2

ABERRATION CURVES IN LENS DESIGN

Donald C. O'Shea

Georgia Institute of Technology
School of Physics
Atlanta, Georgia

Michael E. Harrigan

Harrigan Optical Design
Victor, New York

2.1 GLOSSARY

H	ray height
NA	numerical aperture
OPD	optical path difference
P	petzval
S	sagittal
T	tangential
$\tan U$	slope

2.2 INTRODUCTION

Many optical designers use aberration curves to summarize the state of correction of an optical system, primarily because these curves give a designer important details about the relative contributions of individual aberrations to lens performance. Because a certain design technique may affect only one particular aberration type, these curves are more helpful to the lens designer than a single-value merit function. When a design is finished, the aberration curves serve as a summary of the lens performance and a record for future efforts. For applications such as photography, they are most useful because they provide a quick estimate of the effective blur circle diameter.

The aberration curves can be divided into two types: those that are expressed in terms of ray errors and those in terms of the optical path difference (OPD). OPD plots are usually plotted against the relative ray height in the entrance pupil. Ray errors can be displayed in a number of ways. Either the transverse or longitudinal error of a particular ray relative to the chief ray can be plotted as a function of the ray height in the entrance pupil. Depending upon the amount and type of aberration present, it is sometimes more appropriate to plot the longitudinal aberration as a function of field angle.

For example, astigmatism or field curvature is more easily estimated from field plots, described below. Frequently, the curves are also plotted for several wavelengths to characterize chromatic performance. Because ray error plots are the most commonly used format, this entry will concentrate on them.

2.3 TRANSVERSE RAY PLOTS

These curves can take several different forms, depending on the particular application of the optical system. The most common form is the transverse ray aberration curve. It is also called lateral aberration, or ray intercept curve (also referred to by the misleading term "rim ray plots"). These plots are generated by tracing fans of rays from a specific object point for finite object distances (or a specific field angle for an object at infinity) to a linear array of points across the entrance pupil of the lens. The curves are plots of the ray error at an evaluation plane measured from the chief ray as a function of the relative ray height in the entrance pupil (Fig. 1). For afocal systems, one generally plots angular aberrations, the differences between the tangents of exiting rays and their chief ray in image space.

If the evaluation plane is in the image of a perfect image, there would be no ray error and the curve would be a straight line coincident with the abscissa of the plot. If the curve were plotted for a different evaluation plane parallel to the image plane, the curve would remain a straight line but it would be rotated about the origin. Usually the aberration is plotted along the vertical axis, although some designers plot it along the horizontal axis.

The curves in Fig. 1 indicate a lens with substantial underconnected spherical aberration as evidenced by the characteristic S-shaped curve. Since a change of the evaluation plane serves only to rotate the curve about the origin, a quick estimate of the aberrations of a lens can be made by reading the scale of the ray error axis (y axis) and mentally rotating the plot. For example, the blur spot can be estimated from the extent of a band that would enclose the curve a in Fig. 1, but a similar estimate could be made from the curves b or c, also.

The simplest form of chromatic aberration is axial color. It is shown in Fig. 2 in the presence of spherical aberration. Axial color is the variation of paraxial focus with wavelength and is seen as a difference in slope of the aberration curves at the origin as a function of wavelength. If the slopes of the curves at the origin for the end wavelengths are different, primary axial color is present. If primary axial color is corrected, then the curves for the end wavelengths will have the same slope at the origin. But if that slope differs from the slope of the curve for the center wavelength, then secondary axial color is present.

A more complex chromatic aberration occurs when the aberrations themselves vary with wavelength. Spherochromatism, the change of spherical aberration with wavelength, manifests itself as a difference in the shapes of the curves for different colors. Another curve that provides a measure of lateral color, an off-axis chromatic aberration, is described below.

For a point on the axis of the optical system, all ray fans lie in the meridional plane and only one plot is needed to evaluate the system. For off-axis object points, a second plot is added to evaluate a fan of skew rays traced in a sagittal plane. Because a skew ray fan is symmetrical across the meridional

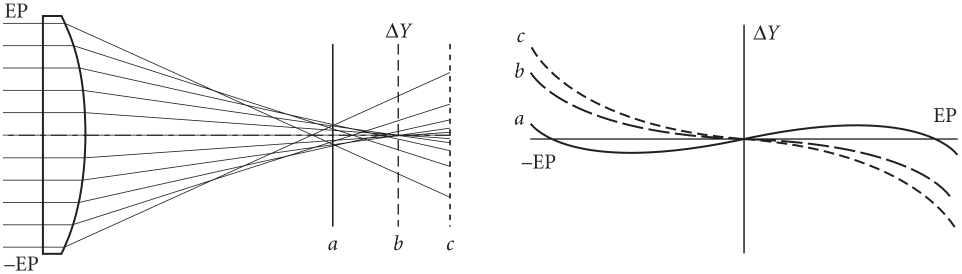

FIGURE 1 (*Left*) Rays exiting a lens are intercepted at three evaluation planes. (*Right*) Ray intercept curves plotted for the evaluation planes: (a) at the point of minimum ray error (circle of least confusion); (b) at the paraxial image plane; and (c) outside the paraxial image plane. (See also color insert.)

