Peter Kalló *

Some Fundamental Relations in the Design of Optical Systems

This paper aims at the systematization of literary information, as well as experiences collected through calculations, suitable for generalization and related to optical system design, with respect to responsibilities of the optical designer. Our interest is focussed, above all, on the tolerances of image formation errors, system selection aspects, and the findings when using third order analysis.

An exact or direct [1] method of optical design has not been developed, as yet, so attention has been directed here, for quite some time now, to the eventual ways and means whereby the design problems might be best solved. Methods and techniques known form literature are to be systematized by starting from design logics as they could obviously this way be derived one from the other.

The subject of optical design is to dimension an optical system quantitatively and qualitatively providing for a transformation specified by the objectives. According to Clark Maxwell [2], an optical system ensures ideal transformation, if each point of the object plane normal to the optical axis is sharply or definitely transformed without distortion (point to point) into the image plane similarly perpendicular to the optical xais. In other words (E. Abbe, 1872), see [3], if the object and image points are each other's collinear transforms.

No real optical system can provide for an ideal transformation In practice, professionals have found a long time ago, and can determine today with good approximation, the extent to which real and ideal transformations may differ without making the solution of the transformation questionable. Transformation quality is known to be influenced, in addition to the deficiencies of the optical system employed for transformation, by diffraction due to the wave character of the light, described first by Airy [4] through the energy distribution E of the object point transformed:

$$E = k\varrho^4\left[1-\frac{1}{2}\left(\frac{m}{2}\right)^2+\frac{1}{3}\left(\frac{m^2}{2^2\,2!}\right)^2-\frac{1}{4}\left(\frac{m^3}{2^3\,3!}\right)^2+\right.$$
$$\left.+\frac{1}{5}\left(\frac{m^4}{2^4\,4!}\right)^2-\ldots+\ldots\right], \qquad (1)$$

*) Invited Lecturer, Department of Precision Mechanics nd Optics, Budapest Technical University.

where k is a constant characteristic of the energy transfer of the transformation system, ϱ is the half aperture diameter of the system,

$$m = \frac{2\pi\varrho}{\lambda}\sin\alpha, \qquad (2)$$

in which λ is the wavelength, α is the angle in cluded by the optical axis and the straight line connecting the exit pupil centre and the image point studied.

On the above basis, the permissible deviation of the real transformation from ideal due, anyway, to diffraction (from the point $\rightarrow$ Airy disc) hasbeen defined by Lord Rayleigh [5], [17] intuitively, which may be considered now as the practical criterion of the quality of real transformation. Accordingly, the image point thus transformed is "appreciably" perfect if, within the optical path nl, the actual wave front deviates from the associated ideal by a maximum of $\lambda/4$. The term "appreciable" is used in the sense of "resoluted" which exists, in the case of two image points, when the intensity maxima are displaced relatively by at least one ring [1]. In the case of a circular aperture, the resolution limit value of Z will be

$$Z = \frac{0.61\lambda}{n\sin\theta}, \qquad (3)$$

where $n\sin\theta$ indicates the numerical aperture size.

In 1942, Banti [5] conducted experiments contentually analogous with the Rayleigh limit, by using $\lambda = 16$ cm radio waves. For the sake of completeness it should be noted here that, according to Conrady, the resolution capacity would not change until the double of the Rayleigh limit, but the contrast should be impaired [5].

From among the aberrations of third order, which are commonly used in practice (though they impair the transformation quality), only the tolerance of spherical aberration, may be put in the form presented in [17] as they represent an orientation benchmark starting from the Rayleigh limit only if small aperture angles are considered, although suitable for the same in the case of larger apertures as well. (This is probably due to the fact that in every source of literature the wave front aberration obtained on the basis of the trigonometrically calculated values of third order image errors is always indicated as the function of the spherical aberration (see [5] and [12]).) In order to overcome this difficulty, ARGENTIERI [5] suggested to maintain the extra-axial aberration value within the order of magnitude of the tolerance obtained for the spherical aberration from the relevant Rayleigh limit. In our opinion, this may be sometimes an overstrict and too general requirement. According to our experiences the quality of transformation could never be objected if the values of third order image formation errors did not exceed those within the tolerance below [6]:

