Optical system assessment for design: numerical ray tracing in the Gaussian pupil

G. W. Forbes

The Institute of Optics, University of Rochester, Rochester, New York 14627

Received December 21, 1987; accepted June 8, 1988

The continuing rapid increase in available computing power has not reduced the importance of efficient methods of optical system assessment for automatic lens design. On the contrary, the new capabilities simply show that truly automatic optical design will eventually be accomplished. It is proposed that the merit of a system-assessment scheme be measured in terms of the accuracy of its estimation of the overall performance of a proposed system as a function of the amount of work done (e.g., number of rays traced). By using this criterion, a number of schemes based on ray tracing are compared, and some highly efficient assessment procedures are developed. As a simplifying approximation, the effects of vignetting and pupil distortion are ignored here. The key to the most-effective methods lies in coupling appropriate coordinates to Gaussian quadrature schemes. Appropriate coordinate systems are those for which the relevant integrands (either wave-front errors or transverse intercept errors) take the form of smooth functions. The resulting methods for system assessment are typically at least an order of magnitude more efficient than comparatively simple schemes.

1. INTRODUCTION

Automatic design of sophisticated systems of any kind (optical or otherwise) usually involves two distinct stages. The initial phase is the formulation of a precise description of the task at hand. This includes the specification of the properties desired of the eventual system and the definition of a single figure of merit that measures, in some appropriately weighted sense, how well a proposed system meets those requirements. The second stage consists of searching through the configuration space of systems in an attempt to uncover the system with the optimal figure of merit. In practice, this stage is usually performed by starting with rough specifications for an appropriate system and switching to an automated procedure, which iteratively modifies the specifications in search of the ultimate system. In such a form, the stage is well described by the term optimization.

Typically, optimization is by far the most computationally demanding stage in the design process and is dominated by the work done in assessing the myriad systems proposed along the way. It is therefore crucial that an efficient scheme be available for system assessment. Virtually all the optimization algorithms presented to date follow a strictly monotonic path in the configuration space; that is, they accept modifications to the system only if the figure of merit is thereby improved. This type of strategy greatly simplifies the process but restricts the final solution to a certain neighborhood of the starting point, yielding a locally optimal solution. Such a strategy cannot discover the globally optimal solution if there is no monotonic connection to the starting point. The design of global optimization schemes is a popular topic of research, and some of these ideas are being applied in the context of optical design.[1,2] Any such strategy clearly adds a new degree of complexity and will generally require significantly more exploration of the configuration space. For this reason, the pressure on the efficiency of the schemes for optical system assessment is now redoubled.

In optical design, the merit function involves a combination of factors, including some account of the feasibility of manufacturing the system (tolerances, cost, etc.), the optical performance of the system, and perhaps penalties to enforce certain constraints (although these may be incorporated at a higher level in the optimization algorithm). In this paper I concentrate solely on the component of the merit function that measures the optical quality of the system. This aspect is usually evaluated by tracing a limited set of rays through the proposed system and measuring transverse errors in the ray intercepts or optical path differences (OPD's) in some way. Recently Robb[3] proposed that a measure of the optical quality of a system could be determined as much as an order of magnitude more efficiently by using a direct computation of aberration coefficients in place of ray tracing. However, Robb had access to only the seventh-order series, and it was found that, owing to truncation errors, this approach was not suitable when ray angles in the system approached 30°. Since that research was done, schemes were devised to compute aberration series to arbitrarily high orders[4,5] and to extend and accelerate their convergence.[6,7] These new results suggest a reexamination of the efficiencies of various schemes for the assessment of the optical quality of proposed systems.

A first step in this direction is the investigation of the efficiency of ray tracing for optical system assessment. The fundamental measure of this efficiency is taken here to be the number of rays that need to be traced in order to determine the value of the overall system performance to various accuracies (say, to 1, 2, or 3 significant digits). Clearly, this issue can be addressed only after the measure of overall system performance is defined. In practice, it is the behavior of the modulation transfer function (MTF) of the system for various patches of the field (together with some idea of the geometrical distortion) that often provides the ultimate measure, and there are a number of efficient means to calculate approximations to the properties of the MTF. For

0740-3232/88/111943-14$02.00

example, the variance of the OPD gives an approximation to the Strehl ratio,[8] which, in turn, is related to the volume under the MTF. Alternatively, the variance of the transverse ray error (i.e., the mean-square spot size) can be seen to be related to a second-order Taylor expansion of the MTF about zero spatial frequency.[9] These two entities, the variance of the OPD and that of the ray intercepts, are simply averages of quantities that are directly available from the ray-trace data, and they are often adopted as bases upon which to build a merit function. In this paper the efficiency of a system-assessment scheme is measured in terms of the accuracy with which the scheme can approximate this type of parameter as a function of the number of rays traced.

The schemes for optical assessment by using ray tracing can be regarded then simply as numerical integration schemes for performing the appropriate averages. Furthermore, since the transverse intercept error is, to a good approximation, proportional to the gradient of the OPD, it is reasonable to expect a scheme that efficiently approximates one of these two types of integral to be also well suited to the other. This is in fact the case. However, the data from a single ray through a system directly yield not only the value of the OPD but also it derivatives. This suggests the possibility of devising more-specialized numerical integration schemes to use this extra information.[10] Since the spot size is more straightforward, I shall first deal with that problem alone.

The problem now takes the form of the evaluation of particular multidimensional integrals, which can be structured as follows. The innermost integration, for averaging over the pupil, ranges over the two-dimensional family of rays transmitted by the system from a point source of a given color and position. Next, it is convenient to integrate over color in order to model a white point source and then to finish with the integration over all positions lying within the desired field. (For a rotationally symmetric system, the last integral can clearly be reduced to a one-dimensional form.) Each of these integrals involves a weighting function, which must be laid down during the first phase of design. The region of interest for the field and color is also specified explicitly at that stage. This significantly simplifies the associated averaging, and it is often the case that a small sample in field and color is all that is required. For the aperture average, on the other hand, the family of rays involved (or the true entrance pupil shape) is almost never known precisely. For this reason, a relatively large number of rays is sometimes needed in sampling the aperture. For simplicity, the Gaussian entrance pupil is often used as a first approximation in determining the rays transmitted by the system. This approximation is also adopted in this paper, and the effects of pupil distortion and vignetting are left as separate (often crucial) issues to be addressed independently. Methods that account for these effects must generally be adopted in at least the final stages of design.

For calculating the spot size of a monochromatic point source to an accuracy of 1%, it was claimed by Foreman[11] that approximately 100–150 rays are required (in one half of the pupil). The integration scheme in the Gaussian pupil was based on sampling on a polar grid with uniform spacing in radius and angle and uniform weighting for each ray. (It was recommended that a radial spacing of at most 1/20 of the aperture radius and an angular spacing on each ring of approximately 30–45° be used.) This figure was reduced by Andersen,[12] who introduced a fourth-order weighting scheme using the same sample points. Andersen expected his routine to yield an accuracy of 1% with about 20 rays and an accuracy of 0.01% with 40–50 rays. However, as reported in Section 3 of this paper, his scheme appears to require approximately 75 and 300 rays, respectively, to attain these accuracies. (For 1% accuracy, Andersen's scheme requires once again a 30–45° angular spacing with about 12 rings across the pupil.) For comparison, if uniform weighting on a square Cartesian grid in the pupil is used, it is found in Section 3 that 500 rays are needed in the half pupil in order to produce reliably an accuracy of 1%, and hundreds of thousands of rays are needed to reduce this figure to 0.01%.[13]

By introducing appropriate variable transformations and more-accurate integration schemes, it is shown that the numbers of rays quoted above can be reduced remarkably. Since high-order integration schemes can be expected to yield higher accuracy only when the integrand is a smooth function (i.e., one resembling a low-order polynomial), the variable transformations mentioned above are used to make the integrand smooth. Likewise, in the subsequent averaging over field and color, it is vital that appropriate variables be adopted. In the context of reducing truncation error for aberration series, variables have been sought for this purpose[6,14] (i.e., attempting to render the aberration function polynomial-like), and the results are used to advantage here together with higher-order integration schemes such as Gaussian quadrature. It is emphasized that the key novel feature of the research presented in this paper lies in the choice of appropriate underlying coordinates that govern the sampling and weighting schemes for approximating the integrals.

The form of the integrals used for the spot-size calculations is given in Section 2, and the ensuing Sections 3, 4, and 5 are devoted to the integrals over aperture, color, and field, respectively. The results in these sections are based on the analysis of a variety of unobscured systems, including a wide-angle system (the hypergon[15]), a Cooke triplet,[16] a double Gauss,[17] a Schmidt camera,[18] and a microscope objective.[19] A discussion of the efficiency of overall integration schemes appears in Section 6 and is followed by a number of concluding remarks.

2. A FIGURE OF MERIT FROM THE RAY INTERCEPTS

It is convenient to define the entities to be computed in subsequent sections in one location. To this end, some notation must first be introduced: Three variables are used to specify a ray (two of which are two dimensional). The field and aperture variables, written here as **f** and **a**, specify the location of the intercept with the object and entrance pupil surfaces, respectively, and the color of the ray is specified by a coordinate written as ω. (If either the object surface or the pupil surface is at infinity, the corresponding variable must be taken to specify a direction.) Since the integrations performed here are for the purposes of averaging over these variables, it is convenient to introduce some shorthand. The average of a function $f(x)$ with weight $\mathcal{X}(x)$ is written as

$$\mathrm{Avg}_x\{f(x)\} := \frac{\int f(x)\mathcal{X}(x)\mathrm{d}x}{\int \mathcal{X}(x)\mathrm{d}x}, \tag{2.1}$$

where the integral can formally be over all values of x if the region of interest is absorbed into the weighting function. As a final piece of notation, in the image space of the system, let $\mathbf{y}(\mathbf{f}, \mathbf{a}, \omega)$ and $\mathbf{Q}(\mathbf{f}, \mathbf{a}, \omega)$ be the functional representations of the intercept and the direction tangents of a ray at some fixed plane in the image space, say, the Gaussian image plane. The intercept, say, $\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega)$, with a plane defocused by Δ is then given by $\mathbf{Y} = \mathbf{y} + \Delta\mathbf{Q}$.

If $L^2(\mathbf{f})$ is the mean-square spot size of the image (in a fixed defocused plane) of a point source at the field point specified by $\mathbf{f}$, the figure of merit for the overall image in that plane, say, $\mathcal{M}$, is taken to be given by

$$\mathcal{M}^2 := \mathrm{Avg}_{\mathbf{f}}\{L^2(\mathbf{f})\}. \tag{2.2}$$

That is, $\mathcal{M}$ is an average measure of the size of the (polychromatic) geometric point-spread function. Note that the associated weighting factor, $\mathcal{F}(\mathbf{f})$, is specified in the initial phase of the design[20] together with $\mathcal{W}(\omega)$, which is needed below. For a symmetric system, $L^2(\mathbf{f})$ depends only on $|\mathbf{f}|$, and the average in Eq. (2.2) can be reduced to a one-dimensional integral.

