescope Advisory Council (STAC) recommended to the ST ScI Director that approximately one third of the competitive, or General Observer (GO), time on HST be reserved for "Key Projects," which would require approximately 300 hours each. This policy would ensure sufficient time to undertake fundamental surveys and in-depth investigations. Following the advice of the STAC, the ST ScI convened working groups to recommend Key Projects in seven different disciplines. The STAC then reviewed the recommendations and chose three programs to be designated as Key Projects in the Call for HST GO Proposals in 1985 The three Key Projects recommended by the STAC, and subsequently selected by the Time Assignment Committee (TAC) in a keen competition for HST GO time, include one envisioned explicitly by the AO. "Distances to Galaxies and the Determination of the Hubble Constant H_0," honors the memory of Hubble and his profound discovery of the expanding universe. The second Key Project, "A Medium Deep Survey," calls for taking WFPC images "in parallel," that is, wherever the telescope is pointing during an exposure by another instrument. The survey's primary goal is to study the shapes and stellar content of galaxies at large lookback times. Another aim is to discover pristine comets beyond Neptune. The Medium Deep Survey is one of the best bets for finding entirely new classes of astronomical objects. The third Key Project, "Quasar Absorption Lines," is aimed at establishing the distribution of gas and galaxies along lines of sight that traverse a large fraction of the observable universe before terminating at a high red-shift quasar.

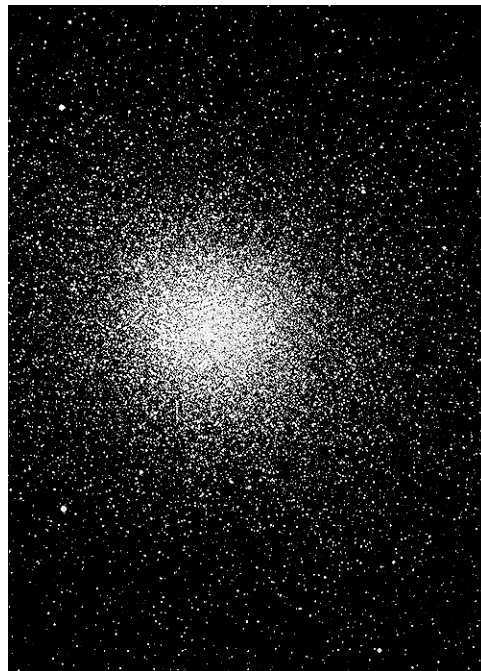The third component of HST science in the queue today consists of the other

Figure 11. A majestic globular cluster of several hundred thousand stars in the constellation Centaurus. X-ray pictures show that there are five X-ray sources in the field of this cluster. Each of these "low luminosity" sources is as bright in X-rays as the entire luminosity of the sun. The X-ray sources are most like close binaries consisting of a normal star and the compressed core of a dead star, i.e., a white dwarf. These binaries are "manufactured" through close encounters between stars and white dwarfs in the dense core of the cluster.

GO proposals submitted by astronomers in the community and selected by the TAC. The high expectations for HST were reflected in the deluge of high quality proposals for HST time: astronomers submitted proposals for 10 times the amount of time available! Only one proposal in six was selected. The selected programs are the most carefully planned and competitively selected proposals ever prepared by astronomers. They embody the very best science we can hope to do.

Much of the GO program is based on new discoveries and insights since the AO in 1977. Many of these discoveries were made possible by the charge-coupled device (CCD) technology developed by NASA for the HST and the Galileo mission. Using CCDs, astronomers have been able to double the redshift at which faint young clusters of galaxies can be found. The HST will be used to understand how these clusters form and change, and to see the processes in these ancient "nurseries" that led to present day galaxies. A very recent and exciting result has emerged from ultra deep images of clusters at modest redshifts. The images show that the clusters and the galaxies within the clusters act as gravitational lenses that amplify the light of more distant galaxies behind the galaxies. This finding raises the possibility that deep images of high redshift clusters will reveal the amplified images of even more distant primeval galaxies that are emerging from the cosmic chaos of the early universe. HST observations will undoubtedly play an important part in these studies, though the ideas are so new there were no proposals for this research in the first year of HST operation.

On another cosmic scale, systematic redshift surveys undertaken during the late 1970s and the 1980s revealed that, on large scales, galaxies and clusters of galaxies are found in large sheets and on the surfaces of enormous "bubbles" that are interspersed with giant holes or voids where there are few if any galaxies.

Our view of the universe as a whole has been profoundly and irrevocably changed by these unpredicted and unexpected findings. One of the GO programs scheduled for the first year will look for the spectral signature of gaseous matter and unborn galaxies in the giant voids by taking spectra of quasars that are behind the voids. Since the 1977 AO, intensive optical, ultraviolet, and x-ray studies have uncovered tantalizing evidence that massive black holes may be at the exact centers of many galaxies. Four GO programs intend to

Figure 12. The image shows 3C 273, the first quasar to have a redshift identified correctly (z = 0.158). The quasar is 10,000 times brighter than a typical galaxy, and the jet that extends 45,000 parsecs from the galaxy is 15 times brighter than an average galaxy.

use the HST to detect and "weigh" black holes by measuring the motion of stars and gas in their gravitational fields. During the 1980s, it became clear that the circumstances that led to the enormous energy release in the centers of galaxies are the presence of a black hole, which can be thought of as a cosmic engine, and gas in the nucleus that funnels into the engine through processes that are only beginning to be understood. The engines release energy through intense radiation, supersonic winds, and jets. Many GO programs plan to use the superb detail in HST images to understand how the powerful jets and flood of ionizing radiation from active galactic nuclei and quasars affect the surrounding galaxy.

Within our galaxy, there are several programs aimed at determining the chemical composition, temperature, density, ionization, and distribution of gas in the disk and halo. These are direct follow ups to insights and questions stemming from NASA's small, but extraordinary successful, International Ultraviolet Explorer. Radio observations during the last decade produced the unexpected result that there are a large number of pulsars, previously thought to be young neutron stars, in globular clusters that contain only ancient stars that were born at the time the galaxy formed. We now understand that the crowded cores of globular clusters are "factories" that produce binary stars through collisions of old stellar remnants, i.e., white dwarfs (a dense star the size of the earth) and neutron stars (an ultra dense star a few miles in diameter), with ordinary stars a little smaller than the sun. Evolution and expansion of the captured star forces gas onto the neutron star or white dwarf, thereby spinning it up and reactivating the pulsar or creating a bright cataclysmic variable star (CV). The HST is ideally suited for searching the dense stellar cores of globular clusters for the optical and ultraviolet counterparts of recycled pulsars and newly minted CVs. Infrared observations with NASA's very successful IRAS satellite and subsequent optical observations have recently detected dusty, gaseous disks around nearby stars. These disks may be the precursors of planetary systems. If so, there likely will be a revolution in our understanding of how and in what circumstances planets form around stars. Observations with the restored HST undoubtedly will be crucial for this emerging field.

Within the solar system, the wealth of images and measurements returned by NASA's grand planetary tours have raised many questions that can be answered only with the HST. Intensive GO studies of the atmospheres of the planets and their satellites will be firmly rooted in the data from the probes that tra-

Figure 13. An image from the ST ScI Guide Star Survey plates showing the spectacular spiral galaxy M51 interacting with the smaller galaxy at the end of one of the spiral arms. Tides induced by the small galaxy force gas to fall into a massive black hole at the center of M51. The gas fuels the black hole, which in turn powers a jet that is inflating a large bubble of ionized gas near the nucleus.

APPENDICES

GUIDE TO THE APPENDICES

The following appendices document the ideas, analysis, and considerations that shaped the HST Strategy Panel's final recommendations. The Panel began with a technical assessment aimed at elucidating the precise cause of the optical aberration in the HST and determining if the source of the aberration was understood with sufficient accuracy to allow us to design optical solutions confidently.

As discussed in the third chapter, *Optical Problem*, comparison of the conic constant for the primary mirror derived by the Allen panel with that derived at the ST ScI from analysis of HST images satisfied us that we knew enough to design corrective optics. The next step was a brainstorming session in which we tried to get all possible optical solutions on the table. Those wide ranging, imaginative ideas are tabulated and summarized in the *Options* section, Appendix A.

Each option has a one page summary that shows the idea in a form close to the way its originator presented it to the Panel. Many of the options quickly "fell on the floor" when critical discussion persuaded either the originator or a majority of the Panel that the idea could not work. Nonetheless, we include all the options here.

Our goal was to find solutions that would maximize the scientific return and that could be implemented with minimum risk to the HST. Several considerations factor into the evaluation of scientific return. The first is the quality of the optical solution. The best options will restore diffraction-limited images to all the instruments from the far ultraviolet (122 nm) to the near infrared (1000 nm). The essay on *Optical Analysis*, Appendix B, provides a framework for measuring the optical performance of the different options. Logical optical criteria related to this framework are used to put similar options into the lists on pages 35, 49, 55, and 67. At the end of the *Optical Analysis* essay, each of the options is evaluated.

Appendix C, *Aperture Masking*, provides an analysis of an option the Panel viewed as a potential insurance policy or "back-up" in the final strategy. As such, it is worth keeping in mind.

The benefits of a particular optical solution also depend on spacecraft pointing, the status of which is discussed in Appendix D, *Pointing Issues*. The value of an otherwise excellent optical solution for an individual instrument would be diminished if spherical aberration increased the pointing jitter in the FGSs or degraded the FGS acquisition limit to a magnitude where few guide stars could be found.

Still another important factor in maximizing scientific return is timeliness. Because the HST has a finite lifetime, early fixes yield greater science dividends than late fixes. The Panel discussed the possible technical problems that might slow the development of each option. These technical problems are often called out under the "Cons" in the summary for each option. Shuttle schedules and NASA's EVA procedures strongly affect timeliness and feasibility of implementing a particular option. These issues are discussed in Appendices E and G, *Shuttle Servicing of HST* and *Implementation Factors*.

The Panel spent considerable time discussing and analyzing the risk that a solution might not work, or worse yet, that implementation of an option might damage the telescope or further degrade the optical performance. The feasibility and risk of implementing the different options are considered in Appendix F, *Risk*

Management, as well as in Appendices E and G.

Although the Panel began with a "clean sheet of paper," the new science instruments currently being developed—and the NASA baseline of correcting spherical aberration within these instruments—were a central part of our discussions. Appendix H, *Second Generation SIs*, gives a brief description of these vitally important instruments.

At the end of our deliberations, we fashioned our final recommendation from a complex weighing of all the factors of benefit and risk. We chose, and this report recommends, the strategy we thought best.

APPENDIX	CONTENT
A	*Options*
B	*Optical Analysis*
C	*Aperture Masking*
D	*Pointing Issues*
E	*Shuttle Servicing of HST*
F	*Risk Management*
G	*Implementation Factors*
H	*Second Generation SIs*
J	*Glossary*

OPTIONS A1 - A12:

FULL-FIELD CORRECTION WITH FULL APERTURE

A1   Mechanical Deformation of the Primary Mirror

A2   Thermal Deformation of the Primary Mirror

A3   Overcoating the Primary Mirror

A4   Full Aperture Correction Plate

A5   Full Aperture Correction Flat

A6   Gas-Filled Correction Lens

A7   Secondary Mirror (SM) Replacement

A8   SM Reconfiguration of HST to an f/13.25 Ritchey-Chrétien

A9   SM Replacement and 2-Plate Corrector in Central Baffle

A10   1-Plate Corrector on SM and 2-Plate Corrector in Stovepipe

A11   Three Aspheric Plates in Central Baffle

A12   Double Cassegrain Relay in Central Baffle

A1 MECHANICAL DEFORMATION OF THE PRIMARY MIRROR

MECHANICAL FORCE is applied to the primary mirror to produce an elastic deformation that restores the mirror to its correct shape. In theory, this approach would correct the spherical aberration fully and introduce no new optical effects.

Currently, there are twenty-four force actuators behind the primary mirror; their original purpose was to fine tune the shape of the mirror on orbit. The actuators can produce a force of only about 10 lbs each, whereas about 200 lbs each would be required to effect a significant correction of the existing spherical aberration. (The HST primary is egg-crate stiff.) Furthermore, the locations of the existing actuators were selected to retouch astigmatism, and they are not good for spherical aberration. Only about one-half of the present error could be corrected even if an *unlimited* deforming force were applied through the existing actuators; limited to 10 lbs. each, they can correct only a few percent of the error.

A new force actuator system could take the form of an inflatable ring located near the outer edge of the mirror—an approach successfully employed on ground-based telescopes to correct spherical aberration.

PROS
 • The solution is achromatic.
 • The problem is corrected at the source, so the stop shift is zero.
 • The HST performance is fully restored.

CONS
 • A negative reaction mechanism is needed to relieve the force.
 • On-orbit access to installation areas is difficult, if not impossible.
 • Structure print-through is a possible danger even if bending *could* be achieved.
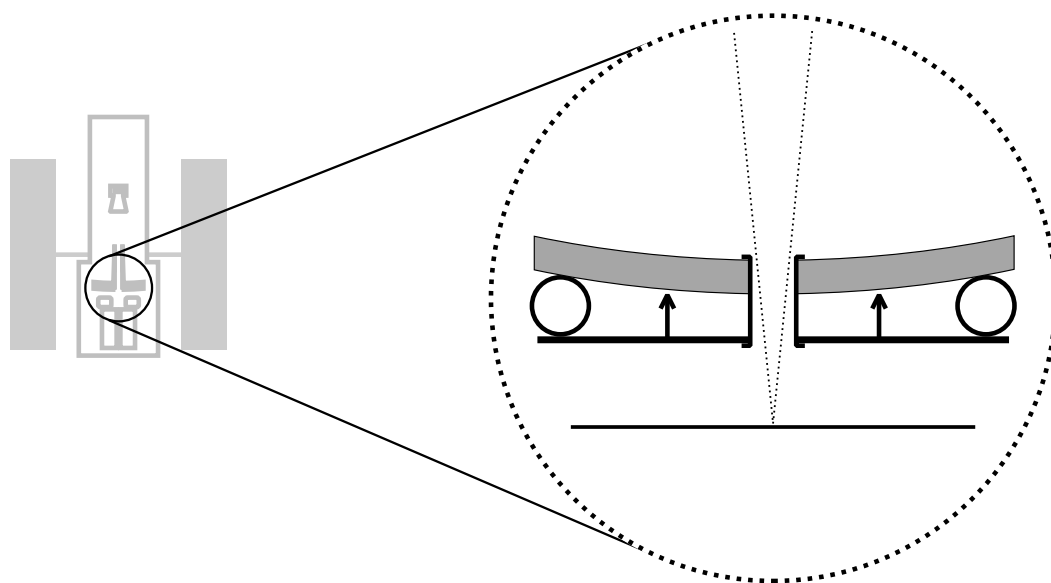
A2 Thermal Deformation of Primary Mirror

THE PRIMARY MIRROR is deformed thermally to compensate for the spherical aberration.

This can be accomplished by (1) changing the bulk temperature of the mirror (2) creating a temperature gradient in the mirror either radialy or axially, or (3) both.

Studies by Hughes Danbury Optical Systems (HDOS) indicate that turning off the heaters completely to change the bulk temperature of the primary does not deform the mirror enough to remove the spherical abberration.

Larger effects could be obtained by introducing an axial or radial temperature gradient across the mirror. Preliminary calculations by HDOS indicate, however, that these gradients must be very large in order to create the proper compensation.

PROS
- The problem is corrected at its source, as in A1.
- HST performance is fully restored.

CONS
- A very large, controlled temperature gradient is required.
- The infrared background is increased.
- A lot of power is needed.
- On-orbit access to installation area is difficult, if not impossible.

ISSUES
- Would the thermal balance of the focal plane be upset?
- What is the risk that facesheets on the primary will delaminate?
- Print-through is a risk as in A1, and may be exaggerated by thermal effects.
- Are variations of expansion coefficients in the blank a problem?

Heaters
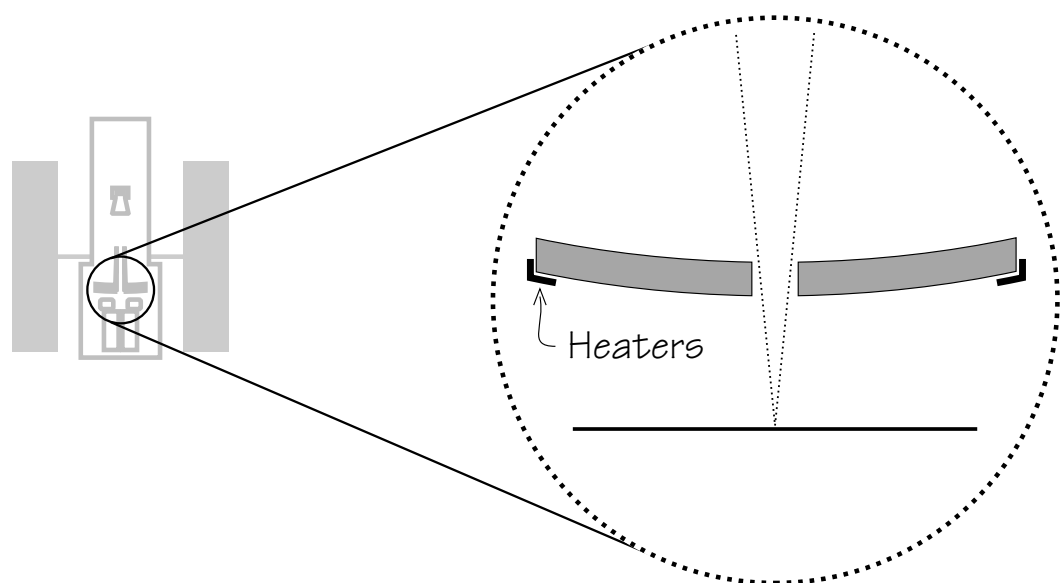
A3 OVERCOATING THE PRIMARY MIRROR

THE PRIMARY MIRROR is too flat at the edge by about two micrometers. A differential deposition of material on the mirror surface could restore the correct figure. A substrate of BaF, CaF, or LaF would be evaporated onto the mirror, followed by thin layers of Al and Mg F_2. This would be achieved by a device with material-laden heating coils mounted temporarily inside the telescope tube, presumably below the secondary.

PROS
- This is a perfect optical solution—as with A1 and A2.
- The process can be tested on the ground.

CONS
- Overcoating is irreversible.
- The likely roughness of thick coats makes UV scattering a problem.
- Peeling off the existing mirror coating is a risk.
- Evaporation requires high power (~36 kW).
- Because a complex device has to be temporarily mounted inside the HST tube, implementation is difficult.

ISSUES
- Is the ambient vacuum at the HST altitude adequate? (1 micro-Torr is required.)
- What is the adhesion, stress, and microroughness of different materials?
- How would the surface be cleaned of dust and prepared for coating (which normally requires a 10 milli-Torr gas discharge)?

Evaporation Heaters
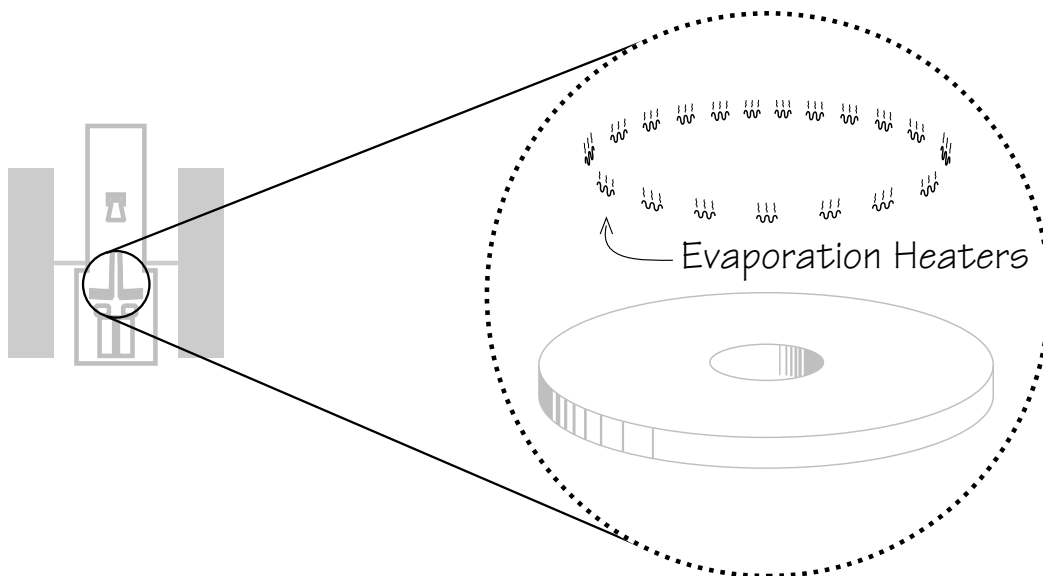
A4 FULL APERTURE CORRECTION PLATE

A FULL-APERTURE LENS is constructed from a mosaic of thin Mg F$_2$ plates, and attached to the OTA optical bench above the secondary mirror, thus correcting the light beam before it reaches the primary mirror.

PROS
- Small (y_{pr}/y) makes this the most favorable of all "plate" solutions. Field coma at 10 arcmin field radius is 0.12 arcsec; astigmatism with this solution is negligible.
- Substantially corrects all SIs and the FGS/OCS in one stroke.
- There are only 2 vacuum/glass interfaces.
- The aperture plate is tolerant to mis-positioning.

CONS
- Fabrication and testing is demanding.
- Segmentation produces diffraction effects.
- The plate is not achromatic; spherochromatism (variation of spherical aberration with wavelength) is appreciable over the entire spectral range.
- Installation is difficult.

ISSUES
- How is sufficient thermal control achieved to avoid thermal gradients across the plates?
- Is birefringence a problem?
- How is equal optical thickness ensured where plates butt together?
- Is the HST thermal background made worse?
- How is the assembly supported and protected during launch?
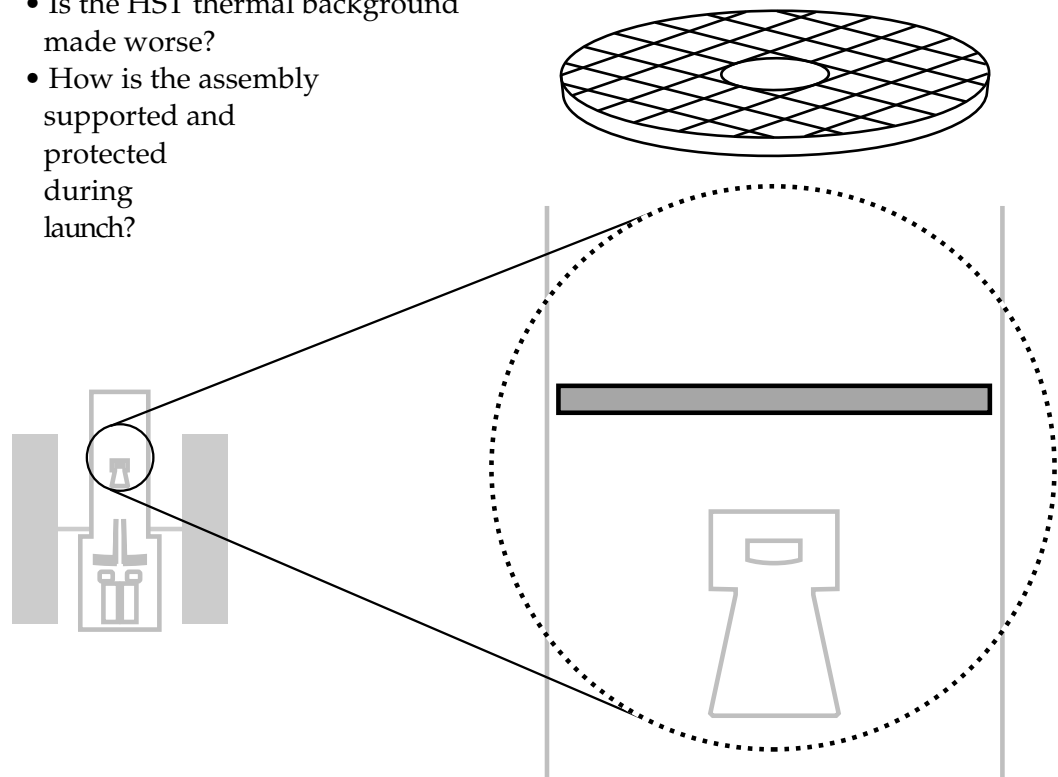
A5 FULL APERTURE CORRECTION "FLAT"

A DIAGONAL MIRROR at the top of the telescope corrects for spherical aberration. (y_{pr}/y) is about twice that of A4, and the field coma at 10 armin is about 0.24 arcsec. This solution could be combined with a coma-correcting plate in the central baffle.

PROS
- This solution is achromatic. (If a coma correcting plate is included, it would be weak and would generate only small chromatic effects.)
- Zero (or 2, with plate) vacuum/glass surfaces.
- The entire telescope is restored to the original optical design with no compromise in the far UV.
- Mirror technology is advanced in the direction needed for HST successors.
- The Flat is external and reversible.