Marginal and zonal spherical aberration:

$$\frac{\lambda}{n'\sin^2\theta'}, \quad \text{and} \quad \frac{6\lambda}{n'\sin^2\theta'}. \tag{4}$$

Offense against the sine condition:

$$\frac{\lambda}{2H'n'\sin\theta'}. \tag{5}$$

Astigmatic difference:

$$\frac{|t-s|}{2} = \frac{0.025}{n'\sin\theta'}\ [\text{mm}]. \tag{6}$$

Axial chromatism:

$$\frac{\lambda}{n'\sin^2\theta'}. \tag{7}$$

Chromatic difference of magnification:

$$\frac{\lambda}{2H'n'\sin\theta'}, \tag{8}$$

where λ is the wavelength of the (mean) light employed,

n' is the refractivity of the image area,

$\sin\theta'$ is the sine of the aperture angle in the image area,

H' indicates the half picture height.

A method for the rapid evaluation of transformation quality was suggested by Conrady [6]. He has classified the pictures intended for visual presentation as excellent, fair, and poor, if the diameters of the aberrational discs produced arround their ideal image points, due to image formation errors, amounted to ~0.025, ~0.1, and ~0.25 mm, respectively.

To sum up, and review what has been stated above from practical design aspects, it may be said that dimensioning conclusions can be arrived at from the requirements set to image quality with respect to either the wave front or the geometrical aberrations. Wave front aberration testing has not led so far to conclusions directly adaptable for practical design in a simple way, but if this will be achieved then, in our opinion, it will render efficient assistance above all for the implementation of fine corrections. Fundamentally, therefore, our further investigations will cover only the variation of geometric aberrations, so we shall deal only with the seven third-order aberrations.

It is remarkable even from the philosophical aspects that, among the third-order aberrations related in an exact form to a realistic optical system, only the absence of spherical aberrations (Herschel condition)

$$\frac{n\sin^2\frac{\vartheta}{2}}{n'\sin^2\frac{\vartheta'}{2}} = \text{const}, \tag{9}$$

the Abbe sine condition usually associated with a coma-free state

$$\left.\frac{n\sin\vartheta}{n'\sin\vartheta'}\right|_{S_1\neq\infty} \rightarrow \left.\frac{nh}{n'\sin\vartheta'}\right|_{S_1=\infty} = \text{const}, \tag{10}$$

and the lack of distortion

$$\frac{n\tan\alpha}{n'\tan\alpha'} = \text{const} \tag{11}$$

are known in the form of expressions [3] where, in addition to the hitherto introduced symbols,

ϑ, ϑ' represent the angle included by the optical axis and the conjugate rays crossing the object and image points, respectively, along this optical axis,

h is the distance of the incident ray at the first surface of the optical system, parallel to the optical axis, therefrom, and

α, α' are the angles included by the principal ray and the optical axis in the object and image areas, respectively.

In the case of other third-order image formation errors (colour aberrations and field curvature), nothing but studies on the location of the rays creating the image point in its narrow range may lead to conclusions (spot diagram).

Specification of the tolerances of image aberrations indicates, from image formation aspects, the necessary and sufficient requirements set to the optical system. In given cases often simplification is feasible without jeopardizing the solution of the problem, that is, the number andor the strict character of the requirements may be reduced.

It follows logically that a given number of requirements (image aberrations to be corrected) can be satisfied by an optical system having a number of degrees of freedom (free degrees) at least as many as the number of requirements specified. Application of the above condition system is impeded only by the fact that, so far, the data representing the free degrees of optical systems could not be defined unequivocally, that is, as verified by practice. This statement can be confirmed by the following examples.

According to certain authors [5], [13], the free degree number of an optical system is given by the integrated number of its curvature radii and air gaps, which is true in the case of triplets, but cannot be generalized. A system consisting of any number but only positive lenses cannot be corrected with respect to spherical aberration, just like a pupil-asymmetric system for distortion or a two-part cemented achromate for, again, spherical aberration if the refractivity difference of its components is too small, etc. In our opinion, large relative aperture or large angle of field type systems require for the compensation of higher order aberrations the provision of "+ free degrees" exceeding the usual extent which, however, is not discussed at all in literature. To sum up, the present state of the art has no theory elaborated on how an optical system structure could optimally solve a given problem. Naturally, in simple cases the knowledge of existing constructions will render sufficient information for the solution of the problem, but these instances ought to be considered today as routine design duties (condensers, aplanatical achromates, oculars, etc.).