$L^2(\mathbf{f})$ is defined in terms of the centroid of the image of the point source. Accordingly, define $\overline{\mathbf{Y}}(\mathbf{f}, \omega)$ to be the centroid of the image for a given color, i.e.,

$$\overline{\mathbf{Y}}(\mathbf{f}, \omega) := \mathrm{Avg}_{\mathbf{a}}\{\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega)\}. \tag{2.3}$$

In all the averages over the aperture variable (which is taken to be the innermost integral here), the weight used for the element d^2a in the pupil is assumed to be proportional to the solid angle subtended by the element at the field point $\mathbf{f}$ and is written as $\mathcal{A}(\mathbf{f}, \mathbf{a})$. If the object is at infinity, $\mathcal{A}(\mathbf{f}, \mathbf{a})$ is taken to be the normal cross-sectional area of the beam from $\mathbf{f}$ that fills the element d^2a. The centroid of the colored image, say, $\overline{\overline{\mathbf{Y}}}(\mathbf{f})$, is then defined to be

$$\overline{\overline{\mathbf{Y}}}(\mathbf{f}) = \mathrm{Avg}_{\omega}\{\overline{\mathbf{Y}}(\mathbf{f}, \omega)\} = \mathrm{Avg}_{\mathbf{a},\omega}\{\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega)\}. \tag{2.4}$$

$L^2(\mathbf{f})$ is now given by

$$\begin{aligned} L^2(\mathbf{f}) &:= \mathrm{Avg}_{\mathbf{a},\omega}\{[\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega) - \overline{\overline{\mathbf{Y}}}(\mathbf{f})]^2\} \\ &= \mathrm{Avg}_{\mathbf{a},\omega}\{([\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega) - \overline{\mathbf{Y}}(\mathbf{f}, \omega)] + [\overline{\mathbf{Y}}(\mathbf{f}, \omega) - \overline{\overline{\mathbf{Y}}}(\mathbf{f})])^2\} \\ &= \mathrm{Avg}_{\omega}\{\mathrm{Avg}_{\mathbf{a}}\{[\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega) - \overline{\mathbf{Y}}(\mathbf{f}, \omega)]^2\} + [\overline{\mathbf{Y}}(\mathbf{f}, \omega) - \overline{\overline{\mathbf{Y}}}(\mathbf{f})]^2\}. \end{aligned} \tag{2.5}$$

It is convenient to define σ^2 to be the monochromatic mean-square spot size, i.e.,

$$\sigma^2(\mathbf{f}, \omega) := \mathrm{Avg}_{\mathbf{a}}\{[\mathbf{Y}(\mathbf{f}, \mathbf{a}, \omega) - \overline{\mathbf{Y}}(\mathbf{f}, \omega)]^2\}, \tag{2.6}$$

then $L^2(\mathbf{f})$ can be written as[21]

$$L^2(\mathbf{f}) = \mathrm{Avg}_{\omega}\{\sigma^2(\mathbf{f}, \omega) + [\overline{\mathbf{Y}}(\mathbf{f}, \omega) - \overline{\overline{\mathbf{Y}}}(\mathbf{f})]^2\}. \tag{2.7}$$

Since averaging is a linear process, it follows that $\overline{\mathbf{Y}}(\mathbf{f}, \omega)$ is exactly a linear function of the defocus, Δ. In an obvious notation, $\overline{\mathbf{Y}}$ is given by $\overline{\mathbf{Y}}(\mathbf{f}, \omega) = \overline{\mathbf{y}}(\mathbf{f}, \omega) + \Delta\overline{\mathbf{Q}}(\mathbf{f}, \omega)$. Similarly, after expanding the arguments in Eq. (2.6), one can see that $\sigma^2(\mathbf{f}, \omega)$ is exactly a quadratic function of Δ. So, in fact, is $L^2(\mathbf{f})$ and hence $\mathcal{M}^2$, and it is convenient to write these in the following form:

$$\sigma^2(\mathbf{f}, \omega) = S_0(\mathbf{f}, \omega) + 2\Delta S_1(\mathbf{f}, \omega) + \Delta^2 S_2(\mathbf{f}, \omega), \tag{2.8}$$

$$L^2(\mathbf{f}) = L_0(\mathbf{f}) + 2\Delta L_1(\mathbf{f}) + \Delta^2 L_2(\mathbf{f}), \tag{2.9}$$

$$\mathcal{M}^2 = \mathcal{M}_0 + 2\Delta\mathcal{M}_1 + \Delta^2\mathcal{M}_2. \tag{2.10}$$

It is easy to show, for example, that $S_0 = \mathrm{Avg}_{\mathbf{a}}\{(\mathbf{y} - \overline{\mathbf{y}})^2\}$, $S_1 = \mathrm{Avg}_{\mathbf{a}}\{(\mathbf{y} - \overline{\mathbf{y}}) \cdot (\mathbf{Q} - \overline{\mathbf{Q}})\}$, etc. The subscripted entities appearing in these three equations are calculated in the following sections and, together with the expressions (linear in Δ) for $\overline{\mathbf{Y}}$ and $\overline{\overline{\mathbf{Y}}}$, can be used to find general surfaces of best focus for white or monochromatic light and the optimally defocused plane for which the mean size, $\mathcal{M}$, is a minimum (viz., $\Delta_{\mathrm{optimal}} = -\mathcal{M}_1/\mathcal{M}_2$, where $\mathcal{M}_{\mathrm{optimal}} = \{\mathcal{M}_0 - \mathcal{M}_1^2/\mathcal{M}_2\}^{1/2}$). Finally, notice that the centroid location is evaluated during this process, and, if needed, some account for distortion can be included in the merit function.

3. AVERAGING OVER THE APERTURE

The efficiency of various schemes for the computation of the monochromatic spot size, i.e., $\sigma(\mathbf{f}, \omega)$ defined in Section 2, is reported here. In practice, $S_i(\mathbf{f}, \omega)$ for $i = 0, 1, 2$ are calculated for the Gaussian image plane, and the errors referred to are typical errors in $\sigma_{\mathrm{Gaussian}} = \{S_0(\mathbf{f}, \omega)\}^{1/2}$ and $\sigma_{\mathrm{optimal}} = \{S_0(\mathbf{f}, \omega) - S_1(\mathbf{f}, \omega)^2/\mathbf{S}_2(\mathbf{f}, \omega)\}^{1/2}$ evaluated at the edge of the field of the system (where the error is usually largest) and near the middle of the spectral region of interest. It is found that the errors for a given integration scheme are, to a good approximation, insensitive to the system type. This rather surprising result holds at least for the range of symmetric systems mentioned at the end of Section 1, except for the microscope objective, which is discussed separately at the end of this section. Since these errors are also insensitive to the value of the color coordinate, one representative value of the error can be quoted for each scheme. Furthermore, once the field point is moved away from the axis to, say, 20–30% of the full extent of the field, the accuracy is found to be roughly independent of movement across the remainder of the field. This means that, in effect, the errors quoted in this section give a good approximation to the errors for a general off-axis field point. (On axis, of course, the aperture average becomes one dimensional for the symmetric systems considered here.) Because of symmetry, it is necessary only to sample one half of the pupil, and the number of rays quoted in this section and elsewhere in this paper is the number of rays that are actually traced.

A couple of standard methods for sampling the pupil appear to be guided by the principle of having a uniform density of points. These include a square Cartesian grid and a polar grid with uniform radial spacing and uniform arc length between adjacent points on a given ring.[22] Examples of these schemes are shown in Figs. 1 and 2, respectively. Notice that the polar scheme samples more densely in the radial direction (the separation in the angular direction is a factor of $\pi/2$ greater than the radial separation), and this is appropriate for the types of functions considered here, since they generally have weaker angular dependence. Cartesian grids can never exploit this property.

If an equal weight is adopted for each point, the accuracy

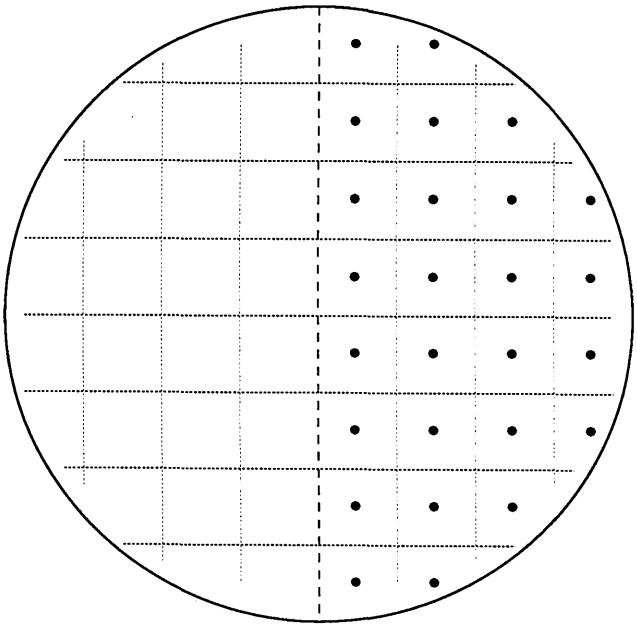

Fig. 1. Example of Cartesian sampling in the pupil. With uniform weighting, this particular grid typically results in 3% accuracy in the calculation of spot size with 26 rays, which is shown to be one of the more-efficient Cartesian schemes.