CONS
- The increase in the moment of inertia may be intolerable due to effects on pointing and slewing.
- The Flat is large, difficult, and expensive to make.
- Introduced coma in FGS flields may be unacceptable.
- Testing the Flat probably requires a liquid mirror.
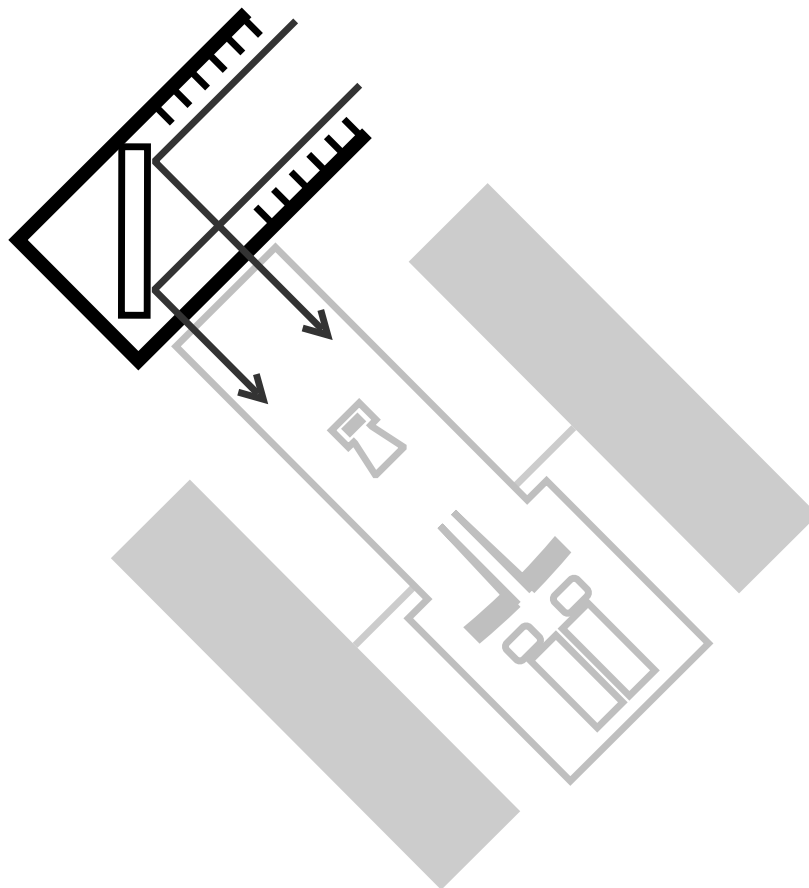
A6 Gas-Filled Correction Lens

THIS OPTION consists of a weak convex lens with gas as the medium. The positive lens is correct for compensating the overcorrected spherical aberration of the primary. The amount of spherical correction is determined by the gas pressure and the shape of the balloon, which must be optimized to minimize higher order corrections. (A spherical balloon is not acceptable due to a very large sixth power dependence on axial distance.)

Adjustment of the gas pressure allows adjustment of the spherical aberration correction for optimization at specific wavelengths.

The material of the balloon is selected for high UV transparency, good optical quality, and longevity. The UV wavelength coverage will be severely limited by the balloon material and by the gas, which must be containable. (N_2 is transparent down to 130 nm. Helium is probably not containable.)

PROS
- The lens is simple to install (and remove).
- Spherical aberration is compensated without introducing other unacceptable aberrations.

CONS
- The gas and walls of the balloon absorb the far UV.
- Unless the wavelength dispersion of the gas is negligible, the gas lens will introduce longitudinal chromatic aberration.
- Spherochromatism proportional to the dispersion ratio of the gas to MgF_2.
- Need to maintain pressurization.

ISSUES
- The walls of the balloon must have uniform thickness to a fraction of a wavelength, otherwise, unacceptable wavefront errors are introduced.
- Strong curvature results in slow convergence of the spherical aberration function. Negligible 5th order aberration requires small incidence angles.

UV Transmitting Balloon

Gas

M_2
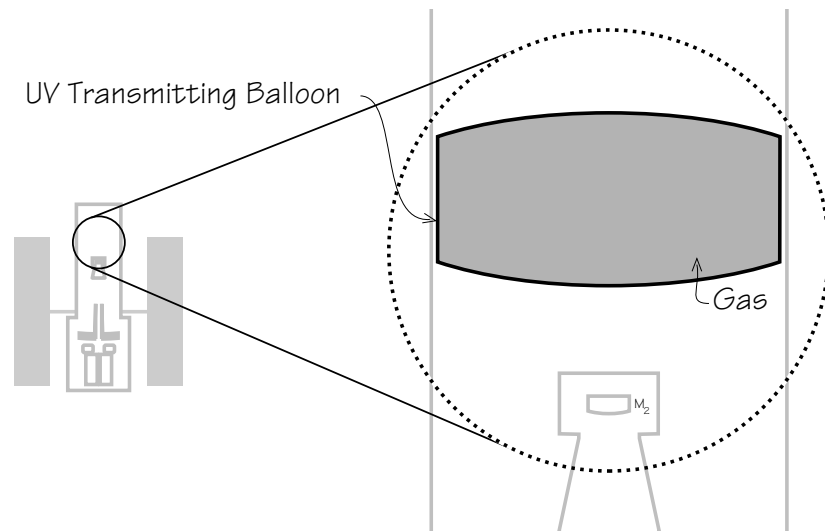
A7 SECONDARY MIRROR (SM) REPLACEMENT

THE SECONDARY MIRROR is replaced with a new one that corrects for spherical aberration. However, the correcting secondary mirror introduces coma—about 1 arcsec at 10 arcmin field radius. The introduced coma would be worse than the present spherical aberration for all instruments except WFPC, and it would be fatal for the FGSs.

PROS
- This reflecting solution is achromatic.

CONS
- (y_{pr}/y) is impossibly large for a single element solution. The field coma is unacceptable.

New
secondary
mirror
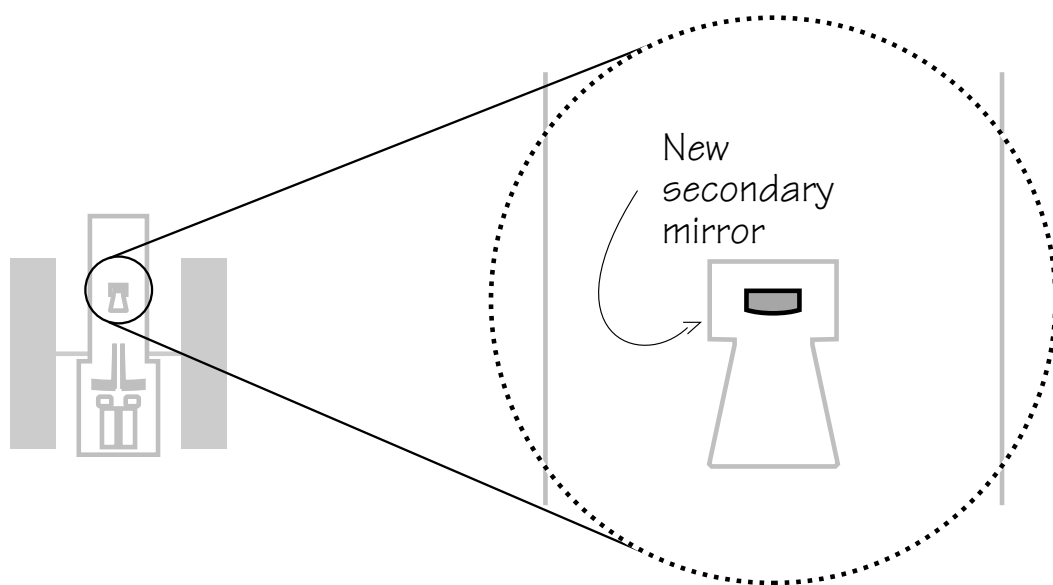
A8 SM Reconfiguration of HST to f/13.25 Ritchey-Chrétien

A NEW SECONDARY MIRROR installed on top of the old one could recover the telescope as a perfect Ritchey-Chrétien optical system. A system focal ratio of f/13.25 requires a hyperboloidal primary of conic constant -1.014, exactly what is in the existing telescope. WFPC II and subsequent new SIs would be designed to work at this focal ratio. Relay optics in front of the existing SIs could feed them at f/24. The existing FGSs would be used initially, but with the losses of an f/13.25 beam. Later, they could be replaced by a new system optimized for the new focal ratio.

The new secondary would have a 20 inch diameter and would be constructed of ultra-lightweight ULE honeycomb. It would be attached to the present secondary by a stage that would provide coarse adjustment. Fine adjustments would come from the existing mechanism.

PROS
- A correct aplanatic optical system is re-created by pure reflection—*this is a perfect optical solution.*
- The faster focus allows direct imaging with a wider field and higher resolution.
- The new secondary is not difficult to make.

CONS
- The new secondary mirror is very difficult to install.
- The new secondary mirror must be deployed with new WFPC, new SIs, and/or corrections for existing SIs.
- The new mirror is difficult to figure without a cross-check against the primary.

ISSUES
- Can the FGSs operate at f/13.25?
- Can an adequate f/13.25 to f/24 relay be designed for the FOC?
- Will the alignment mechanism work?
- Is the f/13.25 beam vignetted as it leaves the telescope?
- Can Barlow "lenses" be used to restore f/24 to the spectrographs?

New Ritchey-Chrétien secondary mirror
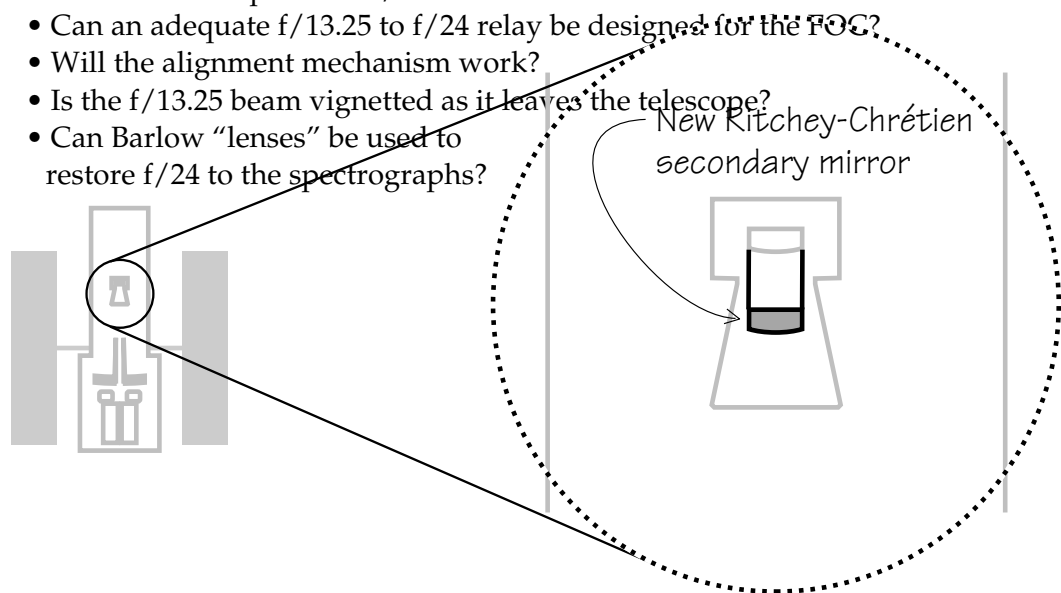
A9 SM REPLACEMENT & 2-PLATE CORRECTOR IN CENTRAL BAFFLE

A NEW SECONDARY MIRROR corrects the spherical aberration achromatically. It is mounted on the existing secondary and utilizes the existing secondary tilt and decenter mechanism.

To compensate for the coma and astigmatism effects resulting from the new secondary, and to match the new degree and sign of astigmatism to that of the HST as designed, spaced CaF_2 and MgF_2 plates are installed in the front of the HST central baffle (stovepipe). If changes in field curvature can be accomodated by SI and FGS changes, a single MgF_2 plate suffices. This would increase UV transmission and reduce ghost images.

PROS
- There is full, achromatic correction over the entire field, including the FGSs.
- The images are very good.
- CaF_2 and MgF_2 are available in the required sizes (~35 cm diameter).

CONS
- The two-plate configuration is opaque at wavelengths less than 130 nm.
- Installation of the SM and the 2 plates is very difficult.
- There are four (or two, for the single plate) glass surfaces, which lose light and produce ghost images.
- The new secondary cannot be tested against the HST primary mirror, and thus may not be figured correctly.

ISSUES
- Can a way be found to insert the new optics?
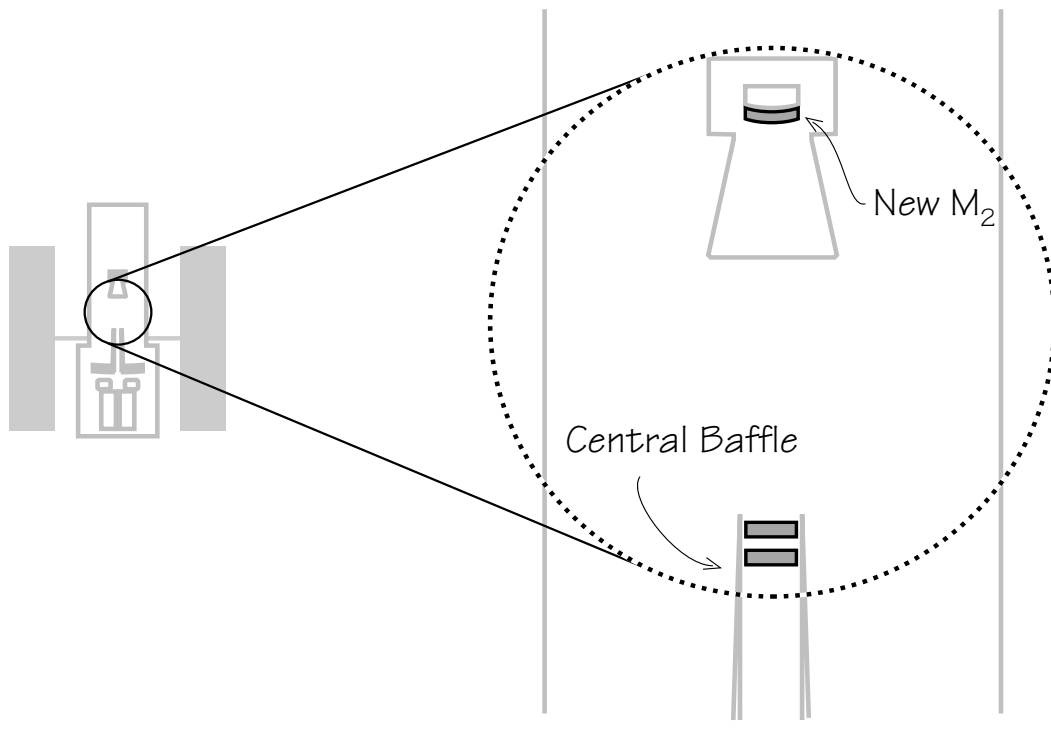
A10 1-Plate Corrector on SM & 2-Plate Corrector in Stovepipe

AN ASPHERIC FLUORIDE PLATE in front of the secondary mirror corrects the spherical aberration. This plate is movable along the optical axis to tune the spherical astigmatism correction to match the real HST spherical aberration at a particular wavelength.

Coma and astigmatism are adjusted by a single or double corrector shell inserted in the central baffle (stovepipe).

PROS
- The full field is corrected.
- Good images are achieved at all wavelengths.

CONS
- There are 8 (or 6 for a single plate in the stovepipe) glass surfaces, which reduce transmission (by factor ~2.5 at 130 nm) and produce 4 (or 3) ghost images.
- Installation of the moving plate mechanism in the secondary mirror baffle will be difficult.
- It will be difficult to install the 2-plate corrector in the central baffle.
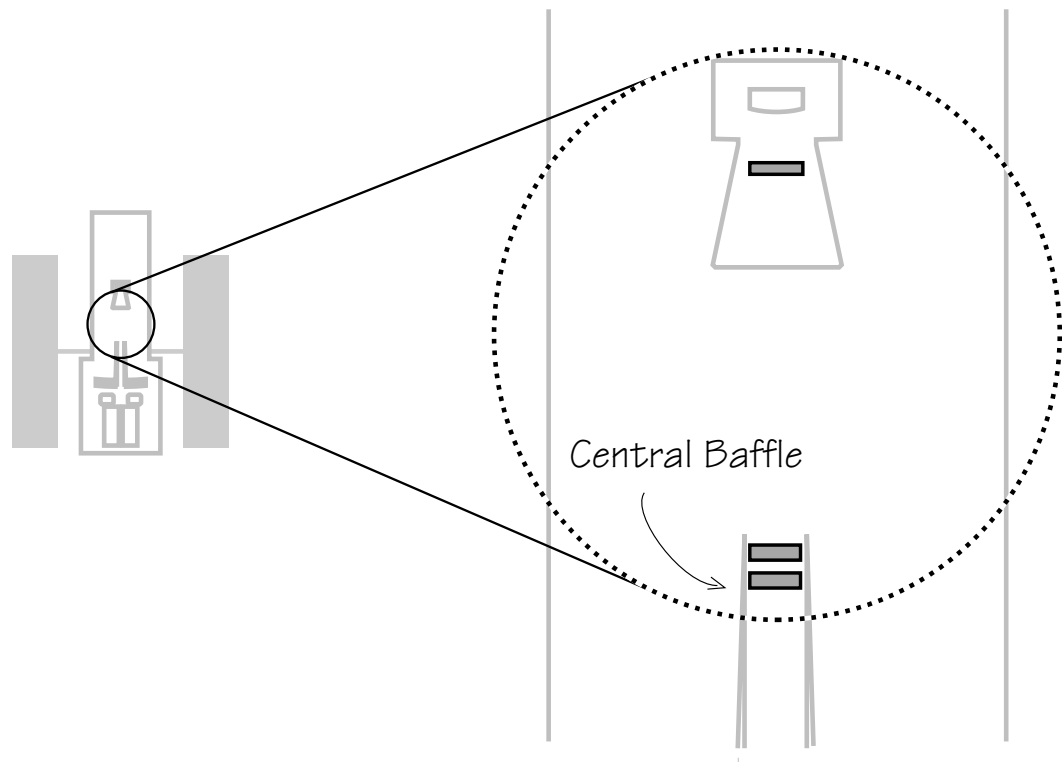
Central Baffle

A11 THREE ASPHERIC PLATES IN CENTRAL BAFFLE

THREE PLATES in the central baffle (stovepipe) can satisfy three optical constraints; consequently, spherical aberration can be corrected without introducing coma and astigmatism.

PROS
- The full field is corrected.
- Modest-size plates of MgF_2 and CaF_2 are available.
- Grinding the aspheric surfaces should require normal technology only.
- The plates can be manufactured, tested, and installed as a single fixed unit.
- The plates are tolerant to positioning errors.
- Standard test procedures can be used to test the complete unit.
- Spherochromatism might be balanced by moving the plates.

CONS
- There are 6 glass surfaces causing reflection losses and ghosts.
- Spherochromatism is the same as a single plate if one material is used. (Spherochromatism can be improved with two materials.)
- The individual plates are highly aspherical because they are relatively near to the focus, so that a group of three plates is required.

ISSUES
- How would the plates be installed in the stovepipe?
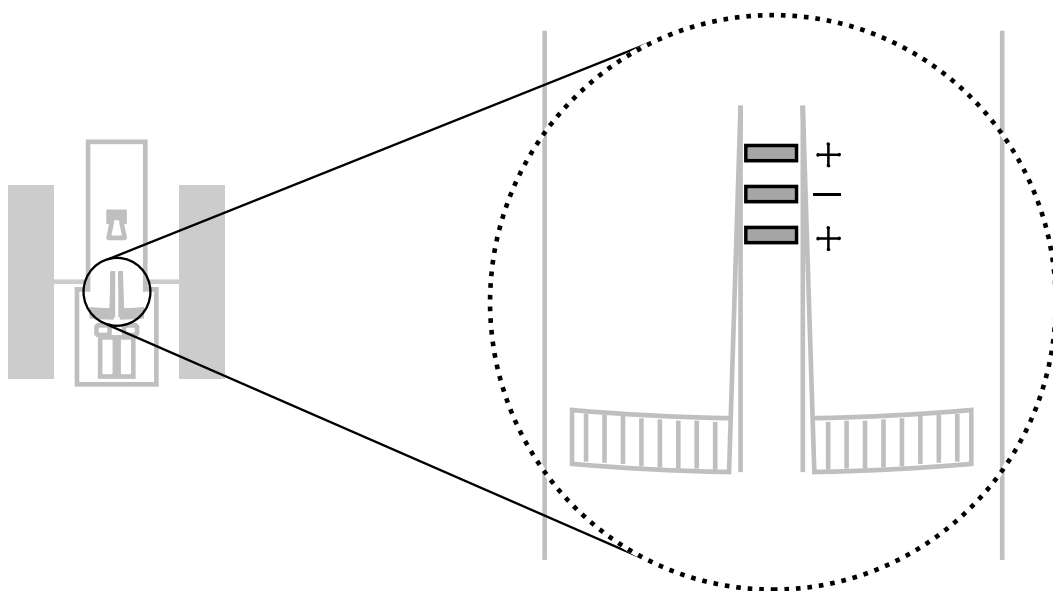
A12 DOUBLE CASSEGRAIN RELAY IN CENTRAL BAFFLE

A BACK-TO-BACK PAIR of Cassegrain optical systems can restore the f/24 beam fully without introducing additional coma and astigmatism. Unfortunately, adequate field requires unacceptable central obscuration. (However, this system may have application to individual instruments.)

PROS
- This is an achromatic solution.
- The double cassegrain is technically conventional.

CONS
- The obstruction/field-of-view problem is insoluble.
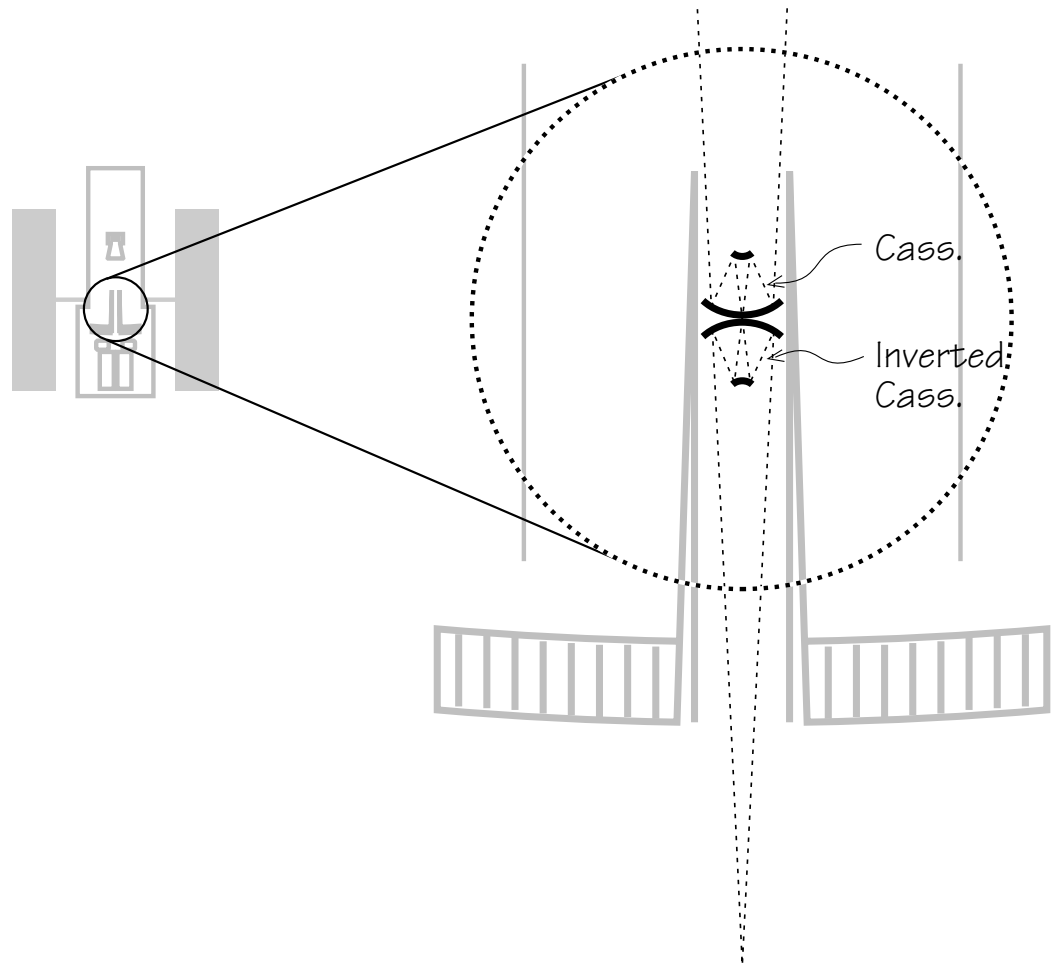
OPTIONS B1 - B4:

PRE-FOCAL PLANE CORRECTION ON THE AXES OF THE SIs

B1   Single Refractive Corrector for Individual SIs

B2   Double Refractive Corrector for Individual SIs

B3   Two-Mirror Reflective Correctors for Individual SIs

B4   Double-Cassegrain Relay for Individual SIs

B1 SINGLE REFRACTIVE CORRECTOR FOR INDIVIDUAL SIS

IN PRINCIPLE, a single, refractive, fourth-order corrector element can compensate for the OTA spherical aberration in each axial SI. Coma, which increases with the distance between the corrective optic and the telescope focus, limits the useable field. Also, tolerance to decentering decreases with the distance from the telescope focus.

