II. Aspheric surfaces: design and optical advantages

J. J. M. Braat

In designing a lens system the unavoidable lens errors for each of the separate components should ideally be given values and signs, positive or negative, such that the sum, calculated over all the components, is as small as possible. If this is done, the system will be capable of giving an image with a minimum of aberrations. It has long been known that this rather subtle balancing out of the individual lens errors can be considerably simplified by including aspheric surfaces in the design.

Until recently the state of technology did not really permit the introduction of aspheric surfaces into optical systems of the highest grade of accuracy as categorized in part I of this article (see also fig. 1), for a number of practical reasons. The only exceptions were the occasional astronomical telescope in which the skilful technicians had managed to produce aspheric surfaces by the classical approach — the difficult and extremely time-consuming method of repeated local polishing and remeasurement.

Dr Ir J. J. M. Braat is with Philips Research Laboratories, Eindhoven.
In the case of simple spherical surfaces the final quality of a system is determined by a carefully considered choice of the radii of curvature of the surfaces, the distances between the separate components and the lens thicknesses, and also the types of optical glass to be used. The deflection angles of the light rays at the lens surfaces should be small, since otherwise the permissible mechanical tolerances for the maker would no longer be realistic specifications. The total strength of such a purely spherical system is for these reasons often distributed over a large number of components (ten or more is not unusual).

The direct consequence of the admissibility of aspheric surfaces is the rigorous suppression of at least one aberration, sometimes more. The balancing out of the remaining aberrations also requires the presence of fewer components in the system than in the comparable classical case with simple spherical surfaces. In practice the reduction amounts to a factor of two or more. Moreover, in spite of the reduction in the number of components, the final image quality of the system can be better than in the classical design.

The extra freedom offered by the admissibility of aspheric surfaces has thus greatly eased the design problem. Now the difficulties that are still encountered are mainly those associated with fabrication. The main features of this have already been indicated in part I of this article. On their part, the designers obviously have to select from the multiplicity of possible aspheric surfaces, the ones that in steepness (the 'magnitude' of the asphericity) and sensitivity to position (permissible decentering, tilt) will present the fewest problems for the maker. Optical designers and mechanical engineers therefore work in close cooperation, especially when they are designing products that must be suitable for quantity production.

Desirable features

The users of an optical system often require it to have a high 'light-gathering power', i.e. a high aperture angle (expressed as the numerical aperture \( N.A. = n \sin u \), where \( u \) is half the maximum angle of a pencil of image-producing rays on the object side and \( n \) is the refractive index in the object space). Users need this light-gathering power as soon as they need a high resolving power in the object plane. At the same time such users usually require the largest possible 'field'. ('Field' here means the part of the selected object plane corresponding to an adequately faithful image in the image surface — usually planar.) Both requirements can be seen as the desire to transfer the maximum number of bits of information (resolvable points) from object plane to image plane.

Unfortunately an optical system becomes more sensitive to aberrations as the N.A. increases. The aberrations can become so serious that further increase in the N.A. does not give an improvement in the resolving power but makes it worse instead. A trade-off then becomes unavoidable. In television cameras, for example, and also in cine and photographic equipment, designers always try to maximize the N.A., because of the essential importance of the illuminance at the image plane (it is proportional to the square of the N.A.). A small loss of resolving power then has to be taken into the bargain, certainly in conventional optical systems. This compromise can turn out better when aspheric elements are used.

The design of aplanatic systems

A good example of better image correction combined with a high N.A. is the achievement of aplanatism in an optical system through the use of two aspheric surfaces. (An 'aplanatic' system is one completely corrected for spherical aberration and coma.) The theory of this problem was first treated in 1948, by G. D. Wassermann and E. Wolf \[8\]. The design formulae that they derived, based on ray tracing, show that the specification of a single aspheric surface can lead to stigmatic image formation at the optical axis. In other words, the spherical aberration of the rays originating from one object point on the optical axis can be completely suppressed, even for systems that admit pencils with a large aperture angle. A lens with one such aspheric surface is rather thicker towards the edge than the equivalent lens with two spherical surfaces. The associated increase in optical pathlength compensates for the spherical aberration and makes the emergent wavefront perfectly spherical. This means that all rays intersect the axis at the same point, whether they are paraxial rays or marginal rays.