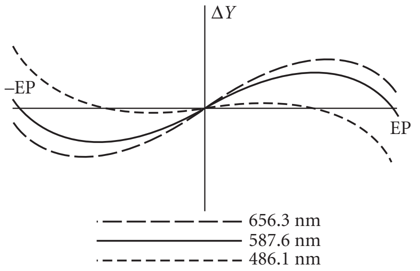

FIGURE 2 Meridional ray intercept curves of a lens with spherical aberration plotted for three colors. (See also color insert.)

plane, only one side of the curve is usually plotted. For all curves the plots are departures from the chief ray location in the evaluation plane (Fig. 3). (In the case of the on-axis point, the chief ray is coincident with the optical axis.) For systems of small-field coverage only two or three object points need to be analyzed, but for wide-angle systems, four or more field points may be necessary.

What can be determined most easily from a comparison between the meridional and sagittal fans is the amount of astigmatism in the image for that field point. When astigmatism is present, the image planes for the tangential and sagittal fans are located at different distances along the chief ray. This is manifested in the ray intercept curves by different slopes at the origin for the tangential and sagittal curves. In Fig. 3 the slopes at the origins of the two curves are different at both 70 percent and full field, indicating astigmatism at both field points. The fact that the difference in the slopes of these two curves has changed sign between the two field points indicates that at some field angle between 70 percent and full field, the slopes are equal and there is no astigmatism there. In addition, the variation of slopes for each curve as a function of field angle is evidence of field curvature.

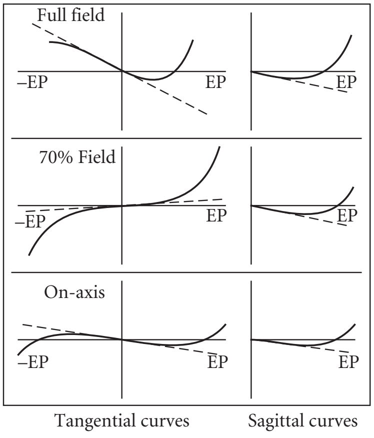

FIGURE 3 Evaluation of a lens on-axis and at two off-axis points. The reduction of the length of the curve with higher field indicates that the lens is vignetting these angles. The differences in slopes (dashed lines) at the origin between the meridional and skew curves indicate that the lens has astigmatism at these field angles. The variation in the slopes with field indicates the presence of field curvature. (See also color insert.)

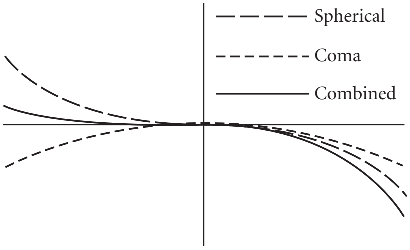

FIGURE 4 Ray intercept curve showing coma combined with spherical aberration. (See also color insert.)

The off-axis aberration of pure primary coma would be evident on these plots as a U-shaped curves for the meridional fan and sagittal fans, the tangential curve being three times larger than the sagittal curve. The "U" will be either upright or upside down depending on the sign of the coma. In almost all cases coma is combined with spherical to produce an S-shaped curve that elongates one of the arms of the "S" and shortens the other (Fig. 4).

The amount of vignetting can be determined from the ray intercept curves also. When it is present, the meridional curves get progressively shorter as the field angle is increased (Fig. 3), since rays at the edges of the entrance pupil are not transmitted. Taken from another perspective, ray intercept curves can also provide the designer with an estimate of how far a system must be stopped down to provide a required degree of correction.

2.4 FIELD PLOTS

The ray intercept curves provide evaluation for a limited number of object points—usually a point on the optical axis and several field points. The field plots present information on certain aberrations across the entire field. In these plots, the independent variable is usually the field angle and is plotted vertically and the aberration is plotted horizontally. The three field plots most often used are: distortion, field curvature, and lateral color. The first of these shows percentage distortion as a function of field angle (Fig. 5).

The second type of plot, field curvature, displays the tangential and sagittal foci as a function of object point or field angle (Fig. 6*a*). In some plots the Petzval surface, the surface to which the image would

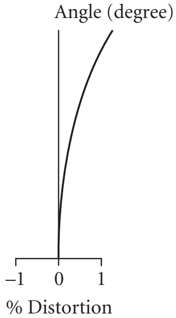

FIGURE 5 Field curve: distortion plot. The percentage distortion is plotted as a function of field angle. Note that the axis of the dependent variable is the horizontal axis. (See also color insert.)

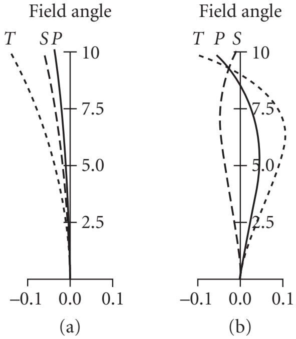

FIGURE 6 Field curve: field curvature plot. The locations of the tangential *T* and sagittal *S* foci are plotted for a full range of field angles. The Petzval surface *P* is also plotted. The tangential surface is always three times farther from the Petzval surface than from the sagittal surface: (*a*) an uncorrected system and (*b*) a corrected system. (See also color insert.)