   With regard to the different SIs, good correction is possible in the HRS (both slits simultaneously) and also in the FOS (separate correctors for the red side and blue side), with the exception of the acquisition fields in the latter. The fields in the FOC are too large to be covered.

PROS
  • This solution is simple.

CONS
  • The field of view is small.
  • The refracting lens limits UV transmission. The end-of-range transmission will be uncertain due to surface effects.
  • Strong spherochromatism produces a halo of light at the two ends of the spectral range (slitless mode) or reduces the throughput at these wavelengths in narrow slits. "Tuning" for specific wavelengths by axial displacement may be necessary.
  • The optical performance is sensitive to decentering. The deployment mechanism must be able to adjust the alignment.

ISSUES
  • Fabrication seems feasible, but will be far from easy.
  • Ghost images from internal reflections may be troublesome.
  • Installation on orbit will be difficult, and may not be reversible.
  • Positioning control will be necessary.
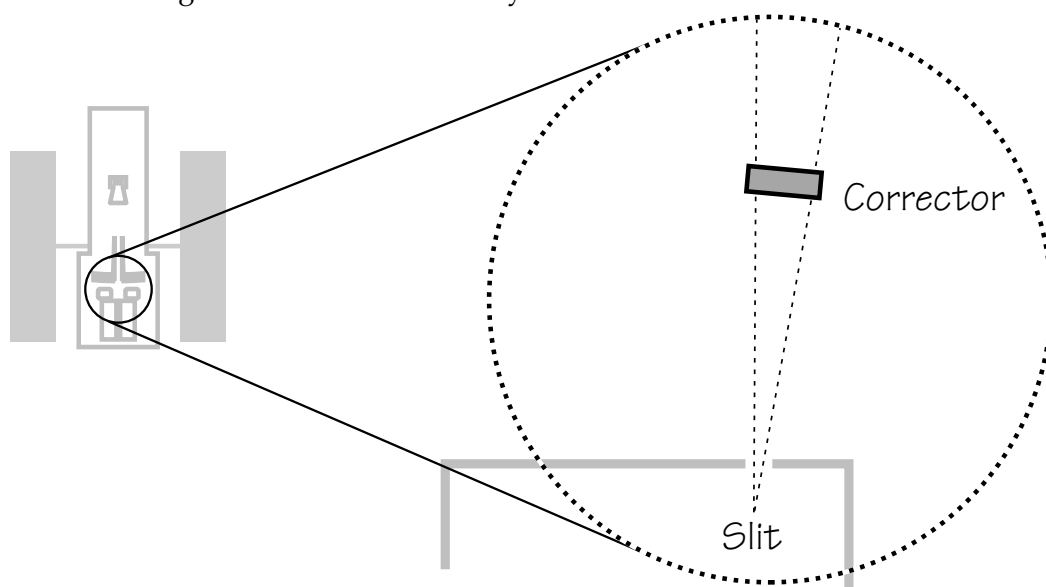
B2 Double Refractive Corrector for Individual SIs

A SECOND REFRACTIVE OPTIC can largely eliminate the field limitations that coma imposes on the single-corrector solution. However, the condition for zero coma drives the asphericities of the two plates to very high values.

PROS
- The double corrector has a wider field than a single corrector.
- The corrector is insensitive to small decentering errors.
- Achromatization is feasible in principle.

CONS
- The correctors have high asphericity.
- There will be excessive spherochromatism without achromatization.
- The two plates limit UV transmission, and there will be uncertainty near the wavelength limits (surface effects enter doubly).
- Achromatization probably will be very difficult in the UV.
- Multiple reflections will cause ghost images.

ISSUES
- Centering is not critical, but precise tilt control is necessary.
- Fabrication is very difficult in view of high asphericities.
- Finding an achromatic combination of materials will be difficult in the far UV.
- Installation on orbit will be difficult, and may not be reversible.
- Positioning control will be necessary.
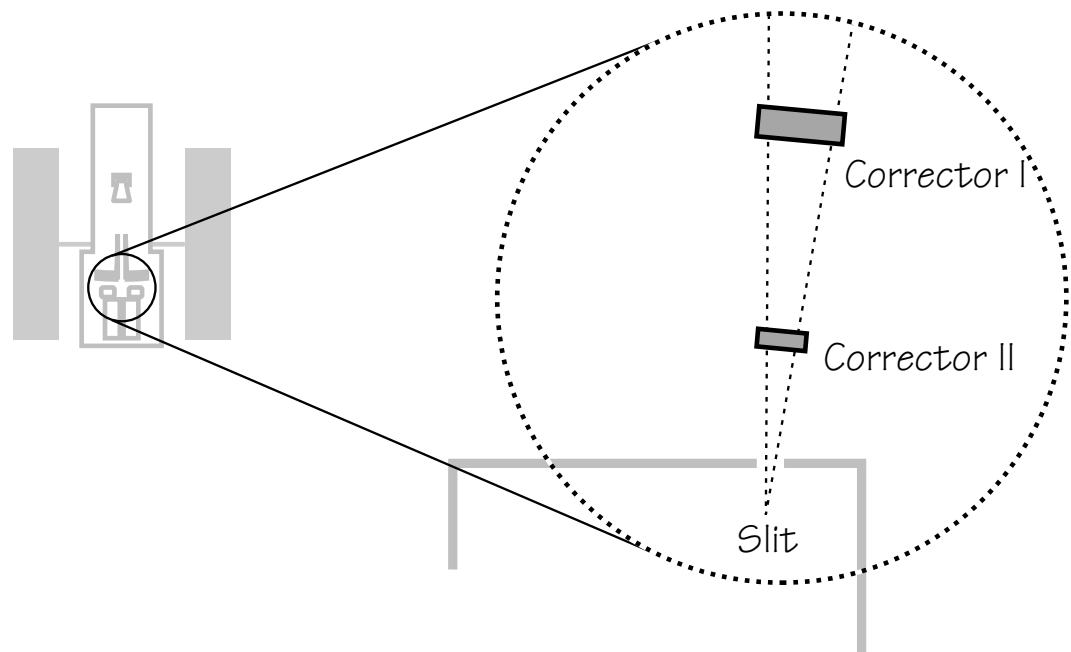
B3 Two-Mirror Reflective Correctors for Individual SIs

A TWO-MIRROR RELAY can compensate achromatically for the OTA spherical aberration in small fields. This is the fix recommended by the HST Strategy Panel.

An infinite range of combinations of asphericities is available to correct the spherical aberration. Among these, the aplanatic combination offers insensitivity to decentering and creates the largest possible field of view. The distance between the mirrors should be as large as possible to minimize residual aberrations.

The two-mirror corrector, optimized for the small slit in the HRS, covers the large slit very well. In the FOS, the "blue" and "red" beams each require a pair of mirrors.

PROS
- The reflective correctors are achromatic and coma-free.
- The two mirrors are insensitive to small decentering errors.

CONS
- The mirrors have high asphericity.
- The optical performance is highly sensitive to errors in the tilt.
- UV light may be scattered from the surfaces of the mirrors.
- The entire FOC field cannot be corrected.

ISSUES
- Fabrication, although considered feasible by at least one expert, is difficult.
- Residual surface roughness may cause UV scattering.
- On-orbit alignment control is necessary.
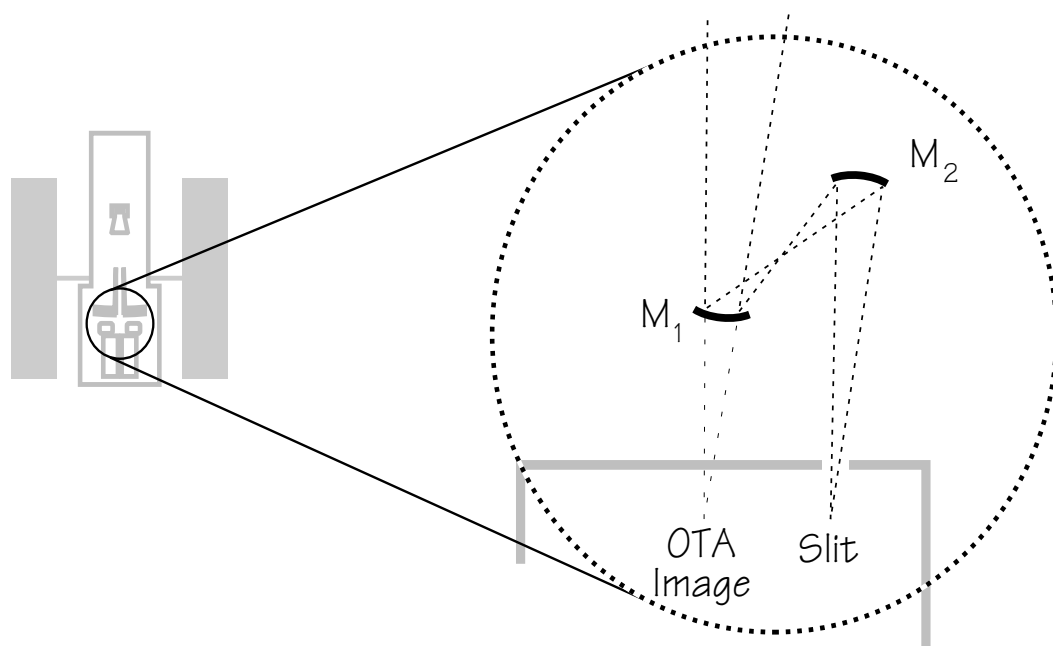
B4 Double-Cassegrain Relay for Individual SIs

A DOUBLE-CASSEGRAIN RELAY offers much wider correction potential than a two-mirror relay. It may be of interest for the FOC but needs more detailed study.

PROS
- The double-cassegrain has a wider field than the two-mirror relay.

CONS
- There are reflection and scattering losses on four surfaces.
- There is a large central obscuration, which will lose light and produce strong diffraction rings.

ISSUES
- There may be strong tilt and decenter sensitivities.
- Installation on orbit will be difficult, and may not be reversible.

Cassegrain I

Cassegrain II

Slit
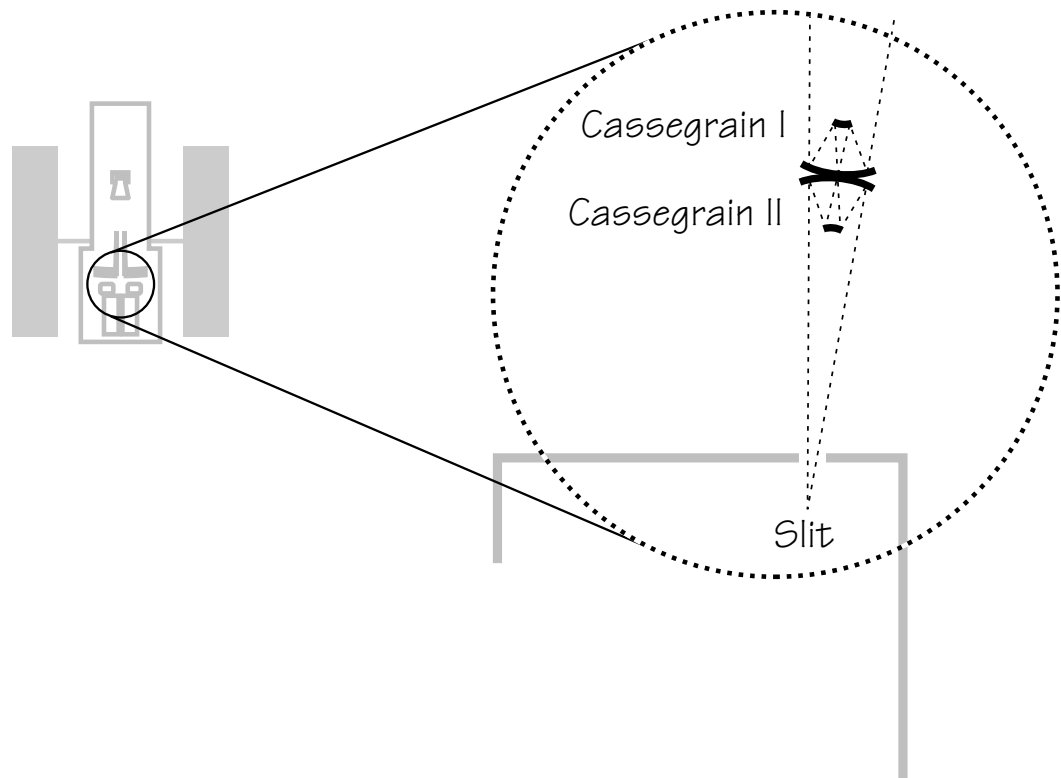
OPTIONS C1 - C9:

POST-FOCAL PLANE CORRECTION IN THE SCIENCE INSTRUMENTS (SIs)

C1   FGS II

C2   Modification of FGS I

C3   Thermal Fixes to SIs

C4   Modify WFPC II SMs

C5   Modification of the FOC: Refractor in Upper Channel

C6   Modification of the FOC: Refractor Near Exit Pupil

C7   Modification of the FOC: Change Optical Head Unit

C8   NICMOS Internal Corrector

C9   STIS Internal Corrector

C1 FGS II

BUILD AND INSTALL in the HST a new FGS system with optics that correct spherical aberration. The new FGS would use CCD quadrant detectors rather than Koesters prisms.

PROS
- The new FGS would substantially fix present pointing problems, except those related to terminator crossing.

CONS
- Installation of 3 FGSs would require multiple shuttle launches.
- Developing new FGSs is a major effort.

Koesters Prism
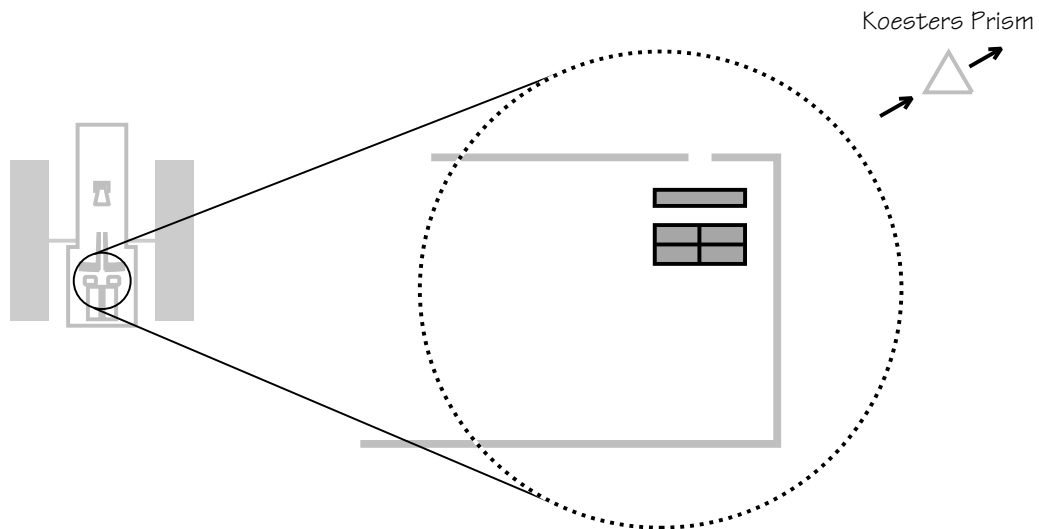
C2 MODIFICATION OF FGS I

ADD A CORRECTOR PLATE to the optical train of the existing FGSs to correct spherical aberration.

PROS
- This is a small, technically simple fix that could be tested on the spare FGS.

CONS
- Installing the correctors requires returning the FGS to the ground.

ISSUES
- No study has been made and the optical performance is uncertain.
- Mechanical feasibility is unknown.
- Is on-orbit replacement feasible?
- Can guiding strategies be developed which will work with only one good FGS? (Either use FHST for roll or use the good FGS in fine lock and the other in coarse track.)
- Will this approach correct the OCS along with the FGS?
- Is FGS performance more dependent on alignment than spherical aberration?
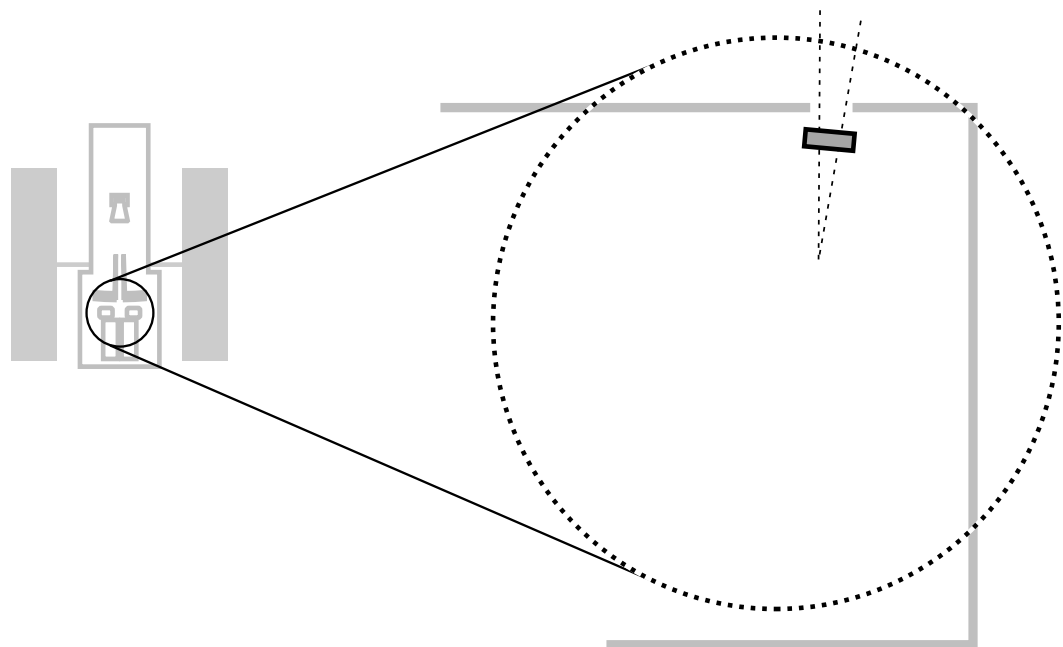
C3 THERMAL FIXES TO SIS

THERMALLY DEFORM the optical elements in the existing SIs in order to correct spherical aberration.

PROS
- Possibly could be done now.
- The risk is low.

CONS
- Does not correct the focal plane.
- No optical element has been identified as a candidate.

ISSUES
- Does this approach even work for spherical aberration correction?
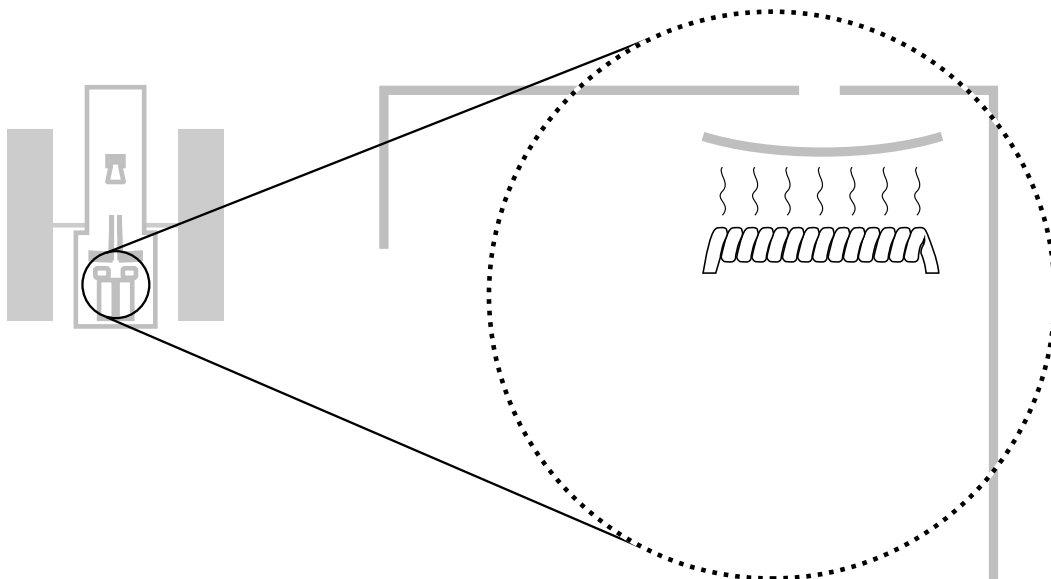
C4 MODIFY WFPC II SMS

REFIGURE THE EIGHT secondary mirrors inside the WFPC II to correct for spherical aberration.

PROS
- This may be simple to implement.
- The entire field of both WFPC II cameras is corrected.
- This solution produces good images at all wavelengths.

CONS
- Only the WFPC is corrected.
- A Baum spot is not possible.

ISSUES
- What is the image quality at the edges of the field?

C5 MODIFICATION OF THE FOC: REFRACTOR IN UPPER CHANNEL

IN ORBIT, the FOC is opened up and a replacement FOC upper channel is installed; an external cover with baffles is equipped with refractive corrective optics.

PROS
- The image quality is fully recovered over a reduced field of view.

CONS
- Coma-like aberration and spherochromatism are present.
- The coronograph and spectrograph modes are not corrected.
- Correction of the High Resolution Apodizer (HRA) is marginal.
- Tolerances and positioning are critical.

ISSUES
- Is removal and replacement of the upper channel feasible in orbit?
- Is the FOC field of view too large for good correction?
- How much UV transmission is lost?
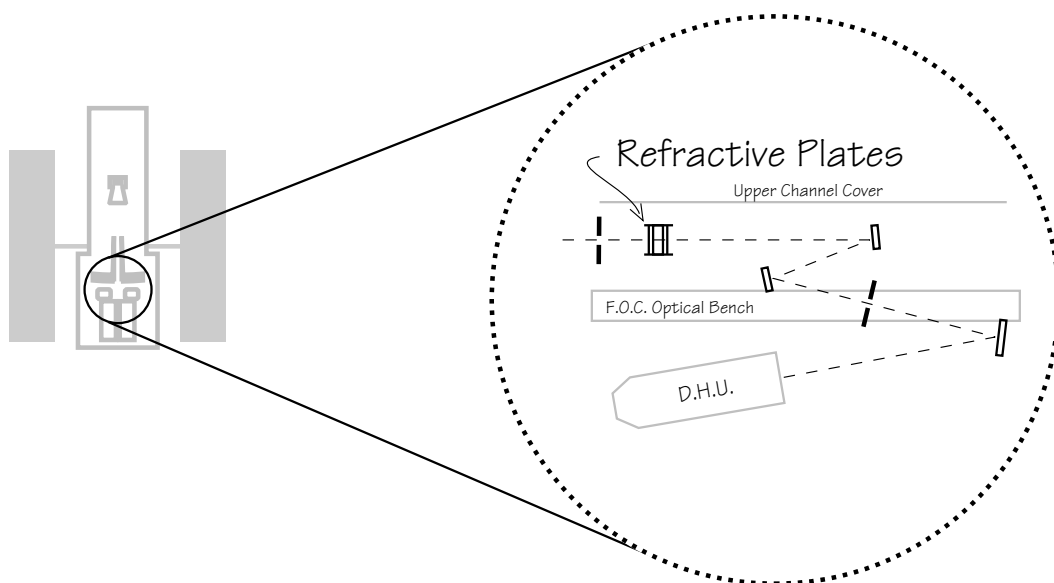
C6 MODIFICATION OF THE FOC: REFRACTOR NEAR EXIT PUPIL

IN ORBIT, the FOC is opened up and a refractive corrector or pupil stop is snapped in place. Alternatively, corrective lenses or pupil stops are added to the filter wheel.

PROS
- The image quality is recovered over the entire field of view.
- The correction can be reversed by rotating the filter wheel to another position.

CONS
- Spherochromatism is present at the ends of the spectral range.
- UV throughput is reduced.
- Implementation is difficult.
- The spectrograph and coronograph modes are not corrected.

ISSUES
- Is it feasible to install a pupil corrector in orbit?

Upper Channel Cover

F.O.C. Optical Bench

D.H.U.