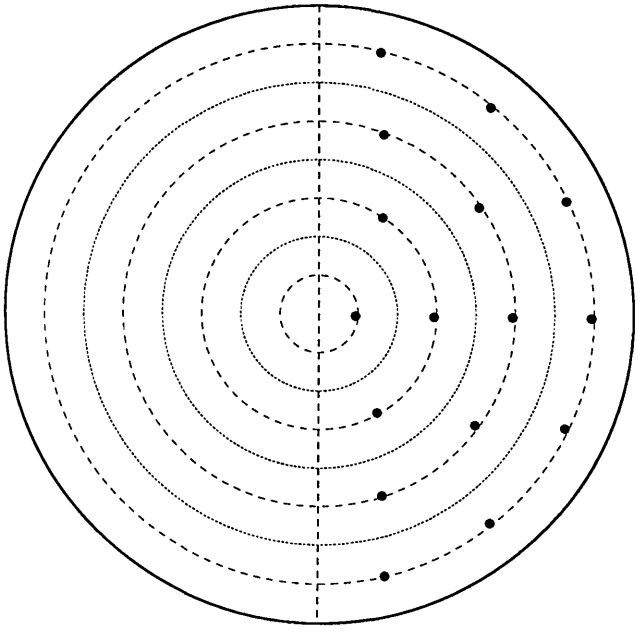

Fig. 2. Example of a uniform polar grid. The sample points are arranged on equally spaced rings with a fixed arc length separating the points on each ring. With uniform weighting, this particular grid yields about 15% accuracy in the calculation of spot size with 16 rays.

of the Cartesian method can be expected to be erratic, as variation of the side length of the lattice causes sample points to fall on or off the disk in an irregular manner. Some idea of the fluctuation can be obtained by considering the results of integrating a constant function over a disk in order to approximate the area. The error in the resulting approximation, together with the number of sample points in half of the disk, is plotted in Fig. 3 as a function of the number of points across a radius, say, n.[23] (The grid shown in Fig. 1 corresponds to $n = 4.0$.) The number of sample points is given roughly by $(\pi/2)n^2$, and the error is on average proportional to $n^{-3/2}$, or $N^{-3/4}$, where N is the number of sample points in half of the disk (as can be deduced simply from the standard results for a random walk, for example). This type of fluctuation is also manifested in the results of spot-size calculations, as is made clear by Fig. 4. The curve plotted

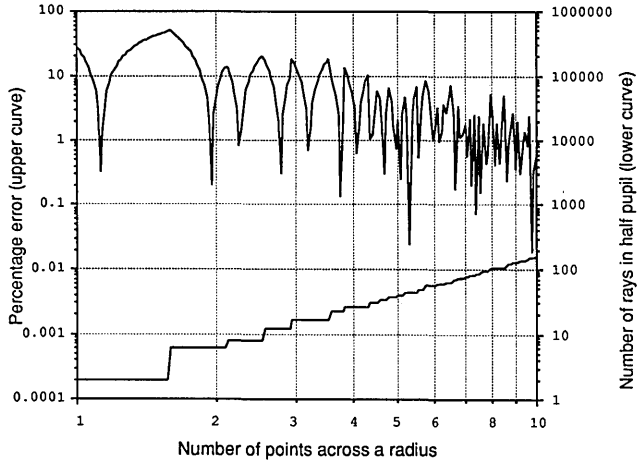

Fig. 3. Efficiency of numerically integrating to determine the area of a disk by using a Cartesian grid with uniform weighting. The downward spikes in the error curve represent continuous changes of sign and locate configurations that exactly integrate a constant over the disk. The turnover points, where the error is at a local maximum, occur when new sample points cross over the edge of the disk (as can be seen by correlating the two curves), and this causes the error to jump discontinuously from a negative value to a positive one.

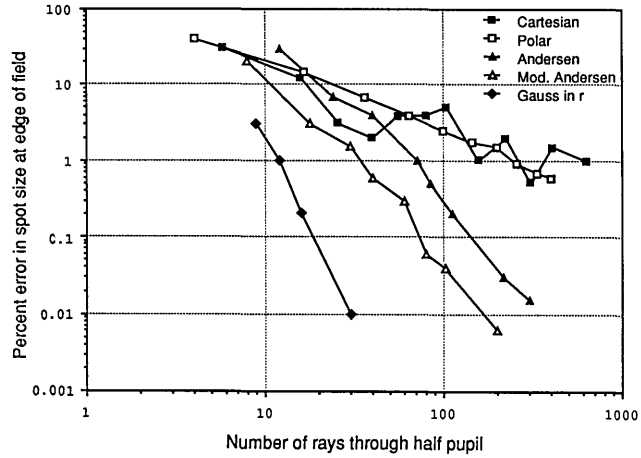

Fig. 4. Efficiencies of a variety of integration schemes for the determination of spot size. The curves are plotted for (■) a Cartesian grid with uniform weights; (□) a polar grid with uniform weights; (▲) Andersen's scheme, which is effectively Simpson's method on a Cartesian grid in the polar coordinate plane (r, θ), where the disk takes the form of a rectangle; (△) a modified version of Andersen's method that uses uniform weighting for the angular integral; and (♦) a scheme that uses Gaussian quadrature in the radial direction. The curves represent typical errors. By chance, a given method may have significantly better accuracy on some occasions, but, for the systems analyzed here, the errors rarely exceeded those reported by more than a factor of 2 or so (a relatively small shift on this log plot).

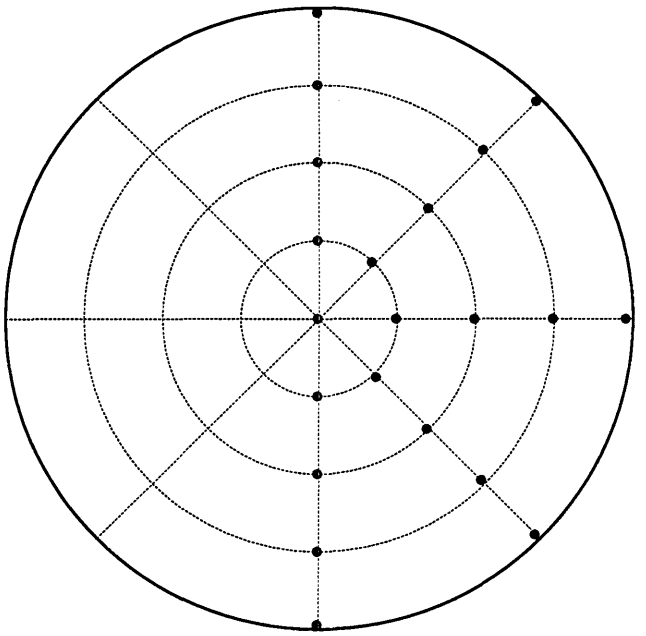

Fig. 5. An example of the type of grid used by Andersen to sample the pupil. With Andersen's weighting, this particular grid yields ≃10% accuracy with 21 rays.

for the Cartesian method uses $n = 2.0, 3.0, 4.0, \ldots, 8.0, 10.0, 12.0, 14.0, 16.0, 20.0$. Notice that the error for $n = 8$ (with 104 rays) is greater than that for $n = 4$ (with 26 rays). It is informative to see where these points lie on the curve in Fig. 3, where it is clear that $n = 4$ and $n = 5$ happen to be close to configurations that can exactly integrate a constant, while $n = 8$ is not. This corresponds well with the oscillations apparent in Fig. 4. In fact, if only those schemes that can exactly integrate a constant are used, the high points on the Cartesian curve in Fig. 4 can be reduced (although not consistently to the level of the straight line through the lowest points, since there remain differences in the errors in integrating higher-order terms).

The polar scheme illustrated in Fig. 2 can clearly be extended consistently to any number of rays by adding more rings to the outside of the disk. The error resulting from this method is also plotted in Fig. 4 and is referred to as the polar method. Clearly, the fluctuations have been removed, and the polar scheme lies between the best and the worst Cartesian schemes. Since the error of the polar scheme is found to be consistently negative, it appears that the best Cartesian methods are superior to the polar schemes because the functions being integrated increase rapidly near the edge of the pupil and the Cartesian methods usually include sample points whose minimum distance to the perimeter is significantly smaller than a lattice side length.[24]

The efficiency of Andersen's method, which uses the sampling scheme illustrated in Fig. 5, is also presented in Fig. 4. The method is not significantly better than those previously discussed until the number of rays exceeds 50 or so, and it can yield 1% accuracy with approximately 75 rays. Andersen expected an error of 1% with only 20 rays, but the error is about an order of magnitude worse at that point (and more than 2 orders of magnitude worse than he expected at 40–50 rays). This is partly due to the fact that when Andersen analyzed the integration over angle, he assumed that all points on a given ring had uniform weight. This is not the case for his method, which uses Simpson's rule and ends up with every other point weighted twice as heavily. This alternating weight in Simpson's rule is an artifact of end effects; yet, for integration around a circle, there are no ends: all points are equivalent.

Replacing Andersen's angular weighting scheme with uniform weighting considerably simplifies the algorithm for the two-dimensional integral, yet it increases the accuracy of the result.[25] (This modification also allows the angular sample points to be rotated by half a subdivision to move all but the central point off the plane of symmetry, thereby further reducing the number of rays.) The error of the resulting scheme is also presented in Fig. 4. It can be seen that the error for a given number of rays is decreased by a factor of 4 or so (and the number of rays required for a given accuracy is roughly halved). It is remarked that when fewer than 100 rays are being traced, it is best to have the number of angular divisions in $(0, 2\pi)$ roughly equal to the number of rings, but when the number of rays increases to approximately 100 and beyond, the number of rings should be a factor of 2 or so higher.

As a final method based on these simple polar coordinates, it is interesting to consider using a Gaussian scheme for the radial integral and retaining the uniform sampling and weighting in angle. The radial integral then takes the form

$$\int_0^R f(r,\theta) r \mathrm{d}r = R^2 \int_0^1 f(Ru,\theta) u \mathrm{d}u \approx R^2 \sum_{j=1}^{N_r} w_j f(Ru_j,\theta), \quad (3.1)$$

where u_j and w_j are the Gaussian sample points and weights chosen to make the method exact for f, a polynomial in r of degree $2N_r - 1$. Uniform sampling and weighting in angle is the optimal Gaussian method for this case involving the interval $[0, \pi]$, with the basis functions being polynomials in $\cos\theta$.[25] Combining these methods yields

$$\begin{aligned} I &:= \int_0^R \int_0^{2\pi} f(r,\theta) \mathrm{d}\theta r \mathrm{d}r \\ &= R^2 \int_0^1 \left\{ 2 \int_0^{\pi} f(Ru,\theta) \mathrm{d}\theta \right\} u \mathrm{d}u \\ &\approx R^2 \sum_{j=1}^{N_r} w_j \left\{ 2 \frac{\pi}{N_\theta} \sum_{k=1}^{N_\theta} f(Ru_j,\theta_k) \right\}, \end{aligned} \quad (3.2)$$

where $\theta_k = (k - \frac{1}{2})\pi/N_\theta$. The efficiency of such a scheme is also presented in Fig. 4 and clearly is a remarkable improvement, yielding ≃1% accuracy with only a dozen rays ($N_r = 4$, $N_\theta = 3$). The Gaussian parameters for these radial schemes are not presented here, since more-accurate schemes are developed below. It is remarked at this point that for cases in which the effects of pupil distortion and vignetting need to be determined (and accounted for) simultaneously with the integration, pure Gaussian schemes are not possible, so other methods are also considered in what follows. (These effects are sometimes approximated by linearly squashing the pupil by a prespecified amount, and, of course, it is also possible to use this type of approach in conjunction with Gaussian methods.)

