An application of the SMS method for imaging designs

Juan C. Miñano 1, Pablo Benítez 1, Wang Lin1, José Infante1, Fernando Muñoz2,
Asunción Santamaría1

1Universidad Politécnica de Madrid, Cedint, Campus Montegancedo, 28223 Madrid, Spain;
2LPI, 2400 Lincoln Avenue, Altadena, CA 91001, USA
jc.minano@upm.es

Abstract: The Simultaneous Multiple Surface (SMS) method in planar geometry (2D) is applied to imaging designs, generating lenses that compare well with aplanatic designs. When the merit function utilizes image quality over the entire field (not just paraxial), the SMS strategy is superior. In fact, the traditional aplanatic approach is actually a particular case of the SMS strategy.

©2009 Optical Society of America

OCIS codes: (080.2740) Geometric optical design; (080.3620) Lens system design; (080.4035) Mirror system design

References and Links
1. Introduction

Classical Imaging design [1–3] is based on maximizing a merit function that measures image quality over a prescribed field of view. This maximization is obtained by parametrically varying the optical surfaces.

When spherical sections were the only surface shape used, the merit function was based on analytical aberration calculations [4]. Later on, multi-parametric optimization techniques were used to maximize much more complex merit functions [5]. Today, multi-parametric optimization is a common tool of any optical design software. The algorithms progress from an initial guess to the final solution. Since the search is local there is no guarantee that the algorithm will find the absolute maximum.

The use of aspheric surfaces has not changed the basic optimization methodology. An aspheric surface is conventionally described by the following function [2]:

$$z = \frac{c \rho^2}{1 + \sqrt{1 - c^2 \rho^2}} + a_4 \rho^4 + a_6 \rho^6 + a_8 \rho^8 + \ldots$$

(1)

where $\rho$ is the distance from a point to the optical axis ($z$ axis), $c$ is the vertex radius of curvature. The remaining parameters ($k, a_4, a_6, \ldots$) describe the “asphericity” of the surface. More powerful ways to describe aspheric surfaces were recently proposed by Forbes [6].

The optimization strategy with aspherics usually starts with a spherical design. The final aspheric surface is not “far” from this initial spherical design, mainly because there are many local maxima where the optimization process finds a solution. Sometimes the optimization process adds new parameters to optimize as the process progresses: first starting with $k$ (with the initial value $k = 0$ for the spherical surface) until a local maximum is found and then following with $a_4$, etc. In practice, the performance of the design does not improve when the number of optimizing parameters is big (>10) because the asphericity of one surface is just cancelling the asphericity of another [3].

The Simultaneous Multiple Surfaces (SMS) method sets up the problem in a different way. Here only 2D designs are covered, the generation of which only uses rays in a plane, hence the name SMS 2D. The ray tracing analysis is done with rotational symmetric systems obtained by rotation of these curved profiles. The SMS method involves the simultaneous calculation of $N$ optical surfaces (refractive or reflective) using $N$ one-parameter bundles of rays for which specific conditions are imposed. When designing an imaging optical system, these conditions comprise the perfect imaging of every one of these $N$ bundles. For example, the $N$ bundles can be rays issuing from $N$ points of the object. The $N$ conditions are that these bundles are perfectly focused at the corresponding $N$ image points. The generative power of the SMS method is in the flexibility of selecting the $N$ points of perfect imaging. We can assess how good is the image formation of each point $P$ by means of the RMS blur radius $\sigma(P)$ [3] of the spot image of each ray bundle. We will use two different versions of the blur radius: $\sigma_{2D}(P)$, for which only tangential rays are considered, and $\sigma(P)$, which considers sagittal rays as well. The function of $\sigma_{2D}(P)$ is used to check the attainment of the design conditions ($\sigma_{2D}$ must be zero for the $N$ selected points).

Assume that $N = 2$, (two surfaces to be designed and two tangential ray bundles that must be perfectly imaged by them) and that these bundles are formed by rays issuing from two symmetric points of the object. In ref [7] some of us proved that a two-surface rotationally symmetric SMS design (i.e., obtained by rotating a SMS 2D design) is equivalent to an aplanatic design when the two off-axis design bundles of the SMS design converge to on-axis one. Thus, aplanatic designs can be viewed as particular cases of two-surface SMS 2D
designs. This fact has been used ever since (most recently in ref 8). Remarkably, it has been recently claimed [9] novelty of a classical SMS design, the RXI [10], in its aplanatic version.