Filter Wheel

Snap-In Refractive Corrector
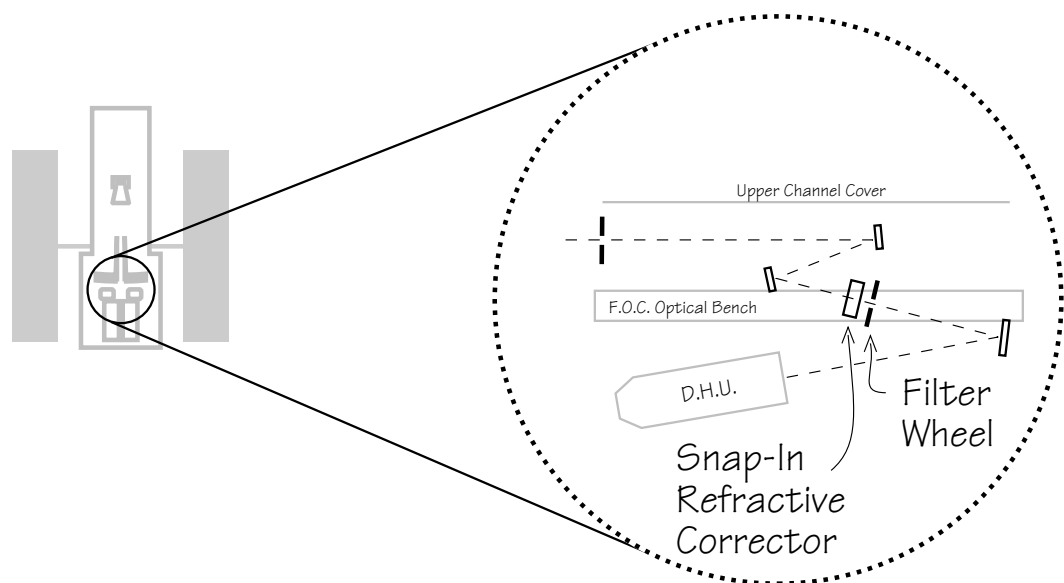
C7 Modification of the FOC: Change Optical Head Unit

After returning the FOC to the ground, the optical head units on the f/48 and f/96 modes are replaced with units that correct the spherical aberration.

Pros

- Image quality is fully recovered without spherochromatism or throughput loss.
- Possibly the spectrographic mode can be corrected.

Cons

- This approach requires return of the FOC to the ground (loss of observing time).

Issues

- Shuttle schedules may cause a long turn-around time.
- The FOC may be contaminated during the return.
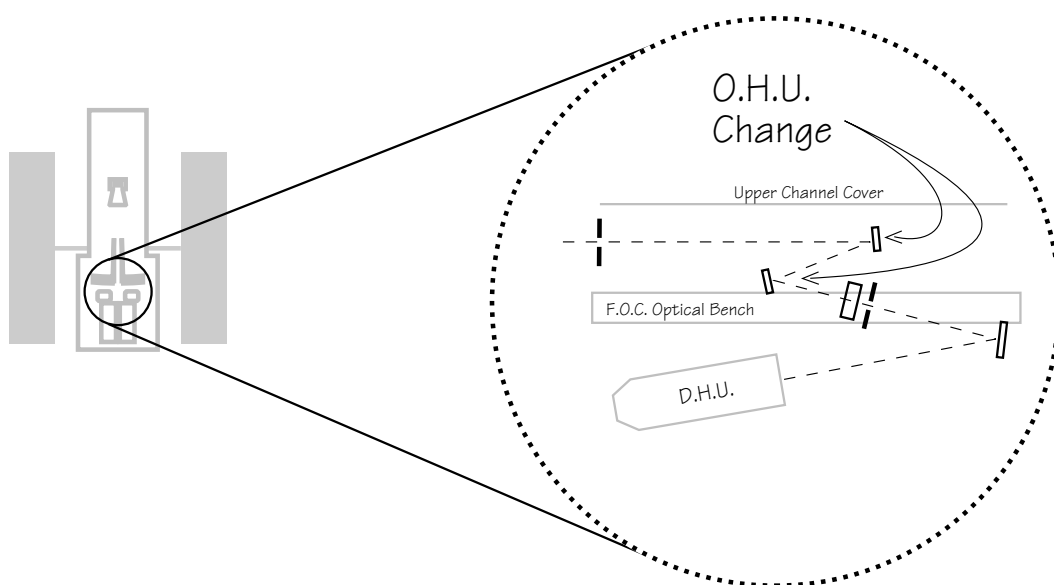
C8 NICMOS Internal Corrector

NICMOS is an ORI and has *ab initio* been designed to have an initial re-imaging stage. The OTA pupil is imaged at a beam-steering mirror. This allows full correction of the OTA spherical aberration, if it is sufficiently well specified. (The conic constant of the OTA primary must be known to better than ±0.0008, equivalent to about 7% wavefront error.) Residual spherical aberration could possibly be corrected by adaptive optics. However, at the present time it does not seem likely that this will be necessary. The present plans for NICMOS do not include this refinement.

PROS
- Full correction of the OTA spherical aberration is feasible.
- No change in the present OTA optics is required.

CONS
- A moderately aspherical beam-steering mirror (diameter 20 mm) must be manufactured.

OTA
Image

Reimaging
Mirror

Articulated
Beamsteering
Mirror and
S.A. Corrector

Reimaged
OTA Field
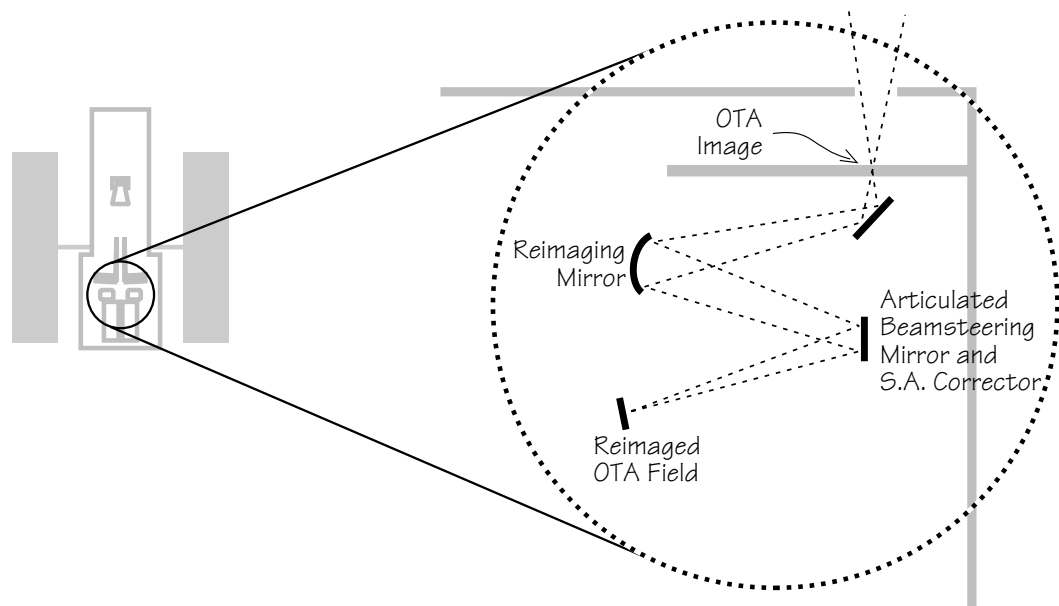
C9 STIS Internal Corrector

A TWO-MIRROR RELAY inside the STIS is planned to correct OTA spherical aberration before the slit. Initial design of this corrector is completed. It is quite adequate for the spectral modes but leaves some residual image degradation in the camera modes. To the extent to which this is acceptable is unclear.

The addition of the relay makes it necessary to rearrange the entire STIS optical train. Associated with this is a reconfiguration of the optical bench and some of the mechanisms.

PROS
- The OTA spherical aberration is corrected at a pupil image.
- No changes in the present OTA optics are necessary.

CONS
- Two mirrors are needed, which reduces UV efficiency.
- The corrector is highly aspherical.
- Matching the corrector to the OTA spherical aberration is critical. (The conic constant of the OTA primary must be known to ±0.00035, i.e. about 3% of the conic constant error or wavefront error.)
- Adaptive optics may be necessary.
- The entire STIS optical train must be reconfigured.

ISSUES
- UV scattering and reflection losses on the correcting mirrors must be considered.
- The correctors add complexity to the instrument.
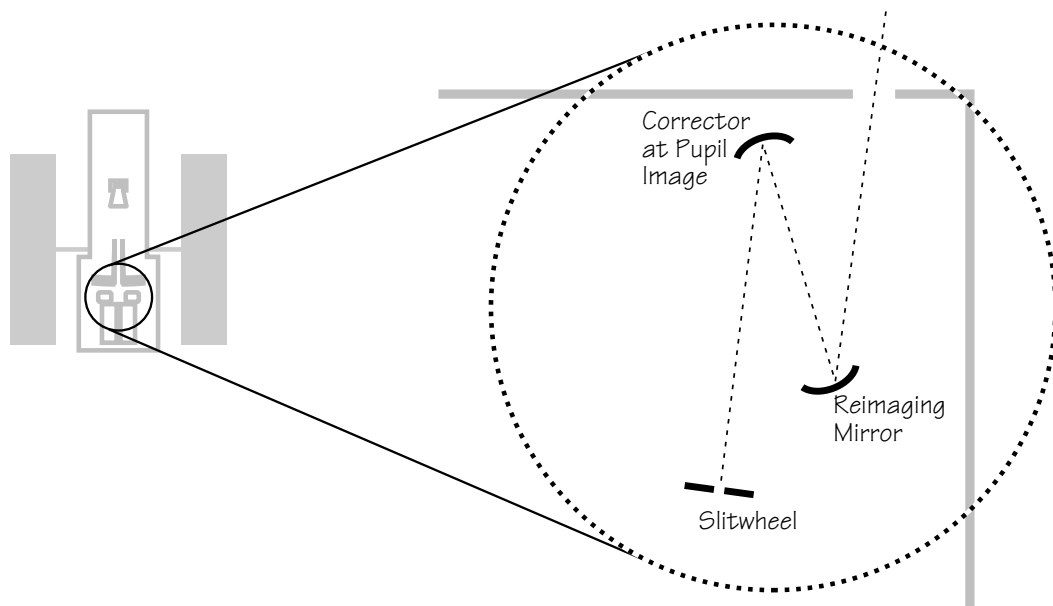
OPTIONS D1 - D3:

MASKING – REDUCTION OF APERTURE WITH FULL FIELD

D1   Aperture Masking

D2   Aperture Door Vignetting

D3   Halo Baffle at Individual SIs

D1 Aperture Masking

AN ANNULAR APERTURE mask(s) is deployed at the top of the telescope tube. For example, a simple mechanism replacing the aperture door would allow a selection of aperture sizes—in addition to a fully closed position.

Spherical aberration causes only the light from a restricted annular region on the primary mirror to be focused on the image plane. Light from the remainder of the primary causes a diffuse halo around the focused image. Annular masking is a low-technology amelioration of spherical aberration that could be implemented in the time period before other optical corrections become available.

Given the amount of OTA spherical aberration, the mask (as specified by the inner and outer fractional radii, α and β) can be optimized for various performance criteria. For example, for 0.5 waves rms spherical aberration, the most compact image core is obtained with $\alpha = 0.76$ and $\beta = 1$, while the most HST-like image (lower side-lobes) is obtained with $\alpha = 0.33$ and $\beta = 0.72$.

Aperture masking will make image deconvolution more straight-forward, because diffraction, rather than aberration, dominates. With appropriate deconvolution methods, superresolution is possible.

PROS
- Aperture masking is achromatic.
- Masking is low technology and relatively cheap.
- Improvements in the point-spread function can be tailored to observational requirements.
- Image deconvolution is easier.
- The mask would be easy to install and easy to remove.

CONS
- The mask blocks up to half the light.

ISSUES
- Is introduced field coma tolerable?
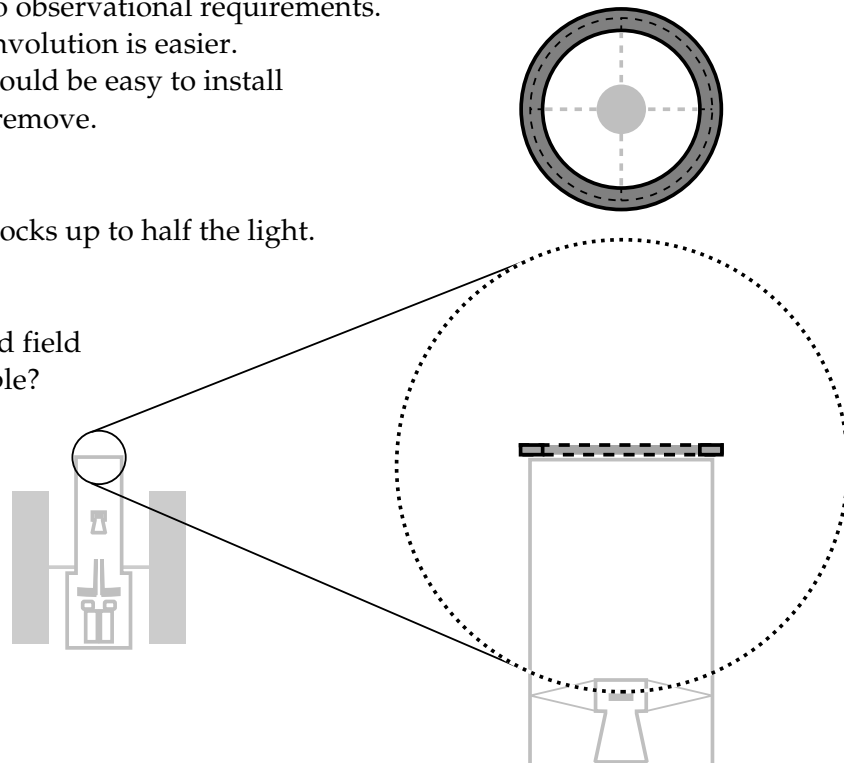
D2 APERTURE DOOR VIGNETTING

THE APERTURE DOOR is commanded shut but stopped part way. The remaining unobscured pupil is corrected partially by tilting the secondary mirror to correct the resulting coma and astigmatism. Some higher-order residuals may be removed with the primary mirror actuators.

PROS
- No shuttle launch is required.
- This could be implemented now, and is reversible (and tunable).

CONS
- Light is lost.
- Spherical aberration is only partially removed
- FGS performance may be affected.

ISSUES
- This option must be studied to know if it is worthwhile.
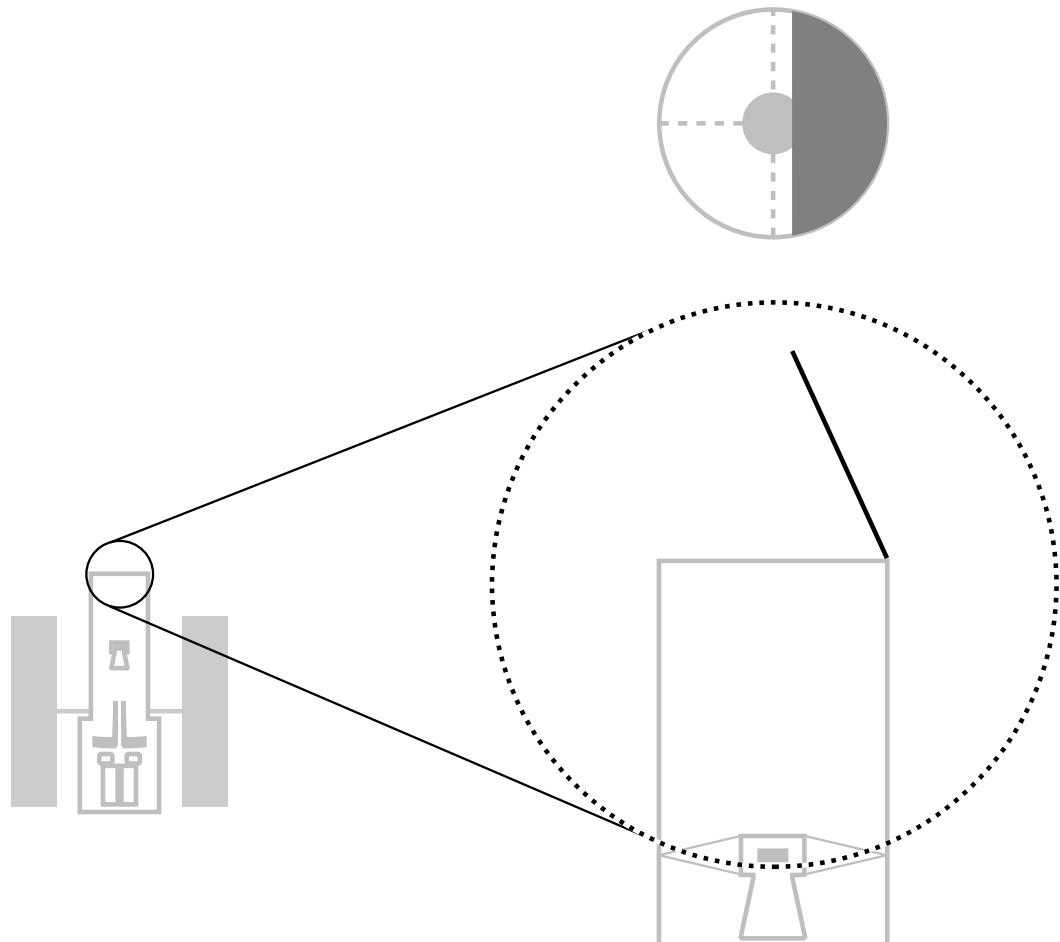
D3 Halo Baffle at Individual SIs

If OTA SPHERICAL ABERRATION is not corrected, there is an optimum focus setting with regard to the energy collected in a narrow slit. This light, it has been found, comes from an area around the central obscuration with about half the radius of the pupil. The remaining area of the pupil contributes only a halo background which introduces noise in crowded fields. A baffle can effectively remove the halo.

The baffle need not fit the unused pupil area very closely, and thus does not need to be centered accurately. Consequently, installation on orbit does not require critical positioning.

PROS
- The image halo is effectively removed.
- The baffles are simple.

CONS
- The OTA aperture diameter is effectively reduced to about 1.2 m.

ISSUES
- Installation in front of the SIs on orbit may be difficult.
- The baffles are not reversibile after installation.
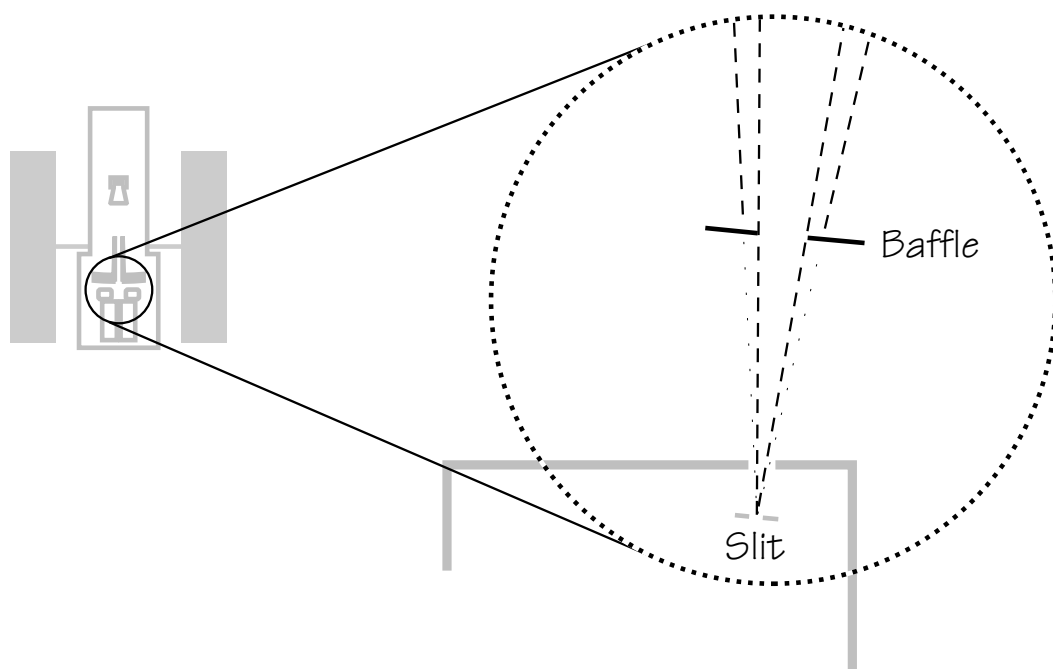
OPTICAL ANALYSIS

1. INTRODUCTION

THIS ANALYSIS IS CONCERNED with the optical aspects of the options available for correction of the spherical aberration in the HST. It will consider the theoretical aspects of the fundamental correction possibilities; with the optical performance of these options, and with the technical aspects of procurement and manufacture in so far as these can be assessed at this stage.

There are three basic reasons why correcting the error in the HST in the general sense, i.e., for the full field and maintaining the full aperture, is difficult:

• **The enormous wavelength bandpass**
There is a serious limitation in the materials available for refracting solutions because of absorption in the UV. Effectively, we are limited to MgF_2, CaF_2 and LiF, whereby there are serious diameter limitations, above all for LiF. Furthermore, even these materials will be approaching the absorption band in the extreme vacuum UV so that the dispersions are far higher than we normally encounter for ground based telescopes. This has important implications for the chromatic performance of options using refracting elements.

• **The very high magnification m_2 of the secondary in the HST**
The HST has a very high telephoto effect with $m_2 = 10.435$. It will be shown that this magnification has serious consequences for a whole class of solutions with regard to field aberrations.

• **The spherical aberration error is apparently mainly on the primary**
This fact has advantages for the use of the telescope in its present uncorrected state, since the error is at the pupil and introduces no field aberrations, above all no field coma. This is of great importance for the FGS which is far less sensitive to symmetrical aberration. However, from the point of view of correction of the error, the primary represents a plane in the system which is inaccessible for most practical correction options. This inaccessibility means that most options apply correction at a significant distance from the pupil which has—combined with the high secondary magnification—negative consequences for the correction in the field.

A correction of the spherical aberration error implies, of course, that the error must be accurately known, in sign and amount. The evidence presented so far suggests that the sign is known with considerable certainty (over-correction, i.e., primary too aspheric). The amount seems to be known within 10% or better. At least 90% seems to be due to the primary. As one would expect from the nature of the tests with null-correctors, the error also seems to be largely *third-order*—the classical lowest term of spherical aberration.

Assume that the longitudinal spherical aberration is 40 mm and is pure third-order. The peak-to-valley wavefront aberration is then

$$W_{ptv} = 4350 \text{ nm.} \tag{1}$$

This gives a third-order spherical aberration coefficient S_I of

$$S_I = 8W_{ptv} = 34800 \text{ nm,} \tag{2}$$

an angular image *diameter* at the Gaussian (paraxial) focus *with 100% geometrical energy* of

$$\delta\alpha_{GF}(\text{diameter}) = \frac{S_I}{y_1} \text{ rad} = 5.98 \text{ arcsec,} \tag{3}$$

and an angular image diameter at best focus (disk of least confusion) of

$$\delta\alpha_{BF}(\text{diameter}) = \frac{S_I}{4y_1} \text{ rad} = 1.50 \text{ arcsec,} \tag{4}$$

where y_1 = the semi-aperture of the telescope = 1200 mm.

If the error is to be compensated on a mirror surface, the deformation for the Gaussian focus (without focus compensation) is of the simple form

$$dz = ay^4, \tag{5}$$

whereby the value at the edge of the aperture beam (irrespective of the size and position of the mirror) would be

$$(\delta z)_m = \frac{W_{ptv}}{2} = 2175 \text{ nm} \cong 2.2 \text{ μm.} \tag{6}$$

This is the physical error on the primary.

To third order accuracy, the form of the mirrors is defined by

$$z = \frac{1}{2r} y^2 + \frac{1}{8r^3}(1 + b_S) y^4 + \dots , \tag{7}$$

in which b_S is the Schwarzschild constant defining the aspheric form. For the theoretical Ritchey-Chrétien telescope, b_S is given for the two mirrors by

$$(b_S)_{1,RC} = -1 - \frac{2(f_1 - d_1)}{d_1 m_2^2} \tag{8}$$

$$(b_S)_{2,RC} = -\left(\frac{m_2 + 1}{m_2 - 1}\right)^2 - 2\frac{f_1}{d_1}\frac{m_2}{(m_2 - 1)^3}, \tag{9}$$

where

$$f_1 \text{ is the focal length of the primary}$$

$$d_1 \text{ is the primary-secondary separation}$$

$$m_2 \text{ is the magnification of the secondary.}$$

With f_1 = 5520 mm, d_1 =4906 mm, m_2 = 10.435, Eq. (8) gives

$$(b_S)_{1,RC} = -1.002299. \tag{10}$$

With $y = y_1$ = 1200 mm, (7) gives

$$\delta z_{RC} - \delta z_p = -0.0004429 \text{ mm,} \tag{11}$$

where δz_p is the asphericity for the parabola. The actual error δz is, from (6),

$$\delta z = -0.002175 \text{ mm,} \tag{12}$$

i.e., nearly 5 times the difference of (11). Since

$$\delta z_{RC} = -0.193075 \text{ mm,} \tag{13}$$

we have

$$\frac{\delta z}{\delta z_{RC}} \cong 1.13\%, \tag{14}$$

a large error giving an actual Schwarzschild constant of

$$(b_S)_A = -1.01359. \tag{15}$$

2. STOP-SHIFT FORMULAE FOR THIRD-ORDER ABERRATIONS

THE "STOP" IN THE HST is also the entrance pupil and is at the primary mirror. The "stop" in an optical system is defined by the diameter of the optical elements relative to the width of the axial beam incident on them. If the secondary is dimensioned to have a diameter surplus which allows the full field to pass without vignetting, then this defines the stop as being at the primary. If we introduce additional correcting elements, they must also be over-dimensioned in the same way if stop-shift (or vignetting) is not to be introduced. In particular, masking techniques at the top of the telescope may effectively shift the stop to that point. Such a "stop-shift" affects the optical performance in a way which depends on the aberrations in the system.