Wassermann and Wolf also showed that two aspheric surfaces are sufficient to extend the stigmatism of the object point on the axis to its immediate environment. This means that the introduction of two aspheric surfaces can ensure that Abbe's sine condition is exactly satisfied, so that there is no coma. In this case a small circular zone perpendicular to the optical axis and with its centre on the optical axis at the object point will give a sharp image in a plane. The validity of this aplanatism is again guaranteed up to high values of the N.A. This was the most important advance resulting from Wasserman and Wolf's theoretical work.


A maximum permissible N.A. is always the quantity that limits the validity of design calculations for optical systems. Such calculations are made at three levels.

The first or 'lowest' level relates to very small values of the N.A., up to about 0.01, or 0.02 at the most. The calculations here are those of 'paraxial' optics; they give results such as the position of the image plane (for a given object plane) and the associated lateral magnification, the focal length, etc. However, they offer no way of determining aberrations, let alone reducing them.

The calculations at the second level are valid for N.A.-values up to almost 0.10. This method, originally devised in 1856 by the German astronomer L. P. von Seidel, can be used to influence five of the different aberrations: spherical aberration, coma, astigmatism, curvature of field and distortion [9]. All these primary aberrations (or Seidel aberrations) correspond to particular wavefront errors, i.e. deviations reckoned from an assumed perfect spherical wavefront (fig. 4). Only if a wavefront remains perfectly spherical will it be able to converge at an ideal point image.

Von Seidel's treatment starts by considering a perfectly spherical lens surface. The contour function $z(y)$ of the lens (its profile curve), a part of a circle, can be represented by:

$$z = r[1 - 1 - (y/r)^4]^{1/2},$$

where $z$ is the coordinate along the optical axis of the lens and $y$ a coordinate perpendicular to the axis (identical to the radial coordi-
point \((0,0)\); it differs from \(r\) at other points. Any aspheric surface with rotational symmetry can then be described by

\[
z = \frac{r}{2} (y/r)^2 + \frac{1}{6} (1 + A)(y/r)^4.
\]

The 'aspheric residue' \((y/r)^4 A\), which corresponds to a departure from a spherical surface (in Von Seidel's approximation), can be used to influence the five primary aberrations.

The calculations at the third level are necessary if the N.A. is greater than about 0.1. Von Seidel's approximation would fail here, essentially because it substitutes the 'third-order' approximation \(u = (1/6)\mu^3\) for \(\sin u\) (\(u\) is the angle of slope of the incident ray). This approximation is completely inadequate for calculating the aberrations to an accuracy typically better than 0.01 \(\lambda\), a value of relevance for lenses of the highest quality. At this highest level the analysis can only be based on an 'exact' ray-path calculation. This is made by considering a pencil of rays leaving an object point and filling the entire entrance pupil of the system; the differences in pathlength between the various optical paths to the image plane are then calculated. These differences in optical pathlength determine the shape of the perturbed wavefront in the image space and hence the aberrations. The description 'exact' refers to the exact application of Snell's law; this must be done at each lens surface, for each light ray that goes through it. 'Shooting' all the rays through a system, and in this 'exact' fashion, was formerly a time-consuming activity. Fortunately, that difficulty has disappeared with the arrival of the computer. This method based on ray-tracing is an analysis of the total system, since it only gives the sum of the individual contributions of the various surfaces to the aberrations found, and hence not the effect of each individual surface — at least not in a readily usable fashion.

The two aspheric surfaces, which have rotational symmetry, can be specified as contour functions \(z(y)\), with \(z\) the axial coordinate and \(r\) the radial coordinate of the surface. They are found here from two simultaneous differential equations of the first order, which are obtained by applying Snell's law consistently to all the reflecting surfaces. A numerical solution of the differential equations provides a number of points on the two desired curves \(z(y)\). The complete curves can be approximated between these calculated points by using an interpolation method.