Aplanatic designs have long been known through the work of Schwarzschild [11], Wassermann and Wolf [12], Welford [13], Mertz [14,15] and others. Recently Lynden-Bell and Willstrop derived an analytic expression of more general aplanats [16,17]. Aplanatic systems are axisymmetric optical designs free from on-axis spherical aberration and linear coma. These two conditions entail two further design conditions: (1) stigmatic on-axis image (note that all the rays involved in this condition are exclusively tangential) and (2) the Abbe sine condition. The contrary statement is also true: a design forming a stigmatic on-axis image and fulfilling the Abbe sine condition is aplanatic. The Abbe sine condition can be derived from the condition of zero linear coma, but this is just one possible point of view. For instance, Clausius, who derived the Abbe condition in 1864, did it from thermodynamic arguments [13] (we can also say that he used the conservation of etendue theorem [18]). In 1884, Hockin gave a proof using path differences along tangential rays (without considering sagittal rays) [13]. Because the Abbe sine condition can be derived by imposing the zero coma condition just on tangential rays, we can say that the aplanatic are a set of 2D design conditions since they can be imposed and checked using just tangential rays. Additionally, aplanatic systems turn out to also be free of linear coma for near-axis sagittal rays, but this fact does not need to be used in the design procedure, it is just a fortunate consequence in spite of the design procedure being purely 2D. This requirement, imposed solely on cones of light corresponding to the on-axis points, affects the quality of imaging of nearby off-axis points [19]). This is why a 2D design procedure as the SMS 2D converges to aplanatic designs (for some particular cases) although it is commonly said that aplanatic imaging requires the third dimension. That is true, but this does not mean that sagittal rays must be used to design an aplanat.

When the two points of the object selected for a two-surface SMS design are not infinitesimally separated on the object and around the axial direction, then the images of these two off-axis points are no longer stigmatic (they are only stigmatic for the tangential rays) and they are no longer free of linear coma. Thus this SMS design is no longer aplanatic. Nevertheless, this SMS design can perform better than the aplanatic design. This happens when the image quality in the whole field of view (and not only in the axial regions) is considered. In this case there is always a two-surface SMS design with image quality better than that of an aplanatic system, as it was first proved in Ref [7]. This does not mean that the image quality is better for all the points of the field, but that once we fix a minimum imaging quality for all points of the field, the SMS design provides wider fields than the aplanat solution.

![Fig. 1. Two-surface SMS design for bundles incoming at ± 2°. Focal length 14.32 mm.(f/1.576) refractive index = 1.5](image)

In the same way as the two-surface SMS design converges to an aplanatic design when the two off-axis design bundles of the SMS design converge to the on-axis one, the four-surface SMS design converges to a second-order aplanatic designs (following Schulz’s definition of aplanatic order [20,21]).

Presented herein are several SMS configurations with two and four refractive surfaces, designed for perfect tangential ray image formation at two and four points respectively. The
merit function for comparison is the field-of-view diameter for a given maximum spot radius \( \sigma_M \). This merit function makes sense when the optical system is going to be applied, for instance, to a focal plane array. In this case, the pixel size establishes the maximum admissible spot size (which is in general slightly smaller than the pixel size, to allow for tolerances). Going to a spot size smaller than the upper bound is in general useless.

Because our goal is to compare two design strategies (general SMS vs. aplanatic) that are only based on Geometrical Optics, then no wavelength-dependent effects will be considered.

The same concepts presented herein, which have been applied to refractive surfaces, can also be applied to reflective surfaces, as well as to combinations of reflective and refractive surfaces.

2. Design procedure

The design procedure for a two-surface SMS design is given in Ref [7]. More information on the SMS method in 2D is found in [18,22–24], and in [25,26] for 3D SMS. Figure 1 shows a two-surface SMS design with perfect focussing (in 2D) of two bundles of parallel rays incoming at \( \pm 2^\circ \) from the axis. The foci are located at \( \pm 0.5 \text{mm} \) from the symmetry axis.

The SMS design procedure has two basic steps: (a) Calculation on the starting conditions and (b) the SMS extension of the curves. In a two-surface SMS design, a possible set of starting conditions comprises the vertex of one of the surfaces and the part near the axis of other one, specifically that part between the two design rays passing through the vertex of the former surface. Another possible set of starting conditions are shown in Ref [7]. The four-surface SMS design procedure differs from the two-surface one at the first step. In the four-surface case, the starting conditions include the parts of the surfaces near the axis, which are calculated with Gaussian optics, but with the additional condition that there are four rays connecting each of the design points on the object with its corresponding image point (see Fig. 2). The minimum initial part fulfilling this condition is selected. The remaining parts of the curves are deleted. In this situation, each one of the four rays (or their symmetric counterparts) crosses 2 of the 8 edges of the 4 curves A, B, C and D. Rays 1-1' and 4-4' are symmetric.