As a matter of fact, the form of optical design process developed so far renders a minimum of bases for the selection of the optimum system. This, like the first design step, is expressed by Nefedov [7] as "when selecting from existing systems on the basis of approximating calculations and various alternatives, or when a new system is adopted, such a version must be developed whereby the subsequent calculation can be expected to, or may lead to success".

According to the still valid statement above all on system selection, the classical approach of optical design is rather an art than a technique: at the beginning of a design both practice and intuition are of critical importance [8]. Correct selection of the system is often decided upon by a good drawing, or on the basis of the designer's experiences [9]. There are, however, certain general regularities which, if adhered to, can make us avoid a number of design dead-ends.

Thus in our opinion, if the subsequent steps of design are reckoned with in advance (third order approximation, correction, etc.), system selection is best performed by surveying the below parameters characteristic of the individual systems, in a tabulated or graphic form.

1. The focal distances of the elements contained by the system should be selected as great as possible since, in the case, the absolute value of the image aberrations will be sufficiently small, the aberration balance providing for correction can be achieved in a stable way, the third-order error equations lead to a much better approximation and, owing to the increased radii of curvature, a system easily adopted for mass production and less sensitive to assembly will be obtained.

2. The distance between the lenses should be selected as small as possible since, thereby, the coma hazard can be reduced. The elements achromatized in themselves are not necessarily needed and, due to the expected lesser height of the incident rays, the third-order approximations will better approachreality.

3. Easy to realize f/number values should be selected.

4. The pupil should be located in its natural position providing, in advance, for a possible correction of any distortion.

5. The feasibility of correcting the Petzval sum and the colour aberration is then to be checked upon, by using the data of glasses exhibiting no peak characteristics (n, ν). In other words, the possibility of eventual glass replacement in the course of correction is provided for [15], [16].

According to the above aspects the possibly most advantageous alternative(s) is (are) selected, the ray courses are indicated in the drawing, just like the partial focus distances, incident ray heights, aperture stop figures, lens distances, object and image distance measures, magnifications, refractivities, Abbe indices and the pupil as well as dimension data, whereby the initial characteristics required for the third--order calculations can then be obtained.

As for the results of the third-order calculations and their adaptability, it seems to be necessary to make some further comments.

"It is well known that Seidel has underestimated his own theory, and regarded it as inapplicable in the practice of dimensioning optical systems" [10]. Nevertheless, the Seidel theory with certain modifications is "still the fundamental aid of optical design, rendering assistance for studies on the possibilities offered by optical systems" [11], and "there is no other method that could lead so rapidly to a comprehensive discussion" [12].

As an alternative of the Seidel theory (Conrady, Coddington-Taylor, Berek, Argentieri, Flügge, etc.), information much more difficult or expensive to obtain can be gained only, if it is taken into consideration that, when evaluating the results of the third -order approximation,

(1) the image aberrations of the optical systems calculated by third-order approximation, in the revelant range, develop almost parallelly to the variation of the trigonometrically calculated image formation errors [1], [14], [15],

(2) the higher order image aberrations of the system elements are small (9), if their third-order image formation errors are similarly small,

(3) there is usually a possibility existing for satisfactory fine correction by reducing the actual error through third-order approximation, and with the variation of the image aberrations trigonometrically calculated taken into account to the desired level, by the use of high-speed iteration, [12], [14], [15], and

(4) the best solution is to graphically illustrate the image aberration values calculated by third--order approximation in the function of the form of the lens; these graphs will then make possible the selection of the solution range providing for the optimum compensation of image aberrations [1], [12], [14], [15].