The original authors, Wassermann and Wolf, working in pre-computer days, used Taylor series for their calculations. We used a series expansion of \(z(y)\) in Chebyshev polynomials, which is much better suited to computer calculations \([10]\), as our interpolation expression. These expansions in Chebyshev polynomials usually converge more rapidly than the corresponding Taylor series. The expansion can be continued until the difference from the original function is no greater than a predetermined value. If this value is taken as 0.001 \(\lambda\), say, we should be able to obtain lenses whose shape ensures the best possible image quality. The only remaining image imperfection is then diffraction at the edge of the lens, provided of course that the surfaces are polished to give the highest degree of optical-quality finish.

Polynomials of the sixth, eighth and even higher degree sometimes contribute appreciably to the series expansion for the interpolation, depending on the N.A. This confirms that Von Seidel's approximation \([9]\) cannot be correct, at least for these high N.A.-values. Von Seidel's approximation 'of the third order' has remained of interest ever since it was first put forward more than a century ago. It provides a simple means of correcting the five primary aberrations, but the N.A.-value must not exceed 0.10.

**Diffraction**

The design analysis described above, for use in the specification of aspheric lens surfaces in high-aperture aplanatic lenses, is purely a matter of geometrical optics; no account can therefore be taken of diffraction effects, although they are present. Fortunately these effects can usually be taken into consideration through the application of a simple two-stage analysis. In the second stage the Fresnel-Kirchhoff diffraction equation \([11]\) is applied to the bounded wavefront as calculated from geometrical optics (the first stage, in which the edge of the lens is simply considered as the abrupt boundary of a refracting zone). In this way the correct distribution of the light intensity in the image plane, which represents the total diffraction image, can be calculated. An example of this is the well-known Airy pattern, consisting of a small bright disc surrounded by alternate dark and light concentric rings; the pattern is formed even when an image of a 'true' point source is produced by a lens that itself could produce an aberration-free image. In this case the diffraction is purely due to the edge of the lens, which behaves as a circular aperture. Such a lens is an example of an 'ideal' element in geometrical optics. The quality of the lens is then diffraction-limited, since the unavoidable diffraction at its edge is the only 'aberration' present.

**Examples of aspheric aplanatic lenses**

The design of lenses for the LaserVision and Compact Disc systems \([12]\) is a good example of practical results that the method of analysis originating from the work of Wassermann and Wolf can provide \([8]\).

A feature of both systems is that the information density on the discs is very high, e.g. 100 Mbit/cm\(^2\) on a Compact Disc, and for read-out they require a microscope objective with an N.A. greater than 0.4 and a useful field diameter varying from at least 100 \(\mu\)m to 500 \(\mu\)m.

E. Hugues and C. Babolat were the first to design a bi-aspheric lens with such a large N.A. \((0.43)\) \([13]\). Their method is somewhat different from the method described in the text: they used a purely...
iterative matching of incident and emergent light rays, with specified conditions for the behaviour in the image space. At the same time a number of specific boundary conditions have to be satisfied. These include tolerances for residual decentring and tilt for each aspheric surface, and these have to be optomechanically acceptable.

The objective that has been designed consists of only one lens, of diffraction-limited quality, with two aspheric surfaces (fig. 5). In the Table in fig. 5 a number of calculated contour parameters for the two desired surfaces are shown. The dimensional differences from perfectly spherical surfaces are not very large — the asphericity is a few tens of μm at most, but their significance is more obvious when they are expressed in wavelengths. This lens (we often call it a 'bi-aspheric'), gives an image free from coma and spherical aberration. Slight astigmatism and field curvature are present, near the edge of the useful field.

