THE ASTROPHYSICAL JOURNAL, 392:760–765, 1992 June 20

OPTIMAL GRAZING INCIDENCE OPTICS AND ITS APPLICATION TO WIDE-FIELD X-RAY IMAGING

CHRISTOPHER J. BURROWS,[1] RICHARD BURG,[2] AND RICCARDO GIACCONI
Space Telescope Science Institute,[3] 3700 San Martin Drive, Baltimore, MD 21218
Received 1990 April 16; accepted 1991 November 4

ABSTRACT

We discuss a class of high-resolution, efficient, and wide-field grazing incidence optics. Although such designs have been discussed in the literature before, they have rarely been considered for a practical application. We have developed optical designs to search efficiently for distant X-ray clusters. A similar approach is applicable to other planned X-ray missions, such as *AXAF*. Previously flown missions such as *ROSAT* would have benefited from these ideas. Indeed, we will show that a rather general procedure exists for the design of an optimal mission if one can provide well-defined scientific goals.

In particular, we wish to resolve clusters of galaxies with an angular diameter of $\sim 5''$ over a field of $\sim 1°$, so that a comprehensive deep search is possible with a payload of 1/10 the linear dimensions of *AXAF*. No design of a standard type is capable of meeting this requirement. Wolter-Schwarzschild and parabola-hyperbola designs yield a useful field that is many times smaller than needed because they suffer from unacceptably large off-axis aberrations, even though they yield perfect images on-axis. By dropping the requirement for perfect on-axis imagery, searching within a suitably general class of telescope design, and optimizing a quantity directly related to the scientific requirement, we have been able to show that satisfactory designs do exist. The resulting telescope is shown to be no more difficult to fabricate than existing mirrors and can be nested.

Subject headings: telescopes — X-rays: general

1. INTRODUCTION

The design of most of the upcoming X-ray astronomy missions of the 1990s is based on the use of grazing incidence X-ray telescopes. It is the use of these telescopes which has made possible the improvements in sensitivity and angular resolution in the last two decades. These improvements have permitted the extension of X-ray observations to all known classes of celestial objects.

Much of the technological development effort in this same period has been devoted to the fabrication of mirrors with higher reflection efficiency, higher on-axis resolution, and larger collecting areas.

The achievement of high-reflection efficiency requires very smooth surfaces (5–10 Å) and the deposition of heavy metal coatings. This is particularly important if one wishes to extend X-ray observations to higher energies than have been feasible in the *Einstein*, *EXOSAT*, and *ROSAT* missions. Fabrication tolerances and mirror coatings for the *AXAF* telescope, for instance are designed to ensure high efficiency up to 10 keV.

High angular resolution can be achieved, on-axis, by use of Wolter or Wolter-Schwarzschild optics configurations, provided fabrication errors on large scales can be kept within very tight limits. Current estimates yield an on-axis resolution of $0''.5$ for *AXAF*.

Finally, the drive to obtain large collecting areas within fixed spacecraft constraints has led in many missions to the adoption of optics designs which are only rough linear approximations to conics and in which collecting area is obtained at the expense of angular resolution. Such missions include *ASTRO-D*, *BBXRT*, and *Spectrum X*.

During the last decade, little attention has been given to the development of better optics designs, partly because it was felt that optimum designs were already on hand and partly because further refinements were not useful given the fabrication errors. An exception is reported by Nariai (1987, 1988) who describes optimized designs to be used in the *Solar-A* mission.

We have encountered a specific research program of X-ray observations which require optics of characteristics different from those adopted in all the current missions. The research program is a survey for distant clusters of galaxies ($Z > 1$) by means of an *Explorer* satellite suitable for launch on a Scout-sized rocket. The requirements for such a mission are described in Giacconi (1989) and can be summarized as follows: We need an X-ray optics design which can achieve high angular resolution (better than $2''.5$ radius rms) over a large field of view (1° diameter).

We were successful in the search for such a design and independently rediscovered the polynomial optics solutions which were first considered by Werner (1977). His work has largely been ignored up to now, in part due to an early misinterpretation of his results, and in part because in his discussion of this work the applicability of the approach to a number of significant problems in astrophysics was not made sufficiently clear. Through our own work, we have become convinced of the following points.

1. Designs exist for relatively wide-field, high-resolution grazing incidence optics which are most suited for survey work (our original research problem).