Suppose the stop is shifted a distance ∂E in the "optical space". Then we can define a parameter which is of great importance for the basic theory of most of our correction options:

$$H \, \partial E \equiv y_{pr}/y \tag{16}$$

where:

H is n'u'η', the Lagrange invariant

∂E is the effective "stop shift"

y_{pr} is the height of the "principal" or "chief" ray at a given plane in the system

y is the height of the aperture paraxial ray at a given plane in the system

n' is the refractive index in the image space

u' is the semi-aperture angle in the image space

η' is the angular image field radius.

Figure 1 shows the ray path of the aperture and principal rays through the telescope. For the different planes shown, one can see qualitatively how the ratio y_{pr}/y increases linearly from zero "upwards" from the primary into the object space; and very rapidly going "downwards" from the primary via the secondary into the image space.
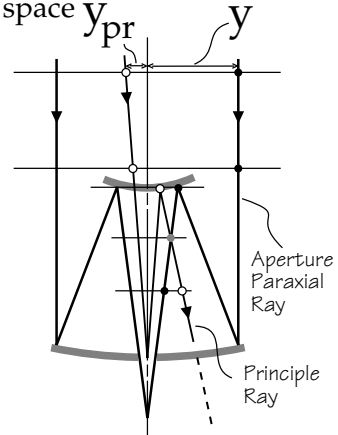
Figure 1. Aperture and principal ray paths through the telescope showing the evolution of (y_{pr}/y) in the different planes in the system.

The effect of such a global stop shift is:

3rd-order spherical $\partial S_I = 0$

coma $\partial S_{II} = (H \, \partial E) \, S_I$

astigmatism $\partial S_{III} = 2(H \, \partial E)S_{II} + (H \, \partial E)^2 \, S_I$ (17)

field curvature $\partial S_{IV} = 0$

distortion $\partial S_V = (H \, \partial E) \, (S_{IV} + 3S_{III}) + 3(H \, \partial E)^2 \, S_{II} + (H \, \partial E)^3 \, S_I$.

Here S_I to S_V are the wavefront coefficients of third-order spherical aberration, coma, astigmatism, field curvature, and distortion respectively, without stop-shift, i.e., ∂E = 0. For our current problem, we are essentially concerned only with the first three terms of (17).

3. THE THEORY OF ASPHERIC "PLATES"
THE STOP-SHIFT FORMULAE (17) assume a particularly simple and important form if we

consider their application to an aspheric corrector plate shifted from the stop. The term "aspheric plate" applies here in a quite general sense to any element, refracting or reflecting, which affects the third-order aberrations without introducing *optical power.* In other words, it refers to any term according to Eq. (5) depending on the *fourth power* of the aperture and has no second power (except, possibly, a minimal balancing term). *It is valid, then, for a thin aspheric plate, or a deformed plane mirror, or a change of deformation on an existing curved mirror.*
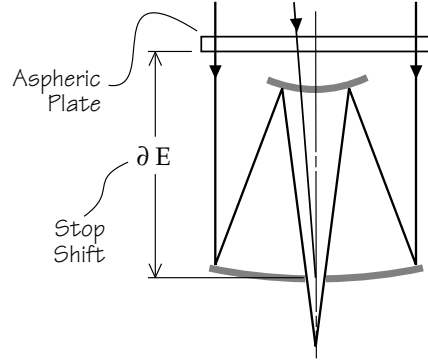
Figure 2. Stop-shift effect for a single aspheric plate shifted from

The stop-shift effect for a single "plate" is shown in Figure 2. Suppose, in the HST, an aspheric "plate" could be placed directly at the primary with the desired amount of spherical aberration ∂S_I to correct the error in the primary. Monochromatically, if this were possible, it would remove the error at its source, at the pupil. An aspheric plate at the pupil is, to third-order accuracy, a "pure" element in its monochromatic function: it only affects S_I, all field effects are zero, so that we have:

$$\partial E = 0: \quad \begin{aligned} S_I &= \delta S_I \\ S_{II} &= 0 \\ S_{III} &= 0 \\ S_{IV} &= 0 \\ S_V &= 0. \end{aligned} \qquad (18)$$

If the plate is now shifted from the stop by ∂E, then substituting (18) in (17) gives:

$$\begin{aligned} \delta S_I &= S_I \\ \delta S_{II} &= (H\,\partial E)\, S_I \\ \delta S_{III} &= (H\,\partial E)^2\, S_I \\ \delta S_{IV} &= 0 \\ \delta S_V &= (H\,\partial E)^3\, S_I. \end{aligned} \qquad (19)$$

These simple formulae enable us to give immediately the monochromatic, third-order effects of the correction δS_I on the field aberrations S_{II} and S_{III}. All we require for a given δS_I is the value of (y_{pr}/y) at the plane of the system in question.

For a single plate, the aberrations are given directly by Eq. (19). Clearly, for two plates *with a significant separation,* we can correct two conditions; with three separated plates we can correct all 3 conditions, S_I, S_{II}, S_{III}.

For an aspheric plate, the spherical aberration is given directly from (2) and (5)

by

$$\delta S_I \;=\; 8(n' - n)\, ay^4, \tag{20}$$

where n = 1 and n' is the refractive index of the plate material, while for a supplementary deformation on a curved mirror, it is given by

$$\delta S_I = -2\,\frac{y^4}{r^3}\,\delta b_S\,, \tag{21}$$

because n' - n = -2 for reflection.

Finally, Eqs. (19) and (20) reveal at once an important consequence concerning the *chromatic* effects of any system of aspheric plates. Setting n = 1 in (20) and differentiating with respect to n', we have:

$$\partial(\delta S_I) = 8ay^4\,\partial n'. \tag{22}$$

Suppose a corrector has three plates of a single material fulfilling the conditions from (19):

$$\Sigma S_I \quad= (S_I)_1 + (S_I)_2 + (S_I)_3 \qquad\qquad\qquad = \delta S_I$$

$$\Sigma S_{II} \quad= (H\,\partial E)_1\,(S_I)_1 + (H\,\partial E)_2\,(S_I)_2 + (H\,\partial E)_3\,(S_I)_3 \quad = 0 \tag{23}$$

$$\Sigma S_{III} \quad= (H\,\partial E)_1{}^2\,(S_I)_1 + (H\,\partial E)_2{}^2\,(S_I)_2 + (H\,\partial E)_3{}^2\,(S_I)_3 \quad = 0,$$

then, because the S_I terms only appear linearly, it follows for the chromatic variations:

$$\partial(\Sigma S_I) \;=\; \partial(S_I)_1 + \partial(S_I)_2 + \partial(S_I)_3$$
$$= \partial(\delta S_I)$$

$$\partial(\Sigma S_{II}) = (H\,\partial E)_1\,\partial(S_I)_1 + (H\,\partial E)_2\,\partial(S_I)_2 + (H\,\partial E)_3\,\partial(S_I)_3$$
$$= 0 \qquad (24)$$

$$\partial(\Sigma S_{III}) = (H\,\partial E)_1{}^2\,\partial(S_I)_1 + (H\,\partial E)_2{}^2\,\partial(S_I)_2 + (H\,\partial E)_3{}^2\,\partial(S_I)_3$$
$$= 0.$$

Eqs. (24) express a simple physical situation for a system of plates of one material, namely the fact that $\delta S_I \neq 0$ (this is the purpose of the system) implies that spherochromatism (Gauss error) is a fixed quantity depending only on (δS_I) and $(\partial n')$. The other monochro-
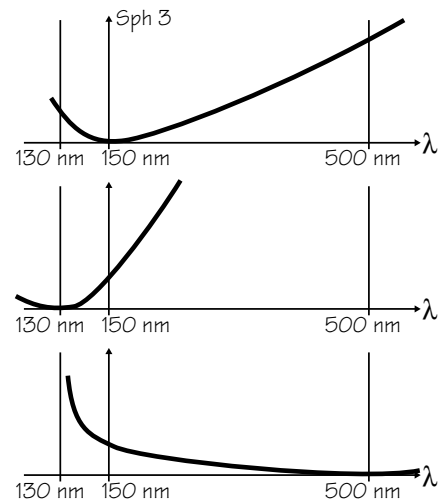

Figure 3. Rebalancing effect of spherochromatism achieved by moving correctors, or elements thereof, in the converging beam.

matic terms ΣS_{II} and ΣS_{III} are zero; then the third order chromatic variations are also zero. Of course, if higher order effects are present which are *not* zero, then there will be chromatic variations.

The above conclusions concerning chromatic effects apply to any system made of *one single material*. If two materials with usefully different dispersions $\delta n / (n_o - 1)$ are available, then direct correction of the spherochromatism may be possible.

Controlling the axial position of the whole corrector—or of an element or elements within it—provides a way to rebalance the dispersion curve of S_I and to improve the spherochromatism of the refracting plate solutions in the converging beam, see Figure 3. If the whole corrector is moved, the effect on S_I will go with the fourth power of the distance from the image since this is the aperture dependence of S_I. Thus a change of axial position from the image by 10% will change S_I by about 40%. Other aberrations will also be affected. If internal elements are moved, the shifts will be less, because individual plates are stronger than the net S_I effect. Also, moving the separate elements will affect other aberrations individually. If all elements are moved optimally, then the original quality is restored, optimized for the new wavelength bandpass.

Of course, such rebalancing is not possible for correctors placed in the parallel incident beam.

4. SOME SIMPLE APPLICATIONS TO POSSIBLE OPTIONS IN THE HST

4.1 STOP-SHIFT EFFECT AT THE SECONDARY MIRROR, M_2

For a field radius of 10 arcmin, the value of our stop-shift parameter is

$$\left(\frac{y_{pr}}{y}\right)_{SM} = 0.1069. \tag{25}$$

From (2) and (19):

$$\delta S_{II} = (H \, \partial E) \, \delta S_I = 3720 \text{ nm}. \tag{26}$$

Now the angular size of the tangential coma patch is

$$(\delta\alpha)_{Coma} = \frac{3 S_{II}}{2 y_1} \text{ rad} = 0.959 \text{ arcsec}. \tag{27}$$

This result effectively rules out the solution of correction of S_I by a new secondary maintaining an $f/24$ output beam. The high coma value for correction at the plane of the secondary is a consequence of the high magnification m_2 giving a small secondary and high telephoto effect. Correction at the secondary gives a departure from the Ritchey-Chrétien solution which is far too large to be acceptable.

A comparison with the ESO New Technology Telescope (NTT) is instructive. Here, there was also a matching error of S_I on the primary with $W_{ptv} \cong 3000$ nm compared with 4350 nm for the HST. Because $m_2 = 5$ for the NTT, the factor (y_{pr} / y) at the secondary is only 0.04298 giving

$$(\delta\alpha)_{\text{Coma}} = 0.182 \text{ arcsec}$$

for a field radius 10 arcmin, a *factor of 5.3 times lower than the HST case.* This coma might have been considered acceptable for a passive telescope, but the effect was totally removed at source by active optics bending of the primary.

The astigmatism induced by a replacement of M_2 is, from (19)

$$\delta S_{\text{III}} = (H \, \partial E)^2 \, S_I = 398 \text{ mm}$$

at field radius 10 arcmin. At best focus, halfway between the astigmatic lines,

$$W_{\text{ptv}} = 199 \text{ nm}$$

and

$$(\delta\alpha)_{\text{ast, mean}} = \frac{\delta S_{\text{III}}}{y_1} = 0.068 \text{ arcsec.} \qquad (28)$$

This illustrates how rapidly the effect on astigmatism becomes negligible if $(y_{\text{pr}}/y) = (H \, \partial E) \ll 1$. Fairly near the image, however, in the stovepipe of the HST, the factor may be comparable to 1 or larger. Then the effects on astigmatism may be larger than on coma.

4.2 EXAMPLE OF A CORRECTOR PLATE IN THE PARALLEL INCIDENT BEAM ABOVE THE SECONDARY

At a distance of 5500 mm above the primary (above the secondary mirror in Figure 1) we have $y_{\text{pr}} = 16.00$ mm and $y = 1200$ mm, giving $(y_{\text{pr}}/y) = 0.01333$. This is almost exactly 1/8 of the value at M_2. For the same field radius of 10 arcmin

$$(\delta\alpha)_{\text{Coma}} = \frac{0.959}{8} = 0.120 \text{ arcsec.}$$

The astigmatism is 1/64 of the value at M_2 and is therefore completely negligible.

4.3 EXAMPLE OF AN APERTURE MASK

Suppose the stop were shifted by ∂E by a masking operation, then (17) shows that we would get a change of coma and astigmatism of

$$\partial S_{\text{II}} = (H \, \partial E) \, S_I$$

$$\partial S_{\text{III}} = (H \, \partial E)^2 \, S_I \, ,$$

since $S_I \neq 0$ (this is the HST error), but $S_{\text{II}} = 0$ because the error S_I was in the original pupil position.

5. LOGICAL CATEGORIZATION OF THE VARIOUS OPTIONS

THE PARAMETER $(y_{pr}/y) = (H\delta E)$ gives us a direct measure of the effectiveness of a given option in correcting the spherical aberration error with a single element. In most cases, this parameter will not be favorable enough and will impose solutions requiring more than one element.

We have considered so far the correction of the HST in the fullest sense: a correction over the full field and maintaining the full aperture which effectively restores the nominal quality of the HST. This leads to the definition of Group A, as shown in Figure 4. Other groups are also shown in Figure 4. Their definitions are as follows.

Figure 4. The basic groups of options according to their positions in the complete system.

• **Group A: Full field correction maintaining full aperture (including FGS and WFS)**

The options are ordered according to (y_{pr}/y), starting with the value zero for correction at the primary. Options in the object space follow, then options from M_2 down towards the image. All corrector options are symmetrical to the OTA axis.

• **Group B: Pre-focal plane correction for individual instruments (excluding FGS and WFS)**

This group has the following global characteristics:
  - The correctors are centered on the individual instrument entrance beam axes.
  - The fields involved are far smaller than for Group A and are simply the useful fields of the instruments.
  - Note that the FGS and WFS are *not* corrected (unless independent action is taken. See Options Cl, C2 and C4).
  - The correction functions for all instrument modes, including spectral, coronagraphic, etc.

The options of this group are ordered according to the number of elements and complexity.

• **Group C: Post-focal plane correction for individual instruments**
  - FGS and WFS are *not* corrected unless options Cl, C2 and C4 are realized.
  - Spectral modes, etc. cannot be corrected in the first generation spectrographs.

In general, correction in the second generation SIs should be relatively easy from an optical viewpoint since the fields are small and re-imaged pupil planes will be more or less accessible.

Options Cl or C2 correct the FGS and thereby remove a major weakness of both Group C and Group B.
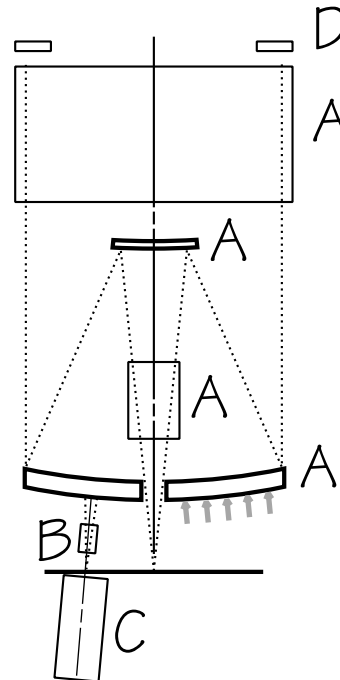
•Group D: Masking - full field correction but with aperture reduction

6. LIST OF OPTIONS WITH BRIEF REVIEW OF THEIR OPTICAL CHARACTERISTICS

•Group A: Full field correction at full aperture

A1. MECHANICAL DEFORMATION OF THE PRIMARY MIRROR
  - Optically ideal as it achieves the correction by reflecting (achromatic) means at the source of the error $(y_{pr}/y) = 0$. This would be the equivalent of the active optics correction of matching error done in the NTT.
  - The HST primary is a stiff egg-crate. The dynamic range of its actuators is almost certainly inadequate for this correction. The actuators were only intended for a re-touch of astigmatism.
  - May be danger of structure print-through if bending could be achieved.

A2. THERMAL DEFORMATION OF THE PRIMARY MIRROR
  - Optically ideal as with A1.
  - May be possible, but temperature gradient control is difficult in practice.
  - Print-through risk as in A1 but may be exaggerated by thermal effects.

A3. OVERCOATING THE PRIMARY MIRROR
  - Optically ideal as with A1.
  - Possibilities of thick Al or a thick coating of other materials with a normal thin coating of Al on top.
  - Roughness problem of thick coats very serious.

A4. FULL APERTURE CORRECTION PLATE
  - (y_{pr}/y) the most favorable of all plate solutions — field coma at 10 arcmin field radius 0.12 arcsec, field astigmatism negligible.
  - 2 vacuum/glass surfaces only.
  - Not achromatic: spherochromatism appreciable for whole spectral range using an MgF_2 plate.
  - Availability of MgF_2 in large sizes not solved at present — segmented solution may be possible but gives diffraction effects.
  - Birefringence, thickness aspects.
  - Large, very difficult object to make.

A5. FULL APERTURE CORRECTION FLAT
  - (y_{pr}/y) about twice that of A4 — field coma at 10 arcmin field radius ca. 0.24 arcsec. Could be combined with a coma correcting plate in the stovepipe.
  - Reflecting (achromatic) solution. If combined with a coma correcting plate this latter would be weak and would generate only small chromatic effects.
  - Zero (or, with plate 2) vacuum/glass surfaces.
  - Large difficult object to make (long axis about 3.5 m).
  - Could be supported passively (rigid), actively (flexible), or semi-actively (fairly rigid).
  - Test probably requires a liquid mirror — difficult.

A6. GAS-FILLED CORRECTION LENS

- (y_{pr}/y) probably similar to A5.
- This is a weak convex lens with nitrogen or helium gas as medium. The positive lens is correct for compensating the overcorrected spherical aberration of the primary.
- The He lens will introduce normal longitudinal chromatic aberration unless the dispersion of He is negligible.
- The spherochromatism will be in the ratio of the dispersions of He to MgF_2 compared with A4.
- The film or plastic (Mylar?) must have uniform thickness to a fraction of λ, otherwise unacceptable wavefront errors.
- The stronger the curvatures the slower the convergence of the spherical aberration function. Negligible fifth order requires small incidence angles.
- An He bag held in a ring may have strong deformation near the ring.

A7. SECONDARY MIRROR (SM) REPLACEMENT

- Reflecting solution — achromatic.
- (y_{pr}/y) impossibly large for a single element solution — see paragraph 4.1. Coma at 10 arcmin field radius *0.959 arcsec*. This would be worse than the present spherical aberration for all instruments except WFPC and fatal for the FGS.
- Difficulty of procurement of new SM without cross-test with primary.
- Inadequate optical solution.

A8. SM RECONFIGURATION OF HST TO AN *f*/13.25 RITCHEY-CHRÉTIEN

- Re-creates a correct aplanatic optical system by pure reflection means – *a perfect optical solution.*
- Difficulty of procurement of new M_2 without cross-check with primary.
- Change from $f/24$ to $f/13.25$ — effect on the instruments, FGS, etc.?
- Passage of $f/13.25$ beam out of telescope — vignetting?
- Possible use of Barlow "lenses" (or reflectors) to restore $f/24$?

A9. SM REPLACEMENT AND 2-PLATE CORRECTOR IN STOVEPIPE

- This option falls in the group of 2- or 3-"plate" solutions discussed in paragraph 3, but where the first "plate" is a deformation on a mirror (a reflecting "plate"). With one additional plate, it has an analogy with A5 (front reflecting "plate" plus possibly an additional plate).
- (y_{pr}/y) not favorable but overcome by a 2- or 3-"plate" solution, of which the first "plate" is a deformation on M_2.
- One additional plate leaves astigmatism uncorrected (reversed sign from telescope), correction possible in front of instruments.
- With 2 additional plates, *an excellent monochromatic solution.*
- Because most of the correction is done at M_2, *the chromatic aberration is far lower than for pure plate solutions — probably less than one quarter.*
- Problem of procurement of new M_2 without cross-check with primary.
- Diameter of aspheric plates favorable compared with A4 (ca. 350 mm). MgF_2 and CaF_2 should be possible?
- 2 or 4 vacuum/glass surfaces — favorable!

- Optically a very good solution, certainly one of the best "plate type" solutions.

A10. 1-PLATE CORRECTOR ON SM AND 2-PLATE CORRECTOR IN STOVEPIPE
- This has a refracting plate in double pass directly in front of M_2. In principle, it is a 2- or 3-plate solution, as treated in paragraph 3.
- (y_{pr}/y) exactly as in A9.
- One additional plate leaves the astigmatism uncorrected exactly as in A9.
- With 2 additional plates, an excellent monochromatic solution as A9.
- Because of double pass in first plate, unfavorable for vacuum/glass surfaces:
  > *6 vacuum/glass surfaces with 1 extra plate*
  > *8 vacuum/glass surfaces with 2 extra plates.*
- With 3 plates of one material, spherochromatism as for a single plate. May be possibility of improvement with combination of MgF_2 and CaF_2. Chromatically less favorable than A9 or A5 (with additional plate) *unless active* rebalancing of spherochromatism is done as described in paragraph 3 and shown in Figure 3. This could be done by shifting the d/p plate axially according to the spectral range. No image analysis is required, the shift is given directly by the wavelength shift. *However, this balance has the consequence for all Group A solutions that the FGS would have variable spherical aberration in its wavelength bandpass!*
- No replacement of M_2 necessary.
- Diameter of aspheric plates favorable for procurement compared with A4.

A11. THREE ASPHERIC PLATES IN THE STOVEPIPE
- This 3-plate solution was treated in paragraph 3.
- (y_{pr}/y) is less favorable than A9 and A10. This leads to higher asphericities and therefore larger fifth order errors. In a preliminary calculation, the fifth order effects seem acceptable. With plausible separations, (23) give asphericities corresponding to +6 (δS_I), -8 (δS_I), +3 (δS_I). The system is like a single-material triplet working in the third-order instead of the first order.
- The chromatic effects are as treated in paragraph 3 and are the same as for A10 with 3 plates: spherochromatism as for a single plate, third-order chromatic coma and astigmatism zero. There will be fifth order chromatic coma and astigmatism proportional to the monochromatic residuals.
- Possibility of partial correction of spherochromatism with plates of MgF_2 and CaF_2.
- Possibility of rebalancing spherochromatism (as in A10) by active control of axial position of whole corrector or single plates. *Same limitation with FGS as with A11!*
- Procurement and test situation as a unit favorable.
- Positional tolerances very uncritical, above all centering since the emergent beam from the telescope (if the error is all on the primary) has only a very weakly defined axis from the field astigmatism of the Ritchey-Chrétien solution.

A12. DOUBLE CASSEGRAIN RELAY IN CENTRAL BAFFLE
- This is generically a 4 "plate" solution in which the "plates" are reflecting deformations and the separations are constrained by the double Cass geometry giving an unchanged $f/24$ beam.
- Reflecting solution — achromatic.

- (y_{pr}/y) situation as with A11.
- No detailed calculation yet done, but the evidence is that a good optical solution, from the point of view of performance in image quality, exists.
- The (y_{pr}/y) situation seems *fatal* from the point of view of central obstruction and full field coverage. Unfortunately, this appears fundamental. The option remains interesting as B4.
- 4 additional reflections.

•Group B: Pre-focal plane correctors on instrument axes

Two comments are relevant to all options in this group:
- The plate theory of paragraph 3 is applicable here exactly as to Group A except that the field effect defined by y_{pr} is referred to the instrument axis, not the telescope. Since the instrument fields are much smaller, this is favorable, but we are very near the image which is unfavorable for (y_{pr}/y).
- Options C1 or C2 would provide a major upgrade of all Group B options because the central weakness of non-correction of the FGS would be removed.