In a conventional design with spherical surfaces some four lenses would be necessary for the same image quality (fig. 5b). Accurate alignment is easier with a single lens than with four, and therefore preferable. (Alignment is the bringing into coincidence of the optical axes of the refracting surfaces. Faulty alignment, the main cause of coma, must be avoided. Coma is the very troublesome aberration that occurs transverse to and to one side of the optical axis and distorts an image point to a 'comet's tail', see fig. 4.)

More recent design work for the LaserVision and Compact Disc systems showed that a single lens with one aspheric surface can sometimes be sufficient (fig. 6). Such lenses cannot of course satisfy Abbe's sine condition exactly. Fortunately the discrepancies remain small provided that an 'equivalent' lens with spherical surfaces gives little coma and in addition the desired field radius in the image space is no larger than 50 to 100 μm. (The word 'equivalent' means that the lens thicknesses and the paraxial curvatures of both lenses are the same.)

---

**Fig. 5.** a) For read-out of the information on an optical disc (DD), in vertical-radial section (shown in part), the microscope objective Ob, with the contour parameters given in the table, was designed as a bi-aspheric aplanat. RAS₁,₂ aspheric refracting surfaces (blue), of diameter 7.32 mm and 4.14 mm respectively. b) A conventional objective for read-out, consisting of four spherical lenses. It is clear that adequate reduction of errors in the alignment of the eight refracting surfaces — to keep the image free of coma — is less simple than for the lens in (a). The width (w) of the pits is about 0.6 μm; the 'pitch' (p) is 1.7 μm.

---

Twenty years before a general theory for the calculation of aspheric surfaces was published, B. Schmidt's correction plate appeared. This was the first really convincing application of an aspheric refracting surface (a surface of revolution, but not obtained from a conic section). This optical element corrected the spherical aberration of a concave spherical mirror. Originally applied in astronomical telescopes, it improves the resolution and the field angle. The device that we now call a Schmidt corrector is also used in projection television.

In the 17th and 18th centuries some theoretical knowledge already existed of the advantages that certain aspheric surfaces (paraboloids, etc., all of them surfaces of revolution obtained from conic sections) could offer in the correction of aberrations, but the manufacturing methods were so poor that in general no improve...
ment in the image quality could be obtained. There was one curious exception, however: the microscopes of Antoni van Leeuwenhoek (1632-1723). Recent measurements have shown that in one of the few surviving microscopes made by him the tiny lens is aspheric. We now know that the resolving power of about 1.3 \( \mu \text{m} \) — unusually good for its day — is due to its shape. The little lens looks like a small glass ‘bead’ about 1 mm thick. Just how Van Leeuwenhoek managed to make it is no longer known. It is almost certain that the lens was blown, and not ground. An empirically based theory (title photograph) has been put forward to explain how Van Leeuwenhoek performed the glassblowing.

Designing for minimal wavefront errors in a large field

Photographic objectives, with their wide field of view, form an excellent example of optical systems whose design can be considerably improved by including aspheric lenses. This should be done not in the same way as described in the previous subsection, but by a direct optimization procedure. Such a procedure is based on general mathematical methods and allows a number of functions — for which wavefront deviations (= errors) have been chosen — to be minimized; this minimization is obtained by adapting the variables of the system. The solution obtained provides no information about the ‘internals’; the relationship between the various image-forming properties and the individual refracting surfaces cannot be recovered, for example. Our earlier approach, making a design analysis that relates to certain optical concepts such as axial stigmatism and aplanatism, seems to be particularly difficult to carry out in the present case where a large field (or aperture angle) and high, although not completely diffraction-limited, image quality are required at the same time.

The method that we have followed for such designs of aspheric elements is a global minimization, a procedure that modifies the coefficients of a polynomial \( z(y) \). This contour function, a series expansion of orthogonal polynomials in \( y \), represents the unknown surface.