2. The polynomial optics are as easy to fabricate as the paraboloid hyperboloid (Wolter type I) optics and easier than the Wolter-Schwarzschild optics.

3. Polynomial optics can be much superior in performance to Wolter or Wolter-Schwarzschild optics, once fabrication

[1] Astrophysics Division, Space Science Department of ESA.
[2] Department of Physics and Astronomy, The Johns Hopkins University.
[3] Operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Aeronautics and Space Administration.

errors are taken into account and once the solutions are optimized for the specific desired performance characteristics.

4. The polynomial optics approach must be taken into account in any trade-off study of telescope designs which seek to maximize scientific mission returns.

We first describe our general approach, and then illustrate it with applications to the required wide-field *Explorer* mission and to *AXAF*.

2. OPTICAL DESIGN

Our approach was to select a general class of X-ray telescopes, and then optimize the associated parameters by the use of a computer search. The general requirement of a reasonably fast, efficient, wide-field telescope leads us to consider two mirror axisymmetric grazing incidence designs with both reflections on the inner surface, for the following reasons.

1. The number of surfaces must be even for grazing incidence optics to yield acceptable off-axis performance. We have verified directly that one- and three-mirror conic designs cannot work. Odd numbers of surfaces violate the sine condition (Wolter 1952a, b).

2. Complex designs with four or more mirrors are probably feasible with modern fabrication and alignment techniques and highly efficient coatings; however, they clearly present a considerably more demanding and costly alternative if a smaller number of mirrors can meet the requirements.

3. Two-mirror telescopes generally consist of two coaxial surfaces of revolution. This is not necessary, but they are likely to be the most efficient for a given area of optics, and there is no known reason to break the symmetry.

4. The lowest overall F/number system for a given grazing angle occurs when both reflections are from the inside surfaces. We must keep the grazing angle small because high grazing angles yield low surface reflectivity at high energies, yet we want the system to be compact for packaging reasons. In addition, such telescopes are easier to build and align and are used almost exclusively in existing and planned missions. A notable exception is *FUSE* which can have the second reflection on the outside surface because it operates at much larger incidence angles.

5. The two-mirror surfaces intersect at a circle which we denote P^*. It is not necessary that they be used all the way up to P^*. If they are not, the filling efficiency on the available aperture will drop. We restrict ourselves to the case where the surfaces are not masked. Werner (1977) treats the more general case.

We need to define certain first-order properties of the telescope, so that the discussion of the conventional and optimized designs can proceed. Any telescope in the above class is broadly characterized by the first-order parameters (ρ_0, F, α, ϵ, $\theta_{\max}$), where

1. The radius of P^* (inner radius of the primary and outer radius of the secondary mirror) is ρ_0. All the other parameters can be dimensionless variables, so any measure of image quality when expressed in angular terms is independent of ρ_0.

2. The parameter F is defined in terms of the distance, z_0, from P^* to the point where rays from an on-axis object cross the axis, by $F = z_0/2\rho$.

3. The angle, α, between the primary mirror tangent at P^* and the optical axis. The secondary mirror tangent is inclined at $\beta = \alpha + \cot^{-1}(2F)/2$.

4. The outer radius of the primary, $\rho_{\max} \equiv (1 + \epsilon)\rho_0$. The length of primary is $Z_1 \approx \epsilon\rho_0/\alpha$, in the reasonable geometrical approximation of replacing the surface by a cone (the approximation is not adequate for optical properties). Given an aperture filling factor derived from efficiency requirements, an acceptable maximum number of nested mirrors, and the thickness of the individual mirrors, the minimum ϵ is fixed purely by geometrical considerations.

5. Maximum field angle $\theta_{\max}$.

Several of these parameters are shown in Figure 1.

An important derived quantity is the ratio of the grazing angles on the two mirrors $\xi = \alpha/(\beta - 2\alpha)$. For fixed $F^{-1} = 2\tan 2(\beta - \alpha)$, the total reflection efficiency is maximized when $\xi = 1$, provided that the scattering efficiency $t(\alpha)$ is a smooth function of α, and that its derivatives satisfy $t''t - t'^2 > 0$. The inequality is satisfied in practice for most materials and grazing angles.