B1. SINGLE REFRACTIVE CORRECTOR FOR INDIVIDUAL SIs
- Since (y_{pr}/y) can only be favorable for very small fields this single plate solution is probably only of interest for very small field spectroscopic applications. Field limitation is field coma.
- Only 2 vacuum/glass surfaces.
- Technically fairly simple, since the plate is small — but it has a fairly steep aspheric function.
- Spherochromatism is as for a single plate. There is also strong chromatic coma since the field coma is uncorrected.
- Spherochromatism could be rebalanced by shifting the plate axially according to the spectral range.
- Difficult centering tolerances.

B2. DOUBLE REFRACTIVE CORRECTOR FOR INDIVIDUAL SIs
- Generically the same as analogous 2- or 3-plate solutions in Group A.
- (y_{pr}/y) will depend on instrument field and the axial depth available.
- Chromatic performance as with A11 *with a single material,* but *with these diameters, possibility of using MgF_2, CaF_2 and LiF to correct spherochromatism. Also, possibility of rebalancing spherochromatism with a one-material corrector by shifting the whole system axially* (this is independent of the FGS system and is more attractive than the equivalent possibility in A12!).
- For optical performance, attractive, particularly for the FOC and combined with C1 or C2.
- 4 (or 6) vacuum/glass surfaces.
- High sensitivity to tilt, reasonable tolerances for lateral decenter.

B3. Two-mirror reflective correctors for individual SIs
- Generically a 2 "plate" reflecting solution with power added for astigmatism correction.
- Reflecting solution — achromatic.
- High asphericities leading to higher-order aberrations which limit the field. Field of FOC (44 arcsec) cannot be covered.
- Good optical solution for instruments with modest field.
- High sensitivity to tilt.

B4. Double Cassegrain relay for the individual SIs
- Reflecting solution — achromatic.
- The same generic characteristics as A12 but the field restriction (even with the FOC) *may* give a better (y_{pr}/y) value for the obstruction/field problem, depending on the axial space available. If so, an optically attractive solution.
- 4 additional reflections.
- Probably very sensitive to tilt.

•Group C: Post-focal plane corrections in instruments

C1. FGS II
- Correction in this way is *a most interesting option as it upgrades the whole of Group B* by removing the weakness that the FGS is not corrected by the individual pre-focal plane instrument options.

C2. Modification of FGS I
- *Same attraction and importance as C1.*
- Since the FGS fields are probably small, a correction should, in principle, not be very difficult. Planes not far from the transferred pupil should be accessible.

C3. Thermal fixes to SI's
- May be attractive in individual cases, but thermal control is usually more difficult than more conventional optical means.

C4. Modify WFPC II Secondary Mirrors
- A simple and reliable means of correcting the spherical aberration.
- Weakness of fix of one instrument in direct imaging mode but upgraded by C1, C2, or C3 combined with instrument fixes in Group B.

C5, C6, C7. Modification of the FOC
- Various reasonable internal possibilities.
- Weakness that spectroscopic and coronograph modes are not corrected – only

applicable to direct imaging.
- Group B solution with C1, C2 or C3 would be preferable.

C8. NICMOS INTERNAL CORRECTOR
- Satisfactory solution seems available.

C9. STIS INTERNAL CORRECTOR
- Satisfactory solution seems available.

•Group D: Masking – reduction of aperture but maintaining full field

D1. APERTURE MASKING
- By sacrificing about half the aperture, the spherical aberration effect might be reduced to about a quarter.
- The masks might shift the pupil to their own plane above the telescope. If so, the stop-shift term (y_{pr}/y) would be similar to option A5 and the field coma may not be negligible, depending on the effective aperture.

D2. APERTURE DOOR VIGNETTING
- By masking on one side, the pupil is sheared and the spherical aberration can be reduced by compensating with a decentering coma term (independent of the field) by translating the secondary.
- Asymmetric masking should not shift the pupil, but the masking compensation will then vary somewhat with field.

D3. HALO BAFFLE AT INDIVIDUAL SIS
- Whereas D1 and D2 do masking at, or near, the pupil, this masks in front of the image caustic and removes the "wings" of the image directly. The corresponding pupil area is removed.

Aperture Masking

THE HST STRATEGY PANEL EXAMINED APERTURE MASKING as an entirely different option for addressing spherical aberration. By selectively removing light rays from the telescope beam, it is possible to improve the PSF considerably for crowded field imaging and spectroscopy, increasing photometric accuracy and reducing source confusion. Potentially, aperture masking could be a low cost and risk, fail-safe, and low-technology option, which could be tailored to favor imaging or spectroscopy and even offer super-resolution in the FOC. However, the Panel concluded that aperture masking, while it improved the current situation, especially in crowded fields, was inadequate as a recovery option. Nevertheless, its low cost and minimal technology complication make it an attractive method, especially as a backup or insurance policy. For this reason, the aperture masking method was described in our final oral presentation as a "prudent option."

In addition to implementation concerns, such as the question of how astronauts might mount the aperture mask in a region of the telescope assembly never envisioned for the task, the Panel recognizes potential problems with earthshine scattering and negative effects on FGS performance. Nevertheless, the Panel's look at aperture masking has sufficient general interest to warrant this brief tutorial.

BASICS

An essential feature of a mirror suffering from spherical aberration is that only the light from some annular region on its surface can be focused at the image plane. The central bright peak of the current HST PSF is made by this annular region. The remainder of the primary is defocused and causes a large background to the diffraction limited signal. Moreover, it also causes destructive interference with the light coming from the focused region and further attenuates the signal. Aperture masking can selectively transmit light in focus and remove light out of focus.

Even in the absence of aberrations, an annular aperture has its own advantage over a fully filled aperture. From the point of view of interferometric imaging, an annular aperture forms a uniformly redundant array. In the parlance of Fourier optics, the u-v coverage provided by an annulus uniformly fills the circle of radius D/λ as opposed to that of a filled pupil, which heavily weights the lowest spatial frequencies. (D is the diameter of the aperture; λ is the wavelength of light.) For ap-

Figure 1. One idea for how an aperture mask could be mounted

erture synthesis imaging, this u-v coverage is almost ideal. A consequence of this difference in weighting is that the central peak of the PSF of a thin annular aperture is about 30% narrower than that of a filled aperture of the same maximum extent.

In applying a non-linear deconvolution algorithm (e.g., CLEAN), it is essential to have a PSF with a pointed central core, which effectively determines the resolution. Thus annular masking is a way to achieve super resolution. One may worry about the relatively large higher order sidelobes that characterize narrow annular apertures—however these fringes also carry source information and do not do much harm as long as deconvolution is applied later.

SPHERICAL ABERRATION AND COLLECTION AREA

We now consider the limitation on the annular mask design set by the spherical aberration and derive the range of inner and outer radii for a given magnitude of aberration and some specified tolerance level. Dimensionless quantities are used throughout the calculation. The starting point is the formula for the orthogonal spherical aberration ϕ for an annular region on the primary measured in units of wavelength. For a given annulus specified by inner fractional radius α and outer fractional radius β, the orthogonal aberration is a function of the fractional radius ρ. When the spherical aberration of the primary mirror is given by $A\rho^4$, the difference between the aberrated wavefront and the best approximate paraboloid (orthogonal aberration ϕ) is,

$$\phi(\rho;\alpha,\beta) = A\left[\rho^4 - (\alpha^2+\beta^2)\rho^2 + \frac{1}{6}(\alpha^4+4\alpha^2\beta^2+\beta^4)\right]$$

$$= A\left[\left(\rho^2 - \frac{(\alpha^2+\beta^2)}{2}\right)^2 - \frac{1}{2}(\beta^2-\alpha^2)^2\right].$$

This expression is obtained by a trivial modification of the spherical aberration formula for a telescope with central obscuration, which is another annulus. As evident from the second line of the above expression, $|\phi|$ takes the maximum value, $|A|(\beta^2-\alpha^2)^2/6$ at the boundaries, $\rho = \alpha$ or β. This maximum value is exactly $\sqrt{5}$ times the rms wave front error. If we require that $|\phi|$ is smaller than t waves,

$$\frac{|A|}{6}(\beta^2-\alpha^2)^2 < t,$$

and

$$\beta^2-\alpha^2 < \sqrt{\frac{6t}{|A|}}.$$

This criterion is directly related to the usable fractional collecting area $(\beta^2-\alpha^2)/(1-\varepsilon^2)$, where e is the fractional radius at the inner useful edge of the primary mirror. For the HST, $\varepsilon = 0.33$. Therefore the spherical aberration of the primary mirror restricts the sensitivity of annular masking (collecting area), but it does not impose

any restriction on resolution (where to choose an annulus on the primary). It should also be noted that the wavefront error is proportional to the square of the collecting area.

AN OPTIMIZED MASK

The rms wave front error for the HST primary mirror is about one-half of a wave. Taking into account the central obscuration, we obtain a value of 8.4 for A, using the standard formula for the rms error. For Rayleigh's criterion, where t = 1/4, we have

$$(\beta^2 - \alpha^2) < 0.42$$

$$(\beta^2 - \alpha^2)/(1 - \varepsilon^2) < 0.47.$$

Thus, 47% of the collecting area should be usable, and we are allowed to choose either α or β to achieve the given maximum collecting area. Thus the type of mask can range from the highest resolution mask ($\beta = 1$, $\alpha = 0.76$), which has the most compact core in its PSF, to the most telescope-like mask ($\alpha = \varepsilon$, $\beta = 0.72$), which does not have strong sidelobes. If we impose a stricter criterion, for example demanding that the maximum wavefront deviation be no more than 1/14 waves, then

$$(\beta^2 - \alpha^2) < 0.23$$

$$(\beta^2 - \alpha^2)/(1 - \varepsilon^2) < 0.26,$$

so that, in this case, the annulus parameters range from ($\beta = 1$, $\alpha = 0.87$) to ($\alpha = \varepsilon$, $\beta = 0.58$).

In conclusion, the HST image can again be made diffraction limited by the use of aperture masking, but at the cost of rejecting so much light that the efficiency of the spacecraft would be significantly lowered. The loss would be more severe at shorter wavelengths. Nevertheless, for some studies like crowded-field observations, aperture masking would improve HST science performance over the current level.

POINTING ISSUES

1 ORIGINAL DESIGN REQUIREMENTS REGARDING POINTING STABILITY

WITH THE IMAGE QUALITY specified at 70% of the total energy of a stellar image in radius 0.1 arcsec, the allowance for image stability was set at 0.007 arcsec and was budgeted to the different systems as follows:

Guiding system (PCS /FGS) and mechanical disturbances within the support system (reaction wheels, solar arrays, antenna, tape recorders):	0.006 arcsec
OTA, FGS thermal/mechanical disturbances:	0.003 arcsec
SI thermal/mechanical disturbances:	0.002 arcsec
Total (combined by root mean square):	0.007 arcsec

This image stability requirement was to be applicable over a 24-hour period, but excluded the first 4 hours after a worst-case slew.

2 DESCRIPTION OF THE GUIDING SYSTEM

HST's attitude is determined on board using a variety of sources: coarse sun sensors, magnetic sensors, fixed head star trackers (FHST), gyroscopes, and fine guiding trackers. During observations, however, fine guiding relies solely on the gyroscopes, with a periodic position update supplied by two FGSs tracking two guide stars in the field of view of the telescope. Guiding corrections are applied to the entire spacecraft ("body pointing concept") by varying the speed of spinning flywheels (reaction wheels).

The FGSs have two operational modes: "coarse track", based on an image scanning system, and "fine lock", based on an interferometric system. As designed, only fine lock was to provide the image stability budget indicated above. Coarse track was considered an intermediate step in the acquisition sequence or a degraded guiding mode for non-demanding observation or as a backup. As an additional backup, guiding can be done with one FGS for pitch and yaw control and one FHST for roll control, or entirely with gyros.

2.1 GUIDE STAR ACQUISITION PROCEDURE

A typical guide star acquisition scenario begins with a slew of the telescope to the predetermined field, usually followed by an FHST update to refine the attitude. Residual attitude errors are on the order of 10 arcsec. One FGS starts searching for the first guide star, using a spiral search pattern with a 5 arcmin square aperture. Upon finding a star of the correct brightness, it stops and tracks that star in coarse track mode, and the other FGS searches for its guide star. When the second star is found, their relative positions provide the final confirmation of attitude. If confirmation is not obtained, the search resumes. If confirmation is obtained, the two FGSs start tracking in coarse track if this is the final desired mode, or attempt to lock on the star in interferometric mode (fine lock), if the ultimate tracking performance is required. If this process fails because any one of the two guide stars cannot be found or tracked on, a second attempt is made on a different pair of guide stars. Up to three pairs can be uplinked. The acquisition cycle typically takes about three min-

utes per pair of guide stars.

2.2 COARSE TRACK MODE

In this mode, the instantaneous field of view of the FGS (a 5 arcsec x 5 arcsec aperture) is commanded to nutate, or move in a circle, around the guide star at the rate of one revolution per second, with a modifiable radius of about 2.7 arcsec. The intensity collected in the aperture is measured by a photomultiplier tube (PMT) at 40 Hz rate. These intensity values are summed in the four quadrants of the nutation (10 for each quadrant) to obtain the star position error signal in two perpendicular directions, as in the traditional "4-quadrant detector" concept.

The coarse track mode is very robust but has a limited accuracy. Pre-launch predictions were for around 20 mas for a 14.5 m_v star.

Figure 1 shows the coarse track mode as seen in the PMT/aperture frame. The
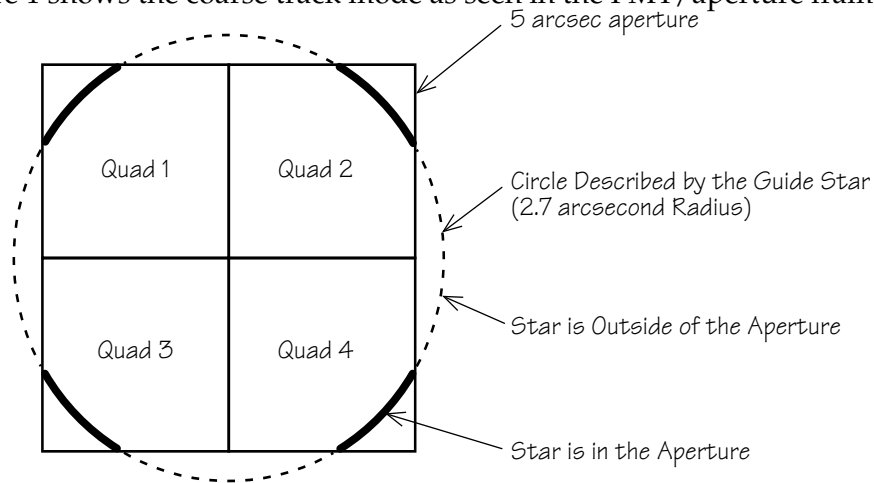
Figure 1: Coarse track mode

guide star is rotated at 1 revolution per second, and intensity is measured in the four quadrants.

2.3 FINE LOCK

In this mode, the collimated light of the guide star is split in two perpendicular directions by a beam splitter, and wavefront tilts in each direction are measured using a Koester prism interferometer.

A Koester prism consists of two halves of an equilateral prism, with a dielectric film sandwiched between the two. The incident beam is divided into two channels and the dielectric coating retards the transmitted beam by $\lambda/4$, while the reflected light is not affected. Intensity in each of the two channels is measured by PMTs (four in all, two for each measurement direction). Figure 2 shows two situations. On the left, the guide star is perfectly centered and there is no tilt in the wavefront, thus each PMT senses the same amount of light. On the right the wavefront has a total of $\lambda/4$ tilt as it hits the prism. The wavefront which is transmitted is retarded by $\lambda/4$, while the reflected one is unaffected. The left channel, A, will experience constructive interference, and the right channel, B, destructive interference, resulting in a greater count in the left PMT than in the right one. The position error sig-

Figure 2: Fine lock mode

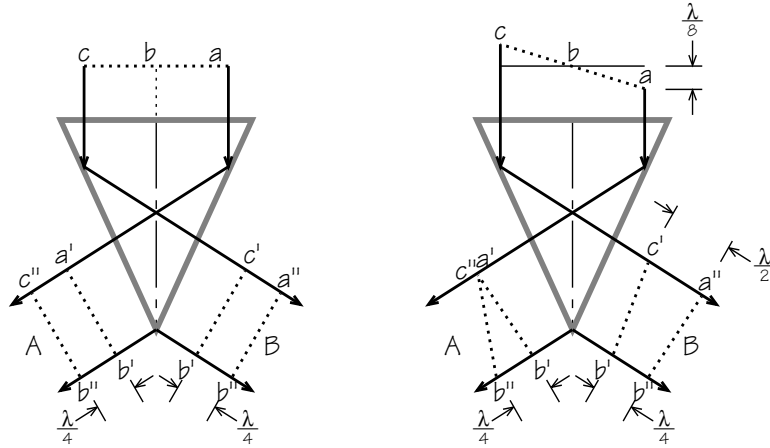nal is formed by combining and normalizing the intensity in the two channels ((A-B)/(A+B)). The resulting white light fringe (S-curve) has a visibility of 0.7 for a peak-to-peak separation of 43 mas, the diffraction limit of the telescope at visible wavelengths.

The advantage of fine lock is its insensitivity to focus and its high error signal gain. However, unlike coarse track, fine lock has a fairly limited dynamic range (about 80 mas). As a result, fine lock is very sensitive to vehicle jitter. Disturbances can throw the system out in the S-curve aprons where the error signal is not sufficient to bring it back into the null. In addition, the fringe visibility is strongly affected by the presence of binaries with angular separations on the order of the telescope resolution.

Figure 3 shows the theoretical transfer functions of the coarse track and fine lock modes. Also shown for reference is the transfer function of an ideal 4-quadrant detector system. The coarse track and 4-quadrant detector curves assume diffraction limited images. Fine lock (interferometric) and the 4-quadrant detector system both achieve the ultimate gain at the null. The lesser gain of the coarse track mode is due to the mode used to "scan" the image. In the nutation system, the intensity in each quadrant of the aperture is not a pure measure of the intensity in the corresponding quadrant of the image, but is contaminated by that of the other image quadrants. A nutation system would approach the gain of a 4-quadrant system for a nutation radius equal to the diagonal of the square aperture, but it would greatly lose in efficiency.
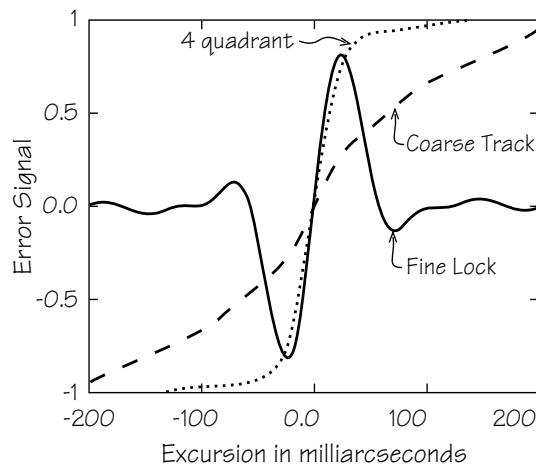
2.4 PCS/FGS INTERFACE

In both the coarse track and fine

Figure 3: Theoretical transfer functions

lock modes the error signal is fed into the FGS control system in order to maintain the star at the null of the transfer functions. In other words, the FGSs are always tracking the guide stars independently of what the vehicle does. The FGS control loop operates at 40 Hz. The telescope positional error detected by the FGS is fed into the spacecraft PCS to correct its attitude. This correction is updated at the rate of 1 Hz.

3 CURRENT PERFORMANCE OF THE GUIDING SYSTEM

Coarse track is much worse than predicted because spherical aberration degrades the image (Figure 4). However, the system is still using settings based on a diffraction limited image and tuning it up should significantly improve its performance.

Simulations show, for example, that the gain should be reduced by a factor of 2 for the degraded image (Figure 4).

In fine lock, tracking is within specifications (3 mas rms) during quiescent periods, and is essentially magnitude independent. However, fine lock cannot be reliably obtained with the full aperture due to the low visibility of the S-curves. Use of the pupil stop (2/3 full aperture) is mandatory on an operational basis. This need is very likely explained by residual aberrations in the system as exemplified by Figure 6. Whether these residual aberrations are due to still imperfect alignment of the telescope or to secondary effects of the spherical aberration in the FGS optics is still unclear at this time. The interferometric system used in the
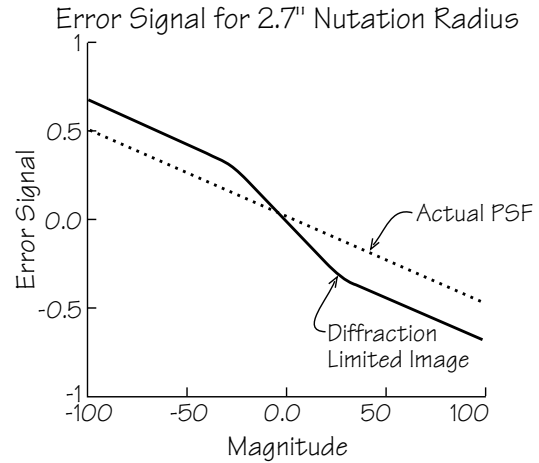
Figure 4: Simulated coarse track
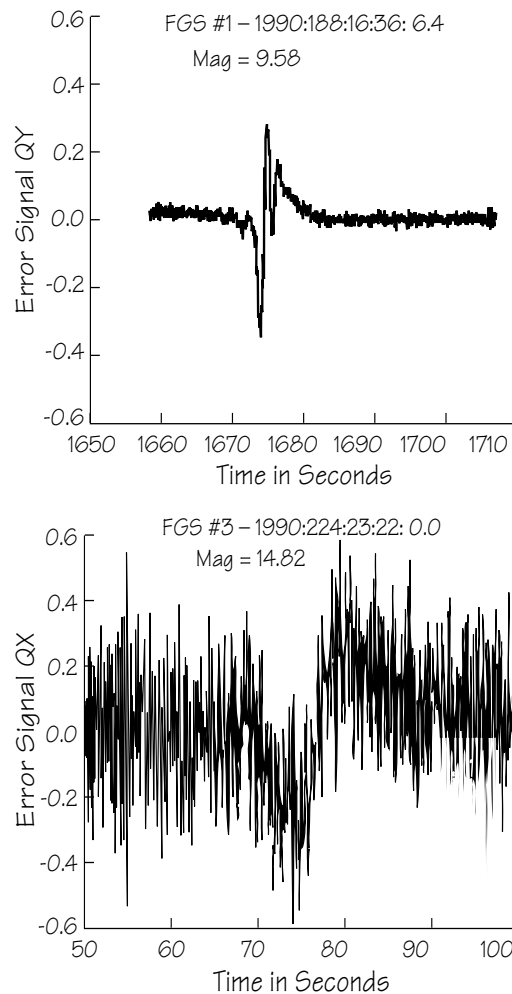
Figure 5: Typical bright-star and faint-star S-curves with 2/3 pupil stop in.

96

fine lock mode is, in theory, unaffected by axisymmetric components of the wavefront such as focus and spherical aberration, but the spherical aberration of the telescope could possibly amplify misalignment effects in the FGSs.

A consequence of the fringe visibility degradation and of the necessity of stopping down the pupil is to limit fine lock to stars brighter than 13 to 13.5 m_v.

Figure 5 shows a typical S-curve for bright (9.58 m_v) and faint stars (14.82 m_v). Both are taken with the pupil stop in. Note the low visibility and the double hump in the case of the bright star even with the pupil stop in, and the very low visibility compared to noise for the faint star.

Figure 6 shows the effect of coma on the S-curve with or without the 2/3 pupil stop in. Minimal amounts of aberration strongly degrade the S-curve visibility with the full pupil. Fringe visibility is almost unaffected by aberrations when the pupil stop is in. Similar curves are obtained for astigmatism.

An unrelated problem of the current guiding performance is the strong jitter and loss of lock induced by the day/night transitions. During the passage of orbital day to orbital night or vice versa, thermal shocks in the solar arrays create oscillations in the vehicle (Figure 7) which cannot be compensated by the PCS, and which almost systematically produce loss of lock when faint guide stars are used. Coarse track with its large dynamic range is unaffected.