For the orthogonal polynomials in this global minimization we again decided to use Chebyshev polynomials; these are orthogonal in the interval \([-1, +1\]). By specifying a maximum optical diameter of a lens surface the distance from the axis of an arbitrary point on
the lens surface can be expressed as a *fraction* of this diameter, and can thus be 'scaled' into the interval \([-1, +1]\). In the global minimization we use a matrix of derivatives that describe the effect of the different variables on the wavefront errors. To 'condition' this matrix appropriately in making a lens surface aspheric, i.e. to ensure that the dependence between the different rows of derivatives is as small as possible, it is highly advisable to choose a series expansion in terms of orthogonal polynomials. This is because it usually gives better and faster *convergence* of the optimization. Also, the numerical 'noise' of the aspheric surfaces thus found is less than if the method is based on a power-series expansion.

The coefficients of the various polynomials function together in a minimization algorithm as a part of the group of variables (as well as other variables, such as distances between surfaces and curvatures of surfaces). This algorithm minimizes a number of functions simultaneously, by finding a suitable set of values for the variables, the desired solution. Usually the minimized functions are the wavefront errors — corresponding to aberrations — of pencils of rays leaving points in the object plane.

The design calculations in such cases thus lead 'automatically' to a better appreciation of the residual wavefront errors as functions of various quantities that are characteristic of the desired optical system. As an example *fig. 7* shows three cases for which the angle of slope of an incident ray and the field radius in the image space are the characteristic optical quantities that are varied in the calculations. The wavefront errors that arise, which correspond to optical path differences below the 'Rayleigh limit' \( \frac{\lambda}{2} \), indicate that the lens design is diffraction-limited. The undesirable effect of possible variations in construction parameters such as the thickness of the system or the alignment of refracting surfaces (in centring and slope) can also be rapidly calculated as extra wavefront errors (*fig. 8*). (Wavefront errors here mean residual local deviations with respect to a perfectly spherical wavefront.)

The optimization procedure suffers from an imperfection: the minimum or maximum it finds is *local*, dependent on the particular starting point selected. In theory many other minima can exist, some of these perhaps giving better image correction; this procedure cannot provide the extra information as to the relative nature or otherwise of a minimum in a simple way, however. Although the computer can give considerable help in optical design, we see here nevertheless that the experience of the user and the sound intuition of the designer are indispensable.

*Fig. 9* shows an example of an optical system that was designed in the way just described. Four of the six lens surfaces are aspheric. This system can be used in projection television. A system with conventional lenses and the same degree of image correction would need seven or eight elements instead of three. This example shows clearly the great advantage of the introduction of aspheric surfaces into lenses: fewer elements giving simpler mounting and alignment, at lower cost.

---

![Wavefront error](image)

*Fig. 8*. Wavefront error \( w_{cp} \), expressed in terms of the mean wavelength \( \lambda \) of the light to be used, as a function of the sine of the angle of slope \( u \) of the incident ray. These three curves apply for rays in the meridional plane, and were also calculated for evaluation of the design for the thick lens referred to in *fig. 7*. Curve I: the effect of increasing the lens thickness by 100 \( \mu \)m. Curve II: the effect of a 25-\( \mu \)m parallel displacement of the optical axis of the front surface of the lens with respect to the axis of the rear surface of the lens. Curve III: the effect of tilting the front surface of the lens. The change in angle is assumed to be 1 mrad with respect to the optical axis. These curves help to give an idea of the extent to which variations in the design parameters lead to an unacceptable total wavefront error, so that the design would no longer be of diffraction-limited quality. In the sagittal plane there is no wavefront error of any significance.

![Optical system](image)

*Fig. 9*. An aspheric lens system designed for projection television. \( RAS_{1,2,3,4} \) aspheric refracting surfaces. \( RS \) spherical refracting surface. \( T \) picture tube. \( S \) projection screen. The distance between the system and \( S \) is about 1.5 \( m \). \( I_{po} \) image points of the object points \( P \) and \( Q \). Magnification 10 times. The lens system was designed by the method based on a global minimization of the wavefront errors (see text).

[13] J. van Zuylen, The microscopes of Antoni van Leeuwenhoek, J. Microsc. 121, 309-328, 1981. The authors would like to thank Dr Van Zuylen for making available the photograph used as the title illustration.