The four colored dots at the bottom edges of the curves indicate which design rays are passing through them. If there is at least one ray (the blue ray in the case of Fig. 2) with a single bottom edge point. If this is not so, then another four rays can be selected that give smaller initial portions of curves A, B, C and D.

The rest of the lens is calculated in step (b) (SMS extension of the curves). As with any other SMS extension, this step comprises the calculation of a sequence of generalized Cartesian ovals. In each of these generalized Cartesian oval calculations, the trajectories of a small fan of rays are known when crossing all but one of the curves (A, B, C, D). The portion of this last curve that is crossed by the fan can be calculated by equating the optical path lengths. For instance, the rays reaching point 4', below the blue-colored ray of Fig. 2, can be traced back through the surfaces B, C and D, as shown in Fig. 3a. These rays should come from point 4. Consequently we can calculate the red portion of curve A, as shown in Fig. 3a.

Once this new portion has been calculated, at least one new fan of design rays is in the same situation as before. In this case, this fan links points 3 and 3' (see Fig. 3b). The rays issuing from 3 can be traced through the portion of curve A just calculated. The rays of 3' can be traced back through the curves C and D. Now we can calculate a new portion of surface B as a generalized Cartesian oval. This procedure can be repeated to calculate more portions of the curves (see Fig. 3c). The order at which each curve progresses is not necessarily the same at each SMS iteration. Note that since this case is symmetric, every time we calculate a new portion of a lens, its symmetric counterpart is also calculated. As in any refractive Cartesian oval calculation, the SMS extension procedure may give curves that do not represent physically possible optical systems, although they fulfill the equality of optical path lengths between corresponding points. In these cases, different sets of initial conditions must be explored to find a more feasible solution.
Figure 4 shows a four-surface SMS design formed by two lenses. This design sharply focuses the bundle of incoming parallel rays with directions ± 2° and ± 6°, respectively at points ± 0.3mm and ± 0.9 mm off axis.
3. Results

Let $\theta$ be a ray’s off-axis angle. We are going to compare aplanatic designs with equivalent two-surface SMS designs. In both cases we assume the object points to be at infinity (i.e., parallel rays). To evaluate the imaging quality of a design we will use the RMS spot radius functions $\sigma(\theta) = \sigma(P)$ and $\sigma_{2D}(\theta) = \sigma_{2D}(P)$, where $P$ is the image point of the bundle of parallel rays incoming at the direction angle $\theta$. Because all design rays are in the meridian plane, an SMS design of $N$ surfaces will have $\sigma_{2D}(\theta) = 0$ for $N$ design directions $\theta$, assuming that a negligible contribution of the initial parts of the four-surface SMS design, which is not theoretically perfect). Then,

$$
\sigma_{2D}(\theta) = A(\theta) \prod_{i=1}^{N} (\theta - \theta_i)
$$

As shown in ref [27], $A(\theta)$ is an arbitrary analytic function of $\theta$ which, in a first approximation can be considered a constant function of $\theta$ (although the value of this constant varies with different designs), i.e., $\sigma_{2D}(\theta)$ is approximated by the absolute value of an $N$-degree polynomial. The ray tracing results of Fig. 5 show that this approximation is good in this case, and that the spot radius $\sigma_{2D}(\theta)$ is the absolute value of a parabolic function of $\theta$ for small $\theta$. Figure 5 left plots the RMS spot radius $\sigma(\theta)$ (dotted lines) and $\sigma_{2D}(\theta)$ (continuous) of three lenses with the same on-axis thickness, the same focal length $f = 14.31$ mm, and the same $f$-number $/1.576$. The leftmost surface acts as stop surface. One of the lenses is designed for $\theta_i = \pm 2^\circ$ (this design will be called SMS $\pm 2^\circ$) and another for $\theta_i = \pm 1^\circ$ (SMS $\pm 1^\circ$), while the last one is the aplanatic case, which can be seen as an SMS $0^\circ$ case with its two off-axis directions become on-axis ($\theta_i \rightarrow 0^\circ$). The value of the constant $A$ turns out to be quite similar in all three cases. The lens profiles of the three lenses considered in Fig. 5 are very
similar to the one shown in Fig. 2 (the maximum deviation between curve profiles of these three lenses is less than 13 microns).