Figure 7 shows a typical day/night transition with oscillations of 100 mas, inducing loss of lock. Recovery is usually obtained within about 5 minutes.
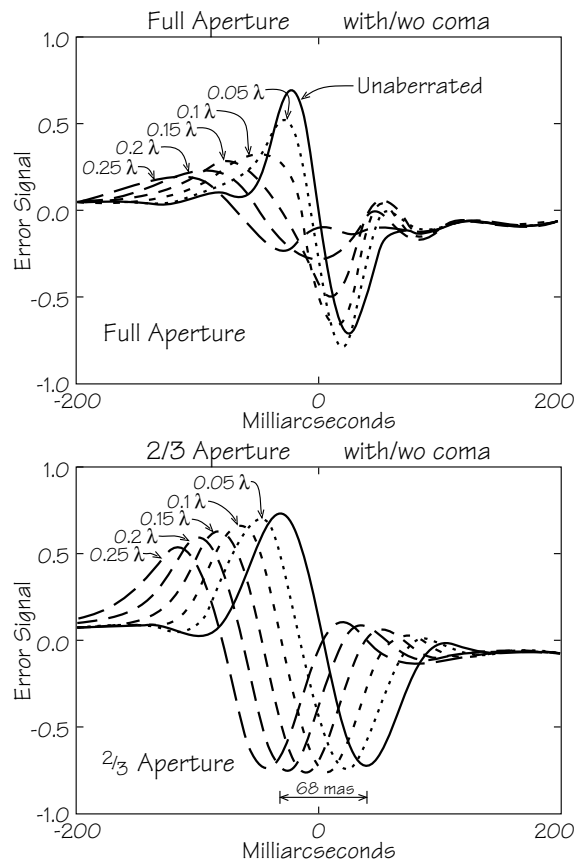
Figure 6: Effect of coma on S-curves with full aperture and 2/3 aperture
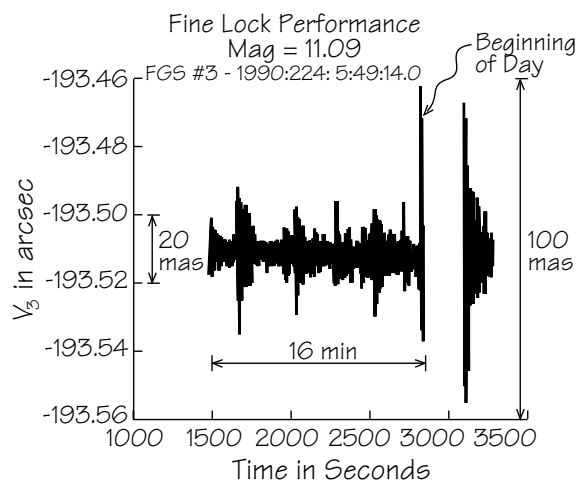
Figure 7: Typical day/night transition

4 PROSPECTS FOR IMPROVEMENT

An active program is currently under way to seek means of returning the guiding capability to values close to the pre-launch expectations. Areas under study include:

- improving coarse track by adjusting the gain
- improving fine lock robustness and ability to lock on faint stars by adjusting walk-down parameters and averaging time
- reducing the effect of day/night transitions by tightening the PCS loop
- reducing residual aberrations by refining the alignment of the telescope optics.
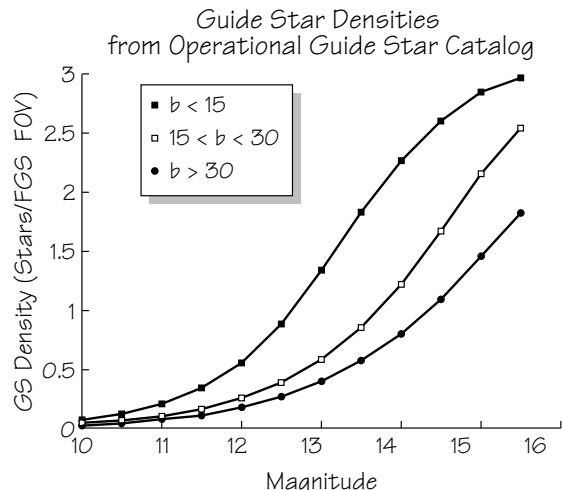
Figure 8: Guide star density vs.

5 GUIDE STAR AVAILABILITY

The HST guiding system had been designed to operate on stars down to 14.5 m_v in order to ensure an 85% probability of finding guide stars at the galactic pole. Figure 8 gives the density of guide stars per FGS field of view as a function of magnitude and Figures 9 and 10 show the effect of various star densities on the ability to
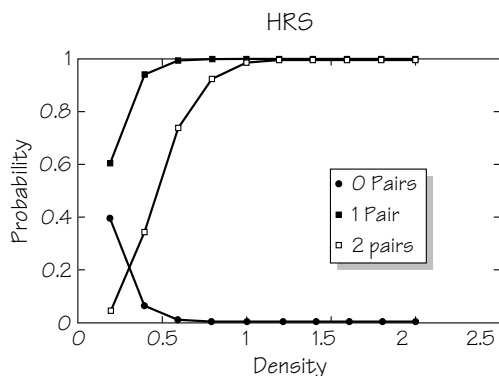
Figure 9: HRS acquisition probability versus guide star density
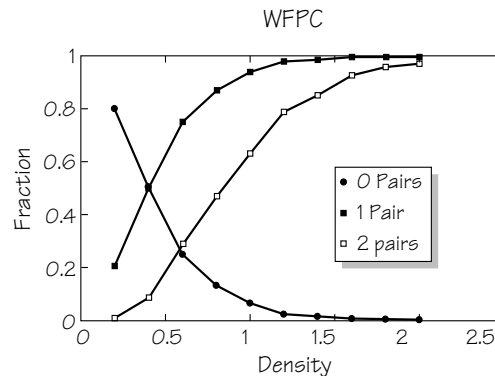
Figure 10: WFPC acquisition probability versus guide star density

find guide stars for a typical axial instrument (HRS) and the WFPC. The axial SIs (which have their aperture off-axis) greatly benefit from rolling the spacecraft around the desired target. This enlarges the total area available within the FGS fields of view, thus increasing the number of possible guide stars. This is not the case for the WFPC, which is essentially on axis. Table 1 shows the effect of limiting the guide stars to 13.5 m_v as opposed to 14.5 m_v for the axial instrument and the WFPC. The impact on the axial instruments is negligible. However, sky coverage for the WFPC is limited to galactic latitudes below 30 degrees. With no improvements, most WFPC pointings at higher latitudes will utilize coarse track and suffer some guiding degradation.

SHUTTLE SERVICING OF HST

THE SPACE SHUTTLE PROGRAM is the only capability in the world for on-orbit maintenance or retrieval of satellites. This is rightfully a source of national pride, and its use for a successful on-orbit fix of the HST spherical aberration problem would be an internationally recognised accomplishment.

The HST Strategy Panel has focused on options for the first HST maintenance mission, which has been scheduled for June 1993 since before the HST launch in April 1990. Even though the HST was designed from the outset to be serviced by astronauts in orbit, in "casting its net widely" the Panel identified many options that lay outside existing servicing concepts. Because the Space Shuttle Program is rigorous and exacting as regards approving and planning missions, as indeed it must be, the Panel has attempted to evaluate its options according to compatibility with the existing Shuttle program. Below is a tutorial, followed by a division of the identified options into broad categories of feasibility based on servicing capabilities.

The payload handling capabilities of the Shuttle are as follows. The Payload Bay is 18.3 m in length and 4.6 m diameter, with a mass capacity in excess of 24,500 kg. The crew can operate payload latches and umbilical mechanisms electrically from inside the crew module. Currently, they can deploy, retrieve, and handle payloads using aids such as the Remote Manipulator System (RMS), and a variety of "dexterous" teleoperators and semi-autonomous robots are under development. Specialized servicing equipment also exists, such as the Flight Servicing System (FSS) for supporting the HST with its aft end in the rear of the Payload Bay. Finally, and of great importance for HST, astronauts can work manually on payloads by means of EVA.

There have been past demonstrations of the Shuttle capability to service payloads in orbit. The Solar Maximum Observatory was repaired on mission STS-41C. Astronauts replaced the Modular Attitude Control System and Coronagraph/Polarimeter main electronics box, both of which had failed, and installed a shield over the Soft X-ray Polychrometer propane vent port. On mission STS-51A, two satellites, Westar-VI and Palapa-B2, were retrieved from useless orbits, returned to earth, and subsequently re-launched on expendable launch vehicles (ELVs). On STS-51L, the SYNCOM IV-3 satellite, an electrical dud, was activated by installing additional electronics that enabled the motor to fire and operations to commence. On STS-50, the first flight of the new Space Shuttle Endeavor, astronauts will attach a new solid rocket motor to INTELSAT-VI F3, which is currently stranded in a low orbit due to incorrect separation circuit wiring on the initial ELV launch.

The singular aspect of the earth-orbital environment is the absence of perceptible relative acceleration—"Zero-g." This is due to the cancellation of the gravitational force by an equal and opposite centrifugal force due to the orbital motion. It is a sort of "eternal free fall" situation. Weightlessness is readily observed, but freedom from gravitational/centrifugal field-induced torques is strictly true only for point masses and other specific classes of mass distributions. Such torques are, however, sufficiently low to present little problem in servicing an object as large as HST. Aerodynamic drag and torque effects exist but are generally insignificant on

the time scale of a maintenance mission.

Zero-g makes the handling of large masses feasible without cranes, forklifts, workstands, or similar devices. Inertia, however, is not affected, so consideration must still be given to initiating and terminating relative movement. A specific troublesome instance of this is the set of requirements on RMS "runaway" loads for payload-attached grapple fixtures. These requirements are derived from the "worst case" forces and torques that could be developed by an RMS that began moving without control—"runaway"—while handling a massive payload. For example, the requirement to be met is 1200 ft-lbs of torque about the grapple fixture probe axis. These requirements may be severe design constraints for smaller, less massive items unless the requirement is waived.

This same phenomenon of zero net gravity means that a crew member cannot "stand" in the conventional sense of the word but needs restraint. If not tethered or mechanically attached to something such as a foot restraint, he will float away. The same statement, of course, applies to tools and equipment; such items cannot be "laid down" without floating away but must be tethered to guard against inadvertent loss. Usually, this is accomplished by fitting the item with Velcro or by using special holders.

Five NASA career astronauts constitute the minimum Shuttle crew. Their activities divide broadly into two categories: Intra-vehicular (IV) or "inside" activities and EVAs. IVAs include vehicle operations, malfunction analysis and mitigation, housekeeping, RMS and other payload unique operations, and EVA support. During an EVA one "inside" crew member is designated as "IV-1." For scenarios including RMS usage, IV-1 is generally either the commander or pilot, and varies from crew to crew. This individual trains with the designated EV crew members and functions on-orbit as valet, safety monitor, procedure reader, and documentary photographer. EVAs include all crew activities in an ambient pressure too low to sustain human life.

There are many constraints and limiting factors on EVAs to be considered. The first is time. Each EVA provides a 6-hour block of time in which two crew members are available for payload related tasks. "Time at vacuum" is limited by consumables, with LiOH (for carbon dioxide removal), being the most inflexible. LiOH capacity is equivalent to a specific number of metabolic BTU's developed by the using crew member, and the temporal capacity can be extended somewhat by easing the physical load on that person, or at least by equalizing the workload between the two EVA crew members. Sublimator water (for cooling) and primary oxygen can be replenished during an EVA by means of the umbilicals in the Airlock. Battery energy can be conserved, but not recharged, by connecting to the same umbilical during a period of inactivity.

A maximum of two scheduled EVAs per mission are permitted unless additional crew members and/or equipment are carried. The capabilities for an additional unscheduled payload EVA, and for an Orbiter contingency EVA are also provided. The first EVA may be scheduled no earlier than the third day after launch (FD4) without a waiver. The concern is over crew recovery from Space Adaptation Syndrome. The last EVA may be scheduled no later than the second day before planned de-orbit. The day before de-orbit is reserved for Orbiter stowage and a contingency EVA, if required.

The pressure suit/life support system demanded by the space environment is an encumbrance. Although continually improving, the gloves significantly reduce tactile feedback, limit dexterity somewhat, and accelerate hand fatigue. The helmet is rigidly locked onto the hard upper torso of the suit and constrains the head from looking significantly upwards, which severely limits the capability for overhead work. The joints in the suit limit the effective two-handed work envelope to a volume about the size of a small beach ball in front of one's chest. The bulk of the suit dictates a 43 inch diameter clear translation path, with reductions evaluated on a case by case basis; *e.g.*, the Orbiter 40 inch diameter hatch. Prolonged rotary motion of simple hand tools is difficult, calling for ratcheting or power tools.

With regard to adhesives and taping, any bonding operations with graphite-reinforced epoxy composites would require qualification of an epoxy that could be mixed and cured in vacuum. Some success was had in developing a room temperature vulcanizing (RTV) material for use in the on-orbit repair of the Orbiter thermal protection system tiles. Only a few tapes are known that can be handled and applied in vacuum and develop any significant adhesion. The historical best is a Kapton tape, used on the Solar Maximum Repair Mission, but even it developed less adhesion than when applied at sea level. When tested in a thermal vacuum chamber at Johnson Space Center (JSC), the "black MLI" tape used on the HST Focal Plane Structure insulation did not adhere well at all.

The Soviets have demonstrated a hand held electron beam welder during EVA, though it is not known to have been used operationally. "Vacuum welding" is theoretically possible but has not been encountered in manned spaceflight.

With respect to astronaut translation, handrails support access to most planned worksites on the Orbiter and HST. The manipulator foot restraint (MFR) is a work platform that is grappled by the RMS and allows a crew member to be positioned by it. The Manned Maneuvering Unit (MMU) provides free flight capabilities within 100 m of Orbiter, but is not currently carried on HST missions. No "soaring" is permitted due to concern that a "miss" might lead to a broken safety tether and a "man overboard."

Worksite restraints provide for both hands to be free for effective work. "Reactionless tools" were rejected long ago in favor of restraining crew member adequately to react against the tool torques. MFRs, portable foot restraints (PFRs), tethers, and unique restraining devices are all available. Velcro "sheds" and thus is a potential source of contamination, but it works very well in vacuum. Use of one crew member to hold the other does not work well at all and is inefficient. Very small quantities of magnetic materials were used in construction of Orbiter and HST, so magnets are useless for restraint.

Safety-related issues include, first, "sharp edges." There are explicit requirements on the minimum radii for two- and three-planar corners, on minimum thickness of sheet metal edges, etc. Most of the foregoing can be waived or inspected to the equivalent of "good shop practices." Projecting ends of safety wire could puncture a suit and must be dressed back in such a manner as to prevent contact with any part of the pressure suit. Projecting screws could snag a suit and must either be capped with an "acorn" nut, cut off flush, or blunted with a glob of epoxy. All EVA-accessible areas on the HST have been inspected and accepted as safe. However, the interior of the Forward Light Shield has not been addressed. The radially

projecting edges of the baffle assemblies would need particular scrutiny if an EVA into this volume is required, both from the standpoint of crew safety and from concerns about chipping off black paint as a result of inadvertent contact. A preliminary look at specifications for the baffle edges shows an approximate radius of 0.005 inches, after painting, while the EVA requirements stipulate that any sheets thinner than 0.02 inches must have rolled or curled edges. The underlying aluminum fin itself has an edge thickness of only 0.002 inches.

Tethers for personnel and tools are key safety items. Since the Space Shuttle pressure suit does not employ an umbilical while working outside of the airlock, a 16.7 m self-tending safety tether system is provided for each EVA astronaut. This links him or her to a slider on one of two dedicated cables, along the port and starboard payload bay door hinge lines. Although the safety tethers can be disconnected at either end, in practice they are used at all times except when flying the MMU or occupying the MFR. Additionally, each EVA crew member is provided an additional 0.86 m long waist tether for use in helping to hold position at a worksite, or in connecting to relatively large loose items.

All loose items must be positively tethered at all times to prevent loss by drifting away. Each EVA crew member has one or more of the wrist tethers used for this purpose. Each item of loose equipment is required to provide a suitable means for connection of a tether hook, although not necessarily on an exclusive basis. Many tools are stowed on "tool boards," and tethered thereto with tiny self-retracting tethers. The boards may be installed either on the MFR or on a crew man's chest-mounted "mini-work station." There is also a "self-tethering connection system" employed in configuring items such as socket wrench assemblies.

The sun introduces thermal safety concerns. The SIs have individual, tight constraints against direct impingement of sunlight on their (black) surfaces to guard against overheating. This necessitates selecting an Orbiter attitude that will allow the entire SI changeout process to take place "in a shadow." Owing to the low temperature at which the MLI in the Support Systems Module (SSM) equipment bays was baked out (about 130 degrees Fahrenheit), these bays are also constrained against the impingement of direct sunlight on their interiors. If this constraint is violated, additional volatile materials may be driven off, with the possibility of re-condensing elsewhere on the HST and causing a problem.

On the day side of the orbit there is enough light scattered into shadowed areas to work comfortably. On the night side, numerous sources of artificial illumination are available. There are six flood lights mounted along the bottom of the payload bay, and a seventh on the forward bulkhead aimed aft. The RMS end effector carries a floodlight that is usable when the MFR is not employed. Each EVA crew member has two small battery-operated, helmet-mounted light assemblies. The HST itself has 28 v.d.c. outlets provided for a portable worklight assembly, which is a 50 watt aircraft type floodlight gimbal-mounted to an EVA clamp assembly. If required, flashlights could be provided, too.

Training is a critical aspect of preparing for any mission to service HST. There is no single "perfect" facility for the simulation on earth of the weightless EVA. Parabolic flight in the KC-135 aircraft provides true free fall conditions, but only for a maximum duration of about 45 seconds—and followed by a two-g pullout! The size and weight of the test set-up is limited by aircraft constraints, and the

coordinate (aircraft) reference frame rotates through about ninety degrees as the "push over" parabola is executed. All things considered, the underwater neutral buoyancy technique is the most satisfactory training facility, however. Large, massive mockups can be used, and the duration of the "run" is limited only by the endurance of the support divers. Such operations may be interconnected with other facilities by television and voice communications links for integrated simulations, although some of the results have been disappointing. On the negative side, the water exhibits viscous drag on moving objects; it is difficult or impossible to ballast small dense items such as tools neutrally; test hardware is subject to corrosion or other water damage; functional electrical items must be specially protected; and the crew member is not "weightless" *within* the pressure suit. Facilities exist at the JSC (WETF = Weightless Environment Training Facility: 10 x 24 x 7.6 m deep), the Marshall Space Flight Center (NBF = Neutral Buoyancy Facility: 23 m diameter x 12.2 m deep), and McDonnell-Douglas (Huntington Beach). A new very large facility (NBL = Neutral Buoyancy Laboratory: 41 x 72 x 18 m deep) is planned at JSC in support of Space Station Freedom development.

SI changeout can be broken down into two categories: those of axial and of radial SIs. The axial SI changeouts are all virtually identical; the only differences arise from the existence of "right-" and "left-handed" configurations and the potential use of cryogenic gas vent ports by future advanced instruments. At a summary level, the following steps are involved in an axial SI changeout once the HST is situated on the FSS.

1. Open the appropriate pair of Aft Shroud access doors.
2. Disconnect wingtab electrical connectors (4), ground strap, and pre-launch purge hose.
3. Open "A" and "B" latches using ratchet (or power) socket wrench and MFR.
4. Using handles on SI and working from the MFR, slide the SI towards the aft (V1 = 100) bulkhead and then extract it from the Aft Shroud.
5. Temporarily stow the removed SI while extracting the replacement unit from its protective enclosure.
6. Reverse the above procedure to install the new SI, using microswitch controlled indicator lights to verify seating in the latches and specific torque values when closing them.

Radial SIs include the WFPC and the FGS's (3). FGS changeout shares some functional attributes from both the axial and WFPC scenarios, but will not be further described herein. WFPC changeout proceeds through the following summary steps.

1. Attach handhold plate/radiator protector using four bolts.
2. Release ganged electrical connectors using socket wrench on the single drive shaft.
3. Disconnect ground strap.
4. Release the "A" latch using socket wrench.
5. Extract the old WFPC using the MFR and stow temporarily.
6. Remove the new WFPC from its protective enclosure and remove the pickoff mirror protective cover.
7. Insert the new WFPC into the radial bay and reverse the above steps, torquing

the latch and the ganged connector drive shaft to specified levels. Microswitch driven status lights may be observed through either of the +V2 bay doors.

It is evident from the foregoing that a significant investment in time and money has already been made in designing hardware, building mockups, validating the design, and developing procedures for the changeout of standardized axial and radial bay modules. Thus, any option that mimics one of the standard modules has several major advantages. Development of the optical portion of the "fix" may begin immediately with confidence that the interfaces are well defined. The approach will be acceptable to the servicing mission crew representatives on the basis of similarity, eliminating the need for an urgent mockup-supported feasibility evaluation. Crew training may be planned to be routine and to follow standard scheduling templates.

If a new procedures or tools were envisioned for the HST repair mission, simulations for development testing should begin as soon as possible. This would maximize the development time and allow recovery if a "show stopper" were to be encountered. It would also yield greater maturity at the various management reviews dealing with mission manifesting, safety, and so forth. Such testing and development can be conducted anywhere; historically much of the HST work was accomplished in the MSFC Neutral Buoyancy Facility. Training hardware, specifically including mockups for underwater use, must be delivered to JSC by Launch minus 12 months. Virtually all training is conducted at JSC. Owing to other priorities on WETF utilization, mission-specific underwater training typically does not begin in earnest until Launch minus 6 months.

There is an understandable conservatism in NASA management regarding new missions, especially near-term, manned missions accompanied with new requirements and intense media attention, such as a repair mission for HST. Obviously, crew safety is sensitive factor. For example, the longer it has been since the last actual EVA the greater is the management-perceived risk. That there is some increase in risk over staying inside the pressurized cabin is indisputable, but the participants generally are more confident. By the time of the next scheduled EVA (STS-37, Secondary Objective on the Gamma Ray Observatory Deployment Mission, scheduled for April or May, 1991), it will have been well over five years since a U.S. EVA. EVA will certainly be allowed on an HST M&R mission, since it is required, but mission rules may be more conservative than if EVAs had been more common recently.

In any Shuttle servicing mission, there are inherent risks to the payload. There is always a "safety of flight" requirement on the Orbiter that it be capable of performing an emergency de-orbit burn within 20 minutes of sustaining an imminently

life-threatening event, such as a large hole in the cabin pressure shell. If this should happen between EVAs, the HST would be unceremoniously redeployed in order to allow the payload bay doors to be closed. On a lesser scale, a major failure in a pressure suit or life support system could obviously require interruption of work at almost any point in the maintenance scenario. A spare suit/life support system is carried on all missions with scheduled EVA; however, there is a good probability of recovering from this latter type of failure. Even without unexpected departures from the planned timeline, there is little time for checking out completed work. About 36 hours is allowed between consecutively scheduled EVAs. The HST redeployment would probably be accomplished near the end of, but during, the last scheduled EVA in order that EVA support is on hand for latch and umbilical disengagement support without requiring another EVA. The HST aperture door would not be re-opened until about 45 hours after re-deployment due to contamination concerns.

EVA Feasibility Assessment

1. Within Baseline Capability.
Definition: The associated EVA procedures and tools are already in existence. Some astronauts have trained on or have experience with them. Where specific ORUs are involved, these have internal upgrades or very minor external differences, by means of which the improvement in HST performance is obtained. Within each category, options are listed roughly in descending order of attractiveness from an EVA feasibility standpoint. Endnotes are indicated by "(n:#)."

  B1: *Single Refractive Corrector for Individual SIs (n:1)*
  B2: *Double Refractive Corrector for Individual SIs (n:1)*
  B3: *Two-Mirror Reflective Corrector for Individual SIs (n:1)*
  B4: *Double-Cassegrain Relay for Individual SIs (n:1)*
  D3: *Halo Baffle at Individual SIs (n:1)*
  C1: *FGS II (n:2)*
  C2: *Modification of FGS I (on Ground) (n:2)*
  C4: *Modify WFPC II Secondary Mirrors (n:2)*
  C7: *Modification of FOC: Change of Optical Head Unit (on Ground) (n:2)*
  C8: *NICMOS Internal Corrector (n:2)*
  C9: *STIS Internal Corrector (n:2)*
  ###: *"COSTAR" (n:2)*
  ###: *Use of Space Hab for Upgrade of FOC (n:2 and 5)*

2. Modest Extension of Capabilities.
Definition: Although the full suite of tools and procedures required for the accomplishment of these options have not been developed, many existing tools and pieces of current procedures would be applicable. Development of the balance of these items appears straightforward based on past on-orbit experience, and there appears to be little risk from the crew operations standpoint that these options could not be successfully completed.

  D1: *Aperture Masking*
  A5: *Full Aperture Correction Flat*

A6: Gas-Filled Correction Lens (n:3)

A4: Full Aperture Correction Plate (n:3)

A3: Overcoating the Primary Mirror (Evaporators outboard of SM) (n:3)

3. INVOLVED/MODERATE RISK.

DEFINITION: These options represent extension of current capabilities into "new territory." There have been insufficient resources within the scope of the HST Strategy Panel to evaluate these options to the depth required for a full assessment of the difficulty involved or the ramifications of undertaking any of them. Accordingly, a moderate programmatic or technical risk to the successful completion was assigned at this stage of maturity.