The primary and secondary surfaces can be expanded as power series of the form

$$\rho_1^2/\rho_0^2 = \sum_{i=0}^{n_1} a_i(z_1/\rho_0)^i ,$$

$$\rho_2^2/\rho_0^2 = \sum_{i=0}^{n_2} b_i(z_2/\rho_0)^i , \qquad (1)$$

where (ρ_1, z_1) and (ρ_2, z_2) are radial and axial coordinates on surfaces 1 and 2, respectively. By definition $a_0 = 1$ and $b_0 = 1$. The parameters $a_1 = -2\tan\alpha$ and $b_1 = -2\tan\beta$ are twice the slope of each surface at P^*.

If a_2 and b_2 are zero, the surfaces are parabolic. Nonzero values give other conic sections. Higher order terms, if present, represent deformations of the conics. Our parameterization

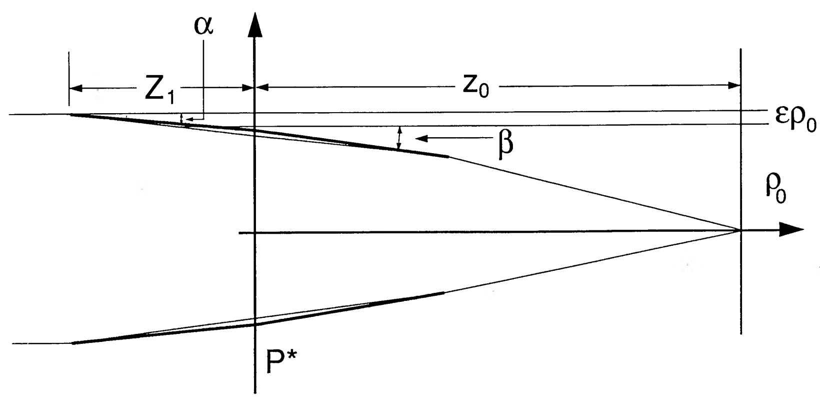

FIG. 1.—First-order parameters used to define the telescope configuration

above differs from that of Werner (1977). He expands ρ as a power series in z, while we expand ρ^2. We feel that the latter choice is preferable because then the series for conic sections contains only three terms. These two descriptions are interchangeable provided that the series are continued to sufficiently high order.

In order to fix the coefficients, and hence completely specify the design, at least three approaches have been used.

1. The simplest design is to make both mirrors cones. This has the benefit that large collecting areas can be constructed at low cost and weight. In our notation this choice corresponds to $a_2 = \tan^2 \alpha$ and $b_2 = \tan^2 \beta$, and all higher order terms zero. Such designs suffer from very large aberrations, even on-axis. The on-axis spot diameter is $2\epsilon \sin 2(\beta - \alpha)/\cos^2 2(\beta - \alpha)$. We do not consider them further explicitly here. They are a member of the class of designs over which we optimized.

2. The paraboloid-hyperboloid class (Wolter 1956; Van Speybroeck & Chase 1972) consists of a paraboloid followed by a confocal hyperboloid. Then $b_2 = 2\rho_0 \tan \beta/(z_0 + \rho_0 \cot 2\alpha)$, and a_2 and all higher order terms are zero. For fixed F and ξ, there are no free parameters to optimize the design. These two parameters are generally tightly constrained by efficiency and packaging considerations. These telescopes have perfect on-axis imagery and off-axis performance that degrades rapidly, primarily due to coma (which grows linearly) for small F and spherical aberration (which grows quadratically) at larger F. The place where the turn-off between coma and spherical-dominated imagery occurs depends on $\theta_{\max}$ and ϵ. When coma-dominated, better results are obtained by increasing the focal length, and when spherical-dominated, by decreasing it. If $\theta_{\max}$ is increased, the optimum F increases. *AXAF* and XMM images will be dominated by spherical aberration off-axis. On the other hand, wide-field telescope designs tend to be in the coma-dominated region of the $F - \theta$ plane, unless they have extremely long focal lengths.

3. The Wolter-Schwarzschild design discussed in Chase & Van Speybroeck (1973) exactly satisfies the sine condition, with zero (linear) coma, and perfect imaging on-axis. Given ξ and F, all the coefficients are determined if the surfaces are expanded as power series, and decrease rapidly in magnitude with increasing index i. These designs are better than the paraboloid-hyperboloid design in the region where the latter are coma-dominated, and worse in the spherical limit.

Instead of applying these criteria, we advocate fixing the coefficients by defining a merit function that is directly related to the scientific goal of the investigation. The parameter space defined by the a_i and b_i can then be searched by computer in order to give an optimum solution.