The accuracy of all numerical integration schemes depends on the smoothness of the integrand. For example, as

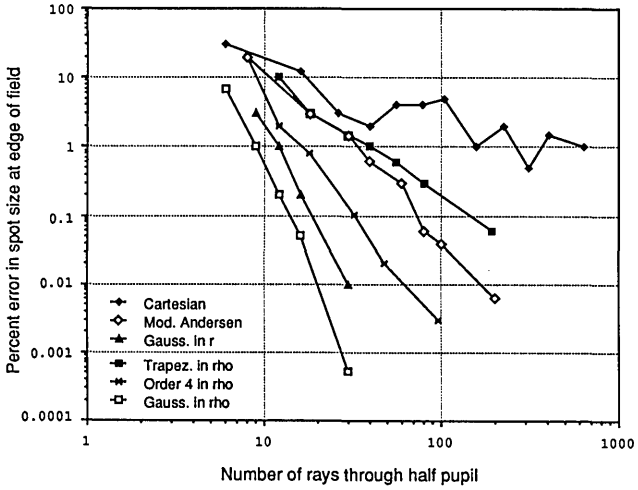

Fig. 6. Comparison of the efficiencies of schemes based on the radial coordinate r with those based on the coordinate ρ, which is just a normalized version of r^2. The first three methods in the legend are described in Fig. 4. The latter three are based on ρ and include (■) a trapezoidal scheme, (×) a fourth-order scheme, and (□) Gaussian quadrature, respectively.

mentioned for the Gaussian scheme, with four rings in the pupil it is possible to integrate exactly a polynomial of degree 7 in r. If the integrand cannot be approximated well by such a polynomial, the accuracy will be poor. In fact, for the case in hand, it is not necessary to be able to integrate odd powers of r. This can be seen as follows: the integrand $f(r, \theta)$ has a plane of symmetry and can be expanded as a series in r^2 and $r \cos\theta$. If the sample points for the angular integral are those specified after Eq. (3.2), the symmetry entails that all odd powers of $\cos\theta$ sum identically to zero in the numerical approximation of the angular integral.[26] It follows then that the radial integral has to deal only with even powers of r. This observation suggests that improvements can be made to the radial components of the integration schemes discussed above.

For example, the modified version of Andersen's method uses Simpson's rule radially, which means that it is exact for cubics in r. However, the linear and cubic terms in r drop out, and only the constant and quadratic terms are integrated exactly. If the radial sampling is changed to uniform sampling in r^2 and a fourth-order weighting scheme (such as Simpson's method or an end-corrected trapezoidal rule) is again adopted, the method becomes exact for cubics in r^2, which means that significantly higher accuracy can be expected. Accordingly, expressing the integral in terms of the variable $\rho = r^2/R^2$ leads to an integration scheme of the form

$$I = R^2 \int_0^1 \left\{ 2 \int_0^\pi f(R\rho^{1/2}, \theta) \mathrm{d}\theta \right\} \frac{\mathrm{d}\rho}{2} \tag{3.3}$$

$$\approx \frac{1}{2} R^2 \sum_{j=0}^{N_r} w_j \left\{ 2 \frac{\pi}{N_\theta} \sum_{k=1}^{N_\theta} f\left[R \left(\frac{j}{N_r} \right)^{1/2}, \theta_k \right] \right\}, \tag{3.4}$$

where θ_k is defined as before and w_j can be chosen in a number of ways. The simplest choice follows from the trapezoidal rule, which specifies in an obvious notation that the weights be given by $\mathbf{w} = \mathbf{w}_T = (1/N_r)\,(1/2, 1, 1, 1, \ldots, 1, 1/2)$. Anticipating the effects of vignetting and pupil distortion, it is interesting to consider this case, since it probably gives a reasonable indication of the accuracy attainable when the boundary must be determined and accounted for with information from the same rays used in the integration. The efficiency of this trapezoidal scheme is plotted in Fig. 6, which shows that this second-order method (which can exactly integrate a function linear in ρ) has an efficiency comparable with that of the modified version of Andersen's method for 1% accuracy or less. The radial weights can alternatively be chosen according to a fourth-order rule: either Simpson's rule, i.e., $\mathbf{w} = \mathbf{w}_s = (1/3N_r)\,(1, 4, 2, 4, \ldots, 2, 4, 1)$, or the more attractive end-corrected scheme in which the weights are equal to those for the trapezoidal case (and hence are uniform away from the ends) but have the last four weights at either end modified by taking $\mathbf{w} = \mathbf{w}_T + \mathbf{c} + \tilde{\mathbf{c}}$, where $\mathbf{c} = (1/48N_r)\,(-7, 11, -5, 1, 0, 0, 0, \ldots, 0)$ and $\tilde{\mathbf{c}}$ is the mirror image of $\mathbf{c}$. These methods give roughly the same accuracy, and their efficiency is also plotted in Fig. 6. The error is typically an order of magnitude smaller than that of the best fourth-order scheme based on r when more than 30 rays are traced and is clearly a significant improvement when more than 10 rays are traced, yielding $\simeq 1\%$ accuracy with 17 rays ($N_r = 4$, $N_\theta = 4$).

Finally, it is possible to devise a superior Gaussian scheme by once again adopting ρ in the radial integral. From Eq. (3.3),

$$I \approx R^2 \sum_{j=1}^{N_r} w_j \left\{ 2 \frac{\pi}{N_\theta} \sum_{k=1}^{N_\theta} f(R\rho_j^{1/2}, \theta_k) \right\} \tag{3.5}$$

can be obtained, where w_j and ρ_j are related simply to the standard Gaussian parameters for integration with a constant weight function and can be found from the zeros of Legendre polynomials (and the values of their first derivatives at those points).[27] The appropriate weights and sam-

Table 1. Gaussian Integration Parameters for the Radial Integral as Approximated in Eq. (3.5)

N_r	j	$\rho_j^{1/2}$	w_j
1	1	0.70710678	0.50000000
2	1	0.45970084	0.25000000
	2	0.88807383	0.25000000
3	1	0.33571069	0.13888889
	2	0.70710678	0.22222222
	3	0.94196515	0.13888889
4	1	0.26349923	0.08696371
	2	0.57446451	0.16303629
	3	0.81852949	0.16303629
	4	0.96465961	0.08696371
5	1	0.21658734	0.05923172
	2	0.48038042	0.11965717
	3	0.70710678	0.14222222
	4	0.87706023	0.11965717
	5	0.97626324	0.05923172
6	1	0.18375321	0.04283112
	2	0.41157661	0.09019039
	3	0.61700114	0.11697848
	4	0.78696226	0.11697848
	5	0.91137517	0.09019039
	6	0.98297241	0.04283112

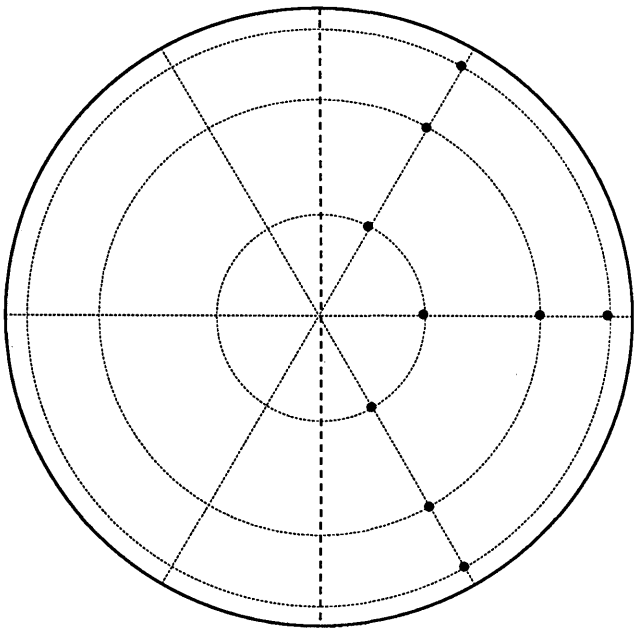

Fig. 7. Sample points for the nine-ray configuration based on Gaussian quadrature. This scheme (when the weights in Table 1 are used for $N_r = 3$) has an accuracy of $\simeq 1\%$.

ple points are presented in Table 1 for $N_r = 1, 2, 3, \ldots, 6$. (Parameters of this sort are typically needed only to two decimal places beyond the expected accuracy of the resulting integral approximation. However, the additional tabulated accuracy is useful for checking purposes.) The efficiency of this method (also presented in Fig. 6) is extremely good, with 1% accuracy requiring only nine rays ($N_r = 3$, $N_\theta = 3$) arranged in the configuration shown in Fig. 7. Some important observations concerning the accuracy of these schemes based on such small ray sets are offered at the end of this section. The error for this method when 30 rays are used ($N_r = 5$, $N_\theta = 6$) is almost 4 orders of magnitude below the error of the Cartesian method, which requires more than 10^6 rays to yield this same accuracy. The fact that the Gaussian pupil is taken as the region of integration should not be forgotten, however, and once the accuracy has reached the limit set by the error incurred by this assumption, it makes no sense to continue any further. (Procedures for handling this problem must be considered separately.) However, notice that even within the 1–10% accuracy range, this Gaussian scheme is unequaled.

When experimenting with these small ray sets, it is interesting to examine other options that may yield some configurations between the six-ray case ($N_r = 2$, $N_\theta = 3$) with 10% accuracy and the nine-ray case of Fig. 7 with 1% accuracy. [Note that the next two configurations plotted in Fig. 6 are ($N_r = 3$, $N_\theta = 4$) and ($N_r = 4$, $N_\theta = 4$).] One obvious possibility is to add one ray at the center of the pupil, thereby changing the radial sampling to a Radau scheme,[28] if maximal accuracy is to be retained. The integral is then approximated by

$$I \approx R^2 \sum_{j=0}^{N_r} v_j \left\{ 2 \frac{\pi}{N_\theta} \sum_{k=1}^{N_\theta} f(Ru_j, \theta_k) \right\}, \tag{3.6}$$

for which the associated weights and sample points are given in Table 2. The seven-point configuration ($N_r = 2$, $N_\theta = 3$) is illustrated in Fig. 8 and is found to yield an accuracy of $\simeq 3\%$, which lies on the lowest curve in Fig. 6 (as all these Radau schemes do) and is significantly better than that of the six-ray configuration mentioned above. With fewer rays than this the results become highly erratic.