A11: Three Aspheric Plates in Central Baffle (n:4)

A12: Double Cassegrain Relay in Central Baffle (n:4)

C5: Modification of FOC: Refractor in Upper Channel (n:5)

C6: Modification of FOC: Refractor near Exit Pupil (n:5)

###: FOC: Replacement of Filter Wheels (n:5)

4. VERY DIFFICULT/HIGH RISK.

DEFINITION: In addition to delving far into "new territory," this group of options requires either that one EVA crew member physically position himself between the primary and secondary mirrors and complete exacting tasks within this very restricted, poorly illuminated, work space or that an elaborate set of tools or teleoperated devices be developed on a tight schedule for the remote accomplishment of these same tasks. In the former case, very significant crew safety issues must be resolved, and methods developed to protect the HST optics from contamination. Very significant technical and programmatic risks also appear to be associated with these options.

A3: Overcoating the Primary Mirror (by evaporators between SM and PM)

A7: Secondary Mirror Replacement

A8: Secondary Mirror Reconfiguration of HST to ƒ/13.25 Ritchey Chretien

A9: Secondary Mirror Replacement and 2-Plate Corrector in Central Baffle

A10: 1-Plate Corrector on Secondary Mirror and 2-Plate Corrector in Stovepipe

5. VIRTUALLY IMPOSSIBLE/VERY HIGH RISK.

DEFINITION: At the current stage of maturity there does not appear to be any way of accomplishing these options on-orbit within known constraints, or there appears to be such a high risk of failure and of degradation of the HST from its current capabilities as to be not worthy of further investigation.

A2: Thermal Deformation of Primary Mirror (by additional heaters installed on-orbit) (n:6)

A1: Mechanical Deformation of Primary Mirror (by inflatable ring installed on-orbit) (n:6)

6. EVA NOT INVOLVED.

DEFINITION: These options require no direct crew involvement, and can be exercised from the ground over existing command links.

C3: Thermal Fixes to SIs (using existing heaters)

D2: Aperture Door Vignetting

A2: Thermal Deformation of Primary Mirror (using existing heaters)

NOTES:

1. Assuming that these options can be effected by mounting the required elements on the WFPC II mirror support arm, or by incorporating them into the "COSTAR." If neither of these is feasible, then the relevant option drops down into Category 2: "Modest Extension of Capabilities," or lower.

2. Effected by on-orbit removal and replacement of axial and/or radial bay modules.

3. Assuming that the EVA crew members were not required to go fully inside the Forward Light Shield, but could complete the tasks by means of long handled tools to be designed for the purpose. If such crew ingress should be required, the relevant option drops down into Category 3: "Involved/Moderate Risk." It should be noted in passing that the RMS plus MFR combination as it exists cannot position a crew member very far into the interior of the Forward Light Shield due to a lack of co-planarity between the RMS shoulder and elbow joints and the V1 axis of the HST as positioned on the FSS.

4. Assuming that a satisfactory deployment mechanism can be developed. It appears that one might be configured to operate temporarily within the WFPC bay for the purpose of making such an installation. Mechanical "hangups" within the interior of the "stovepipe" would be virtually impossible to resolve by direct manual intervention, however, and could be crippling to the entire HST. This would place a very high premium on the reliability and/or the redundancy of the deployment mechanism.

5. These options still await detailed evaluation. Anecdotal information has been received regarding fasteners that fail to come out when disassembly of the FOC is attempted. If the Space Hab module can be made available for use as a pressurized work area these options would be considerably more attractive than if all work were required to be accomplished in vacuum during the limited time of the EVA's.

6. The access required for implementing these options on-orbit does not appear to exist. Both require extensive access to the volume between the back of the primary mirror and the Figure Control Actuator Reaction Plate, which is obstructed by multi-layer insulation, the "Invar ring," and the inner shell of the SSM.

RISK MANAGEMENT

A SIGNIFICANT FACTOR in the evaluation of any fix to the HST is identification of the risk involved in implementing the solution. Once a solution is identified, a risk management plan should be developed to minimize any remaining risk to acceptable levels. In evaluation of risk for a given fix, several areas should be considered:

• **Damage or loss of the HST.** This would include physical damage caused by astronaut or shuttle impact with HST elements, loss or damage of HST by stress induced in landing and re-launch, or unintended interference with existing HST components by newly installed elements.

• **Failure of Solution to resolve problem.** What is the risk that the fix won't solve the problem? Things to consider here include: how accurately the optical prescription must be known, how sensitive to placement the fix is, how well the technology is developed, and how well the fix can be tested and verified on the ground. These considerations lead directly to the concept of reversibility, which refers to how difficult it is to reverse the fix if necessary.

• **Reversibility.** Proposed solution should be evaluated to determine if the fix can be reversed by ground command, on-orbit removal, or ground removal. Some of the fixes identified are not reversible without major structural replacement. For example, overcoating the primary would require removal of the primary mirror. This fix would be impossible on-orbit and difficult on the ground.

• **Contamination.** During the construction of HST, great effort was expended to keep the optics clean. Small amounts of organic contaminates can completely destroy the throughput in the ultraviolet. Ordinary dust on the mirror scatters light and interferes with sensitive observations. Debris, like the little hooks that break off the Velcro straps each time astronauts remove a tool, have the potential to block the very small apertures that are present in each instrument or jam sensitive mechanisms. Each fix must be evaluated as to the possibility of organic contamination to the optics as well as generation of small items such as dust or debris. All fixes will introduce new environmental elements near HST: shuttle generated propellent exhaust, space suit expellant, or in the case of HST return, air ingestion on descent. Once on the ground the pre-launch contamination control would need to be reinstituted.

• **Costs, schedule, performance.** Each of the fixes has associated with it the cost-schedule-performance risk typical of any high technology endeavor. For the fixes identified by the panel, an excellent indicator of risk in these areas is how new the technology is and if the implementation method is a planned or simple extension to a planned capability in the M&R servicing program. Although detailed cost is outside the scope of the Panel, quantification of the relative risk to schedule and performance of each solution must be accomplished.

IMPLEMENTATION FACTORS

KEY CONSIDERATIONS IN THE SELECTION of a repair strategy for the spherical aberration problem on HST are the feasibility and difficulty of the on-orbit operations required. The limiting factors are the construction of the HST, which makes access to some areas difficult or impossible and the capabilities of suited astronauts. Both of these limitations can be overcome to some degree by the development of tools and devices to simplify installation of the components required for the fix.

The options for restoring HST's capabilities can be divided into two groups. First, fixes that can be implemented using the available astronaut EVA planned for the scheduled M&R missions including simple extensions to this baseline set of capabilities. Second, those options which call for new capabilities far in excess of those available and tested. In the second case, access required to implement the fix may not be available since on-orbit maintenance was not envisioned in these locations.

When the options are viewed in this manner, we can readily see that most of the options, nineteen of twenty-six, can be implemented utilizing the baseline M&R capabilities or simple extensions. The remaining seven would require extensive new on-orbit capabilities and at least two of these appear to have severe access difficulties.

POST-FOCAL PLANE
Options C1 through C9, Post-Focal Plane corrections in individual SIs, are all within the baseline M&R capabilities. These options are accomplished by modifications done internal to an SI. The SIs are designed to be removed and replaced on-orbit. The capabilities and facilities needed to perform this type of activity have a high degree of maturity. Astronauts have practiced these activities in water tank simulations and participated in real insertion and extractions of SIs on the flight hardware during development. Options C1, C4, C8, and C9 accomplish the fix by corrective optics in the second generation SIs. These SIs, now in the planning or development stage, would have the modification incorporated during construction and would be installed in HST on a standard servicing mission. Options C2 and C5 to C7 are modifications to existing instruments currently installed in HST. These options would require removing the SI from the HST during a routine servicing mission, ground repair and subsequent reinsertion on orbit. Both removal and reinsertion are baseline capabilities. (For option C5 to C7, Modifications to the FOC, it may be possible to insert the corrective optics into the SI on-orbit. This is not within the baseline capabilities and is being studied to determine if it is feasible to accomplish on orbit. This solution would save the down time associated with repairing the FOC on the ground and reinstalling on a future maintenance mission.) Option C3, if technically feasible, is the simplest in this group. It could be accomplished using ground commands to adjust heater settings to distort the optics with in the SI. While simple to implement, the solution is unlikely to work.

PRE-FOCAL PLANE
Implementation of corrections in the area in front of the SIs, Group B Pre-Focal Plane Corrections, can be implemented using existing baseline capabilities if the corrective optics are built into an axial SI-like package. This package (COSTAR) would be

inserted into one of the axial SI bays on HST on a routine service mission. Ground commands would cause the COSTAR to insert the corrective optics into the optical path required to implement solutions B3, and possibly D3. Both spectrographs (FOS, HRS) and the FOC could be corrected on a single mission. A proposal to place corrective optics on the pickoff arm of the WFPC II does not provide the adjustments necessary to align the optics correctly on orbit. The COSTAR concept has several advantages:

1. Fixes up to four SIs on a single mission.
2. Could be piggy-backed with WFPC II installation.
2. Does not impact future SIs
3. Uses existing technology
4. Eliminates the need for accurate placement of optics by suited astronaut.
5. Uses baseline servicing techniques
6. Allows normal servicing of all SI
7. Allows future adjustments and alignments or removal by ground command.

FULL-FIELD CORRECTIONS WITH FULL APERTURE

The Full-Field Corrections with Full Aperture, Group A solutions, do not seem to be implementable using baseline M&R techniques (with exception of A11 and A12). A1 and A2 require access to the back of the primary mirror. Access to this area appears impractical on orbit. The three options A7 to A10 require access to the secondary mirror. Access to this area by astronauts is not practical. While some sort of tool might be developed to make installation possible, it is considered by both astronauts and engineers to be exceedingly difficult in both design and installation. The next solution in order of difficulty is A4, the Full Aperture Corrector Plate. This option requires developing challenging new M&R techniques. While less difficult than the previous A Group solutions, the mass of the structure combined with the difficulty of attachment lead engineers and astronauts to the conclusion that this is also an exceedingly challenging option to implement. The correction using a Full Aperture Plate, A4, has the advantage of being less massive than A5, however the solution incorporates a number of thin optical plates which will be difficult to handle on-orbit. Option A6, the placement of a Gas-Filled Corrector Lens, if technically feasible, is more than an extension of existing techniques, but perhaps more feasible to accomplish than the other group A techniques. This option assumes that the lens would be held in place by inflation and that its location is not critical. Option A11 and A12 require inserting elements into the Central Baffle of the primary mirror. It appears a tool could be developed that has the form factor of a radial SI (FGS or WFPC). This tool could be inserted using baseline techniques and the element package could be placed and locked into the central baffle to within 1 mm automatically.

MASKING

The solutions in the D group, Masking, range in complexity. Using the existing aperture door to vignette the beam combined with secondary mirror re-centering, can be accomplished using ground commands. Individual instrument baffles in front of each SI could be implemented using the COSTAR. Attachment to the new WFPC II pickoff mirror arm is also an option since placement of a baffle would be less sensitive than placement of optical elements. The most complex of the group is the full aperture mask. The removal of the aperture door, however, is a baseline capability. Using these attachment points for the mask mechanism looks attractive. Preliminary studies indicate this is a solution that both engineers and astronauts

SECOND GENERATION SCIENCE INSTRUMENTS (SIs)

THREE REPLACEMENT SCIENCE INSTRUMENTS (SIs) are under development for on-orbit change out: the Wide Field/Planetary Camera II (WFPC II), the Near-Infrared Camera and Multi-Object Spectrometer (NICMOS) and the Space Telescope Imaging Spectrograph (STIS). The current schedule shows installation of WFPC II in 1993, NICMOS in 1996 and STIS near the end of the decade. The schedule and order of the instrument developments is under review due to the current optical problems with the observatory.

The WFPC II is essentially a clone of the existing WFPC with improvements in the CCD imaging chips, an additional optical filter complement, and refinements in the support electronics.

The NICMOS extends the wavelength coverage of the observatory into the infrared for both imaging and spectroscopy.

The STIS is a replacement for both of the current spectrographs, the High Resolution Spectrograph (HRS) and the Faint Object Spectrograph (FOS). It covers a wider wavelength range than the HRS and FOS combined, and it will take two-dimensional data.

It is important to note that WFPC II and NICMOS can accommodate corrections to the existing spherical aberration with little impact to their planned performance.

WFPC II

Prior to the discovery of the spherical aberration problem, WFPC II was scheduled to be installed during a refurbishment mission approximately three years after launch of HST. The schedule for construction of the new camera at JPL is presently under review, but completion will probably still occur in the latter half of 1992 with installation an orbit in mid-1993.

The baseline HST aberration recovery plan includes modifications to the optics of WFPC II such that the original specification for the OTA/camera system can be met. To achieve this, optics teams have convened at JPL and elsewhere in order to measure the OTA aberrations as-built and to define the changes needed to the WFPC II optics in order to compensate for the OTA problem. In the near term, a suite of HST observations are being performed to characterize the OTA, and then the WFPC II optical components can be figured. On the assumption of a simple spherical aberration in the primary of the OTA, the required changes could be effected in the figures of the Cassegrain repeater secondaries (and possibly the fold mirrors) of WFPC II.

When built, WFPC II will be tested with an optical stimulus which will be modified so as to reproduce the actual performance of the OTA.

Other changes and developments on WFPC II relative to WFPC include:

• Better UV response and quantum-efficiency (QE) stability
  The CCDs used in WFPC II are coated with a lumogen phosphor and biased platinum gate. They have shown no signs of the hysteresis or QE decay observed in the current WFPC devices.

- A linear ramp filter with a 1% bandwidth from 4,000 to 10,000 Å
  The bandpass will be selectable by target positioning. This filter will be for use in the Wide Field mode only. Development of a Wood filter (a thin-film alkali metal filter that provides far-UV transmission without the red leaks of the WFPC filters) is continuing at JPL and a subcontractor.

- Reduced contamination of the camera heads
  Mechanical and materials changes to WFPC II have the goal of reducing this contamination to a level at least three orders of magnitude below that observed during the last thermal-vacuum test of WFPC. Contamination control includes increased venting of electronics bays, baffling of CCDs, changes to materials, and the inclusion of CCD boil-off heaters in order to provide sensitivity down to Lyman alpha.

- An internal flat field capability (in UV and visible)
  This ability will be provided by the inclusion of deuterium and quartz lamps within the volume of the current UV light-pipe.

- Improved electronics
  Electronics improvements will eliminate missing codes from the analog to digital converters and eliminate residual images. Extended registers are provided on each chip for bias determination.

STIS

The Space Telescope Imaging Spectrograph (STIS) is designed to take spectra over a wavelength range—1050 to 11,000 Å—that is wider than the combined ranges of the HRS and FOS instruments. STIS is a two-dimensional spectrograph: it images along the entrance slit, which can produce simultaneous spectra for points along a line on the astronomical source. STIS will offer a selection of spectral resolutions from very low to high. Functionally, it can replace both the HRS and the FOS with no loss in spectrographic capability. In addition, it offers:

- the multiplex advantage of 2-dimensional detectors (260X to 2400X in speed, depending on the mode).

- greater wavelength coverage into the near-IR.

- higher quantum efficiency
  ~10X FOS at 7000 Å, even more at longer wavelengths
  ~3X to 6X FOS at 3000 - 6000 Å
  ~2X HRS and FOS at 1050 - 1700 Å.

- no red leak in the UV (unlike FOS, FOC, and WFPC).

- lower effective sky background limit
  ~2X better than FOS, due to simultaneous sky and source measurements
  ~ 4X better than FOS, because pixels are smaller.

- lower scattered light in echelle mode
  ~8X better than HRS, because pixels are smaller and grating is better.

- better optical design in the UV:

STIS has just two reflecting surfaces for the low and very low resolution modes, and it avoids the grazing incidence technology used in FOS. There are two additional reflections if a corrector is installed.

- benefits from HST's high spatial resolution.

- coronographic mode, not available in HRS.

- improved pointing effectiveness due to camera-quality target acquisition mode.

- one focal plane position for all modes—no repointing.

STIS CAN BE USED IN THREE BASIC SPECTROGRAPHIC MODES:

- Echelle mode, like IUE, which uses a 2-D detector format to capture a long spectrum in strips. Unlike the HRS, STIS observes the spectrum and the inter-order background simultaneously, with no loss of on-target observing time. Furthermore, the HRS uses a 1-D detector with only 512 detecting elements, which can record only a section of one spectral order in a single exposure. STIS records a spectral range 130x bigger than HRS at higher resolution (R=140,000) and with about twice the sensitivity.

- Long slit mode, which takes spectra simultaneously (at medium, low, or very low resolution) at each of 1000 positions along the projected position of the slit on the target. In this mode, STIS surpasses the corresponding mode of the FOC in the UV because the FOC has a significant red leak and because STIS is much more sensitive. (The HRS and FOS have no long slit mode.)

- Slitless spectrograph mode, which takes spectra simultaneously of every point-source in the field of view. In this mode, STIS surpasses the WFPC prism mode by providing useful sensitivity in the UV with no red leak. (The HRS and FOS do not operate in a slitless mode.)

Unlike the present HRS and FOS with their small entrance apertures and 1-D detectors, STIS will operate usefully in parallel with other instruments. In the slitless spectrograph mode, it obtains point-source spectra in a 50 x 50 arcsec field of view. The long-slit spectrograph mode can take the spectra of extended sources falling into the field of view.

The STIS camera mode, designed primarily for target acquisition, is capable of imaging sources in a 50 x 50 arcsec field to a 29th magnitude limit in 1 hour at visible wavelengths, and to 31st magnitude in the near-IR.

A photon time-tagging capability in the STIS UV detector system can be used to recover spatial resolution that might be lost due to spacecraft jitter. HRS, FOS, and WFPC lack this capability.

NICMOS

The Near-Infrared Camera and Multi-Object Spectrometer (NICMOS) will extend HST's capabilities into the infrared. NICMOS has short wavelength coverage down

to 1 μm for overlap with the other HST SIs, but NICMOS also provides diffraction-limited imaging out to 2.5 μm, and spectroscopy out to 3 μm.

Broad-band imaging between 1 and 2.5 μm is background limited for both ground-based telescopes and HST. However, since the background for ground-based telescopes is due to airglow emission while the HST sees primarily scattered zodiacal light, the HST background is hundreds to thousands of times fainter than for the ground-based case. (Note also that the airglow can be highly variable with time.) The lowest background on HST will occur near 1.7 μm (the background here is 10,000 times fainter than from the ground) providing NICMOS on HST with the capability to perform extremely deep surveys for protogalaxies. (Lyman-alpha emission for a galaxy with z = 13 occurs at 1.7 μm.) For wavelengths longward of 2.5 μm, thermal emission from the warm (290 K) HST optics begins to erode the advantage of HST over ground-based telescopes for broad-band imaging applications.

NICMOS has three independent cameras, all covering the same 1 to 2.5 μm window but with different magnifications. Diffraction-limited imaging is available over the entire 1-2.5 μm wavelength region with one of the cameras. Each camera views a separate field in the HST focal plane, but a beam-steering mirror can be used to divert a fixed point in the HST focal plane (called the maximum image quality, or MIQ, position) to any of the cameras. The range of motion available for the beam-steering mirror is approximately 2 arcmin. Each camera also has its own 20 position filter wheel. Thus, a diverse range of scientific programs can be addressed by the appropriate camera/filter combination.

NICMOS also has three independent spectrometers. The Multi-Object Spectrometer (MOS) covers the wavelength range between 1 and 2 μm. Beam-slicing optics map a 16 arcsec by 8 arcsec rectangular region of the HST focal plane onto the long-slit spectrograph. Multiple gratings mounted on a rotatable carousel allow for a variety of spectral resolutions, up to a resolving power of ~60 km/sec.

Two spectrometers cover the 2-3 μm wavelength range: Long Wavelength Spectrometer 1 (LWS 1) covers from 2 to 2.5 μm and Long Wavelength Spectrometer 2 (LWS 2) covers from 2.5 to 3 μm. Both are cooled, cross-dispersed echelle grating spectrometers having 0.2 x 3 arcsec slits and no moving parts. The spectral resolution for each is ~100 km/sec. LWS 1 covers the astrophysically important CO band near 2.3 μm. Although the HST thermal background limits its broadband imaging sensitivity longward of 2.5 μm, the LWS 2 provides extremely sensitive spectroscopy out to its long wavelength cut-off at 3 μm. The region between 2.5 and 3 μm is especially important because severe atmospheric absorption makes ground-based observations virtually impossible. Even airborne observations are significantly compromised in the 2.5 to 3 μm region.

All six NICMOS "functions" (i.e., the three cameras and three spectrometers) use identical 256 x 256 Hg-Cd-Te array detectors (one array per function). (The LWS 2 detector is doped slightly differently in order to extend its wavelength coverage to 3 μm.) Each function can be used in stand-alone mode, or all can be operated simultaneously. NICMOS has its own microprocessors (dual 80386 microprocessors) which can be programmed for complex operations. However, NICMOS operations may be limited by the commanding capability of the HST ground system.

All six NICMOS functions share common "stage 1" optics consisting of three elements: a folding flat, a re-imaging mirror, and the beam-steering mirror. The re-imaging mirror produces an image of the HST pupil on the beam-steering mirror. In the original plan, the re-imaging mirror was spherical and the beam-steering mirror was flat. However, by changing the figure on these two elements, the OTA

GLOSSARY

Å	Ångstrom unit, 10^{-10} m
AO	Announcement of Opportunity
arcmin	minute of arc, or 1/60 of a degree
arcsec	second of arc, 1/3600 of a degree
BTU	British Thermal Unit
Caltech	California Institute of Technology
CCD	charge coupled device; a solid-state, light detecting array
CLEAN	a computer program for deconvolving radio telescope images
cm	centimeter, 10^{-2} m
COSTAR	Corrective Optics Space Telescope Axial Replacement
CV	cataclysmic variable star
deg	degree of arc
ECF	ST European Coordinating Facility
ELV	expendable launch vehicle
ESA	European Space Agency
ESTR	Engineering and Science Tape Recorder
EVA	extra-vehicular activity
FDn	Flight Day n
FGS	Fine Guidance Sensor
FHST	Fixed Head Star Tracker
FOC	Faint Object Camera
FOS	Faint Object Spectrograph
FSS	Flight Servicing System
ft	foot
GO	General Observer
GSC	Guide Star Catalog
GSFC	Goddard Space Flight Center
GTO	Guaranteed Time Observer
HDOS	Hughes Danbury Optical Systems
HRA	High Resolution Apodizer
HRS	High Resolution Spectrograph
HSP	High Speed Photometer
HST	Hubble Space Telescope
Hz	Hertz, cycle per second
IDT	Investigation Definition Team
IR	infrared
IRAS	Infra Red Astronomy Satellite
IUE	International Ultraviolet Explorer
IV	intra vehicular
JHU	The Johns Hopkins University
JPL	Jet Propulsion Laboratory
JSC	Johnson Space Center
K	degree Kelvin
km	kilometer

LAS	Laboratoire Astronomie Spatiale
lbs	pounds
LMSC	Lockheed Missiles and Space Corporation
LWS	Long Wavelength Spectrometer
M&R	Maintenance & Refurbishment, the NASA program to service HST in orbit
m	meter
magnitude	unit of star flux; larger values mean fainter stars
mas	milliarcsecond, 10^{-3} arcsec
MFR	manipulator foot restraint
MLI	multi-layer insulation
mm	millimeter, 10^{-3} m
mm	micrometer or micron, 10^{-6} m
MMU	Manned Maneuvering Unit
MIQ	maximum image quality
MOS	Multi-Object Spectrometer
MSFC	Marshall Space Flight Center
NASA	National Aeronautics and Space Administration
NICMOS	Near-Infrared Camera and Multi-Object Spectrometer
nm	nanometer, 10^{-9} m
NTT	New Technology Telescope
OCS	Optical Control System
ORI	Orbit Replaceable Instrument
ORU	Orbit Replaceable Unit
OS	Observatory Scientist
OTA	Optical Telescope Assembly
PC	Planetary Camera
PCS	Pointing Control System
PFR	portable foot restraint
PM	primary mirror
PMT	photomultiplier tube, a single-channel, vacuum-tube light detector
PCS	Pointing Control System
PSF	point spread function
PWLA	Portable Work Light Assembly
QE	quantum efficiency
QSO	quasi-stellar object, or quasar
R	resolving power
rad	radian (p rad = 180°)
rms	root mean square
RMS	Remote Manipulator System
RSU	Rate Sensor Unit
RTV	room temperature vulcanizing
SAs	solar arrays, the HST power source
SAC	Solar Array Carrier
SI	Science Instrument
SIPE	Science Instrument Protective Enclosure
SM	secondary mirror

SSM	System Support Module
STAC	Space Telescope Advisory Council
STAR	Space Telescope Axial Replacement
STIS	Space Telescope Imaging Spectrograph
STS	Space Transportation System
ST ScI	Space Telescope Science Institute
TAC	Time Allocation Committee
TBD	to be determined
UA	University of Arizona
ULE	ultra-low expansion (high thermal stability)
UV	ultra violet
WFC	Wide Field Camera
WFPC	Wide Field and Planetary Camera
WFS	Wave-Front Sensor
z	redshift