The best merit function of all might be to optimize the probability of detecting (or resolving?) with given observing time a source with background, and energy spectrum drawn from the appropriate probability distributions. One problem in principle with this approach is that we do not know enough about the X-ray sky to characterize these distributions.

The merit function that we have used for this study is to minimize the root mean square spot size over the whole field. That is,

$$M = \sqrt{\left(\int \theta\sigma^2 \, d\theta\right) \Big/ \left(\int \theta \, d\theta\right)} \, . \tag{2}$$

This has the benefit of being simple to understand, and parameter-free. It provides a reasonable approximation to the ideal cited above. Alternative definitions that balance contrast and resolution differently, for example, the average half (or 70%) energy width, are of course possible.

Expected manufacturing and alignment tolerances can be included at this point, for example, by adding an appropriate constant to σ^2 in the above equation. However, the modified merit function will have a maximum at the same parameter values, so the resulting design would be unchanged. Thus, if the manufacturing and alignment errors can be considered to contribute in quadrature to the unperturbed point image variance, the optimum design is unaffected.

If the requirement involves a trade-off (for example, there are two mission goals such as a wide field with good imaging and also a need for optimum imaging on-axis), these can both be included in the merit function with suitable weights. Alternatively, one weight can be varied in the optimization as a Lagrange multiplier until a constraint is satisfied. Such techniques are standard in normal incidence optical design but have not been used routinely in the grazing incidence case.

Our merit function does not explicitly include any measure of efficiency. However, maximum efficiency for a given focal length occurs when $\xi = 1$. We optimize the design at this fixed efficiency, thus fixing α and β. This means that the only free parameters in the optimization are then $a_2, a_3 \ldots$ and $b_2, b_3 \ldots$. We have therefore implicitly included the appropriate Lagrange multipliers to constrain the efficiency. An alternative is to explicitly trade resolution against average efficiency by including appropriately weighted terms in the merit function. In practice, we have found that the optimum resolution designs occur near $\xi = 1$ for the cases we have examined (both in the coma and spherical-dominated regions).

We have written a computer subroutine to evaluate the merit function for any particular design. It operates by explicitly tracing a large number of rays through the system and computing the resulting distribution in the image plane. For most investigations, the image plane was defined as the (curved) surface in which the merit function is a minimum. For our choice of merit function, the required defocus at any field angle is analytically computed from the outgoing ray bundle. The subroutine is used in a simplex algorithm to compute a minimum of the merit function, with specific parameters allowed to vary. It has been found that by optimizing low-order coefficients first, and then using them as initial values when more coefficients are allowed to vary, one can both speed up convergence and avoid convergence to secondary minima. The individual shells in a nested design are probably best optimized separately at least initially, because otherwise the number of parameters becomes large, and this adversely affects the convergence rate. In practice, they can be independently optimized and then combined in a satisfactory manner, provided that an adjustment for focal plane scale is made as described in the next section.

3. RESULTS

In addition to the general investigations described above, we have applied the optimization technique to two specific example missions. These are the proposed Explorer class *Wide Field X-ray telescope* (*WFXT*), which originally motivated our study, and NASA's Great Observatory the *Advanced X-ray Astrophysics Facility*, *AXAF* (Weisskopf 1987). The first-order parameters for the two examples are given in Table 1. These parameters were kept constant for the optimization leading to the results presented in the bottom of the table. The area-