To this point, nothing has been said of the microscope objective. Notice that all the systems discussed so far either have their object at infinity or, in the case of the hypergon (a

Table 2. Radau Integration Parameters for the Radial Integral as Approximated in Eq. (3.6)

N_r	j	u_j	v_j
1	0	0.00000000	0.12500000
	1	0.81649658	0.37500000
2	0	0.00000000	0.05555556
	1	0.59586158	0.25624291
	2	0.91921106	0.18820153
3	0	0.00000000	0.03125000
	1	0.46080423	0.16442216
	2	0.76846154	0.19409673
	3	0.95467902	0.11023111
4	0	0.00000000	0.02000000
	1	0.37384471	0.11155195
	2	0.64529805	0.15591326
	3	0.85038637	0.14067801
	4	0.97102822	0.07185678
5	0	0.00000000	0.01388889
	1	0.31390299	0.07991019
	2	0.55184756	0.12134680
	3	0.74968339	0.13023170
	4	0.89553704	0.10422533
	5	0.97989292	0.05039710

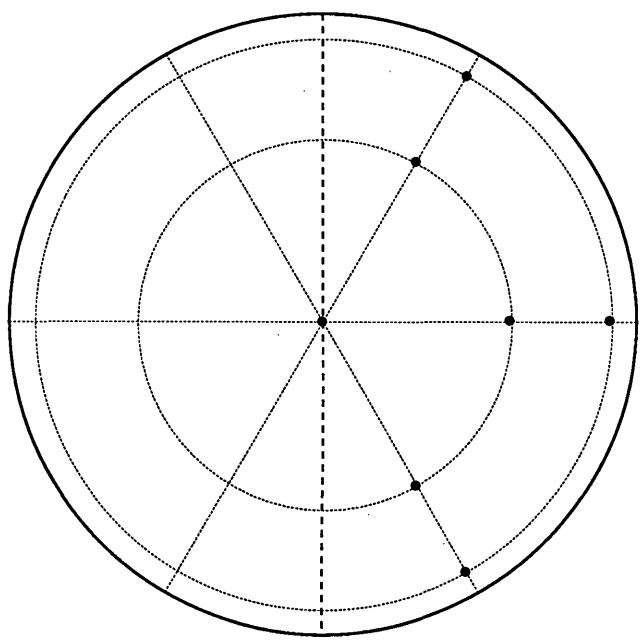

Fig. 8. Seven-ray configuration, with a Radau method used for the radial integral. This sampling scheme, together with the weights from Table 2, for $N_r = 2$, yields $\simeq 3\%$ accuracy for calculating spot-size information.

unit-magnification, $f/30$ system), have small numerical apertures. When the entrance pupil subtends a sizable angle (say, a half-cone angle in excess of 20° or so) the efficiencies of the Gaussian integration schemes discussed above are found to deteriorate significantly. For the microscope objective considered here, this angle, say, U, is approximately 50°. It is usual, however, to treat the object as the image in designing such a system, and in this way these large angles are avoided. (The entrance pupil for the system reversed in this way subtends roughly 1°.) If the system is analyzed in this manner, the results quoted above hold reasonably well. However, if the system is not reversed, the Gaussian schemes perform poorly, with errors sometimes significantly more than an order of magnitude bigger than expected. This unsatisfactory situation is to be expected, since it was found[6] that aberration functions of the form considered here are by no means smooth when U approaches 40°, and a new coordinate was recommended (for both the field and aperture) in place of the regular position vectors that have been used so far.

The new aperture coordinate, written here as $\mathbf{a}$, is just the direction cosine of the line joining the center of the object and the location of the intercept on the pupil. Adopting this variable changes only the radial part of the integration schemes as follows. The variable r^2 must be replaced by $\xi = \mathbf{a} \cdot \mathbf{a} = r^2/(D^2 + r^2)$, where D is the distance from the object to the entrance pupil. This puts the radial integral in the form

$$\int_0^R f(r, \theta) r \mathrm{d}r = D^2 \int_0^S f\left[D\left(\frac{\xi}{1-\xi}\right)^{1/2}, \theta\right] \frac{\mathrm{d}\xi}{2(1-\xi)^2}, \quad (3.7)$$

where $S = \sin^2 U$. The Gaussian quadrature scheme must then be designed by using a weighting function proportional to $(1 - \xi)^{-2}$. Introducing the normalized variable $v = \xi/S$ yields the new expression for the integral I:

$$I = R^2 \int_0^1 \left\{2 \int_0^\pi f\left(R\left[\frac{(1-S)v}{1-Sv}\right]^{1/2}, \theta\right) \mathrm{d}\theta\right\} \frac{1-S}{2(1-Sv)^2} \mathrm{d}v$$

$$\approx R^2 \sum_{j=1}^{N_r} x_j \left\{2 \frac{\pi}{N_\theta} \sum_{k=1}^{N_\theta} f(R s_j, \theta_k)\right\}. \quad (3.8)$$

The associated weights and sample points x_j and s_j are presented in Tables 3–6 for $S = 0.2, 0.4, 0.6, 0.8$ (i.e., $U \approx 27°, 39°, 51°, 63°$). Notice that, from the form of Eq. (3.8), when S is equal to zero, the parameters are identical to those given in Table 1 (i.e., $x_j = w_j$ and $s_j = \rho_j^{1/2}$). It is interesting that, as the aperture subtends larger angles, the sample points are pulled nearer to the center in the plane of the pupil and the weights are shifted more heavily to the outer points. By using these parameters it is possible to obtain the efficiencies found earlier for the Gaussian scheme for low-numerical-aperture systems. Since these parameters change quite slowly with S, it is expected that the tabulated values should be adequate for most purposes when only a small number of rays are to be traced (e.g., N_r not bigger than 3). From my limited experience in using these parameters, it appears to be best to round down the numerical aperture of a particular system when deciding which of the tables to use. It should also be noted that for cases in which the pupil is at infinity, the values for the parameter s_j given here are to be taken as normalized direction tangents in place of normalized radial positions.

Table 3. Gaussian Integration Parameters for the Radial Integral in Eq. (3.8) with $S = 0.2$ (i.e., $U \approx 27°$)

N_r	j	s_j	x_j
1	1	0.69384445	0.50000000
2	1	0.43543723	0.23070844
	2	0.87510456	0.26929156
3	1	0.31232501	0.12232506
	2	0.67840051	0.22109640
	3	0.93304442	0.15657855
4	1	0.24273576	0.07465129
	2	0.54098777	0.15278777
	3	0.79417534	0.17191291
	4	0.95844762	0.10064803
5	1	0.19831273	0.05006067
	2	0.44698355	0.10783627
	3	0.67414583	0.14117340
	4	0.85757794	0.13124438
	5	0.97175498	0.06968528
6	1	0.16756660	0.03583065
	2	0.37990375	0.07919620
	3	0.58067807	0.11097140
	4	0.75762209	0.12141973
	5	0.89583930	0.10163465
	6	0.97957228	0.05094737

Table 4. Gaussian Integration Parameters for the Radial Integral in Eq. (3.8) with $S = 0.4$ (i.e., $U \approx 39°$)

N_r	j	s_j	x_j
1	1	0.67647261	0.50000000
2	1	0.40432614	0.20615699
	2	0.85723961	0.29384301
3	1	0.28334434	0.10279606
	2	0.64005104	0.21639410
	3	0.92016371	0.18080984
4	1	0.21754825	0.06076048
	2	0.49792986	0.13813168
	3	0.76021048	0.18081687
	4	0.94919652	0.12029096
5	1	0.17644476	0.03998941
	2	0.40520854	0.09278801
	3	0.62998719	0.13682736
	4	0.82948791	0.14523767
	5	0.96490002	0.08515755
6	1	0.14837563	0.02827855
	2	0.34105017	0.06602071
	3	0.53356700	0.10141885
	4	0.71692624	0.12446318
	5	0.87287672	0.11660231
	6	0.97432582	0.06321641

In closing this section, some observations are offered concerning the orders of aberrations that are exactly integrated by the Gaussian schemes considered here. This is of importance since, with the small numbers of rays being used, the quoted accuracy of a given scheme can be seriously violated when the behavior of the system is dominated by aberrations of an order higher than that which can be integrated by the

Table 5. Gaussian Integration Parameters for the Radial Integral in Eq. (3.8) with $S = 0.6$ (i.e., $U \approx 51°$)

N_r	j	s_j	x_j
1	1	0.65168244	0.50000000
2	1	0.36130432	0.17288555
	2	0.83010332	0.32711445
3	1	0.24509622	0.07896118
	2	0.58426612	0.20408175
	3	0.89935733	0.21695707
4	1	0.18520696	0.04473904
	2	0.43851128	0.11612775
	3	0.70803535	0.18758867
	4	0.93360380	0.15154454
5	1	0.14883717	0.02875378
	2	0.34963074	0.07296355
	3	0.56590348	0.12571259
	4	0.78438867	0.16168849
	5	0.95299762	0.11088159
6	1	0.12441841	0.02003190
	2	0.29061007	0.04984692
	3	0.46811830	0.08585576
	4	0.65517275	0.12322550
	5	0.83472366	0.13677914
	6	0.96501817	0.08426077

Table 6. Gaussian Integration Parameters for the Radial Integral in Eq. (3.8) with $S = 0.8$ (i.e., $U \approx 63°$)

N_r	j	s_j	x_j
1	1	0.60937257	0.50000000
2	1	0.29185413	0.12205389
	2	0.77964770	0.37794611
3	1	0.18770381	0.04804410
	2	0.48778717	0.17142335
	3	0.85701314	0.28053255
4	1	0.13853492	0.02554845
	2	0.34408819	0.07987852
	3	0.61061112	0.18218416
	4	0.89966525	0.21238888
5	1	0.10989681	0.01587943
	2	0.26589927	0.04527436
	3	0.45695849	0.09808976
	4	0.69444529	0.17597196
	5	0.92577314	0.16478450
6	1	0.09112174	0.01084000
	2	0.21711832	0.02919158
	3	0.36395062	0.05872956
	4	0.54335308	0.10759254
	5	0.75440686	0.16281059
	6	0.94291541	0.13083573

scheme. In the case of the axial field point (where $N_\theta = 1$ yields exact results for the angular integral), if the spherical aberration curve is assumed to be a polynomial of order $2m + 1$ in r, it follows that $N_r = m + 1$ yields the exact spot size. For example, a 7th-order spherical aberration curve, when squared (to determine spot size), includes terms out to 14th order in r (i.e., 7th order in ρ), and this is exactly integrated by a Gaussian scheme with $N_r = 4$. For off-axis points, a similar result holds. If the aberration function is assumed to be of order $2m + 1$, then choosing both N_r and N_θ to be equal to $m + 1$ yields the exact spot size. So, for example, the scheme represented in Fig. 7 exactly determines the spot size for any aberration function up to fifth order. (Strictly speaking, these results are valid only for systems with the object at infinity, since the aperture weight depends on only the field variable in this case. If weighting proportional to solid angle is used for a system with finite conjugates, these results no longer hold exactly, since the aperture weight function remains a factor in the integrand.)

4. AVERAGING OVER COLOR

The second stage in the global average discussed in Section 2 involves an integral over the spectral region of interest. In this section the efficiencies of schemes for integrating over color are examined. The errors quoted are typical errors at the worst point in the field for $L_{\text{Gaussian}} = \{L_0(\mathbf{f})\}^{1/2}$ and $L_{\text{optimal}} = \{L_0(\mathbf{f}) - L_1(\mathbf{f})^2/L_2(\mathbf{f})\}^{1/2}$. From the ideas presented in Section 3, it seems that this is also an ideal case for the application of Gaussian quadrature. This is not new. However, it is clear that the resulting efficiency hinges on the smoothness of the integrand when expressed in terms of the chosen coordinate, and the most immediate coordinates for specifying color (viz., either wavelength or frequency) are in fact poor candidates. This follows from the fact that the refractive index of glass is roughly hyperbolic in wavelength over the visible spectrum (steep in the blue and flat in the red) and cannot be fitted well by polynomials of low order (say, much less than degree 10). Consequently, it can be expected that the entities to be integrated here will not be smooth when expressed in terms of these coordinates. This is in fact found to be the case. In the context of chromatic aberration theory, a coordinate was devised by Buchdahl[29] expressly to render the dispersion of glass as a low-order polynomial. This coordinate plays a key role in recent applications of chromatic aberration theory,[30] yields efficient forms for interpolating dispersion data,[14] and was used to advantage recently in the analysis of data for the interferometric measurement of refractive index.[31] It is applied here in yet another context.