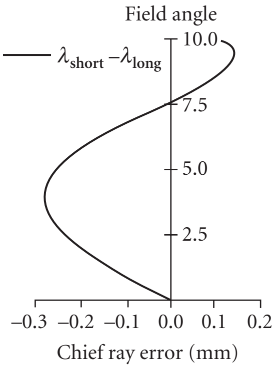

FIGURE 7 Field curve: lateral color plot. A plot of the transverse ray error between red and blue chief ray heights in the image plane for a full range of field angles. Here the distance along the horizontal axis is the color error in the image plane. (See also color insert.)

collapse if there were no astigmatism, is also plotted. This plot shows the amount of curvature in the image plane and amount of astigmatism over the entire field. In cases of corrected field curvature (Fig. 6*b*), this plot provides an estimate of the residual astigmatism between the axis and the corrected zone and an estimate of the maximum field angle at which the image possesses reasonable correction.

The last of the field curves provides information on color error as a function of field angle (Fig. 7). Lateral color, the variation of magnification with wavelength, is plotted as the difference between the chief ray heights at the red and blue wavelengths as a function of field angle. This provides the designer with an estimate of the amount of color separation in the image at various points in the field. In the transverse ray error curves, lateral color is seen as a vertical displacement of the end wavelength curves from the central wavelength curve at the origin.

Although there are other plots that can describe aberrations of optical systems (e.g., plot of longitudinal error as a function of entrance pupil height), the ones described here represent the ensemble that is used in most ray evaluation presentations.

2.5 ADDITIONAL CONSIDERATIONS

In many ray intercept curves the independent variable is the relative entrance pupil coordinate of the ray. However, for systems with high NA or large field of view, where the principal surface cannot be approximated by a plane, it is better to plot the difference between the tangent of the convergence angle of the chosen ray and the tangent of the convergence angle of the chief ray. This is because the curve for a corrected image will remain a straight line in any evaluation plane.[1] When plotted this way, the curves are called *H*-tan *U* curves.

Shifting the stop of an optical system has no effect on the on-axis curves. However, it causes the origin of the meridional curves of off-axis points to be shifted along the curve. In Fig. 8, the off-axis meridional curves are plotted for three stop positions of a double Gauss lens. The center curve (Fig. 8*b*) is plotted for a symmetrically located stop; the outer curves are plots when the stop is located at lens surfaces before and after the central stop.

It is usually sufficient to make a plot of the aberrations in the meridional and sagittal sections of the beam. The meridional section, defined for an optical system with rotational symmetry, is any plane containing the optical axis. It is sometimes called the tangential section. The sagittal section is a plane perpendicular to the meridional plane containing the chief ray. There are some forms of higher-order

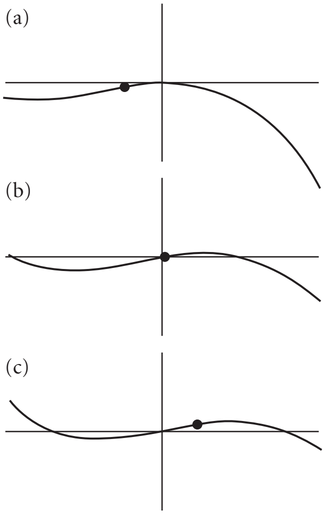

FIGURE 8 The effect of stop shifting on the meridional ray intercept curves of a double Gauss lens. (*a*) Stop located in front of the normal centrally located stop. (*b*) Stop at the normal stop position. (*c*) Stop behind the normal stop position. The dot locates the point on the curve where the origin is located for case (*b*). (See also color insert.)

coma that do not show in these sections.[2] In those cases where this aberration is suspected to be a problem, it may be helpful to look at a spot diagram generated from rays in all sections of the bundle.

For a rotationally symmetric system, only objects in a meridional plane need to be analyzed. Also for such systems, only meridional ray errors are possible for purely meridional rays. To observe certain coma types, it is a good idea to plot both the meridional and sagittal ray errors for sagittal rays. It is possible for the meridional section to show no coma and have it show only in the meridional error component of the sagittal fan,[2] but this aberration is normally small.

In addition to plots of the ray error in an evaluation plane, another aberration plot is one that expresses wavefront aberrations as an optical path difference from a spherical wavefront centered about the image point. These OPD plots are particularly useful for applications where the lens must be close to diffraction-limited.

2.6 SUMMARY

Aberration curves provide experienced designers with the information needed to enable them to correct different types of aberrations. Chromatic effects are much more easily classified from aberration curves also. In comparison to spot diagrams and modulation transfer function curves, the types of aberrations can be more easily seen and quantified. In the case of diffraction-limited systems, modulation transfer functions may provide better estimates of system performance.

2.7 REFERENCES

1. R. Kingslake, *Lens Design Fundamentals,* Academic Press, San Diego, 1978, p. 144.
2. F. D. Cruickshank and G. A. Hills, "Use of Optical Aberration Coefficients in Optical Design," *J. Opt. Soc. Am.* **50**:379 (1960).