TABLE 1

FIRST-ORDER PARAMETERS AND RESULTS FOR *AXAF* AND *WFXT*

		AXAF		*WFXT*	
PARAMETER	SYMBOL	Outer	Inner	Outer	Inner
Focal distance	z_0	10 m	10 m	3.0	3.0 m
Mirror common radius	ρ_0	0.6 m	0.3 m	0.30 m	0.2 m
Maximum field angle	θ_{max}	7′.0	7′.0	0°.5	0°.5
Initial grazing angle	α	0°.86	0°.57	1°.43	0°.95
Grazing angle ratio	ξ	1.0	1.0	1.0	1.0
Used aperture	ϵ	0.021	0.021	0.01	0.01
Length of primary	Z_1	0.84 m	0.84 m	0.12 m	0.12 m
Paraboloid-hyperboloid	M	0″.93	1″.91	5″.62	9″.03
Surface 2	$b_2 \times 10^4$	17.9798	4.49874	49.8445	22.1914
Best conic design	M	0″.50	0″.94	3″.59	4″.37
Surface 1	$a_2 \times 10^4$	0.0729	0.020	1.137	0.405
Surface 2	$b_2 \times 10^4$	17.9624	4.5343	49.7477	22.781
Optimized design	M	0″.46	0″.90	2″.62	3″.78
Surface 1	$a_2 \times 10^4$	0.253	0.0712	3.597	0.6346
	$a_3 \times 10^4$	0.0539	0.00772	2.63	−0.02424
Surface 2	$b_2 \times 10^4$	17.773	4.4789	47.392	22.120
	$b_3 \times 10^4$	0.0627	0.0085	2.162	0.316

NOTE.—For each of three possible designs for the inner and outer shells, we give the merit function M and the corresponding nonzero coefficients (a_i, b_i) (uncertain in the last digit).

weighted rms spot size M defined in equation (2) was used as a merit function for the optimizations and is given in the table. The parameters apply to a single shell of a nested design, and results for that shell in isolation are given. For each design, results for both the inner and outer shells are given. Figure 2 shows the cumulative field of view for the *WFXT* design, compared to the existing *AXAF* and *ROSAT* designs. Figure 3 shows that the best image quality is obtained off-axis in the optimized designs, and that it is comparable to the on-axis image quality of the paraboloid-hyperboloid designs when mirror surface errors are included in the analysis. Figure 4 is the spot diagrams for this design.

We wish to emphasize the factor of at least two improvement in average resolution, obtained with the optimized design compared to the equivalent traditional paraboloid-hyperboloid designs. In particular, the background-limited detection of point sources can be done at a given signal-to-noise level 4 times faster, and the number of sources in the field that can be resolved will be increased. This promises to be increasingly important for future deep X-ray surveys.

Werner (1977) has extensively analyzed the aberrations. As a result, it is possible to better understand what is happening. All the designs suffer from a component of high-order spherical aberration that grows quadratically as one moves off-axis. This component is not significantly reduced by the optimized designs. However, the optimim designs do introduce a large

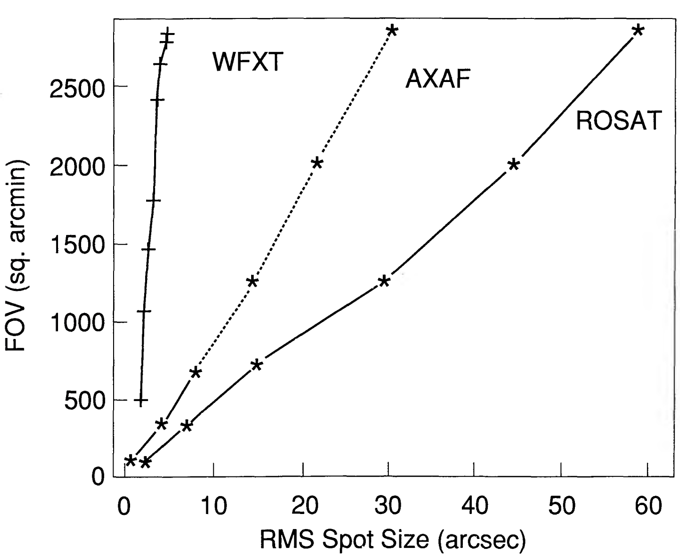

FIG. 2.—The cumulative field of view for *ROSAT* (for a flat Gaussian focal plane), *AXAF* (before optimization), and the *Wide Field X-Ray Telescope*. The total available field area is plotted as a function of the average image quality. The dotted line for *AXAF* indicates field angles which presently planned detectors will not cover.