It will have to be noted here that the calculation and critical analysis of third-order image aberrations will enable us, even after the review of a relatively small number (5–15) of systems, to realize selfevidently on the basis of the equation coefficients whether it is worthy to continue that design, or is it more expedient to modify the initial system, and start working all over again expecting a better approach of the original target (in the Coddington–Taylor parameter system, for example, a source of difficulties discovered in a later phase of design may be when the absolute value of the form of lenses representing the system exceeds 2–3, and that of the constants in the various image aberration equations the figure 5–10).

With the above principles reckoned with, independently of whether the fine correction is accomplished on the basis of third-order approximative equations or by means of an electronic computer, the successful result is only a matter of patience and systematic work [12].

Our conclusions may be summarized as follows [8]:

1. Application of an electronic computer will accelerate the design process, and open up a new horizon for the designer. However, the computer cannot replace the designer but, on the contrary, will set stricter requirements to his work.

2. In optical design the role played by the computers will continue to increase, although no fully automated design technique can be expected to develop in the near future.

3. As for the advancement of the optical design methods, a better understanding of geometrical optics and lens characteristics seems to be a fundamental precondition (see, for example [14]), or else optical design would remain almost entirely nothing but successive empirical approximation.

4. Everybody starting the design of optical systems will be compelled to collect his own direct experiences, while the selection of the required methods will be influenced often randomly, or by personal bias [12].

Relations élémentaires utilisées dans la construction des systèmes optiques

On se propose, dans cet article, de systématiser les informations bibliographiques ainsi que les expériences acquises au cours des calculs, sous une forme se prêtant aux généralisations et correspondant en même temps aux besoins de la construction de systèmes optiques. L'intérêt de l'auteur s'est concentré avant tout sur les tolérances des erreurs dans la formation de l'image et sur le problème du choix des types des ensembles en utilisant l'analyse du troisième ordre.

Некоторые фундаментальные зависимости в проектировании оптических систем

Цель статьи — систематизировать литературные данные и опыт, приобретенный при расчетах, в виде, пригодном для обобщений и отображающем потребности проектировщиков оптических систем. Внимание автора сосредоточено в первую очередь на допуске погрешностей при формировании образа и вопросах выбора типа системы при использовании анализа третьего порядка.

References

[1] Smith W. J., Modern Optical Engineering (The Design of Optical Systems), New York 1966.

[2] Emsley H. H., Aberration of Thin Lenses, London 1956.

[3] Hardy A. C., Perrin F. H., The Principles of Optics, New York 1932.

[4] Wolfe W. L., Handbook of Military Infrared Technology, Washington 1965.

[5] Argentieri D., Ottica Industriale, Milano 1954.

[6] Jacobs D. H., Fundamentals of Optical Engineering, New York–London 1943.

[7] Nefedov B. J., Metody resheniya zadatch po vytchislitelnoy optike, Moscow–Leningrad 1966.

[8] Stavroudis O. N., Automatic Optical Design. Advances in Computers 1967, p. 227–255.

[9] Gaj M., *Über ein quasisymmetrisches Spiegelobjektiv.*, Optik 1959, p. 234–244.

[10] Romanov D. D., *Raschet zerkalno-linzovykh sistem po formulam aberacy tretevo poryadka*, IZVUZ Geodesya i aerofotosyemka 1962, p. 123–128.

[11] Rusinov M. M., Tekhnicheskaya optika, Moscow–Leningrad 1961.

[12] HAVLIČEK F. J., Einführung in das Korrigieren optischer Systeme, Stuttgart 1960.
[13] MAJOROS S., *The new principal determinant of optotechnics*, (In Hungarian), Kép- és Hangtechnika 1961, p. 180–185.
[14] UJVÁRY I., *Approximative method to design two-part cemented simplets through modification of the Coddington-Taylor equations* (In Hungarian), Kép- és Hangtechnika 1962, p. 129–134.
[15] KALLÓ P., *A method to design a triplet* (In Hungarian), Kép- és Hangtechnika 1970, p. 15–23.
[16] KALLÓ P., *Symmetrical erecting lens system* (In Hungarian), Kép- és Hangtechnika 1972, p. 135–137.
[17] PŁUŻANSKI M., *Obliczeniowe kryteria oceny jakości układów optycznych*, Optyka 1972, p. 28–34.

Received, November 28, 1972