At the outset, it is necessary to specify a spectral weight. It is found that the efficiencies of the schemes discussed here are effectively independent of the details of the weighting function, and two illustrative examples are considered here. The first is a uniform weight in wavelength from $\lambda_B = 400$ nm to $\lambda_T = 700$ nm, and the second, plotted in Fig. 9, is of the form $\mathcal{L}(\lambda) = \exp[-k(\lambda - \lambda_0)^2]$ over the same interval, where the values of the parameters appearing here are taken to be $k = 1.5e-4$ nm^{-2} and $\lambda_0 = 550$ nm.

Three integration schemes are compared here. The first uses uniform sampling at the midpoints of subintervals (i.e., $\lambda_i = \lambda_B + (i - 1/2)(\lambda_T - \lambda_B)/N_c$, for $i = 1, 2, \ldots, N_c$) and weighting given by the spectral weight evaluated at each point. The other two methods are Gaussian integration schemes; one uses λ as the fundamental coordinate, and the other uses Buchdahl's coordinate, the difference being that the weight function in one case picks up a Jacobian.[20]

The relative efficiencies of these three schemes are presented in Fig. 10. As before, the curves represent typical

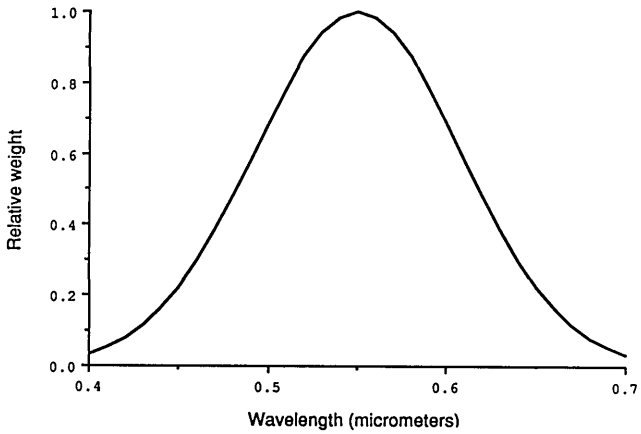

Fig. 9. Relative spectral weight adopted as an example for the color average.

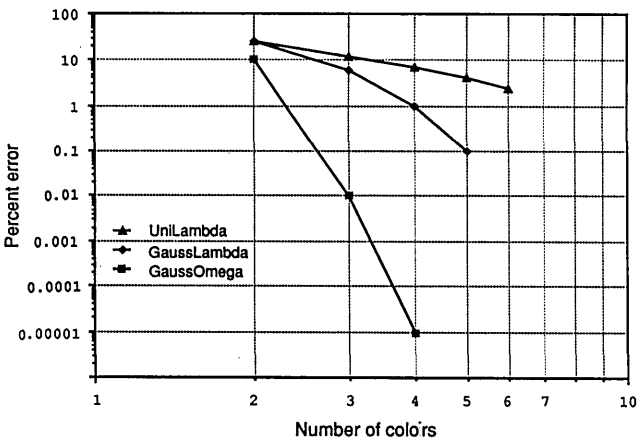

Fig. 10. Comparison of the efficiencies of a number of schemes for averaging over color. The schemes include (▲) uniform sampling in wavelength; (♦) Gaussian quadrature, assuming that the integrand is approximated well by a polynomial in wavelength; and (■) Gaussian quadrature, assuming that the integrand is approximated well by a polynomial in Buchdahl's chromatic coordinate.

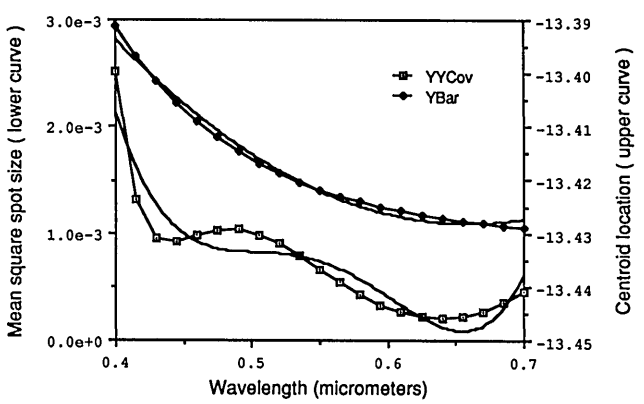

Fig. 11. Chromatic dependence of the spot size and the centroid location for the double gauss halfway out in the field. (The unit of length is chosen to be 1/100 of the focal length.) The plane curves are the quadratic and the quartic of the best fits to the upper and lower curves, respectively, with wavelength used as the variable. Note that these curves can be fitted so closely by a quadratic and a quartic, respectively, in Buchdahl's chromatic coordinate that the results are indistinguishable from the originals on this scale.

Table 7. Gaussian Integration Parameters for Integration over the Visible Spectrum with Uniform Weight in Wavelength and with the Gaussian Weight Plotted in Fig. 9[a]

		Value of Integration Parameter (μm)			
		Uniform Weight		Gaussian Weight	
N_r	j	λ_j	w_j	λ_j	w_j
1	1	0.528944	0.300000	0.541368	0.1433636
2	1	0.440626	0.106793	0.471604	0.0443355
	2	0.608925	0.193208	0.584579	0.0990281
3	1	0.418886	0.049456	0.434658	0.0099821
	2	0.505546	0.123853	0.518983	0.0782096
	3	0.644536	0.126692	0.614795	0.0551719
4	1	0.410836	0.028202	0.418309	0.0032488
	2	0.460136	0.071776	0.477524	0.0316589
	3	0.554033	0.112203	0.552481	0.0795135
	4	0.662979	0.087820	0.637241	0.0289423

[a] It is emphasised that these parameters were determined by using Buchdahl's chromatic coordinate as the underlying variable, which entails that the weight function pick up a Jacobian.

maximum errors. (As before, the errors may occasionally exceed the plotted values by a factor of 2 or so.) For one color, the error is consistently beyond 30% or so (and sometimes as much as 100%) regardless of the scheme used. The errors of the two-color methods often exceed 10% (except in the case of the hypergon, which has no color correction and for which the better Gaussian scheme gets an accuracy of about 0.01%). In view of the form of plots such as that given in Fig. 11, in which the chromatic dependence of the spot size of the double gauss is presented, it is clear that the integration schemes must at least be able to handle quartics. The Gaussian schemes with two points can exactly integrate only cubics. With three sample points, the accuracy of the Gaussian scheme based on Buchdahl's coordinate is orders of magnitude better than the other two methods. The associated sample points and weights for both spectral weighting functions mentioned above are presented in Table 7. These parameters were calculated by using the general-purpose chromatic coordinate presented in Ref. 14 and are also of value for checking purposes.

In practice, lens design is often done by the so-called $D-d$ method, which yields the first derivative of the OPD with respect to color. This is in sharp contrast to the recommendation that all rays be traced in three colors, but it is clear that such a procedure is necessary to determine the overall properties to more than approximately one decimal place. Since the control of the chromatic properties of a system is limited to the selection among discrete glasses, it may be that a simpler scheme is all that is required, but this is not at all obvious and would require further investigation.

5. AVERAGING OVER THE FIELD

The efficiency of averaging over the field variable to generate the final figure of merit in the form presented in Eq. (2.10) is reported in this section. In accordance with the preceding sections, the errors quoted are relative errors in $\mathcal{M}_{\text{Gaussian}}$ and $\mathcal{M}_{\text{optimal}}$. Since there is symmetry between the field and aperture variables, the general form of the integra-

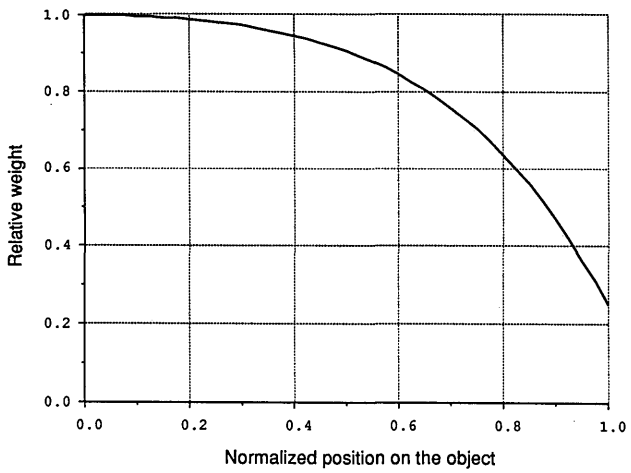

Fig. 12. Weight function, adopted as an example for averaging the mean-square spot size over the field. The function is of the form $W(h) = 1.0 - a_1h^2 - a_2h^4$, where h is the normalized object position and the coefficients a_1 and a_2 are taken to have the values 0.25 and 0.5, respectively.

Table 8. Gaussian Integration Parameters for Averaging over the Field by Using the Weight Function Presented in Fig. 12

N_r	j	s_j	x_j
1	1	0.64168895	0.35416667
2	1	0.43031871	0.20629209
	2	0.85315256	0.14787458
3	1	0.31999532	0.12240883
	2	0.67815443	0.16406851
	3	0.92442742	0.06768933
4	1	0.25381393	0.07923447
	2	0.55436991	0.13417354
	3	0.79663522	0.10415890
	4	0.95517022	0.03659975
5	1	0.21003785	0.05503557
	2	0.46615087	0.10422657
	3	0.68857948	0.10580761
	4	0.86125653	0.06664974
	5	0.97070391	0.02244717
6	1	0.17903360	0.04031148
	2	0.40109978	0.08121040
	3	0.60226192	0.09485103
	4	0.77138936	0.07809421
	5	0.89995791	0.04464430
	6	0.97948621	0.01505525

tion methods developed in Section 3 is also valid here. For symmetric systems, the spot size is independent of rotation about the axis, so only the radial component of the integral is needed. Just as with the color average, there is a system-dependent weighting function that enters into the integrals. It is found, however, that the errors are reasonably insensitive to the form of this weighting function.