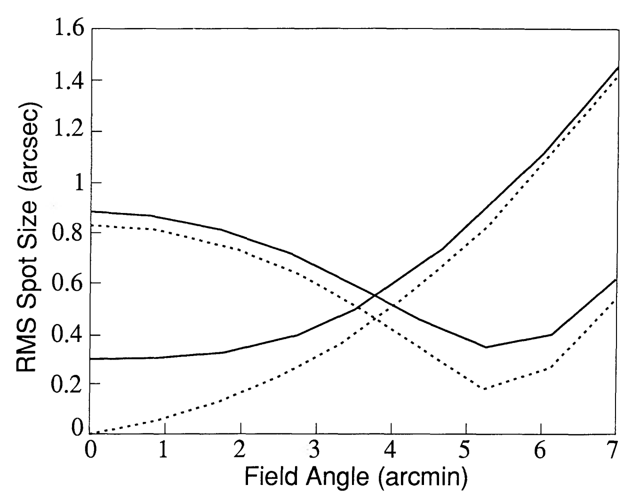

FIG. 3.—The rms image spot size for the outer shell of *AXAF*, in the existing and optimized designs. The solid curves have an assumed image blur of 0″.5 caused by surface roughness added in quadrature to the dotted curves which represent the raytrace results. The optimized design has best image quality near 5″. It remains sub-arcseond over the full field. The existing paraboloid-hyperboloid design is best on-axis, but when averaged over the field is about a factor of 2 worse than the optimized design.

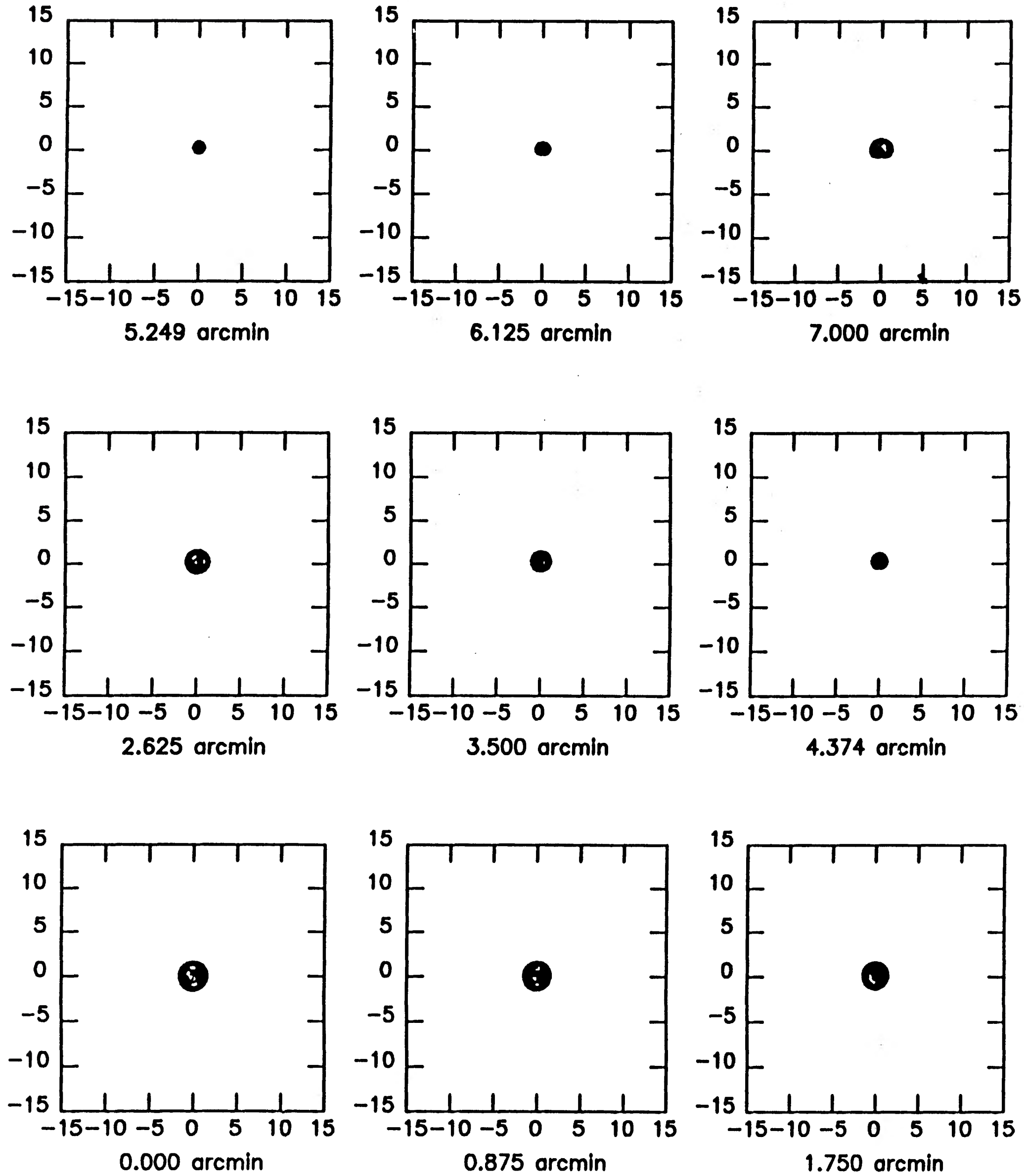

FIG. 4.—Spot diagrams for the optimized *AXAF* design. In the existing design a larger corrected area is obtained by placing the best image about 5′ off-axis. The units for each figure are arcseconds projected back on the sky.