The spot diagrams are shown in Fig. 6 (aplanatic case), in Fig. 7 (SMS ± 1°), and in Fig. 8 (SMS ± 2°) for different incidence angles from 0 to 4 deg. The rays forming these spots are traced from the nodes of a square grid on a plane parallel to the lens aperture. The diagrams clearly show that the minimum spot sizes occur near the SMS design angle (this is 0° for the aplanatic case) and that the SMS designs are not aplanatic unless the design angle is 0°. These results are similar to those shown in [7] and in [28].

The merit function for comparing these results is the field of view (FOV) diameter for a given maximum RMS spot radius \( \sigma_M \), i.e., the problem is to maximize the FOV when the maximum allowable spot size in the FOV is limited to a given value. Figure 5-left shows that no matter the value of \( \sigma_M \), there is always an SMS design with a wider range of incidence angles (i.e., a wider FOV) than the aplanatic design of equal focal length and \( f \)-number. This best SMS design is the one having the maximum allowable spot size at both normal incidence and the edge of the field. For example, when the maximum RMS spot radius is \( \sigma_M \approx 11.1 \) µm, the SMS design having the widest FOV diameter (which is \( \approx 5.9° \)) is the one designed for incidence angles \( \theta_i = \pm 2° \). Included in this comparison is the SMS ± 0° design, which is simply the aplanatic design. Figure 5-right shows the maximum RMS spot radius \( \sigma_M \) vs. FOV diameter for the aplanatic case and for these SMS designs. This result does not, however, prove that the present SMS designs reach the absolute maximum value of that merit function. In ref [28] Shatz and Bortz applied a global optimization procedure to improve an SMS 2D design and proved that there are other designs, very close to the SMS 2D one, that perform better. In these designs the spot diagrams near the SMS design angle are very similar to a
circular spot [28], and the variation of its radius with the incidence angle is less pronounced than the ones shown in Fig. 5-left. Because the optimized design results very close to the SMS design we conclude that the SMS designs are an excellent starting point for the optimization procedure.

Figure 5 also shows that \( \sigma(\theta) \) and \( \sigma_{2D}(\theta) \) of these two-surface SMS designs are quite similar, which means that the selection of the tangential ray bundles as design rays gives good control of the function \( \sigma(\theta) \), i.e., of the RMS spot radius for all the rays.

Figure 9 shows the RMS spot radius \( \sigma_{2D}(\theta) \) calculated by ray tracing of the four-surface SMS design of Fig. 4 (f/2.241, focal length of 8.59 mm). Although the design of Fig. 4 shows equally spaced incoming directions \( \theta_i = \pm 2^\circ, \pm 4^\circ \), this is not necessary in an SMS design. Equally spacing \( \theta_i \) gives approximately a constant ripple of \( \sigma_{2D}(\theta) \), as Fig. 9 shows. When a maximum spot radius \( \sigma_M \) is fixed, the constant ripple, constrained to \( \sigma_{2D}(\theta) \leq \sigma_M \), maximizes the angular field of view for the function \( \sigma_{2D}(\theta) \). Again the function \( \sigma(\theta) \) (dotted line in Fig. 9) is quite close to \( \sigma_{2D}(\theta) \) in this four-surface design, showing that the selected design ray bundles give good control of the function \( \sigma_{2D}(\theta) \) (as expected) and also for the RMS spot radius of all the rays \( \sigma(\theta) \), which is the relevant function for the evaluation of the image.
quality. The selection of the design ray bundles for an optimum control of $\sigma(\theta)$ is a subject of our ongoing research, because the control of the RMS spot radius for all the rays by means of designing for only a limited subsets of tangential rays is not always as good as that shown in Fig. 9.

4. Conclusions

We have shown a new design strategy for imaging applications of the SMS method, in which we choose as many couples of input and output tangential-ray bundles as there are surfaces of the optical system to be designed. The SMS method ensures that the input and output tangential-ray bundles become fully coupled by the optical system. In the particular examples shown the input bundles are sets of parallel rays with directions $\theta_i$, and the output bundles are their focal points. Once a maximum spot radius in the field is fixed, the SMS design has a wider FOV than that of the equivalent aplanatic designs for the two-surface design (the FOV diameter of the best SMS design is about 1.42 times that of the equivalent aplanat), which is a trivial result when we note that aplanatic designs are just particular cases of SMS designs (when the SMS design directions $\theta_i$ coincide with the optical axis).

In the four-surface SMS design case, we have shown that the selection of design rays among the tangential rays of the FOV give good control of the image formation quality of all the rays and not only for tangential rays.

The flexibility to select the design rays in the SMS method enables it to be applied to non-flat sources or images. This ability (not shared with aplanatic designs) is very interesting for nonimaging applications where often emitters or receivers are non-flat.

Acknowledgement