If each patch of the field of view has an equal weight, the parameters presented in Section 3 can also be used here. For many applications, however, the center of the field is of greater importance, and new Gaussian parameters must be calculated that incorporate the appropriate weight function. Again, when the field angle becomes large, it is better to use schemes such as those devised for systems with large numerical apertures, although for the cases considered here the difference is not so significant as in the aperture average. The parameter S of Section 3 must now be taken to represent the square of the sine of the half-angle subtended by the object at the center of the pupil. For the sample weight function presented in Fig. 12, the associated Gaussian and Radau quadrature parameters (for $S = 0$) are presented in

Table 9. Radau Integration Parameters for Averaging over the Field by Using the Weight Function Presented in Fig. 12

N_r	j	u_j	v_j
1	0	0.00000000	0.10518293
	1	0.76531973	0.24898374
2	0	0.00000000	0.04958584
	1	0.56484166	0.20324952
	2	0.89399859	0.10133130
3	0	0.00000000	0.02870831
	1	0.44198632	0.14150922
	2	0.74279304	0.13411331
	3	0.94168944	0.04983582
4	0	0.00000000	0.01868944
	1	0.36144719	0.10008973
	2	0.62563117	0.12268764
	3	0.83161892	0.08406164
	4	0.96374493	0.02863823
5	0	0.00000000	0.01312624
	1	0.30516908	0.07346590
	2	0.53709792	0.10245377
	3	0.73253887	0.09202967
	4	0.88203922	0.05472700
	5	0.97548572	0.01836408

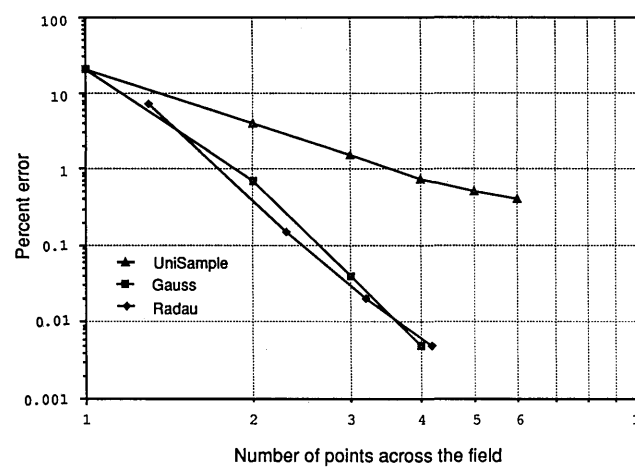

Fig. 13. Comparison of the better methods for averaging over the field. The first method involves uniform sampling in the object position (or direction tangents) squared and corresponds to the trapezoidal scheme used for the aperture average (although here the samples are taken to be the midpoints of the subintervals). The other two methods involve Gaussian and Radau quadratures of the same form as the most-efficient methods developed for the aperture average.

Tables 8 and 9. As before, if the object is at infinity the sample points are to be interpreted as normalized direction tangents. Typical results for the better schemes are presented in Fig. 13. These curves are not valid for the microscope objective, which is dominated by low-order field aberrations, and the corresponding errors are significantly lower. Notice that, for an accuracy of $\simeq 1\%$, only two field points are typically required with a Gaussian scheme. For the Radau schemes, the axial point is always included, and, since $N_\theta = 1$ is all that is needed there, this point is counted as about a third to a fifth of an off-axis sample point (depending on the accuracy).

6. GLOBAL AVERAGING

After the results are collected for the variety of integration schemes considered in Sections 3–5, it is straightforward to generate comparisons of global averaging methods. For example, a number of representative methods are compared in Fig. 14. Note that, even though the most elementary (and inefficient) methods that use uniform sampling in field and wavelength are not included in this plot, the improvement from best to worst in those given is typically measured in orders of magnitude. A number of the most-efficient ray configurations deserve individual mention. To obtain about 10% accuracy overall, only 18 rays are needed: two field points (selected according to the Radau scheme so that the axial field point is included) are used; for the axial point, three rays are traced in the aperture (by using either the Radau or the Gaussian scheme with $N_r = 3$); and for the off-axis point, six rays are traced (by using the Gaussian scheme with $N_r = 2$ and $N_\theta = 3$). Each of these nine rays is traced at two colors (by using the Gaussian scheme based on Buchdahl's coordinate), and the resulting error is typically $\simeq 10\%$. A general ray set that yields $\simeq 1\%$ accuracy can be made up of 54 rays as follows: again, two field points are needed (but this time not including the axial point); at each of these points, 9 rays are traced through the pupil ($N_r = 3$ and $N_\theta = 3$), and these 18 rays are each traced at three colors. For even higher accuracies, using 111 rays yields about 0.1% with a Radau scheme for the field average (using two off-axis points), where $N_r = N_\theta = 4$ for the two off-axis points, $N_r = 5$ for the axial point and, as before, these 37 rays are traced at three colors.

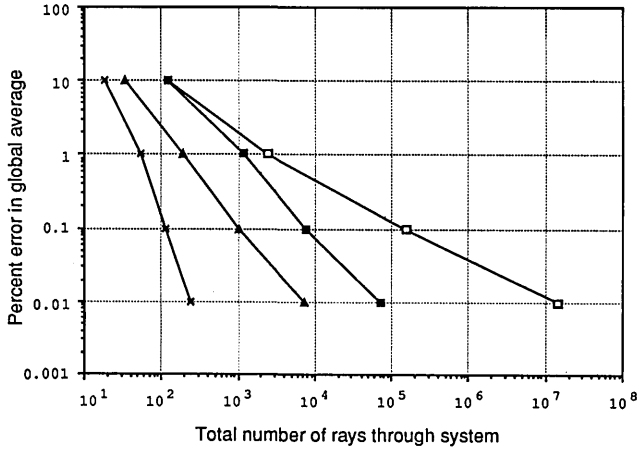

Fig. 14. Efficiencies of a number of global averaging schemes (determined from the results of Sections 3–5) are compared in this plot. Curves: □, Cartesian sampling in the pupil, a Gaussian method based on wavelength for the color average and the trapezoidal rule based on the square of the field variable for the field average; ■, a similar method, except for the aperture average, for which Andersen's method is used (these two curves serve as starting points by which any alternative scheme can be measured); ▲, a trapezoidal scheme based on the square of the aperture variable and Gaussian methods in the color and field averaging based on Buchdahl's chromatic coordinate and the square of the field variable, respectively (this curve is expected to give some indication of the limits of numerical ray tracing when vignetting plays a significant role and is to be accounted for directly); ×, the most-effective Gaussian schemes for each average.

The three general-purpose ray sets specified above can, of course, be fine tuned and/or tailored to particular systems. For example, the hypergon never needs more than two colors, and the microscope objective requires only one off-axis field point. It is emphasized that the quoted accuracies for these ray sets were determined by analyzing only those systems mentioned in Section 1. It follows from the remarks at the end of Section 3 that the mean spot size reported by the first two ray sets will be unreliable for systems in which the dominant aberrations are of seventh order or higher. Also notice that once the weighting functions for the field and the spectrum are specified for a given system, it is necessary to generate the specific Gaussian sample points and weights for that particular problem. For the aperture average, on the other hand, the tables of parameters presented in Section 3 should be adequate for most purposes. It is expected that these methods are also efficient for the calculation of rms wave-front errors.

7. CONCLUDING REMARKS

In much of the lens-design community, system assessment is invariably carried out by tracing small ray sets of sizes comparable with those specified in Section 6 (except for the sampling in color, as mentioned in Section 4). However, the configurations of the rays are often chosen by the designer specifically to control the aberrations at given positions or zones and not to represent the overall performance of the system. It is often found that optimization may significantly reduce the value of this sort of merit function and yet, in doing so, ruin the overall system performance, a phenomenon sometimes referred to as the toothpaste-tube effect (squeeze it here and it simply pops up somewhere else). This phenomenon forces the designer to intervene at regular intervals during optimization to request a thorough system analysis and determine whether the merit function need be redefined. In fact, unless the accuracy with which the merit function represents the overall performance is known, it is impossible to specify how far the optimization should proceed. The question about whether an $X\%$ change in a particular figure of merit is significant (i.e., indicates an improvement in overall performance) is usually unanswerable; optimization is often pushed far into regions of insignificance.

The ray sets introduced here are chosen specifically to give a measure of the overall performance. It can therefore be expected that these ray sets will be significantly less susceptible to the toothpaste-tube effect just mentioned and will require less intervention by the designer.[32] In fact, the remarkably rapid convergence of the merit functions based

on Gaussian quadrature permits automatic monitoring of reliability to be performed efficiently by generating a prescription to start the optimization with a simple ray set and, for each dimension, occasionally to use an extra sample point to determine when it is appropriate to change to a ray set of higher accuracy. (In the case of global optimization, it would also be appropriate to allow for reduction in the accuracy of the ray set.) Such a procedure would trace the minimum number of rays required at each stage, and this efficiency is crucial, especially for global optimization.

ACKNOWLEDGMENTS

I would like to thank John Rogers for countless valuable discussions. From those encounters I was made aware of practical realities and ideas of which I had been completely ignorant, and I learned alternative ways of approaching many optical problems. I am also grateful for the time that he invested in designing the microscope objective used as an example in this paper. I also thank Doug Sinclair for a constructive comment concerning a serious weakness of an earlier version of this paper. This research was funded partly by the University Research Initiative program of the U.S. Army Research Office and represents one piece of a body of work that would not have been possible without the efforts of Duncan Moore, the director of the Institute of Optics, University of Rochester, who succeeded in arranging a semester in which I was free to concentrate on my research.