and *constant* combination of the same high orders with opposite sign. The result is that the aberrations of the system are almost exactly cancelled at $\theta_{max}/2^{1/2}$ This simplified model then predicts that the overall merit function M will be reduced by exactly a factor of 2. We can see, as one might expect, that this is true enough in the spherical-dominated case of *AXAF*. The gains for the *WFXT*, where coma is also significant, are larger.

Nariai (1987, 1988) considers hyperboloid-hyperboloid designs and derives analytical designs with excellent off-axis performance. He introduces on-axis spherical aberration as suggested by Werner and finds that the best performance occurs when the coma at the edge of the mirror is cancelled. However, his choices are largely arbitrary, as he does not define an explicit merit function, and he restricts himself to conic sections.

The designs are nestable in the same way as the conventional designs. There is however an important consideration that must be included in a detailed design: the images from different shells must be superposed, but each shell has differing distortion and plate scale in the focal plane. Taking into account the contribution of distortion, which decreases these corrections at larger field angles, a detailed numerical calculation then shows that the images are blurred by 1″1 at 7′0. This is not negligible for the existing design and more serious for the optimized approach because the best resolution is obtained at larger field angles. Fortunately there is a simple remedy. The different shells need to be moved relative to one another axially so that the average plate scale remains the same. In principle, the shifts should be determined so that the merit function is optimized, but a first-order correction is sufficient to adequately suppress the effect as noted by Nariai (1988). It corresponds to an offset of about 2 cm between the P^* planes for the inner and outer shells of *AXAF*.

4. CONCLUSIONS

A detailed study of the applicability of currently planned missions to cluster research and large-scale structure has led us to believe that the presently planned instruments have limitations that are not related to collecting area but are caused by limited field of view (Burg, Burrows, & Giacconi 1990). In a survey mission, aperture size and angular field of view are equivalent in their contribution to sensitivity and categorization of sources, but mission cost goes up rapidly with collecting area. At high redshift, the only unambiguous identification of clusters is as extended objects in the X-ray; the optical images of the galaxies are too faint. Studies of clusters of galaxies (and large-scale structure) at redshifts greater than unity are presently frustrated by the lack of high-redshift samples. *AXAF* is an observatory uniquely capable of detailed studies of such phenomena, but in the absence of a high-redshift cluster catalog, it will not be able to solve definitively the problem of cluster formation and evolution. We have therefore studied X-ray observatories that could successfully carry out such surveys. We have designed optics that achieve high spatial resolution (better than 2''.5) over large fields of view (of at least 1°). We intend to design in detail a Scout-class *Explorer*, utilizing these optical designs, to carry out a program to study clusters and large-scale structures at high redshifts.

This work was sponsored in part by NASA contract NAS5-26555 and NASA grant NAG8-530.

REFERENCES

Burg, R., Burrows, C. J., & Giacconi, R. 1990, in Proc. Taos Workshop on X-Ray Astronomy into the 21st Century, ed. P. Joss (NY: AIP), 318
Chase, R. C., & Van Speybroeck, L. P. 1973, Appl. Opt. 12, 1042
Giacconi, R. 1989, in Imaging X-Ray Astronomy, ed. M. Elvis (Cambridge: Cambridge Univ. Press), 208
Nariai, K. 1987, Appl. Opt., 26, 4428
Nariai, K. 1988, Appl. Opt., 27, 345
Van Speybroeck, L. P., & Chase, R. C. 1972, Appl. Opt., 11, 440
Weisskopf, M. C. 1987, Astro. Phys. Lett. & Commu., 26, 1
Werner, W. 1977, Appl. Opt., 16, 764
Wolter, H. 1952a, Ann. Phys., 10, 94
———. 1952b, Ann. Phys., 10, 286