REFERENCES AND NOTES

1. V. K. Viswanathan, I. O. Bohachevsky, and T. P. Cotter, "An attempt to develop an 'intelligent' lens design program," in *1985 International Lens Design Conference*, W. H. Taylor, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **554,** 10–17 (1985).
2. I. O. Bohachevsky, V. K. Viswanathan, and G. Woodfin, "An 'intelligent' optical design program," in *Applications of Artificial Intelligence I*, J. F. Gilmore, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **485,** 104–112 (1984).
3. P. N. Robb, "Lens design using optical aberration coefficients," in *1980 International Lens Design Conference*, R. E. Fischer, ed., Proc. Soc. Photo-Opt Instrum. Eng. **237,** 109–118 (1980).
4. T. B. Andersen, "Automatic computation of optical aberration coefficients," Appl. Opt. **19,** 3800–3816 (1980).
5. G. W. Forbes, "Weighted truncation of power series and the computation of chromatic aberration coefficients," J. Opt. Soc. Am. A **1,** 350–355 (1984).
6. G. W. Forbes, "Extension of the convergence of multivariate aberration series," J. Opt. Soc. Am. A **3,** 1376–1383 (1986).
7. G. W. Forbes, "Acceleration of the convergence of multivariate aberration series," J. Opt. Soc. Am. A **3,** 1384–1394 (1986).
8. See, for example, M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980), Sec. 9.1.3.
9. This result follows on differentiation of the autocorrelation expression for the MTF. (It is interesting that the same result follows trivially for the so-called geometrical MTF from the standard relations between the moments of a function and the Taylor series of its Fourier transform.) It follows then that a system designed for minimum spot size will generally have better low-spatial-frequency response than a system designed for minimum OPD, which will have superior response to higher spatial frequencies.
10. In the case of Gaussian quadrature, which is generally well suited to this type of problem, it is not so straightforward to design a successful method that uses the derivative, since Gaussian schemes that use both the function value and its first derivative at the sample points typically require samples out in the complex plane (even when the region of integration is a finite interval on the real axis). This is a serious drawback in this context, in which tracing rays with complex-valued field, aperture, and color variables is relatively costly. This information can be used in integration schemes of lower order, however, as shown, for example, in S. A. Comastri and J. M. Simon, "Aberration function dependence on field—a way to better obtain profit from ray tracing," Optik **69,** 135–140 (1985). The methods of polynomial fitting discussed by Comastri and Simon seem well suited to the calculation of MTF's; however, it is not clear that such an approach is optimal for the determination of the variance of the OPD; Gaussian quadrature methods reveal that, if it is the integral that is required, it is possible to do better than fitting polynomials. For example, with two sample points, it is possible to integrate exactly any cubic over a finite interval, but only a linear function can be fitted with this data.
11. J. W. Foreman, Jr., "Computation of rms spot radii by ray tracing," Appl. Opt. **13,** 2585–2588 (1974).
12. T. B. Andersen, "Evaluating rms spot radii by ray tracing," Appl. Opt. **21,** 1241–1248 (1982).
13. Some particular Cartesian configurations can have somewhat better performance than those quoted; however, the numbers of rays specified here are sufficient to guarantee that the error limits will typically not be violated. Although the fractional error of the Cartesian scheme dies, on average, as the inverse of the number of rays to the power 3/4, the performance is highly erratic (as one might expect with patching a square grid to a round hole), and it is difficult to guess the accuracy of a given configuration by relating it to another. For example, halving the side length of the grid (which approximately quadruples the number of rays) typically reduces the accuracy of the result for some configurations. This is discussed in detail in Section 3.
14. G. W. Forbes, "Chromatic coordinates in aberration theory," J. Opt. Soc. Am. A **1,** 344–349 (1984).
15. The hypergon has a half-field angle of 65° and operates at $f/30$. The specifications can be found in U.S. patent 706,650 (August 12, 1902).
16. The Cooke triplet used here has a half-field angle of 20° and operates at $f/5.6$. The specifications were taken from H. A. Buchdahl, *Optical Aberration Coefficients* (Dover, New York, 1968), Sec. 37, p. 60.
17. The double Gauss used here is among the sample lens specifications provided with ACCOS V (the lens-design program available from Scientific Calculations, Fishers, N.Y.). It has an f number of $\simeq 2$ and a half-field angle of 15°.
18. The specifications of the Schmidt camera used here can be found in Table 3 of Ref. 4. It is designed to operate in the UV, has a half-field angle of 5°, and operates at $f/1.09$.
19. The microscope objective used here was designed by J. R. Rogers of the Institute of Optics, University of Rochester. The half-cone angle at the object is 50° (numerical aperture ≈ 0.766), and the magnification is 50×. The specifications are available from him on request.
20. Note that the form of the weighting functions is dependent on the choice of variables. For example, the weight for color is different if frequency is used as the coordinate in place of wavelength. If a change of variables is needed, a Jacobian must be included to find the new form of the weighting function. So, for example, in changing variables from wavelength to frequency, $\mathcal{L}(\lambda)$ is replaced by $\mathcal{N}(\nu) = \mathcal{L}[\lambda(\nu)]\, d\lambda/d\nu$. It is also significant that, in Eq. (2.2), a mean-square length is averaged so that if the spot size in a region near the center of the field should be weighted k times more heavily than a region near the edge, $\mathcal{F}(f)$ should be an extra factor of k^2 higher at the center (over and above the Cartesian components of any Jacobian that is picked up by using variables other than the position vector).
21. It is worth remarking, at this stage, that expressions of the form $s^2 = \mathrm{Avg}_x\{[f(x) - \bar{f}]^2\}$, which appear in Eqs. (2.6) and (2.7), are evaluated more easily if they are reformulated as follows. It is usual to expand the argument of the average operator and to reexpress s^2 as $s^2 = \mathrm{Avg}_x\{f^2(x)\} - \bar{f}^2$ in order to allow $\bar{f}$ and s^2 to be calculated simultaneously. However, this typically increases the numerical noise owing to cancellation. This is especially the case for the computations indicated in Eqs. (2.6) and (2.7) in which s^2 may be 8 or more orders of magnitude smaller than $\bar{f}^2$. A more convenient expression can be obtained by first writing $\{f(x) - \bar{f}\}$ as $\{[f(x) - f(c)] - [\bar{f} - f(c)]\}$ before expanding the argument, and in this way $s^2 = \mathrm{Avg}_x\{[f(x) - f(c)]^2\} - [\bar{f} -$

$f(c)]^2$ is obtained, where c is taken to be some fixed value near the center of the region of interest.

22. Sampling on a square grid is used, for example, by CODE V (the lens-design program available from Optical Research Associates, Pasadena, Calif.), and the polar grid can be found in SIGMA (the lens-design program from Kidger Optics, UK). These programs do not necessarily use the same weighting adopted here or try to calculate the same entities; I simply take these sampling schemes as a starting points for comparison purposes.
23. If the points on a square lattice (of side length $\delta = R/n$, where R is the radius of the disk) in one quadrant of the disk are located at

$$\mathbf{r}_{ij} = [\delta(i - 1/2), \delta(j - 1/2)]$$
$$\text{for } i = 1, 2, \ldots \langle (n^2 - 1/4)^{1/2} + 1/2 \rangle$$
$$\text{and } j = 1, 2, \ldots \langle [n^2 - (i - 1/2)^2]^{1/2} + 1/2 \rangle,$$

it can be seen that the percentage error in approximating the integral of a constant over the disk by using a simple sum over these points is given by

$$E(n) = 100\{S(n) - \pi R^2\}/\pi R^2,$$

where

$$S(n) = 4\delta^2 \sum_i \langle [n^2 - (i - 1/2)^2]^{1/2} + 1/2 \rangle.$$

$\langle x \rangle$ denotes the integral part of x, and the range for the sum over i is, of course, just that indicated for the placement of points. It is remarked that scaling the overall result by a constant to ensure that all schemes integrate constant functions exactly will have no effect for the work reported in this paper, since all the integrals are for the purposes of averaging and any constant multiplying factor will cancel when the integrals are normalized to obtain the desired average. This particular sampling scheme is identical to that used in CODE V, and some observations are in order. This program has a parameter referred to as DEL, which is just the inverse of n. The default value of this parameter is DEL = 0.385, which corresponds to $n = 2.60$, which, from Fig. 3, can be seen to use 12 rays and to have an error in excess of +10% when integrating a constant (if uniform weighting is used). When the routines of Section 3 are used, this scheme is found to overestimate spot sizes consistently by 20–50%. When a value of $n = 2.764$ (corresponding to DEL = 0.362, the location of the zero on the plot in Fig. 3 with the same number of rays), is adopted, the error in the determination of spot size is reduced to $\simeq$10%, an improvement by a factor of 3 to 5. This improvement is appreciated when it is recalled that the error is, on average, dropping as the number of rays to the power −3/4, so the gain realized by this minor change is equivalent to that typically obtained by increasing the number of rays by a factor of 4 to 8. For the interest of CODE V users, it is noted that the locations of a number of the zeros of the curve in Fig. 3 that seem to give relatively good integration schemes are found to be $n = 2.2568$ (8 rays, 10–30% error), $n = 2.7639$ (12 rays, 4–12% error), $n = 3.7424$ (22 rays, 3–6% error), and $n = 5.9708$ (56 rays, 1–3% error).
24. The polar scheme would probably benefit from sampling at the endpoints of the radial subdivisions rather than at the midpoints. However, there are significantly better schemes available, so this minor issue is not pursued further.
25. With m uniformly distributed samples in $[0, 2\pi)$, say, at $\theta_k = k2\pi/m - \gamma$ for $k = 1, 2, \ldots, m$, it can easily be shown, by using a geometric series, that $\sum_k \cos q\theta_k$ vanishes for q not equal to a multiple of m. Now, since $\cos^p \theta$ can be written as a linear combination of $\{\cos q\theta; q = p, p - 2, p - 4, \ldots, 0 \text{ or } 1\}$, it follows that uniform weighting with m uniformly distributed points can be used to integrate exactly $\cos^p \theta$ for $p = 0, 1, 2, \ldots, m - 1$. The use of alternating weights can be regarded as a superposition of a scheme with $2n$ points, and one with only half that number (skipping every other one) and can be used to integrate exactly $\cos^p \theta$ for $p = 1, 2, \ldots, n - 1$, where uniform weights yield exact results for $p = 1, 2, \ldots, 2n - 1$.
26. This result also holds for the sample points $\theta_j = j\pi/N_\theta$, $j = 0, 1, \ldots, N_\theta$, which includes points on the line of symmetry (i.e., meridional rays), although the same accuracy is now obtained with one extra point on each ring.
27. For interest, it is remarked that since the Legendre polynomials are simply related to the rotationally symmetric Zernike polynomials, it can be seen that the sample points presented in Table 1 in fact correspond to the radial locations of the zeros of these Zernike polynomials. The values of the weights follow simply from the relation between Gaussian quadrature and orthogonal polynomials, which is presented in H. S. Wilf, *Mathematics for the Physical Sciences* (Dover, New York, 1962), Sec. 2.9, pp. 61–64. A convenient form of the recurrence relations for generating the orthogonal polynomials from which the parameters for Gaussian quadrature can be found is presented in W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge U. Press, Cambridge, 1986), Sec. 4.5.
28. For a simple description of the derivation of the Radau integration methods, see, for example, R. W. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd ed. (McGraw-Hill, New York, 1973), Sec. 19.7, pp. 328–330.
29. H. A. Buchdahl, *Optical Aberration Coefficients* (Dover, New York, 1968) Sec. 87, pp. 150–154.
30. See, for example, P. N. Robb, "Selection of optical glasses. I: Two materials," Appl. Opt. **24**, 1864–1877 (1985).
31. J. R. Rogers and M. D. Hopler, "Conversion of group refractive index to phase refractive index," J. Opt. Soc. Am. A **5**, 1595–1600 (1988).
32. This somewhat vague expectation can be given a more definite meaning and verified as follows. The monic polynomial of degree m that has the smallest mean-square value over a given interval can be shown to be proportional to the mth-order orthogonal polynomial over that interval, say, ϕ_m. A Gaussian quadrature scheme with m sample points has roots at the locations of the zeros of this very polynomial. Such an integration scheme is clearly unable to determine the mean-square value of ϕ_m: it exactly integrates polynomials of degree less than or equal to $2m - 1$, whereas the square of ϕ_m is of degree $2m$. Nevertheless, if used as a merit function, the Gaussian integration scheme reports a mean-square value of zero for an mth-order polynomial if and only if the polynomial is a multiple of ϕ_m. This means that, in the sense of minimizing the mean-square value, the mth-order term is optimally balanced by the terms of lower order. In practice, this entails that when the Gaussian merit functions discussed in this paper seriously underestimate, say, a mean-square wave aberration owing to the dominance of aberrations of higher order than those that the scheme can account for, the balancing of the unseen terms will not be far from optimal. In this sense, the toothpaste tube is being squeezed at just those points that guarantee that the smallest possible volume is left inside in the event that the thickness at each point is reduced to zero.